DISS. ETH NO. 29900

Geospatial and disease dynamics modelling against infectious diseases: antimicrobial resistance and COVID-19

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

Presented by

CHENG ZHAO

MSc Environmental Sciences, University of Zurich, Switzerland

born on 20.05.1995

Accepted on the recommendation of:

Dr. Thomas P. Van Boeckel
Dr. Shengjie Lai
Prof. Dr. Ulf Magnusson
Prof. Dr. Robert Finger

2024

Acknowledgements

All of the works presented in this thesis are a result of collaborative effort. First and foremost, I would like to thank Thom for supervising me through this journey of discovering and communicating science. I have benefited a lot from his insightful feedback, mentorship, and encouragement at every stage of my research. He has helped me to grow not only academically, but also personally and professionally. I also want to thank him for his support during the second half of my PhD, when I decided to expand my family. I have been given the privilege to work flexibly to achieve a balance between work and family. This PhD would not have been possible without his support. I would like to also thank my second advisor Lai for providing important feedback on the work and helping to establish external collaborations. I thank Ulf and Robert for being part of my PhD committee. I also thank the Branco Weiss Fellowship and the Swiss National Science Foundation for their financial support.

I would like to thank my colleagues and collaborators who have shared their time and knowledge with me. Many thanks especially to Nico, who shared the same office and this whole PhD journey with me. We shared each other's happy moments and supported each other through hard times. I thank Ranya, Alex, Joao, Katie, Laura, Giulia and Riccardo for all the interesting discussions, enjoyable lunches, and fun parties together.

I want to thank my friends and family who brightened up my private life during the PhD. Many thanks to my Chinese friends here in Switzerland for celebrating festivals together with delicious traditional cuisine. I have my biggest thanks to my husband Zhong who supported me unconditionally in every possible way, and to my daughter Qingrui who has brought us so much joy in life.

Abstract

Infectious diseases are a major public health burden in low- and middle-income countries (LMICs) where 84% of the world's population live. Antimicrobial treatments, which once revolutionized the battle against infectious diseases, are now being threatened by the increasing antimicrobial resistance (AMR) among pathogens. Globally, the majority (73%) of antimicrobials are consumed by food animals, making them an important reservoir of AMR. Increasing evidence shows that AMR determinants and pathogens can be passed from animals to humans and cause infections that are difficult – if not impossible – to treat. The crisis of AMR is a slow-moving "silent pandemic" that spans decades. In comparison, the COVID-19 pandemic was an acute health crisis with sudden and far-reaching impacts. Both cases can lead to substantial public health consequences if no actions are taken, and this thesis uses geospatial and disease dynamics modelling to guide interventions against these two health crises.

Three main objectives are addressed in this thesis. First, we map trends of AMR in food animals using geospatial models, in combination with a literature review of point-prevalence surveys (PPS) of AMR prevalence. PPS represent an alternative source of information on AMR, in the absence of systematic AMR surveillance systems in food animals in most LMICs. Chapter 2.1 focuses on China, and expands the previous data collection efforts to also include PPS conducted in the Chinese language. We predict that eastern China as well as part of central and northwestern China are hotspots of AMR. Chapter 3 combines PPS conducted in European countries with existing country-level systematic AMR surveillance data, to spatially disaggregate AMR prevalence at 10x10 km resolution in Europe. We show that major geographic heterogeneities in AMR exist within countries. Finally, Chapter 4 further refines maps based on a summary metric of AMR across drug-pathogen combinations used in the previous two chapters, and produces maps of AMR for 7 individual antimicrobial classes.

Second, looking into the future, we explore at what locations and for which antimicrobials classes should AMR surveillance efforts be targeted, given limited resources available in LMICs. In Chapter 2.1, we show that future surveillance efforts of AMR in China could be targeted at western and southern China, using an approach that aims at minimizing the uncertainty of predicted AMR trends. We show that such an approach is more effective than other approaches that may be more commonly applied in practice, such as distributing samples randomly across geographic areas, or evenly between administrative units. In Chapter 4, we predict the antimicrobials with highest probability of its resistance exceeding 50% in the future. We show that in Africa and South America, future surveillance could focus on penicillins or tetracyclines; while in Asia, the target antimicrobial classes are penicillins or sulphonamides.

Third, during the COVID-19 pandemic, we explored how local-scale modelling of disease dynamics can play a role in the operational aspect of disease management. An online platform *icumonitoring.ch* was built to display hospital surveillance data of COVID-19, and to provide projections of COVID-19 occupancy in the intensive care units 3- and 7-days ahead. We show

that an integration of local surveillance data with disease dynamics modelling can facilitate decision making during a pandemic.

The findings and their policy implications are discussed, along with reflections on the place of science and epidemic models in policy making.

Zusammenfassung

Infektionskrankheiten stellen in Ländern mit niedrigem und mittlerem Einkommen, in denen 84 % der Weltbevölkerung leben, eine große Belastung für die öffentliche Gesundheit dar. Antimikrobielle Behandlungen, die einst den Kampf gegen Infektionskrankheiten revolutionierten, sind nun durch die zunehmende antimikrobielle Resistenz (AMR) von Krankheitserregern bedroht. Weltweit wird die Mehrheit (73 %) der antimikrobiellen Mittel von Nutztieren verzehrt, was sie zu einem wichtigen AMR-Reservoir macht. Es gibt immer mehr Belege dafür, dass AMR-Determinanten und -Pathogene vom Tier auf den Menschen übertragen werden können und Infektionen verursachen, die schwer - wenn nicht gar unmöglich - zu behandeln sind. Die AMR-Krise ist eine langsam voranschreitende "stille Pandemie", die sich über Jahrzehnte erstreckt. Im Vergleich dazu war die COVID-19-Pandemie eine akute Gesundheitskrise mit plötzlichen und weitreichenden Auswirkungen. In beiden Fällen kann es zu erheblichen Folgen für die öffentliche Gesundheit kommen, wenn keine Maßnahmen ergriffen werden. In dieser Arbeit wird die Modellierung von Geodaten und Krankheitsdynamik eingesetzt, um Maßnahmen gegen diese beiden Gesundheitskrisen anzuleiten.

In dieser Arbeit werden drei Hauptziele verfolgt. Erstens kartieren wir Trends der AMR bei Lebensmitteltieren mit Hilfe von räumlichen Modellen in Kombination mit einer Literaturübersicht über punktuelle Prävalenzerhebungen (PPS) der AMR-Prävalenz. PPS stellen eine alternative Informationsquelle für AMR dar, da es in den meisten LMICs keine systematischen AMR-Überwachungssysteme für Nutztiere gibt. Kapitel 2.1 konzentriert sich auf China und erweitert die bisherigen Datenerhebungen, um auch chinesischsprachige PPS einzubeziehen. Wir gehen davon aus, dass Ostchina sowie Teile Zentral- und Nordwestchinas Hotspots für AMR sind. Kapitel 3 kombiniert die in europäischen Ländern durchgeführten PPS mit bestehenden systematischen AMR-Überwachungsdaten auf Länderebene, um die AMR-Prävalenz in Europa mit einer Auflösung von 10x10 km räumlich aufzuschlüsseln. Wir zeigen, dass es innerhalb der Länder große geografische Heterogenitäten bei AMR gibt. Kapitel 4 schließlich verfeinert die Karten auf der Grundlage einer zusammenfassenden Metrik der AMR über Arzneimittel-Erreger-Kombinationen, die in den beiden vorangegangenen Kapiteln verwendet wurden, und erstellt Karten der AMR für sieben einzelne antimikrobielle Klassen.

Zweitens untersuchen wir mit Blick auf die Zukunft, an welchen Orten und für welche Antibiotikaklassen die AMR-Überwachungsbemühungen angesichts der begrenzten Ressourcen in LMICs ausgerichtet werden sollten. In Kapitel 2.1 zeigen wir, dass künftige AMR-Überwachungsmaßnahmen in China auf West- und Südchina ausgerichtet werden

könnten, wobei ein Ansatz verwendet wird, der darauf abzielt, die Unsicherheit der vorhergesagten AMR-Trends zu minimieren. Wir zeigen, dass ein solcher Ansatz effektiver ist als andere Ansätze, die in der Praxis häufiger angewandt werden, wie z. B. die zufällige Verteilung von Proben über geografische Gebiete oder gleichmäßig zwischen Verwaltungseinheiten. In Kapitel 4 sagen wir die antimikrobiellen Mittel voraus, bei denen die Wahrscheinlichkeit, dass ihre Resistenz in Zukunft 50 % übersteigt, am höchsten ist. Wir zeigen, dass sich die künftige Überwachung in Afrika und Südamerika auf Penicilline oder Tetracycline konzentrieren könnte, während in Asien Penicilline oder Sulfonamide die Zielantimikrobienklassen sind.

Drittens untersuchten wir während der COVID-19-Pandemie, wie die Modellierung der Krankheitsdynamik auf lokaler Ebene eine Rolle für den operativen Aspekt des Krankheitsmanagements spielen kann. Es wurde eine Online-Plattform icumonitoring.ch eingerichtet, um sowohl die Überwachungsdaten von COVID-19 im Krankenhaus anzuzeigen als auch Prognosen über die COVID-19-Belegung der Intensivstationen für drei und sieben Tage im Voraus zu erstellen. Wir zeigen, dass die Integration von lokalen Überwachungsdaten mit der Modellierung der Krankheitsdynamik die Entscheidungsfindung während einer Pandemie erleichtern kann.

Die Ergebnisse und ihre Auswirkungen auf die Politik werden erörtert, zusammen mit Überlegungen zum Stellenwert von Wissenschaft und Epidemiemodellen in der Politikgestaltung.

Table of Contents

Chapter 1

Introduction

1.1 Burden of infectious diseases

Infectious diseases are a major burden to public health, and leading causes of deaths in low- and middle-income countries (LMICs), where 84% of the world's population live[1]. In 2019, the top 10 causes of deaths in low-income countries included 5 associated with infectious diseases (Figure 1A), while in middle-income countries, 2 were associated with infectious diseases (Figure 1B)[2].

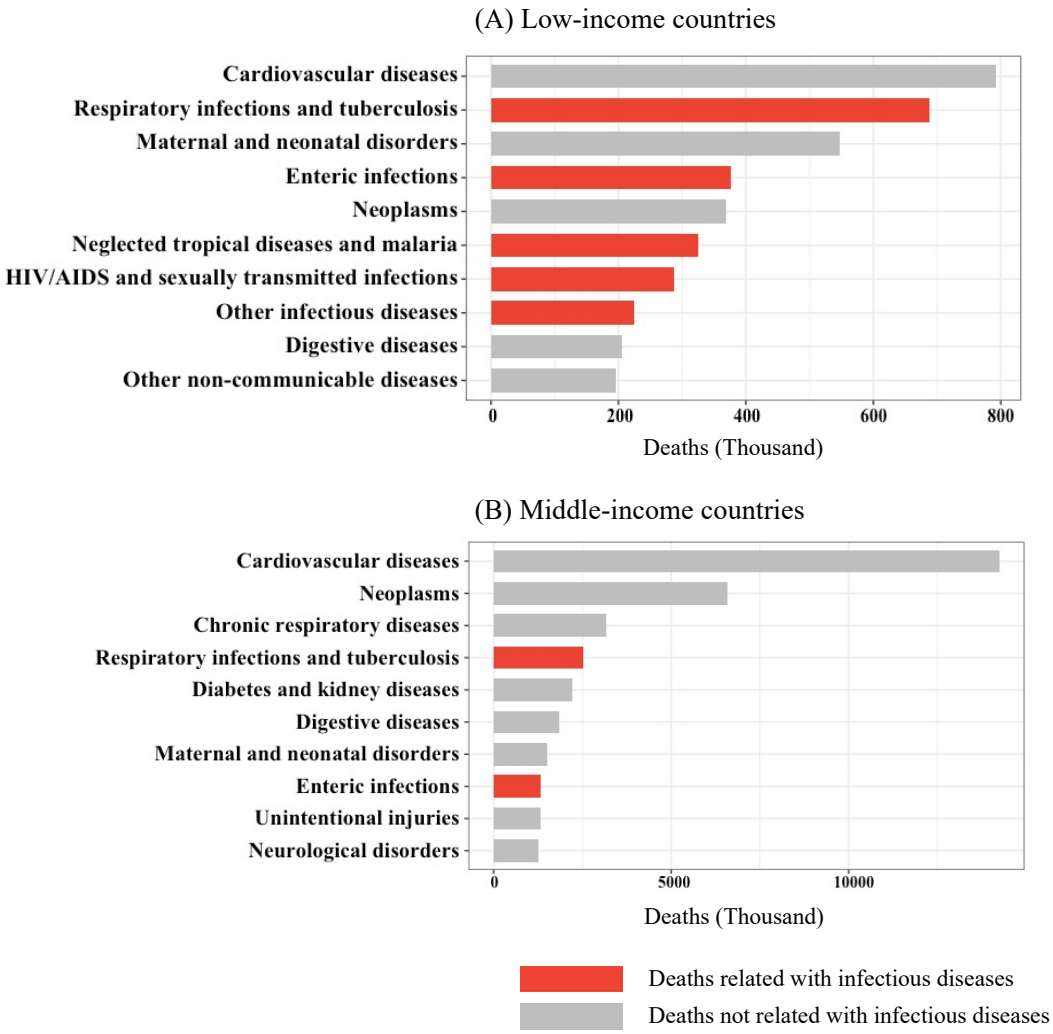


Figure 1. Top 10 causes of deaths in low-income countries (A) and middle-income countries (B) in 2019. Causes of death that are related with infectious diseases are coloured in red. Data source: IHME, Global Burden of Disease (2019).

The association between income and burden of infectious diseases is complex and multifaceted. First, limited resources in LMICs leads to weaker health systems compared with high-income countries: nearly 2 billion people in LMICs do not have access to basic medicines[3], and the coverage of essential vaccines such as the measles vaccine is much lower in low-income countries (66%) than in high-income countries (93%). Second, sanitation facilities are comparatively less developed in LMICs, thereby facilitating disease transmission. In 2015, the World Health Organization published a report on water, sanitation and hygiene in health

facilities across 54 LMICs, and showed that 35% healthcare facilities did not have water and soap for handwashing[4]. Third, insufficient nutrition and education also increases the population's susceptibility to infectious diseases[5]. The percentage of stunted children under the age of 5 in LMICs is 24.3%, which is 5 times higher than in high-income countries (4.0%)[6]. Fourth, the higher prevalence of extensive and small-sized backyard farming in LMICs leads to comparatively more frequent contact between humans and animals than in high-income countries, resulting in a higher risk of zoonotic pathogen spillover[7]. In 2014, work by Adamopoulos et al. showed that in the 20% richest countries, nearly 70% of farms had sizes bigger than 20 hectares, while in the 20% poorest countries, nearly 70% of farms had sizes smaller than 2 hectares, which are often backyard farms[8].

Zoonotic diseases account for approximately 26% of the burden of infectious diseases in humans in low-income countries[9], and are the major contributor to emerging infectious diseases (60.3% between 1940 and 2004)[10]. In 2010, pathogens in animal products were estimated to have caused 168 Disability Adjusted Life Years (DALYs) per 100,000 population, including non-typhoidal *Salmonella enterica*, *Taenia solium*, and *Campylobacter* spp.[11]

Infectious diseases are not only a burden for human health, but also for animal health and the agricultural economy. A striking recent example was the African swine fever outbreak that affected China between 2018 and 2019. This outbreak caused an economic loss amounting to 0.78% of the country's total gross domestic product, approximately 111 billion USD[12]. Similarly, in 2001, the foot and mouth disease outbreak in the UK led to 7% of all cattle and 15% of all sheep being culled[13]. Between 2003 and 2004, the outbreak of highly pathogenic avian influenza in Asia led to a loss of 18% of the poultry population in Vietnam and 15% of the poultry population in Thailand[14]. As a result, these outbreaks negatively impacted the economy and livelihood of farm workers in countries where animal husbandry is an important contributor to the economy. For example, in Thailand – the world's third largest poultry exporter, smallholders used to make up 99% of poultry producers before the 2004 avian influenza outbreak[15]. The outbreak caused a $3 billion USD loss from Thailand's poultry industry, and has made it challenging for smallholders to sustain their businesses, with significant movements out of the sector in 2006 and 2007[16].

1.1.1 Antimicrobial resistance

In the fight against infectious diseases, a major advance was achieved in the beginning of the 20th century with the discovery and subsequent mass production of antimicrobial drugs, which have saved millions of lives. Penicillin, the world's first mass produced antimicrobial, was discovered in 1928 by Alexander Fleming. Followed by Selman Waksman's systematic study of microorganisms as potential sources of antimicrobial agents in the 1930s, multiple drug classes were discovered during the Golden Age of antimicrobial discovery from the 1940s to the 1970s[17]. The relatively fast discovery of new antimicrobial classes during this period led to a surge of its use both in humans and in animals, which plateaued in high-income countries by 2000, but continues to increase in LMICs. Between 2000 and 2015, in humans, antimicrobial consumption rate decreased slightly by 4% in high-income countries, while it increased by 77% in LMICs during the same period[18]. In animals, antimicrobial consumption continues to increase across LMICs, while decreases have been reported in a few high-income countries such as Denmark, France, and the US between 2015 and 2017[19]. If the consumption rates remain constant, it was projected that global antimicrobial use will increase by 15% between 2015 and 2030 in humans[18], and increase by 11.5 % between 2017 and 2030 in food animals[19].

Under repeated exposure of antimicrobials, microorganisms are able to adapt and evolve to withstand the effect of antimicrobial treatment, a phenomenon known as antimicrobial resistance (AMR). Although AMR is an ancient and natural phenomenon that results from the interaction of microorganisms and their environment, the overuse and misuse of antimicrobials by humans and animals is accelerating its development. AMR can occur through four main biochemical mechanisms[20]. First, bacteria can inactivate a drug by modifying or destructing it. This is the predominant mechanism of aminoglycoside resistance, with aminoglycoside modifying enzymes inactivating the drug. Second, bacteria can limit drug intake through the cell wall. The outer membrane of gram-negative bacteria can prevent the influx of antimicrobials such as vancomycin. Third, for some antimicrobials able to penetrate the bacteria's membrane, efflux pumps can effectively remove the compound. This is the mechanism that makes some bacteria such as *E. faecalis* intrinsically resistant to compounds such as streptogramin A[21]. Lastly, bacteria can protect or modify the target sites to avoid the effect of antimicrobials. For example, a tetracycline resistance determinant Tet(O) interacts with the ribosome, and removes the drug from its binding site[22].

The danger of rising AMR through inappropriate use of antimicrobials was already highlighted by Alexander Fleming close to a century ago. In his Nobel Lecture in 1945, he foresaw that "the time may come when penicillin can be bought by anyone in the shops. Then there is the danger that the ignorant man may easily underdose himself and by exposing his microbes to non-lethal quantities of the drug make them resistant". Indeed, just 12 years after its discovery, resistance to penicillin was reported[23]. By the late 1960s, more than 80% of *Staphylococcus aureus* strains were found resistant to penicillin[24]. Increasing resistance has been reported for many medically important antimicrobials such as third-generation cephalosporins[25] and fluoroquinolones[26]. Figure 1 provides a timeline from when an antimicrobial was introduced to when its resistance was detected[27]. It suggests that evolution of AMR is accelerating over time. One important reason behind the acceleration is the increasing detection of resistance determinants on mobile genetic elements, making it easier for them to spread through horizontal gene transfer. For example, genes encoding New Delhi Metallo-β-lactamase (NDM-1) carrying resistance to β-lactam antimicrobials is often found on plasmids. It was first detected in 2008 in a Swedish patient returning from India, and by 2010, NDM-1 had already spread to other countries worldwide, such as China, Canada and Belgium[28]. Another resistance mechanism, plasmid-mediated colistin resistance MCR-1, was first reported in *E. coli* in 2016, with evidence suggesting its origin in food animals due to the use of colistin as feed additives[29]. By 2020, MCR-1 was already reported in more than 70 countries both in humans and in animals[30].

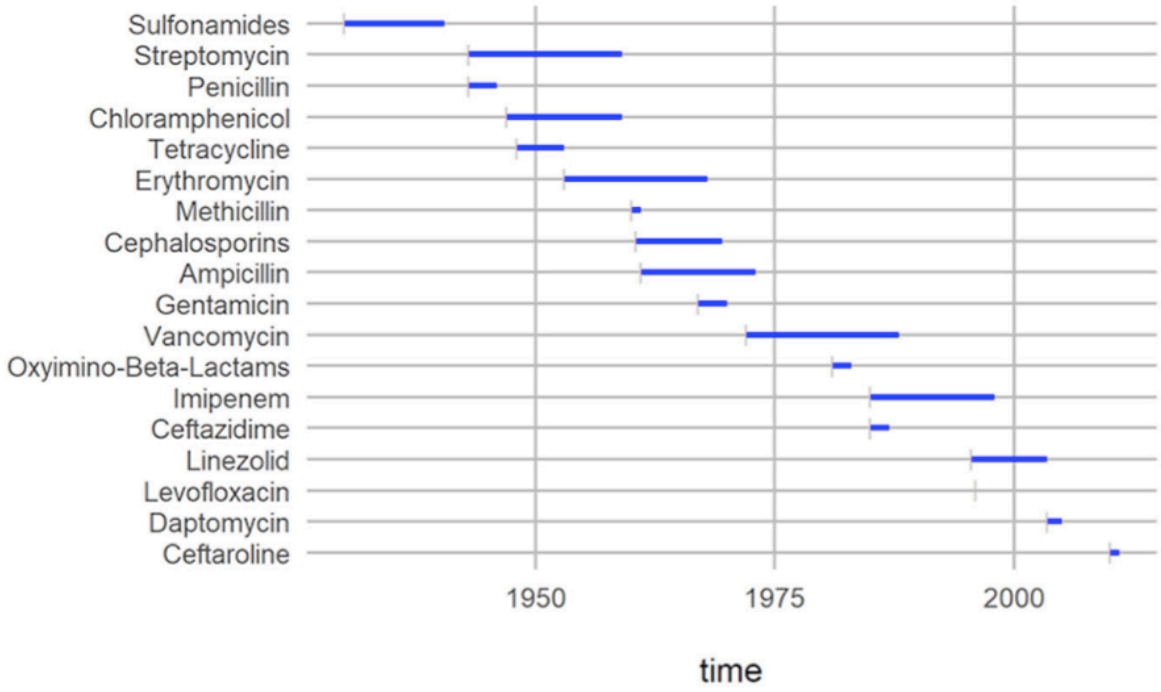


Figure 2. Timeline of antimicrobial introduction (beginnings of the bars) and detection of surmised evolved resistance (ends of the bars). This figure was published as Figure 1a in Witzany et al. 2020[27].

1.1.1.1 Antimicrobial use in food animals

Although concerns over AMR are centred around treatment of human infections, the majority of antimicrobials are consumed not by humans, but rather by animals raised for food[31]. In 2017, it was estimated that globally, 73% of antimicrobials were used in agriculture, making it an important reservoir of AMR. In food animals, antimicrobials are not only used for disease treatment, but also as growth promoters and surrogates for insufficient biosafety measures, which are crucial for reducing the introduction and spread of infectious diseases.

Like for humans, the use of antimicrobials in food animals started with the purpose of treating infections. In the 1930s, synthetic sulphonamides were amongst the first antimicrobials sold for veterinary use (e.g. treating mastitis, an inflammatory condition of the udder, in dairy cows) [32]. Simultaneously, at a time of increasing agricultural demand, experiments were carried out in food animals to search for feed additives such as vitamins that can decrease the feeding and profit ratio. In one of these experiments carried out in the late 1940s, the growth promoting effect of antimicrobials was discovered by chance[33]: the researchers fed waste generated from aureomycin production containing vitamin B12 to chickens, and the resulting weight gain far exceeded the effect of vitamin B12 alone, and was traced to the antimicrobial leftover in the waste.

The promise of producing affordable animal protein in massive amount soon led to the popularization of antimicrobial growth promoters in the animal industry. The booming sale of the vitamin B12/antimicrobial feed additive profited both the animal industry and the pharmaceutical companies, and increasingly more antimicrobial additives were licensed for use as growth promoters. Penicillin, oxytetracycline, and chlortetracycline were amongst the first licensed growth promoters in many countries including the Netherlands, UK, and France[34]. With antimicrobial feed additives, it was observed that animals grew 5% to 10% faster while

consuming 10% less feed[35]. In the context of animal husbandry, an industry known for its often small and fluctuating profitability[36], this boost in productivity was significant. This practice subsequently spread across continents. South Africa waived the requirement of veterinary prescription for many antimicrobials in 1947; Japan licensed antimicrobial feed additives in 1953; Routine antimicrobial use became common in China in the 1980s[34].

Antimicrobials, along with other factors such as the industrialization of agriculture and selective breeding, facilitated the intensification of the global animal industry. Chicken production, in particular, expanded greatly in the second half of the twentieth century. For example, in the UK, antimicrobials helped to decrease the feed-conversion rate in poultry from 4.2 before the Second World War to 2.4 in 1960, and poultry went from comprising only 1% of the total meat consumption in 1950 to comprising 25% of the market in 1980[35]. In Brazil, 90% of poultry were raised in confined settings by 2010[37]. The global pig industry also intensified. A notable example was in China, where the annual pork production increased from 8 Mt 1978 to more than 50 Mt in 2010[38].

1.1.1.2 Rising concern over AMR in food animals

Reports of AMR in food animals emerged soon after the growth promoting effect of antimicrobials was discovered. In 1951, streptomycin resistance was first reported in turkeys[39]. In the following years, resistance to other antimicrobials such as tetracyclines were also reported[40]. In addition, increasing prevalence of AMR following the introduction of the corresponding antimicrobial growth promoters in the market was documented. For example, following the introduction of fluoroquinolones in poultry production in Spain in 1993, the prevalence of fluoroquinolone resistance in *Campylobacter* increased sharply from <10% to more than 80% in 1996[41].

Rising AMR in food animals can lead to therapy failure, and thereby negatively impact animal health and the agricultural economy. Several antimicrobials that were once first-line choices of veterinary medicine are no longer effective against the targeted diseases. For example, in cattle, penicillin has been used to treat mastitis caused by *S. aureus*, as well as pneumonia caused by *P. multocida* and *M. haemolytica*[42]. Today, in some regions, high levels of AMR against penicillin has made it no longer an empirical first-line choice of treatment for these infections[42]. Finding alternative drugs can be difficult. First, antimicrobials that are more recently developed can be more expensive, resulting in increased costs for animal husbandry. Second, some alternative drugs, such as fluoroquinolones and third-generation cephalosporins, are critically important antimicrobials reserved for human use[43]. The convergence of these factors has led to fewer alternatives for treating certain diseases. For example, swine dysentery is a disease caused by *B. hyodysenteriae* that causes mortality and reduced growth in pigs. Tylosin and lincomycin were traditionally used to treat the infection, yet their resistance has become widespread[44]. Resistance to the alternative treatment option, pleuromutilins, has also been reported in some farms, making it difficult to control this disease[44].

Despite the direct impacts of AMR on animal health and agricultural economy, it was the accumulating evidence of AMR's transmissibility from animals to humans that prompted the reforms of agricultural regulations concerning AMR risk. Before the 1950s, scientists had widely believed that bacteria were only capable of acquiring resistance through mutation or by vertically inheriting resistance genes. In 1965, a *Lancet* article showed that horizontal transfer of AMR genes was possible between bacteria in animal and human populations[45]. The first experiment showing such transmission on farms was conducted by Levy et al. in 1974[46]. In the

experiment, tetracycline was added into the feed of chickens that were never previously exposed to antimicrobials. Within two weeks, resistant bacteria could be found in the faeces of 90% chicken fed with tetracyclines. In the next weeks, resistant bacteria also appeared in the faeces of people living on the farm, as well as chicken that were not exposed to tetracyclines.

AMR determinants as well as resistant bacteria can be transmitted from animals to humans via direct contact[47], through handling and consuming contaminated food[48], or through indirect routes via the environment[49]. There has been diverging evidence on the fraction of human infections attributable to antimicrobial use in animals[50]. For example, a review of 45 studies on AMR transmission of *E. coli* between animals and humans found that 18% of studies supported AMR transmission from animals to humans, 56% suggested transmission without showing directionality, while 26% did not support transmission[50]. The divergence in evidence is partly due to differences in the discriminative powers of the molecular typing technique applied. For example, whole genome sequencing is able to detect difference at high resolution that cannot be captured by traditional methods such as pulsed-field gel electrophoresis[51]. Future studies will need to use high-resolution techniques such as whole genome sequencing, and adopt appropriate sampling regimes ensuring good coverage on the geographical and temporal dimensions of all AMR reservoirs, in order to draw robust conclusions on the directionality and severity of AMR transmission between animals and humans.

Despite challenges in quantifying the number of human infections attributable to animals, evidence of AMR transmission is increasingly reported. Methicillin-resistant *Staphylococcus aureus* (MRSA) was one example of concern. MRSA infections were initially mainly acquired within hospitals or healthcare facilities. Since the 1990s, community-acquired MRSA began to increase, and was followed by the detection of prevalent livestock associated MRSA from the nasal swabs of both farm animals and farm workers[47]. There is growing evidence that livestock is an important reservoir of MRSA that can potentially lead to community-wide transmission[52].

Evidence of AMR transmission has been reported for other drug-pathogen combinations as well. For example, poultry workers were found to be 32 times more likely to carry gentamicin-resistant *E. coli*[53]. Genomic analysis of *Klebsiella pneumoniae* showed that the strains isolated from retail meat and urinary tract infections were closely related, indicating a recent transfer of the strains[54]. The emergence of plasmid-mediated colistin resistance *mcr-1* was first reported in *E. coli* in 2016, with evidence suggesting its origin in food animals due to the use of colistin as feed additives[29]. By 2020, *mcr-1* was already reported in more than 70 countries both in humans and in animals[30].

Although rising AMR has been recognized as a leading public health threat for years, until recently, little was known about the actual global burden of AMR in either humans or animals. There were scattered studies on human incidence and deaths associated with AMR targeted at select pathogens (e.g. *Escherichia coli* and *Klebsiella pneumoniae*[55]) and at select locations (the European Union[56]). In 2022, the first comprehensive assessment of global burden of AMR in humans was published by Murray et al[57]. They estimated that 4.95 million deaths were associated with AMR, while 1.27 million deaths were attributable to AMR in 2019[57]. If AMR would have been included as a single "disease" in the rank of causes of deaths in 2019, it would have ranked higher than HIV and malaria[58]. In comparison, in animals, the knowledge on the burden of AMR remains limited, and very heterogeneous across countries.

1.1.2 Impact of the COVID-19 pandemic

The crisis of AMR is known as the "silent pandemic". Though not a disease itself, it is increasingly making the treatment of diseases fail. This somewhat abstract concept has struggled to gain the global attention it warrants. In comparison, the COVID-19 pandemic with its sudden and far-reaching impact on global health, received immediate and comprehensive attention from all sectors.

Since the report of the first case in Wuhan, China in December 2019, COVID-19 has spread across all continents, and was declared a global pandemic by the World Health Organization in March 2020. By October 2023, COVID-19 caused 6.9 million deaths globally[59]. The COVID-19 pandemic also had a profound impact on society and the global economy. More than half of students globally were affected by school closures and many workers lost their jobs[60]. Approximately 71 million people were pushed back into extreme poverty in 2020, defined as living on less than 1.90 USD per person per day[61]. Food insecurity also increased: a survey in the US showed that the number of food-insecure households with children more than doubled after the onset of the pandemic[62]. In addition, routine childhood vaccination was interrupted with a 11% decline in LMICs[63].

The trade-off between protecting public health and mitigating the economic and societal costs associated with pandemic containment measures was taken differently between countries. China for example, prioritized limiting the number of infections and adopted strict lockdown measures. Sweden went largely the other direction and adopted a voluntary approach promoting social distancing and hygiene. In comparison, Switzerland took an in-between position, and adopted the strategy of "flattening the curve", where measures were implemented as necessary to keep hospitals from becoming overwhelmed by patients.

In Switzerland, the first COVID-19 case was reported on February 25, 2020, in the southern canton of Ticino which borders Italy. This person possibly got infected while attending an event in Italy 10 days earlier. Within two days, COVID-19 cases were also reported in cantons Geneva, Graubünden, Zürich, Aargau, and Basel. In response, on February 28, 2020, the Swiss government cancelled large gatherings involving more than 1,000 people, including carnivals in Bern and Basel. On March 27, a nation-wide partial lockdown was put in place. All gathering places, with the exception of those essential for living such as pharmacies and grocery stores, were closed. The resulting decrease in COVID-19 cases led the government to gradually release the measures from April 16, 2020, following a three-step procedure. First, shops and businesses such as beauty salons and day care centres were re-opened. Second, schools and other shops were re-opened. Lastly, other places such as higher education institutions and libraries were also opened. With mandatory mask wearing and social distancing in place, the number of cases remained relatively low between June and September.

The ease in measures saw a rapid increase of COVID-19 cases in October 2020, which marked the start of the second wave of infections. In response, the government increased public health measures. However, these measures, e.g., limiting public gatherings to 15 people, were less strict compared with the first wave. Correspondingly, there were more cases and fatalities during the second wave. In January 2021, the government imposed additional measures requiring restaurants and other non-essential shops to be closed. In additional, the national vaccination campaign was rolled out. COVID-19 cases dropped from 22,523 per week at the beginning of January to 754 cases in the middle of June 2021 (source: corona-data.ch). Since September 2021, only a valid COVID-19 certificate was required for fully vaccinated people with no other restrictions, and almost all restrictions were lifted by February 2022.

1.2 Infectious diseases surveillance

The cornerstone for preventing and managing all kinds of infectious diseases is surveillance, defined by the World Health Organization as "the systematic ongoing collection, collation and analysis of data for public health purposes and the timely dissemination of public health information for assessment and public health response as necessary"[64]. Surveillance can act as early warning systems for disease outbreaks, be used to evaluate the impact of interventions, and to plan and prioritise public health policies[65].

The idea of systematically collecting facts to inform actions may date back to Hippocrates (460 B.C. – 370 B.C.) who was known as the father of medicine[66]. However, the first public health action related to surveillance was first recorded in the early fourteenth century, when three guardians of public health were appointed by the Venetian Republic to detect and exclude ships containing people infected with bubonic plague[66]. Counts of morbidity and mortality have traditionally formed the core of surveillance data. In 1685, John Graunt suggested fundamental principles of disease surveillance that were centred around disease-specific death counts and death rates. In the USA, disease reporting at the national level started in 1850 by publishing mortality statistics, and systematic reporting of morbidity followed in 1874[67].

With the advancement of the quality and capacity of surveillance systems, increasingly diverse sources of data other than mortality and morbidity have been incorporated into surveillance of infectious diseases. Notable examples include population health and point-prevalence surveys, laboratory reports, animal reservoir distribution data, and hospital statistics[68]. Population health surveys use random or stratified sampling techniques to collect data from a representative sample of the population, including demographic information, health-related behaviours (e.g., smoking, physical activity, diet), and the presence of diseases. These surveys may be conducted cross-sectionally – providing a snapshot of health at a specific point in time, or longitudinally – tracking changes in health and behaviours over time. While population health surveys are typically conducted on large scales and cover a wide range of health-related topics, point-prevalence surveys (PPS) focus on a specific point in time and are usually on smaller scales. PPS are often used to assess the burden of a specific disease or condition within a short timeframe. Apart from surveys, laboratory reports facilitate the accurate identification of infectious agents causing diseases. Animal reservoir distribution data help assess risks of zoonotic disease spillover. Hospital statistics, such as patients' length of stay in the intensive care unit, provide valuable information on the progression of diseases in the most severe stages.

Developments in public health surveillance also extend to the use of geolocated data. The earliest presentation of geolocated data were spot maps, which used dots or other symbols to visually illustrate the residential or exposure locations of individual cases, and can be particularly useful for identifying disease clusters. Examples include yellow fever maps in the US in the late 18th century[69], and a Cholera spot map in the UK in the mid-19th century[70]. In comparison, thematic maps aggregate data points into predefined geographic units. In the mid-20th century, thematic maps were used for showing cancer prevalence across counties in the UK[71]. The third type of maps useful for disease mapping was isoline maps, which use continuous lines or curves to connect data points on the map. In the mid-20th century, isoline maps were used to show the dates of arrival of the Black Death in west and central Europe in the 14th century[72]. Recent decades have seen increasing integration of Geographic Information System (GIS) in disease surveillance systems[73]. For example, in 2021, the Food and Agriculture Organization built a web-based global animal disease information system that displays

locations of reported diseases such as avian influenza and African swine fever on a global map[74]. During the COVID-19 pandemic, near-real time interactive dashboards and maps tracking the spread of the virus, vaccination coverage, and healthcare resource allocation have also been widely used to inform the public and policymakers[59].

1.2.1 AMR surveillance

Surveillance of AMR for animals has been established for decades in multiple high-income countries. In Europe, in 1995, the Danish Integrated Antimicrobial Resistance Monitoring and Research Program was the first program to systematically monitor AMR in humans, animals, and food. Following this, other countries including Norway, Sweden, and the Netherlands also established their national surveillance programs. A European wide AMR surveillance system – EARS-Net – was established by the European Commission in 1998[75]. The European Food Safety Authority has been publishing yearly report on AMR in zoonotic and indicator bacteria from humans, animals and food since 2004[76]. In the US, the National Antimicrobial Resistance Monitoring System for Enteric Bacteria was established in 1996. Over time, the system expanded from testing only human isolates, to also testing animal and retail meat samples, and including whole genome sequencing into the surveillance system[77].

In contrast, very few systematic AMR surveillance systems exist in LMICs, including Colombia, Thailand, and China. In 2007, Colombia launched an integrated surveillance program of AMR[78]; In 2008, China launched a surveillance program on AMR of animal origin, although the database is not currently available to the public; since 2017, Thailand publishes a One Health Report on antimicrobial consumption and antimicrobial resistance both in humans and animals[79]. In LMICs, important obstacles for establishing such systems include limited government expenditure on public health services, insufficient resources for conducting laboratory testing, and lack of proper training for public health staff. AMR surveillance in LMICs is also less developed in animals compared with humans. A review identified that only 25% of the world's AMR surveillance systems include animal samples for systematic testing[80]. Surveillance of animals essentially evolved from existing programs for humans. It is often focused on pathogens potentially transmissible to humans such as methicillin-resistant *Staphylococcus aureus* and extraintestinal pathogenic *E. coli*, and bacteria that are common causes of foodborne diseases such as *Salmonella* and *Campylobacter*[81].

The lack of information on trends of AMR in food animals in LMICs is an important obstacle for planning interventions in places where antimicrobial use may be high. For example, evidence suggests that antimicrobials are sold and consumed illegally in the markets in India, China, and Vietnam, potentially resulting in high levels of AMR and threatening the sustainability of antibiotic use, public health, and food safety. In the absence of publicly available systematic surveillance data, point-prevalence surveys (PPS) conducted by independent veterinary and public health institutes could serve as an alternative source of information, and collectively form event-based surveillance to shed light on patterns of AMR. In 2019, I joined a collaborative effort to build a repository of PPS of AMR in food animals in LMICs. We launched a platform *resistancebank.org* to display and collect PPSs of AMR[82]. The data collected in this effort are the basis for my work in Chapter 2 to 4, where I mapped the geographic distribution of AMR in LMICs, in China, and in the European Union, and developed methods to guide future surveillance of AMR.

1.2.2 COVID-19 surveillance

Timely and accurate surveillance played a pivotal role for containing the impact of the COVID-19 pandemic[83]. At the beginning of the pandemic when rapid tests for COVID-19 were not widely available, hospital-based surveillance was the most reliable source of information on the status of the pandemic. It provided important clinical and epidemiological information on COVID-19, such as case fatality rates, and admission rate and duration of stay at intensive care units (ICUs)[84]. In addition, when the virus was first introduced in a country, active surveillance was often applied to trace close contacts of the infected people, and measures such as quarantine was used to avoid the subsequent transmission of COVID-19. As COVID-19 continued to spread widely across countries, the more widely applied type of surveillance was case-based routine surveillance, where cases of COVID-19 cases were reported within 48 hours of identification. The reports were the aggregated by institutions such as the World Health Organization and made public[83].

In Switzerland, measures to contain COVID-19 were less strict in comparison with other European countries, hospitals were operating with increased load over long periods, and the healthcare systems were challenged by the sudden increase of patients[85]. Surveillance data from hospitals was therefore important in informing policy makers on adjustments of measures to make sure that hospital capacities were not exceeded. The Coordinated Sanitary Service in Switzerland manages the Information and Operations System (IES system) which collects data on the occupancy of over 150 hospitals and emergency departments (www.babs.admin.ch/de/aufgabenbabs/ksd/ies.html). The system assists hospitals to record their emergency capacities, and can in theory serve as basis to facilitate the transfer of patients between hospitals. In preparation of a possible overload of ICUs in hospitals, up to 8,000 soldiers were made available, including 3,000 in the medical field, to prepare for the allocation of additional ICU beds and other medical resources across the country with the help of the IES system.

In the IES system, each emergency department uploads twice a day information of the following types: hospital beds, ICU beds, intermediate care unit beds, ICU beds equipped with extracorporeal membrane oxygenation (ECMO), and ICU beds equipped with mechanical ventilation. For each of these types, the following numbers were recorded: the total number of beds available, beds currently in use, and beds currently in use for COVID-19 patients.

With the purpose of optimizing the geographic allocation of medical resources, particularly doctors and hospital beds, in preparation of potential hospital overload, I collaborated with the Swiss Armed Forces to use the IES system to forecast ICU beds occupancy at the hospital level. A collaborative effort was established between our research group and other researchers from ETH Zurich as well as the Swiss Tropical and Public Health Institute, to utilize a range of modelling methodologies for predicting COVID-19 cases in ICUs. I built one of the models and generated predictions of COVID-19 cases, based on work of Althaus et al. (https://ispmbern.github.io/covid-19/swiss-epidemic-model/), and using epidemiological information of COVID-19 estimated by other researchers, such as the biological characters of the virus, its transmission routes, and epidemiological parameters such as the incubation period and infectivity period[86]. Although an abundance of disease dynamics models has been developed during the pandemic, it is important to note that in such times of emergency, the most useful models may not necessarily mean more complex or accurate models, but rather those able to generate plausible results in a timely fashion. First, data in the early phase of the pandemic is often too limited and uncertain to support the parameterization of complex models. Second, an effective model should be one that is ready to support timely decision-making before the pandemic reaches a critical stage, rather than one with higher accuracy but in an ad-

hoc manner. In addition, complex models may suffer from reduced interpretability, leading to a lack of clarity in the messages they can provide regarding intervention strategies.

We built a platform *icumonitoring.ch* to display the current and projected ICU occupancy at the regional, cantonal, and hospital levels. From a scientific aspect, I adapted a compartmental model to make canton level predictions of COVID-19 cases in ICUs. From an operational aspect, I coordinated the platform update twice a week, to inform policy makers on the near-real-time trends of COVID-19 in hospitals.

1.3 Geospatial and disease dynamics modelling as tools to prevent infectious diseases

1.3.1 Exploring the spatial dimension of infectious diseases

Infectious diseases are distributed heterogeneously across regions. The geographic patterns of infectious diseases result from the interaction of multiple factors, including their origin, transmission pathways, the presence of susceptible populations and disease vectors, and climate. For example, Ebola may have originated from African fruit bats, and can be transmitted to human through bushmeat activities. Therefore, its distribution is concentrated in Sub-Saharan Africa, where the reservoir bat species and bushmeat activities are the most important[87]. In comparison, the occurrence of Salmonellosis is nearly ubiquitous, because *Salmonella* resides in the intestines of farm animals which are widespread across the globe. Understanding these distinct geospatial patterns that underlie disease distribution can help policy makers design interventions, target resources, and develop effective public health strategies.

Investigations of the spatial patterns of infectious diseases, in their early stages, were conducted through visual inspections of descriptive maps. Perhaps the oldest such maps with records until today was Filippo Arrieta's maps of the plague in the late 17th century[88]. The maps showed areas that were most affected by the plague and the geographic boundaries of quarantine areas. A well-known example in history that explored spatial patterns of cases to uncover disease patterns was John Snow's investigation on cholera deaths in the mid-19th century[70]. By investigating where the cholera cases lived and their sources of drinking water, Snow was able to conclude that the contamination originated from a water pump. His investigation led to the removal of the handle of that pump, and this intervention successfully ended the cholera epidemics. Back then, Snow's hypothesis of the origin of contamination was confirmed by drawing stacks of bars on a map to show the location and number of fatalities. With the advancements in geospatial modelling – defined as the application of statistical techniques on geolocated data to analyse and predict spatial patterns, geolocated disease surveillance data are increasingly used to identify clusters of cases, estimate trends, plan and evaluate interventions, and support public health actions[89].

Early efforts of geospatial modelling of infectious diseases started in the mid-20th century, and focused either on identifying potential risk factors linked with disease occurrence, or on spatial clustering. Published in 1966, one of the earliest examples for linking risk factors with animal disease was Ollerenshaw's forecast model, which used climate conditions for predicting acute outbreaks of *Fasciola hepatica* in sheep in Wales[90]. The model used an index calculated with

rainfall and potential evapotranspiration data to categorise disease risk into four groups. In the same year, Yeoman also examined the correlation between infestation rates of ticks in farm animals and variation of rainfall[91]. On the side of spatial clustering analysis, in 1973, "spatial diffusion" was applied to model a Newcastle disease epidemic[92]. By 2000, several techniques for cluster analysis were developed and used in veterinary epidemiology, including nearest neighbour, join count statistics, Moran's I test, and the Knox test[93]. One example was the application of Moran's I test to detect clusters of bovine anaplasmosis and babesiosis in Costa Rica[94].

An important advance that proved successful in combining the contribution of risk factors and spatial clustering together for geospatial modelling was Gaussian process regression (GPR). GPR was developed in the mid-20th century due to the mining industry's growing interest in calculating areal or block averages for ore reserves. In combination with the expansion of remotely sensed data of remarkable spatio-temporal resolution, recent decades have seen increasing side use of GPR for disease risk mapping, such as for malaria[95], African swine fever[96], and avian influenza[97]. However, one pitfall of GPR is that the risk factors are combined in a linear fashion for predicting disease risk, while disease distribution is known to be driven by complex interactions between environmental and anthropogenic factors.

Efforts of disease mapping are increasingly harnessing non-linear algorithms developed in the field of species distribution modelling, such as trees[98], networks[99], and non-parametric models such as k-nearest neighbours[100]. Developed in the 1990s, generalised additive model transforms linear models to non-linear models, by using spline functions as the coefficients for the model[101]. The model has high flexibility while also maintaining the interpretability of a linear model. Generalised additive models have been used to model meningitis outbreaks[102] and malaria risk[103]. Developed in 2008, boosted regression trees is another example of non-linear statistical models with improved predictive performance. It harnesses the strengths of both regression trees that use recursive binary splits to relate covariates to the response variable, and boosting that combines different trees for improved predictions[98]. Boosted regression trees have been used to map the risk of dengue[104] and the distribution of malaria vectors[105]. Artificial neural network, another well-known non-linear model, has gained increasing attention since it won the ImageNet Large Scale Visual Recognition Challenge in 2012[106], and has revolutionized the field of artificial intelligence. Inspired by how the human brain functions, it uses a layered structure comprising interconnected nodes or neurons to predict complex patterns in data. Layers of a neural network are organized into an input layer, one or multiple hidden layers, and one output layer. Neurons between the layers are interconnected through weights that are trained using iterative optimization processes. Neural network has been used for predicting the HIV epidemic[107], COVID-19 outbreak[108], and monkeypox outbreak[109], amongst others.

In 2017, Bhatt et al. developed an approach to combine the predictive power of non-linear algorithms within GPR (Figure 3). Named Gaussian process stacked generalization, this method was shown to improve prediction accuracy of disease mapping, and will be used to produce several maps presented in the coming chapters of the thesis.

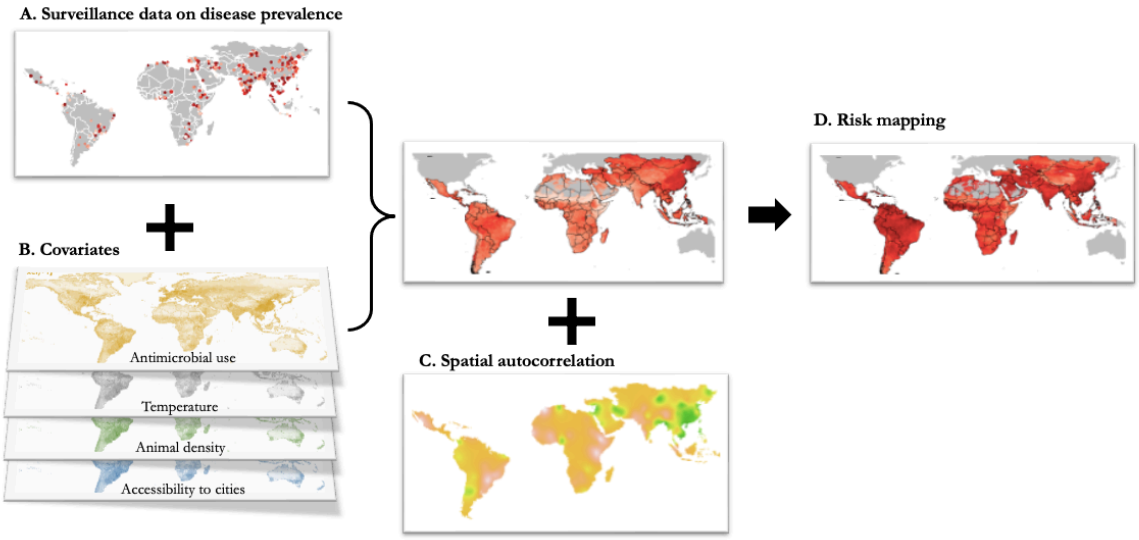


Figure 3. Illustration of geospatial mapping procedure.

The purpose of Gaussian process stacked generalization is to transform geo-referenced (e.g., with coordinates; Figure 3A) surveillance data (e.g., cases of diseases) into spatially continuous maps of health-related metrics (e.g., risk of diseases; Figure 3D). Here, we consider only the situation where disease surveillance data are provided as values at geographic points, as opposed to values in geographic areas. The mapping procedure is conducted in two steps. In the first step (Figure 3B), a set of risk factors (e.g. temperature) are collected in the format of pixel maps as covariates, and values of risk factors are extracted at the locations associated with reported disease surveillance data. Multiple non-linear algorithms developed in the field of species distribution modelling e.g., boosted regression trees, are then used to quantify the relation between risk factors and disease occurrence. Maps of disease risk can then be generated by applying the trained statistical models on the maps of risk factors, these comprise the "trend" component of our final disease map.

Modelling disease occurrence using only risk factors often does not explain all the variation in the data. As Tobler's law states, "everything is related to everything else, but near things are more related than distant things", also referred to as spatial autocorrelation. Failing to incorporate spatial autocorrelation in geospatial modelling can lead to the violation of a key assumption in statistical modelling – the independence of residuals. Therefore, the second step of Gaussian process stacked generalization is to combine the "trend" component with the residual spatial autocorrelation (Figure 3C). A spatial covariance function dependant on the geographic distance between two samples is included in the model. The function can be linear, exponential, squared-exponential, polynomial, or take other forms of radial basis function. The Matern covariance function is a generalization of the Gaussian radial basis function useful for spatial modelling[110]. Here, predictions of disease risk trained with non-linear algorithms are used as covariates in the Gaussian process regression, and therefore combines an ensemble of model predictions with residual spatial autocorrelation.

The training of Gaussian process models, either through maximum likelihood or Markov chain Monte Carlo, can be computationally demanding[111]. The increasingly larger datasets used for disease modelling necessitate efficient approximation approaches to train such models. The integrated nested Laplace approximation (INLA) is an effective approximation method for the

Bayesian inference of Gaussian process models. INLA can be used for modelling spatial data by representing the Matern covariance function as a Gaussian Markov random field and using stochastic partial differential equation (SPDE) approach for the inference[110]. INLA-SPDE has been widely used for mapping diseases such as malaria[112] and Rift Valley fever[113].

1.3.2 Disease dynamics modelling for simulating epidemics in time

Species distribution modelling and Gaussian process regression, as mentioned in the previous section, are based on correlations of risk factors and disease occurrence. Their primary function is to map disease distribution by delineating the suitability of environmental and anthropogenic conditions for the survival and transmission of diseases. However, geospatial models that do not reflect mechanisms of disease transmission may not be optimal for predicting disease progression. In comparison, in compartmental models, transmission of diseases between subsets of the population is explicitly reflected in the model structure. Compartmental models are therefore particularly suitable for modelling the temporal dimension of infectious diseases. In addition, effects of different containment measures on disease progression can be modelled by modifying the relevant transmission parameters in the model, thereby allowing scenario analyses to guide policy intervention.

The first compartmental disease dynamics model, known as the Susceptible–Infectious–Removed (SIR) model, was developed by Kermack and McKendrick in 1927[114]. The model discretizes the population into three compartments based on their states of infection (Figure 4a): susceptible individuals that can become infected through contact with the infectious individuals, infectious individuals that are capable of infecting others, and removed individuals who are deceased, have recovered and developed immunity, or are isolated or in quarantine. The links between compartments are represented as ordinary differential equations with respect to time. Given a specified starting condition of these compartments, the model can simulate time series of all compartments.

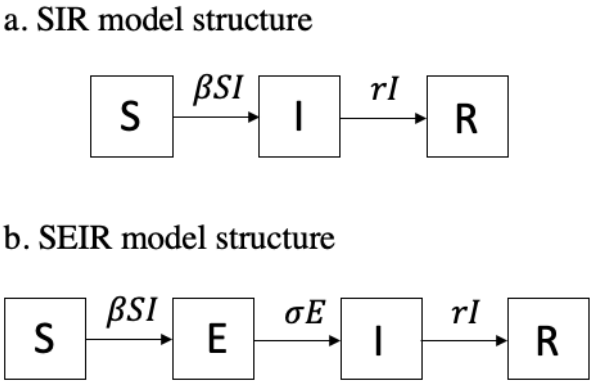


Figure 4. The model structures of a basic SIR model (a) and a basic SEIR (b) model. Squares represent compartments of the population; S: susceptible individuals, I: infectious individuals, E: exposed individuals, R: recovered or deceased individuals. Arrows represent components of differential equations that link the compartments together. β: infection rate, σ: inverse of the average latent time. r: recovery/mortality rate.

The simplest form of an SIR model is represented in Figure 4a. Two fundamental processes of disease dynamics are reflected in the model structure. The first process is infection, reflected

as a decrease of βSI per unit time in the susceptible compartment that is transferred to the infected compartment. Here, β is the infection rate, defined as the number of susceptible individuals that an infective individual infects per time unit. The second process is mortality or recovery from the infection, reflected as a decrease of rI per unit time in the infectious compartment that is transferred to the recovered/deceased compartment. SIR models have successfully predicted disease dynamics such as measles epidemics[115].

However, one important limitation of the SIR model is that it neglects the latent period of an infection. Taking this into consideration, the SEIR model has been developed to include an additional compartment E – exposed individuals who are infected, but not yet infectious[116] (Figure 4b). When susceptible individuals become infected, they are first transferred into the exposed compartment, which are then transferred to the infectious compartment with a rate σ, calculated as the inverse of the average latent time. The SEIR model was widely used for predictions during the COVID-19 pandemic in e.g., India[117], Italy[118], and Indonesia[119]. In Chapter 5, I will present models that extend the SEIR model to provide short-term forecasts of COVID-19 hospitalizations in the intensive care units in Switzerland.

1.4 Thesis structure

In Chapter 2, I conducted a literature review of point-prevalence surveys (PPS) of AMR in food animals in China including surveys published in English and Chinese. I applied geospatial modelling on this collection of PPS to produce maps of AMR in China. I then investigated a topic often overlooked in the field of geospatial modelling – using maps of prediction uncertainty to inform field sampling campaigns. I developed a method to optimize the geographic allocation of future surveillance efforts, and applied the method on the map of AMR in food animals in China, as well as a map of AMR in aquaculture developed in collaboration with my colleague Dr. Dan Schar. I showed that surveillance on AMR in farm animals could be intensified in southwestern and northeastern China, while surveillance on AMR in aquaculture in Asia could be intensified in eastern China and India.

In Chapter 3, together with my colleague Ranya Mulchandani, we conducted literature review of PPS on AMR in food animals in European countries. I then produced a map of AMR at 10x10 km resolution, and adjusted the predictions based on the existing country-level systematic AMR surveillance data in Europe. We showed major geographic heterogeneities in AMR within countries, which could not otherwise be captured by the data from systematic surveillance program that was aggregated at the country level.

In Chapter 4, I further refined maps based on a summary metric of AMR across drug-pathogen combinations used in the previous two chapters, and produced maps of AMR for 7 individual antimicrobial classes. I then developed a method that used co-resistance patterns between antimicrobials, as well as environmental and anthropogenic covariates, to predict the antimicrobial with the highest probability of its resistance exceeding a critical level (50%) in the future. The predicted antimicrobials could be a "priority antimicrobial" for future surveillance. I showed that the priority antimicrobials were penicillins or tetracyclines in most locations in Africa and South America, and penicillins or sulphonamides in most locations in Asia.

Chapter 5 presents the platform *icumonitoring.ch* that we built to display and project COVID-19 hospital occupancy in the intensive care units in Switzerland. During the COVID-19 pandemic, I coordinated the bi-weekly update of contents on the platform. This chapter also presents a collaborative work between me and Dr. Riccardo Delli on improving the disease dynamics model used for the predictions on *icumonitoring.ch*.

1.5 References

1. United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects 2019: Highlights (ST/ESA/SER.A/423). (2019).
2. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results.
3. World Health Organization. *The World Medicines Situation*. (2004).
4. World Health Organization. *Water, Sanitation and Hygiene in Health Care Facilities: Status in Low and Middle Income Countries and Way Forward*. (World Health Organization, 2015).
5. Manderson, L., Aagaard-Hansen, J., Allotey, P., Gyapong, M. & Sommerfeld, J. Social research on neglected diseases of poverty: continuing and emerging themes. *PLoS Negl. Trop. Dis.* **3**, e332 (2009).
6. United Nations Children's Fund (UNICEF), the World Health Organization & the International Bank for Reconstruction and Development/The World Bank. Levels and trends in child malnutrition. (2023).
7. Plowright, R. K. *et al*. Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510 (2017).
8. Adamopoulos, T. & Restuccia, D. The size distribution of farms and international productivity differences. *Am. Econ. Rev.* **104**, 1667–1697 (2014).
9. Grace, D., Gilbert, J., Randolph, T. & Kang'ethe, E. The multiple burdens of zoonotic disease and an ecohealth approach to their assessment. *Trop. Anim. Health Prod.* **44**, 67–73 (2012).
10. Jones, K. E. *et al*. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
11. Li, M. *et al*. Global disease burden of pathogens in animal source foods, 2010. *PloS One* **14**, e0216545 (2019).
12. You, S. *et al*. African swine fever outbreaks in China led to gross domestic product and economic losses. *Nat. Food* **2**, 802–808 (2021).
13. Rushton, J. *et al*. Economic analysis of vaccination strategies for foot and mouth disease in the UK. *R. Soc. Inq. Infect. Dis. Livest.* (2002).
14. McLeod, A., Morgan, N., Prakash, A. & Hinrichs, J. Economic and social impacts of avian influenza. in 7–9 (2005).
15. Rushton, J., Viscarra, R., Bleich, E. G. & McLeod, A. Impact of avian influenza outbreaks in the poultry sectors of five South East Asian countries (Cambodia, Indonesia, Lao PDR, Thailand, Viet Nam) outbreak costs, responses and potential long term control. *Worlds Poult. Sci. J.* **61**, 491–514 (2005).
16. Heft-Neal, S., Kahrl, F., Otte, J. & Roland-Holst, D. Assessment of smallholder indigenous poultry producer viability in Thailand. *Mekong Team Work. Pap. No 9* (2009).
17. Hutchings, M. I., Truman, A. W. & Wilkinson, B. Antibiotics: past, present and future. *Curr. Opin. Microbiol.* **51**, 72–80 (2019).
18. Klein, E. Y. *et al*. Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci.* **115**, E3463–E3470 (2018).
19. Tiseo, K., Huber, L., Gilbert, M., Robinson, T. P. & Van Boeckel, T. P. Global trends in antimicrobial use in food animals from 2017 to 2030. *Antibiotics* **9**, 918 (2020).
20. Munita, J. M. & Arias, C. A. Mechanisms of antibiotic resistance. *Virulence Mech. Bact. Pathog.* 481–511 (2016).

21. Singh, K. V., Weinstock, G. M. & Murray, B. E. An Enterococcus faecalis ABC homologue (Lsa) is required for the resistance of this species to clindamycin and quinupristin-dalfopristin. *Antimicrob. Agents Chemother.* **46**, 1845–1850 (2002).
22. Dönhöfer, A. *et al.* Structural basis for TetM-mediated tetracycline resistance. *Proc. Natl. Acad. Sci.* **109**, 16900–16905 (2012).
23. Abraham, E. P. & Chain, E. An enzyme from bacteria able to destroy penicillin. *Nature* **146**, 837–837 (1940).
24. Lobanovska, M. & Pilla, G. Focus: drug development: Penicillin's discovery and antibiotic resistance: lessons for the future? *Yale J. Biol. Med.* **90**, 135 (2017).
25. Asensio, A. *et al.* Trends in yearly prevalence of third-generation cephalosporin and fluoroquinolone resistant Enterobacteriaceae infections and antimicrobial use in Spanish hospitals, Spain, 1999 to 2010. *Eurosurveillance* **16**, (2011).
26. Goettsch, W. *et al.* Increasing resistance to fluoroquinolones in Escherichia coli from urinary tract infections in the Netherlands. *J. Antimicrob. Chemother.* **46**, 223–228 (2000).
27. Witzany, C., Bonhoeffer, S. & Rolff, J. Is antimicrobial resistance evolution accelerating? *PLoS Pathog.* **16**, e1008905 (2020).
28. Nordmann, P., Poirel, L., Toleman, M. A. & Walsh, T. R. Does broad-spectrum β-lactam resistance due to NDM-1 herald the end of the antibiotic era for treatment of infections caused by Gram-negative bacteria? *J. Antimicrob. Chemother.* **66**, 689–692 (2011).
29. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).
30. Shen, C. *et al.* Dynamics of mcr-1 prevalence and mcr-1-positive Escherichia coli after the cessation of colistin use as a feed additive for animals in China: a prospective cross-sectional and whole genome sequencing-based molecular epidemiological study. Lancet Microbe. 2020; 1: e34–43.
31. Van Boeckel, T. P. *et al.* Reducing antimicrobial use in food animals. *Science* **357**, 1350–1352 (2017).
32. Lesch, J. E. *The First Miracle Drugs: How the Sulfa Drugs Transformed Medicine.* (Oxford University Press, 2007).
33. Waibel, P. E., Cravens, W. & Snell, E. E. The effect of diet on the comparative activities of pyridoxal, pyridoxamine and pyridoxine for chicks. *J. Nutr.* **48**, 531–538 (1952).
34. Kirchhelle, C. Pharming animals: a global history of antibiotics in food production (1935–2017). *Palgrave Commun.* **4**, (2018).
35. Godley, A. & Williams, B. Democratizing luxury and the contentious "invention of the technological chicken" in Britain. *Bus. Hist. Rev.* **83**, 267–290 (2009).
36. Fraser, D. G. *Animal Welfare and the Intensification of Animal Production: An Alternative Interpretation.* vol. 2 (Food & Agriculture Org., 2005).
37. Silbergeld, E. K. *Chickenizing Farms and Food: How Industrial Meat Production Endangers Workers, Animals, and Consumers.* (JHU Press, 2016).
38. Zhang, S. *et al.* Pork production systems in China: A review of their development, challenges and prospects in green production. *Front Agric Sci Eng* **8**, 15–24 (2021).
39. Starr, M. P. & Reynolds, D. M. Streptomycin resistance of coliform bacteria from turkeys fed streptomycin. *Am. J. Public Health Nations Health* **41**, 1375–1380 (1951).
40. Elliott, S. D. & Barnes, E. M. Changes in serological type and antibiotic resistance of Lancefield group D streptococci in chickens receiving dietary chlortetracycline. *Microbiology* **20**, 426–433 (1959).
41. Silbergeld, E. K., Graham, J. & Price, L. B. Industrial food animal production, antimicrobial resistance, and human health. *Annu Rev Public Health* **29**, 151–169 (2008).

42. Bengtsson, B. & Greko, C. Antibiotic resistance—consequences for animal health, welfare, and food production. *Ups. J. Med. Sci.* **119**, 96–102 (2014).
43. World Health Organization. Critically important antimicrobials for human medicine. (2019).
44. Aarestrup, F. M., Duran, C. O. & Burch, D. G. Antimicrobial resistance in swine production. *Anim. Health Res. Rev.* **9**, 135–148 (2008).
45. Anderson, E. S. & Datta, N. Resistance to penicillins and its transfer in Enterobacteriaceae. *Lancet* 407–9 (1965).
46. Levy, S. B., FitzGerald, G. B. & Macone, A. B. Changes in intestinal flora of farm personnel after introduction of a tetracycline-supplemented feed on a farm. *N. Engl. J. Med.* **295**, 583–588 (1976).
47. Smith, T. C. *et al.* Methicillin-resistant Staphylococcus aureus (MRSA) strain ST398 is present in midwestern US swine and swine workers. *Plos One* **4**, e4258 (2009).
48. Voets, G. M. *et al.* Identical plasmid AmpC beta-lactamase genes and plasmid types in E. coli isolates from patients and poultry meat in the Netherlands. *Int. J. Food Microbiol.* **167**, 359–362 (2013).
49. Deng, Y. *et al.* Dissemination of IncFII plasmids carrying rmtB and qepA in Escherichia coli from pigs, farm workers and the environment. *Clin. Microbiol. Infect.* **17**, 1740–1745 (2011).
50. Muloi, D. *et al.* Are food animals responsible for transfer of antimicrobial-resistant Escherichia coli or their resistance determinants to human populations? A systematic review. *Foodborne Pathog. Dis.* **15**, 467–474 (2018).
51. Parkhill, J. Antimicrobial resistance exchange between humans and animals: why we need to know more. *Eng. Beijing China* (2022).
52. David, M. Z. & Daum, R. S. Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. *Clin. Microbiol. Rev.* **23**, 616–687 (2010).
53. Price, L. B. *et al.* Elevated risk of carrying gentamicin-resistant Escherichia coli among US poultry workers. *Environ. Health Perspect.* **115**, 1738–1742 (2007).
54. Davis, G. S. *et al.* Intermingled Klebsiella pneumoniae populations between retail meats and human urinary tract infections. *Clin. Infect. Dis.* **61**, 892–899 (2015).
55. Temkin, E. *et al.* Estimating the number of infections caused by antibiotic-resistant Escherichia coli and Klebsiella pneumoniae in 2014: a modelling study. *Lancet Glob. Health* **6**, e969–e979 (2018).
56. Cassini, A. *et al.* 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. *Lancet Infect. Dis.* **19**, 56–66 (2019).
57. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).
58. Vos, T. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* **396**, 1204–1222 (2020).
59. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
60. Van Lancker, W. & Parolin, Z. COVID-19, school closures, and child poverty: a social crisis in the making. *Lancet Public Health* **5**, e243–e244 (2020).
61. United Nations. Sustainable Development Goals Report 2020. (2020).
62. Bauer, L. The COVID-19 crisis has already left too many children hungry in America. (2020).

63. Cardoso Pinto, A. M. *et al*. Disruptions to routine childhood vaccinations in low-and middle-income countries during the COVID-19 pandemic: A systematic review. *Front. Pediatr*. **10**, 979769 (2022).
64. World Health Organization. *International Health Regulations (2005)*. (World Health Organization, 2008).
65. World Health Organization. Surveillance in emergencies.
66. Choi, B. C. The past, present, and future of public health surveillance. *Scientifica* **2012**, (2012).
67. M'ikanatha, N. M., Lynfield, R., Julian, K. G., Van Beneden, C. A. & Valk, H. de. Infectious disease surveillance: a cornerstone for prevention and control. *Infect. Dis. Surveill.* 1–20 (2013).
68. Declich, S. & Carter, A. O. Public health surveillance: historical origins, methods and evaluation. *Bull. World Health Organ.* **72**, 285 (1994).
69. Stevenson, L. G. Putting disease on the map: the early use of spot maps in the study of yellow fever. *J. Hist. Med. Allied Sci*. 226–261 (1965).
70. Tulchinsky, T. H. John Snow, cholera, the broad street pump; waterborne diseases then and now. *Case Stud. Public Health* 77 (2018).
71. Howe, G. M. A national atlas of disease mortality in the United Kingdom. *Geogr. J.* **130**, 15–22 (1964).
72. Carpentier, E. Autour de la peste noire: famines et épidémies dans l'histoire du XIVe siècle. in vol. 17 1062–1092 (Cambridge University Press, 1962).
73. Fayisa, W. Review on the importance of Geographic Information System (GIS) in epidemiology: In prevention and control of animal disease. *Int J Vet Sci Res* **6**, 47–51 (2020).
74. Welte, V. R. & Terán, M. V. Emergency prevention system (empres) for transboundary animal and plant pests and diseases. the empres-livestock: an fao initiative. *Ann. N. Y. Acad. Sci*. **1026**, 19–31 (2004).
75. Tacconelli, E. *et al*. Surveillance for control of antimicrobial resistance. *Lancet Infect. Dis*. **18**, e99–e106 (2018).
76. European Food Safety Authority & European Centre for Disease Prevention and Control. The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019. *EFSA J*. **19**, e06490 (2021).
77. Karp, B. E. *et al*. National antimicrobial resistance monitoring system: two decades of advancing public health through integrated surveillance of antimicrobial resistance. *Foodborne Pathog. Dis*. **14**, 545–557 (2017).
78. Donado-Godoy, P. *et al*. The establishment of the Colombian Integrated Program for Antimicrobial Resistance Surveillance (COIPARS): a pilot project on poultry farms, slaughterhouses and retail market. *Zoonoses Public Health* **62**, 58–69 (2015).
79. Thai Working Group on Health Policy and Systems Research on Antimicrobial Resistance (HPSR-AMR). Thailand's one health report on antimicrobial consumption and antimicrobial resistance in 2020. (2021).
80. Diallo, O. O. *et al*. Antibiotic resistance surveillance systems: A review. *J. Glob. Antimicrob. Resist*. **23**, 430–438 (2020).
81. Simjee, S., McDermott, P., Trott, D. & Chuanchuen, R. Present and future surveillance of antimicrobial resistance in animals: principles and practices. *Antimicrob. Resist. Bact. Livest. Companion Anim*. 595–618 (2018).
82. Criscuolo, N. G., Pires, J., Zhao, C. & Van Boeckel, T. P. resistancebank. org, an open-access repository for surveys of antimicrobial resistance in animals. *Sci. Data* **8**, 189 (2021).
83. World Health Organization. *Global Surveillance for COVID-19 Caused by Human Infection with COVID-19 Virus: Interim Guidance, 20 March 2020*. (2020).

84. Fröhlich, G. M. *et al*. Hospital outcomes of community-acquired COVID-19 versus influenza: Insights from the Swiss hospital-based surveillance of influenza and COVID-19. *Eurosurveillance* **27**, 2001848 (2022).
85. Miller, I. F., Becker, A. D., Grenfell, B. T. & Metcalf, C. J. E. Disease and healthcare burden of COVID-19 in the United States. *Nat. Med.* **26**, 1212–1217 (2020).
86. Ashcroft, P. *et al*. COVID-19 infectivity profile correction. *ArXiv Prepr. ArXiv200706602* (2020).
87. Pigott, D. M. *et al*. Updates to the zoonotic niche map of Ebola virus disease in Africa. *Elife* **5**, e16412 (2016).
88. Wadhwa, V. Cartographies of disease: Maps, mapping, and medicine. *Vis. Stud.* **24**, 272–276 (2009).
89. Garcia-Abreu, A., Halperin, W. & Danel, I. Public health surveillance toolkit: a guide for busy task managers. (2002).
90. Ollerenshaw, C. The approach to forecasting the incidence of fascioliasis over England and Wales 1958–1962. *Agric. Meteorol.* **3**, 35–53 (1966).
91. Yeoman, G. Field vector studies of epizootic East Coast fever. II.-Seasonal studies of R. appendiculatus on bovine and non-bovine hosts in East Coast fever enzootic, epizootic and free areas. *Bull. Epizoot. Dis. Afr.* **14**, (1966).
92. Gilg, A. W. A study in agricultural disease diffusion: the case of the 1970-71 fowl-pest epidemic. *Trans. Inst. Br. Geogr.* 77–97 (1973).
93. Carpenter, T. The spatial epidemiologic (r) evolution: a look back in time and forward to the future. *Spat. Spatio-Temporal Epidemiol.* **2**, 119–124 (2011).
94. Perez, E., Herrero, M. V., Jimenez, C., Carpenter, T. E. & Buening, G. B. Epidemiology of bovine anaplasmosis and babesiosis in Costa Rica. *Prev. Vet. Med.* **20**, 23–31 (1994).
95. Gething, P. W. *et al*. A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malar. J.* **10**, 1–16 (2011).
96. Bollen, M. *et al*. Managing African swine fever: Assessing the potential of camera traps in monitoring wild boar occupancy trends in infected and non-infected Zones, using spatio-temporal statistical models. *Front. Vet. Sci.* **8**, 726117 (2021).
97. Kent, C. M. *et al*. Spatiotemporal changes in influenza A virus prevalence among wild waterfowl inhabiting the continental United States throughout the annual cycle. *Sci. Rep.* **12**, 13083 (2022).
98. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
99. Deneu, B. *et al*. Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. *PLoS Comput. Biol.* **17**, e1008856 (2021).
100. Maltamo, M. & Kangas, A. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. *Can. J. For. Res.* **28**, 1107–1115 (1998).
101. Hastie, T. & Tibshirani, R. Exploring the nature of covariate effects in the proportional hazards model. *Biometrics* 1005–1016 (1990).
102. Dukić, V. *et al*. The role of weather in meningitis outbreaks in Navrongo, Ghana: a generalized additive modeling approach. *J. Agric. Biol. Environ. Stat.* **17**, 442–460 (2012).
103. Thawer, S. G. *et al*. Spatio-temporal modelling of routine health facility data for malaria risk micro-stratification in mainland Tanzania. *Sci. Rep.* **13**, 10600 (2023).
104. Messina, J. P. *et al*. The current and future global distribution and population at risk of dengue. *Nat. Microbiol.* **4**, 1508–1515 (2019).

105. Sinka, M. E. *et al*. The dominant Anopheles vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic précis. *Parasit. Vectors* **4**, 1–46 (2011).
106. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, (2012).
107. Sibanda, W. & Pretorius, P. Artificial neural networks-a review of applications of neural networks in the modeling of HIV epidemic. *Int. J. Comput. Appl.* **44**, 1–9 (2012).
108. Niazkar, H. R. & Niazkar, M. Application of artificial neural networks to predict the COVID-19 outbreak. *Glob. Health Res. Policy* **5**, 1–11 (2020).
109. Manohar, B. & Das, R. Artificial neural networks for the prediction of monkeypox outbreak. *Trop. Med. Infect. Dis.* **7**, 424 (2022).
110. Lindgren, F. & Rue, H. Bayesian spatial modelling with R-INLA. *J. Stat. Softw.* **63**, (2015).
111. Rasmussen, C. E. & Williams, C. K. *Gaussian Processes for Machine Learning*. vol. 1 (Springer, 2006).
112. Moraga, P. *et al*. Bayesian spatial modelling of geostatistical data using INLA and SPDE methods: A case study predicting malaria risk in Mozambique. *Spat. Spatio-Temporal Epidemiol.* **39**, 100440 (2021).
113. Tumusiime, D. *et al*. Mapping the risk of Rift Valley fever in Uganda using national seroprevalence data from cattle, sheep and goats. *PLoS Negl. Trop. Dis.* **17**, e0010482 (2023).
114. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* **115**, 700–721 (1927).
115. Roberts, M. & Tobias, M. Predicting and preventing measles epidemics in New Zealand: application of a mathematical model. *Epidemiol. Infect.* **124**, 279–287 (2000).
116. Brauer, F. Mathematical epidemiology: Past, present, and future. *Infect. Dis. Model.* **2**, 113–127 (2017).
117. Pandey, G., Chaudhary, P., Gupta, R. & Pal, S. SEIR and Regression Model based COVID-19 outbreak predictions in India. *ArXiv Prepr. ArXiv200400958* (2020).
118. López, L. & Rodo, X. A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: simulating control scenarios and multi-scale epidemics. *Results Phys.* **21**, 103746 (2021).
119. Annas, S., Pratama, M. I., Rifandi, M., Sanusi, W. & Side, S. Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia. *Chaos Solitons Fractals* **139**, 110072 (2020).

Chapter 2

Trends in antimicrobial resistance in food animals in China and the proposal of a geographically-targeted surveillance approach

Chapter 2.1 Geographically-targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China

Authors:
Cheng Zhao[1], Yu Wang[1], Katie Tiseo[1], Joao Pires[1], Nicola G. Criscuolo[1], and Thomas P. Van Boeckel[1,2,*]

Affiliations:
[1]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland.
[2]Center for Disease Dynamics, Economics & Policy, Washington DC, United States.

*Correspondence to: thomas.vanboeckel@env.ethz.ch

Published in:
Nature Food 596–602 (2021). DOI: 10.1038/s43016-021-00320-x
This chapter contains minor revisions compared to the published article.

Contribution remarks:
C.Z., T.P.V.B. and J.P. developed the analytical framework. C.Z., Y.W. and K.T. conducted the literature review. C.Z. conducted the analysis and wrote the first version of the manuscript. N.G.C. incorporated all data into resistancebank.org. All authors contributed to the final version of the manuscript.

Abstract

The rise of antimicrobial resistance in animals is fueled by the widespread use of veterinary antimicrobials. China is the largest global consumer of veterinary antimicrobials – improving AMR surveillance strategies in this region could help prioritize intervention and preserve antimicrobial efficacy. Here, we mapped AMR rates in pigs, chickens and cattle in China using 446 surveys of event-based surveillance between 2000 and 2019 for foodborne bacteria, in combination with geospatial models to identify locations where conducting new surveys could have the highest benefits. Using maps of uncertainty, we show that eastern China currently has the highest AMR rates, and southwestern and northeastern China would benefit the most from additional surveillance efforts. Instead of distributing new surveys evenly across administrative divisions, using geographically targeted surveillance could reduce AMR prediction uncertainty by 2-fold. In a context of competing disease control priorities, our findings present a feasible option for optimizing surveillance efforts – and slowing the spread of AMR.

Introduction

Antimicrobials are used in agriculture as disease treatments, prophylactically to prevent infections in healthy animals, and to increase productivity [1]. However, the routine use of antimicrobials as surrogates for good hygiene practices on farms [2,3] is driving a rise in antimicrobial resistance (AMR), with increasingly serious consequences for animal health [1,4], and potentially human health [5,6].

Globally, 73% of antimicrobials are used in animals [7], with China being the largest consumer of antimicrobials in absolute terms (41,967 tons in 2017) and the second largest consumer in relative terms with 200 milligram used per kilogram of animal (mg/kg)[8] (Supplementary Figure 1A, 1B). In comparison, Denmark and the Netherlands use respectively 39 mg/Kg, and 56 mg/Kg [9] – while maintaining a productive livestock sector. Multiple factors may contribute to antimicrobial overuse in China. Meat production has grown by 560% since 1979 (FAOSTAT, http://www.fao.org/faostat/en/#data/QL), which could have made farmers reliant on antimicrobials to prevent infections. Veterinary antimicrobials are reportedly accessible without prescriptions [10] and are sold at low prices in comparison to other countries [11]. As in many other low- and middle-income countries (LMICs) [12,13], farmers predominantly obtain antimicrobials from local drug stores where vendors also provide medical advice without veterinary training [10,14]. Additionally, enforcing the existing regulations [10] on the compounds authorized in animals, or the recently announced ban on growth promoters [15] remain a formidable challenge in a country where 360 million people are active in agriculture (World Bank, https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS). In the last 5 years, China has reported multiple first emergence of resistance genes to last-resort antimicrobials such as colistin and tigecycline [16,17] and a recent global analysis suggested that China may have become one of the largest hotspots of resistance among LMICs [4], ranking 8th in relative terms, and 1st in absolute terms, for animal-associated burden of AMR amongst LMICs (Supplementary Figure 1D, 1E).

In high-income countries, epidemiological evidence collected by surveillance systems guides AMR responses and provides a baseline for evaluating policy targets. The US Food and Drug

Administration collects meat samples from retail and slaughterhouses to monitor AMR levels (https://www.fda.gov/animal-veterinary/antimicrobial-resistance/national-antimicrobial-resistance-monitoring-system); the European Food Safety Authority (EFSA, https://www.efsa.europa.eu/en/topics/topic/antimicrobial-resistance) serves a comparable role by amalgamating the surveillance efforts of its member states. To the best of our knowledge, the majority of LMICs – including China – either lack systematic surveillance systems or do not publicly report data from animal AMR surveillance[4]. Despite these challenges, China could act as a leader for guiding the international response to AMR – because its domestic policies may have far-reaching benefits for neighboring countries, and its numerous trading partners [18].

Point-prevalence surveys (PPS) published independently by veterinarians constitute an alternative source for documenting AMR trends (Criscuolo *et al*. submitted) – and inferences can be made to map AMR using a large collection of PPS[4]. However, adapting this approach to the Chinese context requires building a critical mass of PPS, including surveys in Chinese to train geospatial models. Accurate maps of disease prevalence have been generated [19–22], but few used the associated uncertainty maps to inform field sampling campaigns [23,24]. In particular, as prediction uncertainty grows with distance from existing surveys, an uncertainty map can help identify the location where conducting new surveys could be most valuable to improve the confidence level of a prevalence map. Repeating this process iteratively can guide long-term surveillance efforts.

Here, we used event-based surveillance data to map trends in AMR in animals and associated uncertainty levels. We identify regions where future surveillance efforts could be intensified to reduce uncertainty on the geographic distribution of AMR in China. In a context of competing disease control priorities, our approach helps optimally target the limited resources dedicated to event-based surveillance of AMR.

Results

Data

We identified 446 point-prevalence surveys (PPS) reporting antimicrobial resistance (AMR) in food animals in China between 2000 and 2019 (Supplementary Text S1). This corresponds to one survey per 470,177 tons of food animals annually (28th rank amongst low- and middle-income countries (LMICs); Supplementary Figure 1C). We collected data on four common indicator bacteria: *Escherichia coli* (184 PPS), nontyphoidal *Salmonella* spp. (131 PPS), *Staphylococcus aureus* (131 PPS), and *Campylobacter* spp. (33 PPS). The 446 PPS included 6,295 resistance rates. We defined a composite metric of AMR to summarize trends in resistance across multiple drugs, and bacteria. For each survey, we calculated the proportion of antimicrobial compounds with resistance higher than 50% (P50; Supplementary Figure 2).

Temporal trends

In pigs, between 2000 and 2019, P50 increased significantly in *E. coli* (+59%), *Salmonella* (+148%), and *S. aureus* (+85%) (Figure 1A, 1B, 1C). In contrast, in chicken, P50 was stable in *E. coli*, *Salmonella*, and *S. aureus*, with mean P50 of 60%, 42%, and 37%, respectively

(Figure 1D, 1E, 1F). In cattle, P50 increased significantly in *E. coli* (+167%; Figure 1G), and was stable in *Salmonella* and *S. aureus*, with mean P50 of 23% and 31%, respectively (Figure 1H, 1I).

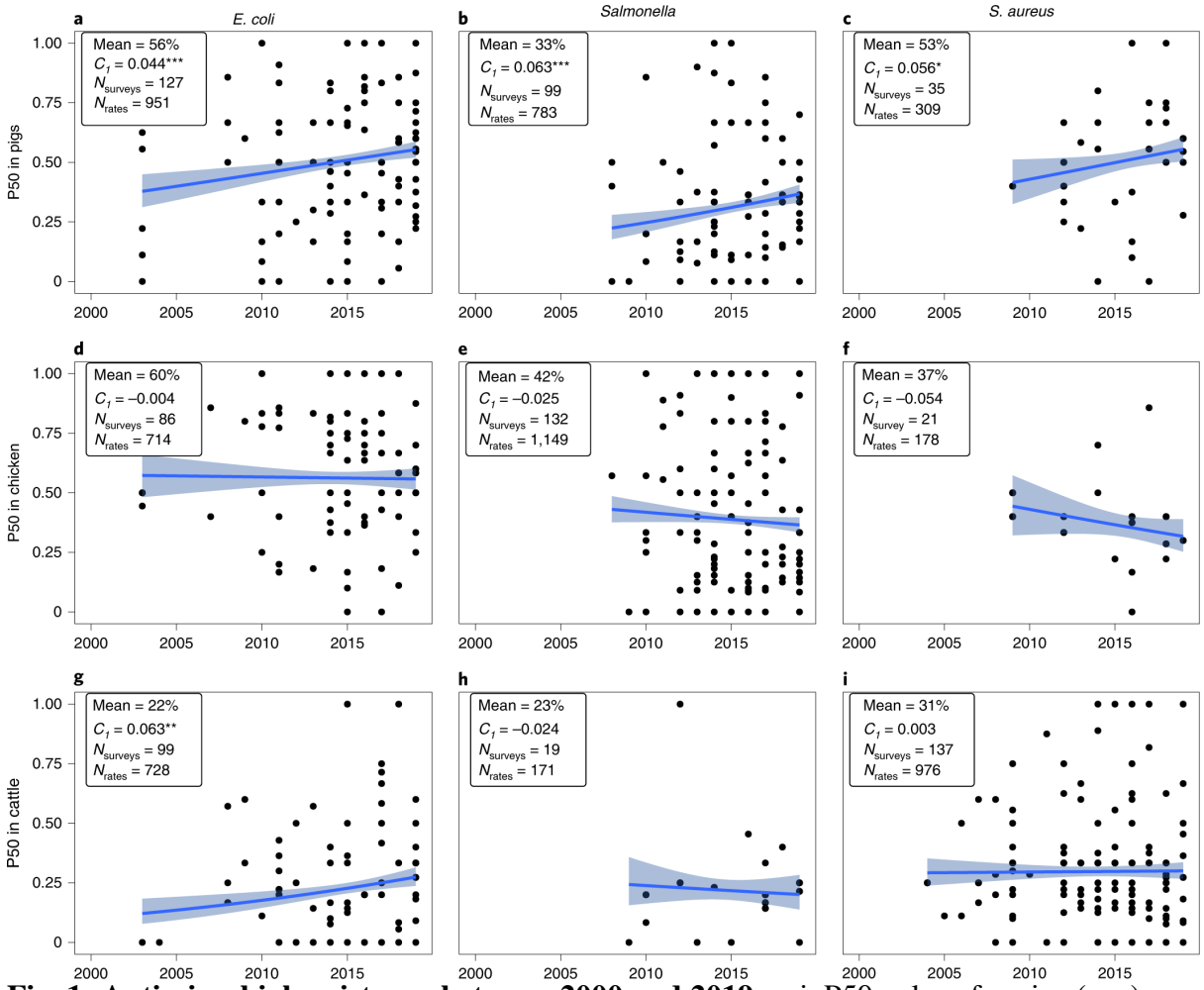


Fig. 1: Antimicrobial resistance between 2000 and 2019. a–i, P50 values for pigs (a–c), chicken (d–f) and cattle (g–i), showing resistance to *E. coli* (a,d,g), *Salmonella* (b,e,h) and *S. aureus* (c,f,i). Mean refers to the mean P50 value of all surveys. C_1 is the coefficient associated with the temporal trend in a logistic regression model weighted by $\log_{10}$-transformed sample size in each survey. Shaded areas indicate 95% confidence intervals. $^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$. $N_{surveys}$ is the number of surveys and N_{rates} is the total number of resistance rates reported in the surveys. Surveys conducted at multiple locations in the same publication are considered multiple surveys.

Prevalence of resistance across antimicrobial classes

For each drug-bacteria-animal combination, we estimated the prevalence of resistance (R%), and calculated the center of mass of the probability density distribution of the prevalence of resistance across PPS (Methods; Figure 2). Prevalence of resistance of tetracyclines, sulfonamides, and penicillins was high across all tested bacterial species between 2010 and 2019 (R% > 25%). In comparison, prevalence of resistance has remained at low levels in polymyxins and cephalosporins (R% < 10% for at least one bacterial species tested in one animal species). For all antimicrobial classes, prevalence of resistance in *E. coli* in chicken and pigs increased after 2010, except tetracyclines with already high prevalence of resistance

(R% > 90%) in pigs before 2010. In *Salmonella*, increase in the prevalence of resistance after 2010 was observed in penicillins in chicken, as well as in sulfonamides, penicillins, and tetracyclines in pigs.

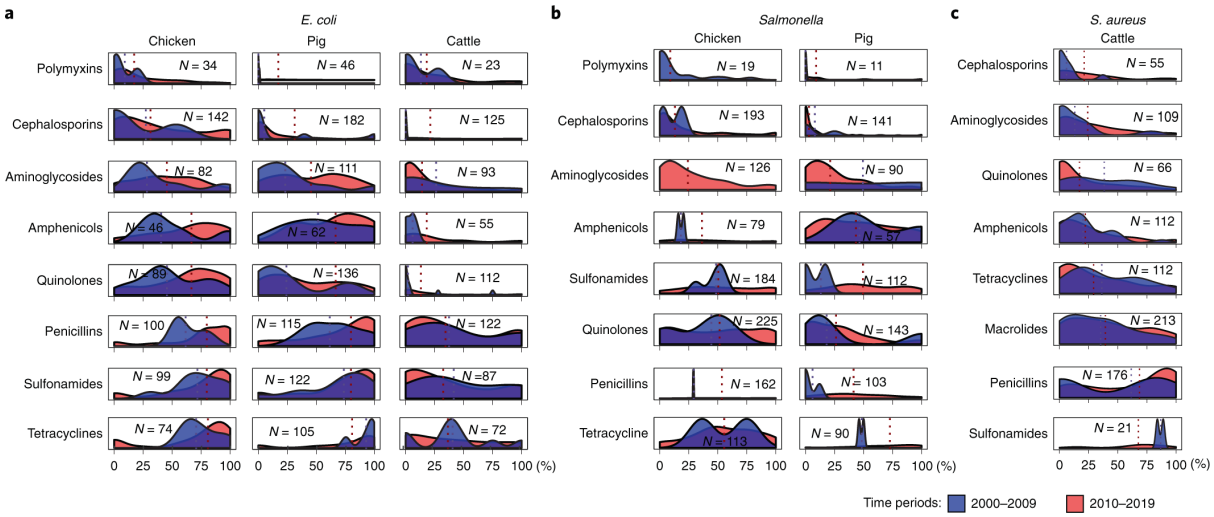


Fig. 2: Prevalence of resistance per antimicrobial class. a–c, Resistance to *E. coli* in chicken, pigs and cattle (a), to *Salmonella* in chicken and pigs (b) and to *S. aureus* in cattle (c). In each panel, the *x* axis represents resistance rates and the *y* axis represents the probability density. The area under the curve between two resistance rates represents the probability that resistance rates fall within the interval. *N*, number of resistance rates used to calculate the density distribution. Dashed lines represent the centre of mass of each distribution.

The prevalence of resistance in *E. coli* was higher than *Salmonella* for all antimicrobial classes (Figure 2A, 2B). Across drug classes, the prevalence of resistance in *E. coli* was 18% higher than the prevalence of resistance in *Salmonella* in chicken, and 16% higher than the prevalence of resistance in *Salmonella* in pigs. Prevalence of resistance for individual antimicrobial classes differed between chicken and pigs. For *E. coli*, cephalosporins and quinolones had respectively 20% and 27% higher prevalence of resistance in chicken compared with pigs, while prevalence of resistance in other antimicrobial classes differed by < 6% between chicken and pigs (Figure 2A). For *Salmonella*, quinolones had 25% higher prevalence of resistance in chicken compared to pigs, while for other antimicrobial classes, the difference in the prevalence of resistance between chicken and pigs was smaller than 12%. (Figure 2B). This comparison was largely influenced by the relative abundance of serotypes of *Salmonella* in different animal hosts (Supplementary Figure 3). However, an in-depth investigation on its influence on resistance trends was challenged by the fact that 70% of the surveys on *Salmonella* (93 out of 131 surveys) did not report the prevalence of resistance broken down by serotypes.

Geographic Distribution of Resistance

We used a geospatial model (Supplementary Text S2) to map P50 at 10 Km resolution, and combined information from PPS with environmental and anthropogenic covariates (Supplementary Table 1). Hotspots of AMR – regions where more than 40% of drugs have resistance levels above 50% (P50 > 40%) – were found in 1) eastern China in the areas of Heilongjiang, western Jilin, western Liaoning, southern Hebei, Shandong, eastern Jiangsu,

southern Anhui, Fujian and Taiwan, 2) central China in the areas of northern Shaanxi, central Hunan and southeastern Sichuan, and 3) the northwestern Xinjiang Uyghur Autonomous Region (Figure 3). Low levels of AMR (P50 < 30%) were found in Tibet Autonomous Region, northwestern Sichuan, and southern Guangxi (Figure 3). We measured the association between P50 and covariates, using the decrease in area under the receiver operator curve (AUC) by sequential permutation of each covariate (Supplementary Text S2). The most important covariates associated with P50 values were: the travel times to cities [25] (-16% AUC), the minimum monthly temperature [26] (-15% AUC), and cattle population density [27] (-13% AUC; Supplementary Figure 4).

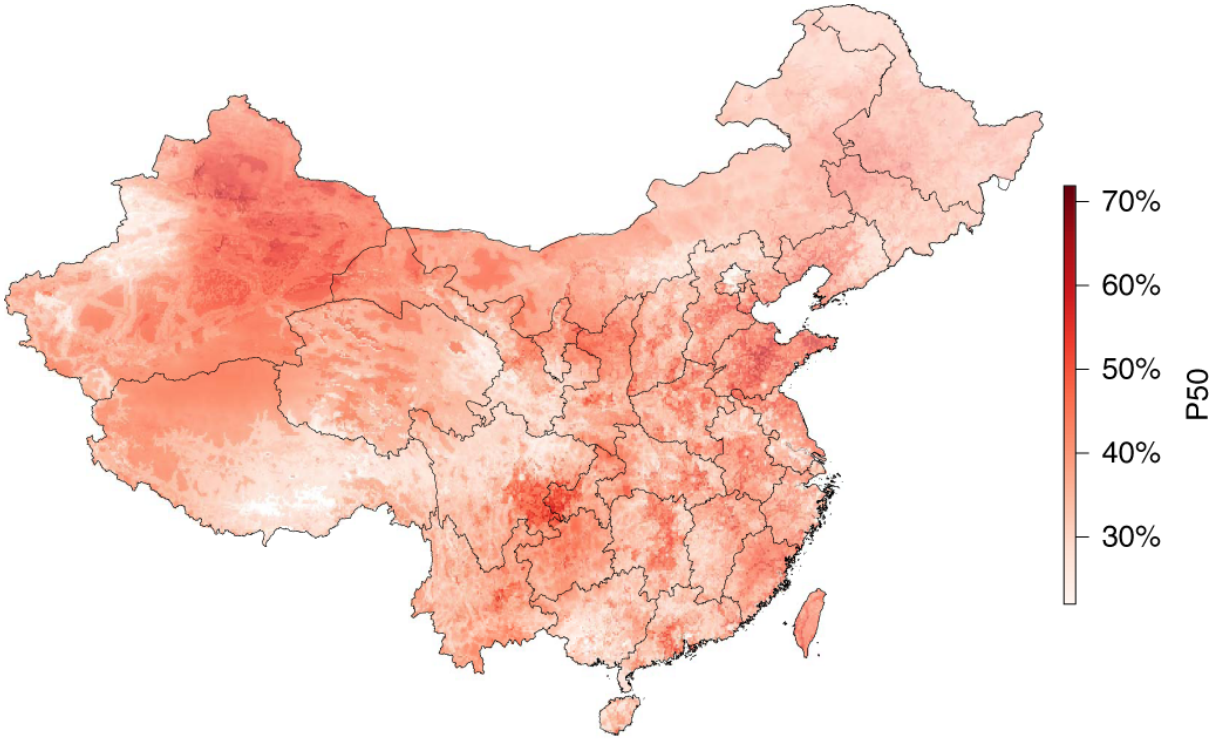


Fig. 3: Geographic distribution of antimicrobial resistance. Colour shading represents the P50 level.

Optimal location for future event-based surveillance efforts

We identified the locations of 50 hypothetical surveys to be conducted in China such that these would minimize uncertainty on the current map of AMR. The uncertainty was quantified using a map of "necessity for additional surveillance" (NS) – the product of the kriging variance (a metric of interpolation uncertainty) and the population density (Methods). The 50 locations for the hypothetical surveys were identified such as to minimize the mean of NS_i across all pixels i in the NS map.

We compared four approaches to distribute hypothetical future surveys (Methods): first a 'greedy' approach that tested all possible locations for new surveys but was associated with high computational cost. Second, an "overlap approach" based on mutual zones of exclusions for consecutive surveys to be conducted. This approach was a computational approximation to the greedy approach. Third, an 'administrative' approach where surveys were distributed equally across administrative divisions. These three approaches were compared with a "null model" consisting of randomly distributing 50 surveys across China. The greedy 'optimal'

approach achieved the greatest reduction of the mean necessity for additional surveillance (*NS*) (Figure 4B, dark red). The greedy approach reduced *NS* by 56% more than the null model (Figure 4B, blue). However, the greedy approach was associated with a considerable computational burden (Figure 4C, 4.5×10^5 CPU minutes). The overlap approach reduced the mean *NS* by 44% more than the null model (Figure 4B, blue), thus achieving near-optimal reduction of NS, but with a considerably lower computational burden than the greedy approach. The overlap approach also outcompeted the administrative approach (Figure 4B, green): it reduced the mean *NS* by 104% more than if surveys had been distributed equally between administrative divisions.

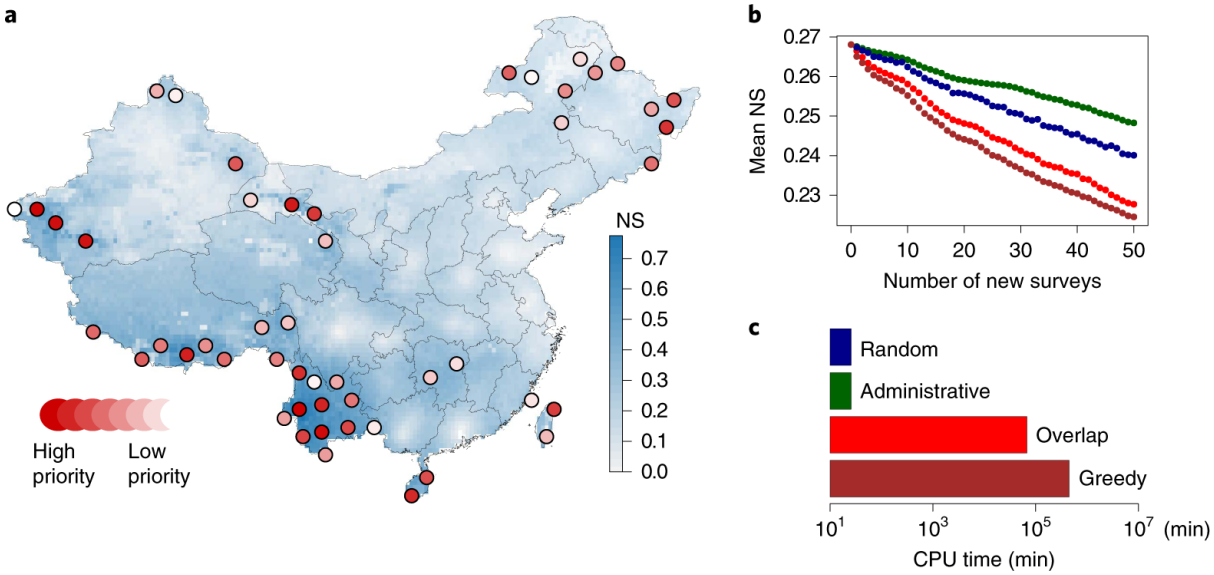


Fig. 4: Predicted locations for future surveys. a, Predicted optimal locations for future surveys using the 'overlap approach'. The background colour represents the 'necessity for additional surveillance' (NS): the product of the kriging variance and animal population density (standardized from 0 to 1). b, Reduction in the mean NS with 50 hypothetical additional surveys. The 50 additional survey locations were identified using the greedy approach (dark red), the overlap approach (red), the administrative approach (green) and the random approach (blue). c, Total central processing unit time for computing the four approaches ($\log_{10}$ scaled).

The overlap approach predicted locations for a large number of new surveys in the southwest (21/50 surveys) and northeast (11/50 surveys) of China. The surveys were predominantly distributed in Yunnan Province (10 surveys), Tibet Autonomous Region (9 surveys), Xinjiang Uyghur Autonomous Region (7 surveys), and Heilongjiang Province (5 surveys) (Figure 4A). These locations were determined using animal population densities as the metric of exposure (Methods). Additionally, we calculated the locations by province for individual animal species, respectively (Supplementary Figure 5). The locations were mainly distributed in Yunnan and Heilongjiang (exposure by chicken or pigs), and in Tibet Autonomous Region (exposure by cattle). If human population was considered to determine exposure (Supplementary Figure 6), then the locations predicted by the overlap approach to conduct new surveys were mainly distributed in Heilongjiang Province (8 surveys), Xinjiang Uyghur Autonomous Region (8 surveys), Yunnan province (7 surveys), and Inner Mongolia Autonomous Region (5 surveys).

Discussion

We identified geographical gaps in event-based surveillance of food animal AMR in China, using a map of AMR derived from 446 PPS, and its associated map of uncertainty, identifying where surveillance scale-up would be the most valuable to reduce uncertainties in the current trends of AMR.

Trends of AMR across animals and bacteria

Between 2000 and 2019, in pigs, P50 doubled in *Escherichia coli*, *Salmonella*, and *Staphylococcus aureus*. This increase in AMR occurred in a period of considerable intensification of pig production in China, and the number of pigs slaughtered in China increased by 45% [28]. Traditional backyard systems were gradually replaced by large-scale intensive farms to support the growing domestic demand for pork [29]. However, as in other countries currently transitioning from extensive to intensive farming, improvements in biosecurity may have lagged behind improvements in productivity [30]. Future improvements in biosecurity may reduce farmers' dependency on antimicrobials for disease prevention, and have potentially indirect benefits for managing AMR in the long-term. Future biosecurity improvements can reduce the risk of diseases introduction through strict hygiene requirements for personnel who enter the farms, appropriate carcass management, and reducing the spread of diseases inside the premises through establishing pig compartments, and regular cleaning and disinfection [31].

Between 2000 and 2019, in chicken, P50 remained stable in *E. coli*, *Salmonella*, and *S. aureus*, albeit at high levels. In 2000, P50 in *E. coli*, *Salmonella*, and *S. aureus* in chicken were already at 58%, 48%, and 56%, respectively, double the levels of resistance in pigs (35%, 15%, and 30%). This suggests that the intensification process (and the routine use of antimicrobials for production) occurred earlier and faster in the poultry sector than for pigs [32]. Excessive use of quinolones (e.g. norfloxacin and ofloxacin) and cephalosporins (e.g. ceftriaxone) in chicken [10] may have caused much higher resistance rates of these two antimicrobial classes in chicken, compared with pigs (Figure 2). Our analysis suggests that the antimicrobials that maintained low prevalence of resistance in chicken are expensive and are seldom available on the Chinese market (Supplementary Text S3), impeding overuse and also preventing the further AMR increases.

Prevalence of resistance for *E. coli* were higher than for *Salmonella* in pigs and chicken (Figure 2), possibly influenced by commensal *E. coli* being associated with lower resistance levels than pathogenic *E.coli* [33]. However, due to the non-systematic nature of the PPS sampling schemes (event-based surveillance), disentangling how resistance rates differ between bacteria exhibiting commensal or pathogenic behavior remains challenging. We attempted to mitigate this potential bias by focusing our analysis exclusively on bacteria isolated from healthy animals.

Resistance levels (P50) in cattle were lower than in chicken and pigs (Figure 1). However, P50 in *E. coli* grew by 81% between 2000 and 2019, while globally the P50 in cattle was stable over the same period [4]. This may be associated with the increasing demand for cattle product in China – cow milk production increased by 261% from 2000 to 2019 [28]. Despite this rapid expansion, the current per capita consumption of dairy products in China is still only one fifth of the dairy consumption in the US and the EU [28] – leaving room for further

expansion. Thus, a window of opportunity may exist at the current stage to slow the rise of AMR in cattle, while resistance rates are still low (22% in *E. coli*) – and immediate action could help secure a sustainable dairy intensification.

Improved maps of AMR in China

Currently, AMR levels in animals are the highest in the east (43%), moderately high in the northwest (40%), and lowest in the southwest (34%; Figure 3). These geographical trends are in agreement (Pearson correlation coefficient 0.48) with the previous attempts to map AMR in China [4]. However, the present map is considerably more robust because it is exclusively based on surveys conducted in China (446 surveys, including 318 publications in Chinese). In comparison, previous maps were produced with just 101 surveys from China supplemented by surveys from other LMICs [4]. The revised maps of AMR help identify hotspots of AMR (Figure 3) where intervention could be targeted immediately as part of domestic policies [34]. Travel time to cities was the factor with the highest influence on resistance levels [25]. The clustering of intensive farms in major consumption centers during industrialization [35], and the ease of access to drug stores in peri-urban areas [10,36] may drive AMR level upwards[37]. High AMR levels were also associated with high minimal monthly temperature [26] – high temperatures may lead to increased stress and conflicts among animals, with risk of animal injuries requiring antimicrobial treatment [38].

Key locations for conducting event-based surveillance

Amongst LMICs, China ranks 28th for the number of surveys in event-based surveillance per kilogram of food animals (population corrected units of food animals; PCU), and 36th for the number of surveys per PCU relative to average resistance level (P50) per country (Supplementary Figure 1). We identified locations where additional surveys on AMR in animals could be conducted in the future to minimize uncertainty associated with the geographical trends in AMR – representing a gain in information given the resources spent on event-based surveillance. Current patterns resulting from event-based surveillance are "sub-optimal" as surveys are clustered around veterinary institutes, mainly in the east (Supplementary Figure 7), where sampling to investigate AMR in their vicinity is easier (Supplementary Text S4) – and may have contributed to geographical information gaps on AMR trends in the southwest and northeast. Cross-provincial efforts between institutes are needed to coordinate future event-based surveillance efforts into these regions, which may be far from existing institutes, but where the gain in information by additional surveys would be the highest.

Our approach for assigning future surveys works by minimizing an index of "need for additional surveillance (*NS*)" based on a map of AMR (China for this example). However, testing exhaustively all possible location for future surveys ("greedy approach") incurs considerable computational cost. We developed an 'overlap approach' which is a rapidly implementable approximation of a greedy approach. The overlap approach achieved 93% of the reduction of the uncertainty in AMR trends achieved by the greedy approach, albeit using just 15% of total computation time required by the greedy approach (Methods). This not only makes the approach faster but also applicable with limited computational resources, and was developed in the context of event-based surveillance, which was abundant in China with 446 PPS and served as a proof of concept. However, the approach could also be used with systematic surveillance data or in other countries with event-based surveillance. In addition, the 'overlap approach' is flexible with respect to exposure. In this analysis we used animal

densities as metric of exposure, but this variable could easily be substituted by other criteria that are relevant for epidemiological or environmental assessments.

Limitations and Future Directions

Although steps are taken (Supplementary Text S1) to ensure comparability between surveys, there remain potential sources of bias in variations in the accuracy of susceptibility testing. These include potential difference in laboratory equipment, and compliance to analysis protocols across regions in China. The World Health Organization assesses the quality of antimicrobial susceptibility testing across countries [39], but to the best of our knowledge, such within-country assessment is not currently available to account for laboratory practices that could lead to variations in the accuracy of susceptibility testing. These sources of experimental variation between surveys may influence the accuracy of the spatial distribution of P50. Inherent to event-based surveillance, a subjective summary metric "P50" was used in the absence of publicly available systematic surveillance data. P50 could be affected by the different antimicrobials subject to susceptibility testing in each survey. The potential bias was reduced by using the drug-bacteria combinations recommended by the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance[40] to calculate P50. Insufficient and irregular geographic coverage of data points may affect the accuracy of the estimations of model parameters. The risk of local overfitting is attenuated by using spatial cross-validation in the models. Finally, future mapping efforts could integrate surveys on AMR in aquaculture, because aquatic animals are important food animals in China, with at least 20 antimicrobials involved in their production [41]. Complementary to phenotypic resistance, AMR surveillance could be expanded to include genomics data, through metagenomic analysis of wastewater [42] from farms, although issues about harmonization remain an active field of analysis (Pires *et al*. submitted). In addition, the current analysis was focused on exploring the potential benefits of conducting additional PPS. Future analysis could be expanded to encompass variations in the costs associated with PPS. These costs may involve accessing the predicted locations, human resources, sampling and laboratory expenses, and coordination between institutes. Incorporating these cost considerations into the analysis will contribute to the development of a more comprehensive strategy for the allocation of resources to PPS.

The health challenges that China currently faces are multifaceted and burdensome, both in humans (e.g. COVID-19 [43]), and in food animals (e.g. African Swine Fever [44]). With limited resources to allocate between competing priorities for disease surveillance, our approach identifies locations where conducting new surveys of AMR in animals could have the highest benefits, particularly in southwestern and northeastern China. Timely policy intervention could curb AMR in China, as illustrated by the significant reduction in colistin resistance after the colistin withdrawal policy [45]. Our analysis helps to optimally deploy the limited resources dedicated to event-based surveillance of AMR – improving chances for successful intervention for curbing AMR development and providing data to inform policy.

Methods

Data

We reviewed point-prevalence surveys (PPS) reporting rates of antimicrobial resistance (AMR) in healthy animals, and animal food products in China between 2000 and 2019 (Supplementary Text S1). We focused on three common food animal species, including chicken, pigs, and cattle. Here, dairy cattle and meat cattle were pooled in this study, in consistency with the categorization adopted in the maps of livestock created by the Food and Agriculture Organization[27]. The review focused on four common foodborne bacteria: *Escherichia coli*, nontyphoidal *Salmonella*, *Staphylococcus aureus* and *Campylobacter*. We recorded resistance rates reported in PPS, defined as the percentage of isolates tested resistant to an antimicrobial compound. In addition, we extracted the anatomical therapeutic chemical classification codes of the drugs tested, the year of publication, the guidelines used for susceptibility testing, the latitude and longitude of sampling sites, the number of samples collected, and the host animals. We recorded sample types for each survey, including live animals, slaughtered animals, animal products, and fecal samples. Each sample was taken from one animal or animal product. These sample types were pooled in the current analysis. 10,747 rates of AMR were extracted from 446 surveys (Supplementary Figure 8), including 318 surveys from China's National Knowledge Infrastructure (CNKI), the leading Chinese-language academic search engine. All data extracted in the review are available at https://resistancebank.org.

Two steps were taken to ensure comparability of the resistance rates extracted from the surveys. First, the panel of drug-bacteria combinations extracted from each survey was that recommended for susceptibility testing by the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR) [40]. This resulted in the extraction of 6,295 resistance rates for 76 drug-bacteria combinations. Second, resistance rates were harmonized using a methodology[4] accounting for potential variations in the clinical breakpoints used for antimicrobial susceptibility testing (Supplementary Text S1). There are two major families of methods used for susceptibility testing in this dataset -diffusion methods (e.g. disc diffusion) and dilution methods (e.g. broth dilution). Previous works have shown good agreement between the two approaches in measuring resistance in foodborne bacteria[4,46]. For each family of methods, variations of breakpoints may result from differences between laboratory guidelines systems (European Committee on Antimicrobial Susceptibility Testing; EUCAST vs Clinical and Laboratory Standards Institute; CLSI), or from variations over time of clinical breakpoints within a laboratory guidelines system (CLSI or EUCAST). Here, we accounted for both situations using distributions of minimum inhibitory concentrations, and inhibition zones obtained from *eucast.org* (Supplementary Text S1).

Trends in Antimicrobial Resistance

We defined a composite metric of AMR to summarize trends in resistance across multiple drugs and bacterial species. For each survey, we calculated the proportion of antimicrobial compounds with resistance higher than 50% (P50). For each animal-bacteria combination, we assessed the significance of the temporal trends of P50 between 2000 to 2019 using a logistic regression model, weighted by the log10 transformed number of samples in each survey.

For each bacteria-drug (antimicrobial class) combination, we estimated prevalence of resistance by calculating a curve of the distribution of resistance rates across all surveys (Figure 2). The analysis was conducted for surveys published between 2000 and 2009, and between 2010 and 2019, respectively. The distribution was estimated at 100 equally spaced intervals from resistance rates of 0% to 100%, using kernel density estimation. We used the center of mass of the density distribution to estimate prevalence of resistance. The calculation

was conducted for six animal-bacteria combinations. This included *E. coli* in chicken, pigs and cattle, as well as *Salmonella* in chicken and pigs, and *S. aureus* in cattle. The remaining animal-bacteria combinations were excluded due to limited sample size, only represented in 32 out of 446 PPS. The analysis was restricted to antimicrobial classes represented by at least 10 resistance rates. In addition, we estimated the association between resistance rates and the ease of obtaining antimicrobials from the market, using data from online stores (Supplementary Text S3).

Geospatial modelling

We interpolated P50 values from the survey locations to create a map of P50 at a resolution of 10 x 10 Km across China. The approach followed a two-step procedure [47]: In step 1, three 'child models' were trained using four-fold spatial cross validation to quantify the relation between P50 and environmental and anthropogenic covariates (Supplementary Text S2; Supplementary Table 1). In step 2, the predictions of the child models were stacked using universal kriging (Supplementary Text S2). This approach combined the ability of the child models to capture interactions and non-linear relationships between P50 and environmental and anthropogenic covariates, as well as the ability to account for spatial-autocorrelation in the distribution of P50.

The outputs of the two-step procedure were: a map of P50 (Figure 3), and a map of uncertainty on the P50 predictions (Supplementary Figure 9, Supplementary Text S2). The overall accuracy of the geospatial model was evaluated using the area under the receiver operator curve (AUC). The contribution of each covariate was evaluated by permuting sequentially all covariates, and calculating the reduction in AUC compared with a full model including all covariates (Supplementary Figure 4). The administrative boundaries used in all maps were obtained from the Global Administrative Areas database (http://www.gadm.org).

Identifying (optimal) locations for future surveys on AMR

We identified the locations of 50 hypothetical new surveys – the rounded average number of surveys conducted per year (54 surveys/year) between 2014 and 2019 in China. The location of each new survey was determined recursively such that it minimized the overall uncertainty levels on the geographical trends in AMR across the country. This process took into account the locations of existing surveys, as well as the location of each additional hypothetical survey. The objective of this approach was to maximize gain in information about AMR given the resource invested in conducting surveys.

The map of uncertainty consisted of the variance in the child model predictions $Var(P_{BRT}, P_{LASSO-GLM}, P_{FFNN})$ (step 1) across 10 Monte Carlo simulations, and the kriging variance Var_K (step 2):

$$Var_{total} = Var(P_{BRT}, P_{LASSO-GLM}, P_{FFNN}) + Var_K$$

In this study, the location of hypothetical surveys was solely based on Var_K, instead of the sum of both terms. This approach was preferred, because including both terms would have required to hypothesize P50 values associated with the surveys to be conducted in the future, adding an additional source of uncertainty that cannot be quantified. In any case, the

uncertainty attributable to Var_K was 4.1 times of $Var(P_{BRT}, P_{LASSO-GLM}, P_{FFNN})$ (Supplementary Text S2).

The allocation of new surveys was based on a map of "necessity for additional surveillance" (NS), defined as:

$$NS = Var_K \cdot W$$

Where Var_K reflected the uncertainty of the spatial interpolation, and W is log10 transformed population density of humans [48], animals[27] in total, as well as in chicken, pigs, and cattle, separately, which reflected exposure (Supplementary Figure 10). Here, animal population density was calculated as the sum of population corrected units (PCU) of pigs, chicken and cattle, using methods described in Van Boeckel et al. 2017 [7]. We adjusted the values of W, such that its density distribution equals that of Var_K. Concretely, for each pixel i, we calculated the quantile of Wi on the map of W, and replaced the value by the corresponding value of Var_K at the same quantile. Var_K and W were both standardized to range [0,1], thus given each term equal weight in the need for surveillance.

Four approaches were used to distribute 50 surveys across China based on the map of NS. The reduction in uncertainty on AMR level associated with each of the four spatial configurations of the hypothetical surveys was evaluated, by calculating the reduction in the mean values of NS across 7,857 possible pixels on the map of China.

First, we used a "greedy" approach where all possible locations for additional surveys were tested. Concretely, the first hypothetical survey was placed at each of the 7,857 possible pixel locations, and a revised map of $NS_{(+1\ survey)}$ was calculated for each of the placements. The survey was eventually placed in the pixel that led to the largest reduction in $NS_{(+1\ survey)}$. The map of NS was then revised to account for the reduction in uncertainty in the neighborhood of the new survey. The process was repeated recursively for the next hypothetical surveys (2nd -50th). This approach, by definition, yields the optimal set of locations to reduce uncertainty, but it also bears a considerable computational burden, because every possible location is tested (N_{pixels}= 7,857) by the geospatial model for each hypothetical survey.

The second approach developed was a computational approximation to the greedy approach, hereafter referred to as the 'overlap approach'. This approach exploits a key feature of the kriging procedure: the decrease of the kriging variance (Var_K) with increasing proximity to existing survey locations. Each additional survey reduces the variance of the geospatial model at its own location, but also in its surrounding area (Supplementary Figure 11). The 'overlap approach' selects an optimal set of locations that reflect a compromise between high local NS and distance to other surveys. It iteratively selects new locations based on the highest local NS penalized by the degree of overlap between the hypothetical new surveys and existing surveys (Supplementary Figure 12). The first survey was placed at the location X_p, Y_p with the highest local NS (Supplementary Figure 12, Part 1). Then, the value of NS at each pixel location X_i, Y_i was recalculated as (Supplementary Figure 12, Part 2):

$$NS_{(+1\ survey)\ X_i,Y_i} = NS_{X_i,Y_i} \times \left(1 - \frac{overlap\ area}{neighborhood\ area}\right)$$

Where the neighborhood area was the circular area of decreased kriging variance around a new survey, and its radius was the distance until which NS decreased due to this new survey; "Overlap area" is the shared area of the neighborhoods of location X_p, Y_p and of location X_i, Y_i. The radius of the neighborhood was determined using a sensitivity analysis, optimized by approximate Bayesian computation (sequential Monte Carlo) [49] (ABC-SMC; Supplementary Text S5). The optimal neighborhood radius was chosen such as it minimizes reduction in NS across all pixels. The procedure (Supplementary Figure 12, Part 1 and Part 2) was repeated recursively for the hypothetical surveys (2^{nd} -50^{th}).

The third approach tested consisted of distributing surveys equally between provinces, to reflect a common approach to disease surveillance based on equal allocation of resources between administrative entities. Here, 22 provinces with the highest human population were assigned 2 surveys, and the remaining 6 provinces were assigned one survey per province. The exact location of each survey was randomly selected inside a province. Finally, all approaches were compared with the fourth approach (the random approach) as a 'null-model', in which the 50 hypothetical surveys were located randomly across the country without any geographic weighting criteria. The reduction in NS associated with the third and fourth approach, which was compared to the greedy approach and overlap approach, was the average over 50 simulations.

References

1. Silbergeld, E. K., Graham, J. & Price, L. B. Industrial food animal production, antimicrobial resistance, and human health. *Annual Review of Public Health* **29**, 151–169 (2008).
2. Coyne, L. *et al*. Characterizing Antimicrobial Use in the Livestock Sector in Three South East Asian Countries (Indonesia, Thailand, and Vietnam). *Antibiotics* **8**, 33 (2019).
3. Mohsin, M. *et al*. Excessive use of medically important antimicrobials in food animals in Pakistan: a five-year surveillance survey. *Global Health Action* **12**, 1697541 (2019).
4. Van Boeckel, T. P. *et al*. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, 1266-+ (2019).
5. Aarestrup, F. M. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140085 (2015).
6. Nordstrom, L., Liu, C. M. & Price, L. B. Foodborne urinary tract infections: a new paradigm for antimicrobial-resistant foodborne illness. *Frontiers in microbiology* **4**, 29 (2013).
7. Van Boeckel, T. P. *et al*. Reducing antimicrobial use in food animals. *Science* **357**, 1350–1352 (2017).
8. Tiseo, K., Huber, L., Gilbert, M., Robinson, T. P. & Van Boeckel, T. P. Global trends in antimicrobial use in food animals from 2017 to 2030. *Antibiotics* **9**, 918 (2020).
9. European Medicines Agency. *Sales of veterinary antimicrobial agents in 31 European countries in 2017*. (2019).
10. Xu, J., Sangthong, R., McNeil, E., Tang, R. & Chongsuvivatwong, V. Antibiotic use in chicken farms in northwestern China. *Antimicrobial Resistance & Infection Control* **9**, 1–9 (2020).

11. Van Boeckel, T. P. *et al*. Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences* **112**, 5649–5654 (2015).
12. Carrique-Mas, J. J. *et al*. Antimicrobial usage in chicken production in the Mekong Delta of Vietnam. *Zoonoses and public health* **62**, 70–78 (2015).
13. Ström, G. *et al*. Antimicrobial use and antimicrobial susceptibility in Escherichia coli on small-and medium-scale pig farms in north-eastern Thailand. *Antimicrobial Resistance & Infection Control* **6**, 75 (2017).
14. Dyar, O. J. *et al*. Knowledge, attitudes and practices relating to antibiotic use and antibiotic resistance among backyard pig farmers in rural Shandong province, China. *Preventive Veterinary Medicine* **175**, 104858 (2020).
15. Xiao, Y. & Li, L. China's national plan to combat antimicrobial resistance. *The Lancet Infectious Diseases* **16**, 1216–1218 (2016).
16. He, T. *et al*. Emergence of plasmid-mediated high-level tigecycline resistance genes in animals and humans. *Nature Microbiology* (2019) doi:10.1038/s41564-019-0445-2.
17. Liu, Y.-Y. *et al*. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infectious Diseases* **16**, 161–168 (2016).
18. Hoffman, S. & Behdinan, A. Towards an international treaty on antimicrobial resistance. *Ottawa Law Review* **47**, (2016).
19. Bhatt, S. *et al*. The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
20. Gething, P. W. *et al*. A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria journal* **10**, 378 (2011).
21. Gilbert, M. *et al*. Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4769–4774 (2008).
22. Kraemer, M. U. *et al*. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. *elife* **4**, e08347 (2015).
23. Kabaghe, A. N. *et al*. Adaptive geostatistical sampling enables efficient identification of malaria hotspots in repeated cross-sectional surveys in rural Malawi. *PLoS One* **12**, (2017).
24. Pacheco, R. A. *et al*. Finding hotspots: development of an adaptive spatial sampling approach. *medRxiv* (2020).
25. Weiss, D. J. *et al*. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333-+ (2018).
26. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315 (2017).
27. Gilbert, M. *et al*. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data* **5**, 180227 (2018).
28. Food and Agriculture Organization of the United Nations. FAOSTAT Statistical Database. (2020).
29. Bai, Z. H. *et al*. Changes in Pig Production in China and Their Effects on Nitrogen and Phosphorus Use and Losses. *Environmental Science & Technology* **48**, 12742–12749 (2014).
30. Saksena, S. *et al*. Evidence for the convergence model: the emergence of highly pathogenic avian influenza (H5N1) in Viet Nam. *PLoS One* **10**, e0138138 (2015).
31. Dewulf, J. & Van Immerseel, F. *Biosecurity in animal production and veterinary medicine: from principles to practice*. (Acco, 2018).
32. Gilbert, M. *et al*. Income disparities and the global distribution of intensively farmed chicken and pigs. *PLoS One* **10**, (2015).

33. Chantziaras, I. Antimicrobial resistance prevalence among pathogenic and commensal Escherichia coli from food-producing animals in Belgium (vol 5, pg 232, 2014). *VLAAMS DIERGENEESKUNDIG TIJDSCHRIFT* **83**, 325–325 (2014).
34. Ying, G.-G. *et al*. China Must Reduce Its Antibiotic Use. *Environmental Science & Technology* **51**, 1072–1073 (2017).
35. Steinfeld, H. *et al*. *Livestock's long shadow: environmental issues and options*. (Food & Agriculture Org., 2006).
36. Chen, J., Wang, Y., Chen, X. & Hesketh, T. Widespread illegal sales of antibiotics in Chinese pharmacies–a nationwide cross-sectional study. *Antimicrobial Resistance & Infection Control* **9**, 1–8 (2020).
37. Huber, L. *et al*. Geographic drivers of antimicrobial use and resistance in pigs in Khon Kaen province, Thailand. *Frontiers in Veterinary Science* **8**, 331 (2021).
38. Diana, A., Manzanilla, E. G., Calderón Díaz, J. A., Leonard, F. C. & Boyle, L. A. Do weaner pigs need in-feed antibiotics to ensure good health and welfare? *PloS one* **12**, e0185622 (2017).
39. World Health Organization. External Quality Assurance System (EQAS) of the WHO Global Foodborne Infections Network. (2015).
40. World Health Organization. Integrated surveillance of antimicrobial resistance in foodborne bacteria: Application of a One Health Approach. (2017) doi:https://www.who.int/foodsafety/publications/agisar_guidance2017/en/.
41. Liu, X., Steele, J. C. & Meng, X.-Z. Usage, residue, and human health risk of antibiotics in Chinese aquaculture: a review. *Environmental Pollution* **223**, 161–169 (2017).
42. Hendriksen, R. S. *et al*. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature communications* **10**, 1124 (2019).
43. Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *The Lancet* (2020).
44. Zhou, X. *et al*. Emergence of African Swine Fever in China, 2018. *Transboundary and Emerging Diseases* **65**, 1482–1484 (2018).
45. Wang, Y. *et al*. Changes in colistin resistance and mcr-1 abundance in Escherichia coli of animal and human origins following the ban of colistin-positive additives in China: an epidemiological comparative study. *The Lancet Infectious Diseases* (2020).
46. Bengtsson, S., Bjelkenbrant, C. & Kahlmeter, G. Validation of EUCAST zone diameter breakpoints against reference broth microdilution. *Clinical Microbiology and Infection* **20**, O353–O360 (2014).
47. Golding, N. *et al*. Mapping under-5 and neonatal mortality in Africa, 2000-15: a baseline analysis for the Sustainable Development Goals. *Lancet* **390**, 2171–2182 (2017).
48. Doxsey-Whitfield, E. *et al*. Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Papers in Applied Geography* **1**, 226–234 (2015).
49. Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. H. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* **6**, 187–202 (2009).

Chapter 2.2 Twenty-year trends in antimicrobial resistance from aquaculture and fisheries in Asia

Authors:
Daniel Schar[1,*], Cheng Zhao[2], Yu Wang[2], D. G. Joakim Larsson[3,4], Marius Gilbert[1,4], Thomas P. Van Boeckel[1,5]

Affiliations:
[1]Spatial Epidemiology Laboratory, Université Libre de Bruxelles, Brussels, Belgium
[2]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland
[3]Center for Antibiotic Resistance Research, University of Gothenburg, Gothenburg, Sweden
[4]Fonds National de la Recherche Scientifique, Brussels, Belgium
[5]Center for Diseases Dynamics, Economics, and Policy, New Delhi, India

*Correspondence to: dlschar@gmail.com; thomas.vanboeckel@env.ethz.ch

Published in:
Nature Communications 12, 5384 (2021). DOI: 10.1038/s41467-021-25655-8

Contribution remarks:
D.S., M.G. and T.P.V.B. designed the research; D.S., C.Z. and Y.W. collected the data; D.S., C.Z., D.G.J.L., M.G. and T.P.V.B. analyzed data; D.S., C.Z., D.G.J.L., M.G. and T.P.V.B. wrote the paper.

Abstract

Antimicrobial resistance (AMR) is a growing threat to human and animal health. However, in aquatic animals—the fastest growing food animal sector globally—AMR trends are seldom documented, particularly in Asia, which contributes two-thirds of global food fish production. Here, we present a systematic review and meta-analysis of 749 point prevalence surveys reporting antibiotic-resistant bacteria from aquatic food animals in Asia, extracted from 343 articles published in 2000–2019. We find concerning levels of resistance to medically important antimicrobials in foodborne pathogens. In aquaculture, the percentage of antimicrobial compounds per survey with resistance exceeding 50% (P50) plateaued at 33% [95% confidence interval (CI) 28 to 37%] between 2000 and 2018. In fisheries, P50 decreased from 52% [95% CI 39 to 65%] to 22% [95% CI 14 to 30%]. We map AMR at 10-kilometer resolution, finding resistance hotspots along Asia's major river systems and coastal waters of China and India. Regions benefitting most from future surveillance efforts are eastern China and India. Scaling up surveillance to strengthen epidemiological evidence on AMR and inform aquaculture and fisheries interventions is needed to mitigate the impact of AMR globally.

Introduction

Aquaculture and fisheries contribute a growing share of nutrition for the global population[1]. Aquatic animals provide 20% of animal protein to the human diet for over 40% of the world, with consumption growth outpacing rates for all other sources of animal protein combined[1]. Driven by increasing demand, global fish production is experiencing rapid growth. From 1960 to 2018, aquatic animal production for human consumption increased from 21.8 to 156.4 million tons[1,2]. Asia contributes the largest share—69% in 2018—with China alone representing 35% of global production[1].

As capture fisheries production has plateaued since the early 1990s, aquaculture production has risen commensurate with global demand and now accounts for the majority (52%) of aquatic food animal production with continued expansion expected through 20301. Growth in cultured aquatic animals averaged 5.3% annually since 2001 with select countries—notably Indonesia and Bangladesh—exceeding 9% annual growth.

The rapid growth in animal protein production has been facilitated by a transition from extensive to intensive farming, which in terrestrial food animal sectors has historically been accompanied by the increasing use of antimicrobials[3,4]. As a result of this global shift in animal production, and growing demand for animal-source foods, the terrestrial and aquatic food animal production industries have emerged as the largest consumer by volume (73.7% and 5.7%, respectively) of antimicrobials globally[3,4,5]. In aquaculture, some species of fish, such as catfish, are associated with antimicrobial use rates per kilogram that exceed those in terrestrial animals and humans[5].

Antimicrobial use exerts selective pressures driving antimicrobial resistance. In terrestrial animals, a growing body of evidence has linked AMR with productivity loss and resistant infections carrying harmful consequences to animal and human health[6,7,8,9]. In aquaculture and fisheries—although more limited than in terrestrial animals—the evidence concerning antimicrobial resistance has also expanded over the last decade[10,11,12,13]. Documenting the movement of resistance determinants in aquatic settings presents challenges, and the body of evidence from which to draw conclusions on antimicrobial resistance transference between aquatic animals and humans remains very limited. Yet the aquatic food animal supply chain may be an under-appreciated route for transmission of antimicrobial-resistant bacteria and resistance genes from aquatic animals and their environment to humans[14,15,16,17,18]. Mobile genetic elements carrying resistance genes of human clinical significance have been associated with aquaculture and the aquatic environment[14,18,19]. Aquatic food animal supply chains are highly globalized[1], facilitating the distribution of locally generated resistance at a global scale[20,21]. In addition, compared with other animal source foods, aquatic animal products are more likely to be consumed raw, increasing the risk of pathogen transmission.

Rising AMR rates are expected to disproportionately affect low-income and middle-income countries, jeopardizing development gains in vulnerable communities, widening economic inequality, and contributing to a rise in extreme poverty by 2030[22]. And resistance in pathogens of production significance may reduce treatment options in commercial aquaculture, with potential implications for food security and nutrition[13,23,24].

Strengthening surveillance to guide AMR interventions is a challenge across sectors. This challenge is particularly acute in the aquaculture and fisheries industry. Presently, even the most heavily consumed aquatic animals globally—freshwater and marine fish—are generally not subject to systematic foodborne pathogen surveillance. Point prevalence surveys provide evidence at discrete geographic and temporal scales and have been used to characterize the global distribution and burden of infectious disease in humans[25] and AMR in terrestrial animals[26]. These surveys may serve as a surrogate in the absence of routine, systematic surveillance, collectively providing a mosaic portrait of antimicrobial resistance trends. Enhanced documentation of AMR trends could then inform targeted surveillance programs and interventions in the world's most productive aquaculture and fisheries region[20,23,27].

Here, we summarize current evidence on AMR in aquatic food animals in Asia using a systematic review of point prevalence surveys. We map AMR levels at 10 km resolution, provide a baseline to monitor future AMR trends, and identify regions where future surveillance efforts should be prioritized.

Results

The systematic review identified 749 point prevalence surveys reporting antimicrobial resistance rates in aquatic food animals in Asia published between 2000 and 2019. We extracted 12,698 resistance rates representing 11,289 isolates and 45 bacterial genera (Supplementary Figs. S1 and S2; Supplementary Data 1). Eastern Asia accounted for 50.6% ($n = 379$) of surveys; Western and Southern Asia, 30.7% ($n = 230$); and South-eastern Asia, 18.7% ($n = 140$). China, India, and Turkey together contributed nearly two-thirds of all point prevalence surveys across Asia. China alone represented 37.9% ($n = 284$) of surveys, a fraction that expanded over the last decade. India and Turkey each contributed 12.5% ($n = 94$) of surveys (Supplementary Figs. S3–S5).

AMR trends from aquaculture and wild-caught fisheries

Between 2000 and 2018, the percentage of antimicrobial compounds with resistance exceeding 50% (P50) in each survey plateaued in cultured aquatic animals at 33% [95% confidence interval (CI) 28 to 37%], and decreased sharply in wild-caught aquatic animals from 52% [95% CI 39 to 65%] to 22% [95% CI 14 to 30%] ($p = 0.003$) (Fig. 1). Across all years, the median P50 of surveys from cultured aquatic animals (31%; $n = 558$) was lower than surveys from wild-caught aquatic animals (44%; $n = 81$) ($p = 0.059$) (Supplementary Fig. S7).

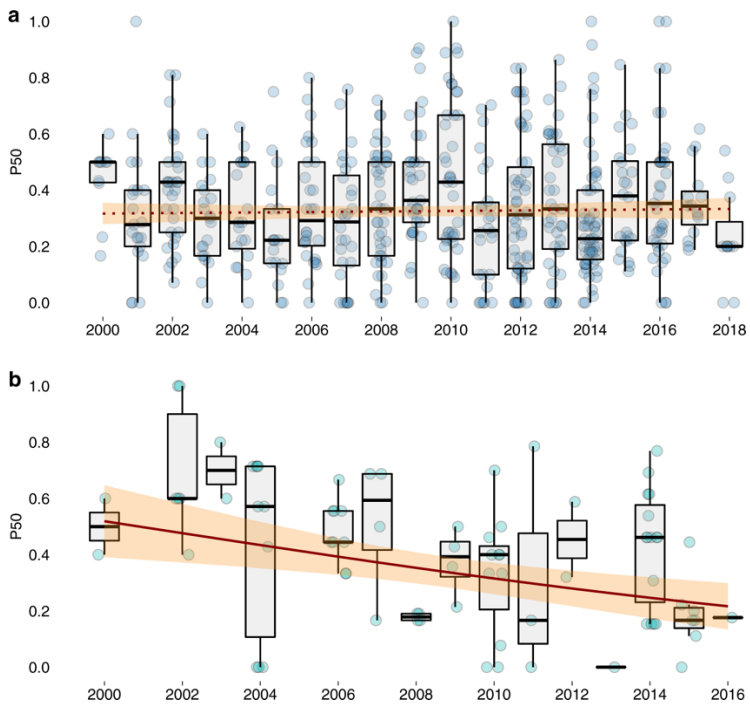


Fig. 1: Annual trends in the proportion of drugs with resistance greater than 50% (P50) in each survey. (a) P50 for cultured aquatic animals (n = 558); and (b) wild-caught aquatic animals (n = 81). The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile −1.5× the interquartile range and the third quartile +1.5× the interquartile range. Each survey is represented by a dot with horizontal jitter for visibility. Regression lines are fitted using generalized linear models, with a solid line indicating statistical significance (p = 0.003); 95% confidence intervals are shown in shaded areas.

AMR profiles in foodborne pathogens

The five most frequently isolated bacteria genera identified in our review—*Vibrio*, *Aeromonas*, *Streptococcus*, *Edwardsiella*, and *Escherichia (E.coli)*—together accounted for 68.5% of surveys. *Vibrio* spp. ($n = 191$) and *Aeromonas* spp. ($n = 174$) contributed nearly half of all surveys.

In foodborne pathogens, we calculated the pooled prevalence of resistance from individual pathogen-drug resistance rates (see "Methods" section). Amongst these foodborne pathogens, resistance was highest to penicillins (60.4%), macrolides (34.2%), sulfonamides (32.9%), and tetracyclines (21.5%) (Fig. 2). Although highly variable, mean resistance rates to the highest priority critically important antimicrobials for human medicine[28] were highest for macrolides (34.2%, 95% CI 33 to 35%) followed by third-generation and fourth-generation cephalosporins (17.5%, 95% CI 17 to 18%) and quinolones (16.2%, 95% CI 15 to 17%). Mean resistance to third-generation cephalosporins in *E.coli* was 27.1% (95% CI 25 to 29%) across surveys. Resistance to compounds in the reserve group of last-resort antimicrobials for human medicine[29] varied by pathogen: fosfomycin resistance in *E.coli* was 10.3% (95% CI 6 to 14%); polymyxin B resistance in *E.coli* was 19.4% (95% CI 7 to 32%), in *Vibrio* spp. was 39.1% (95% CI 33 to 44%) and in *Aeromonas*spp. was 71.5% (95% CI 67 to 76%); and colistin resistance in *E.coli* was 5.2% (95% CI 2 to 8%), in *Vibrio* spp. was 42.7% (95% CI 38 to 47%) and in *Aeromonas* spp. was 51.5% (95% CI 46 to 57%). In Gram-negative bacteria across all surveys, the mean colistin resistance was 41.3% (95% CI 39 to 44%).

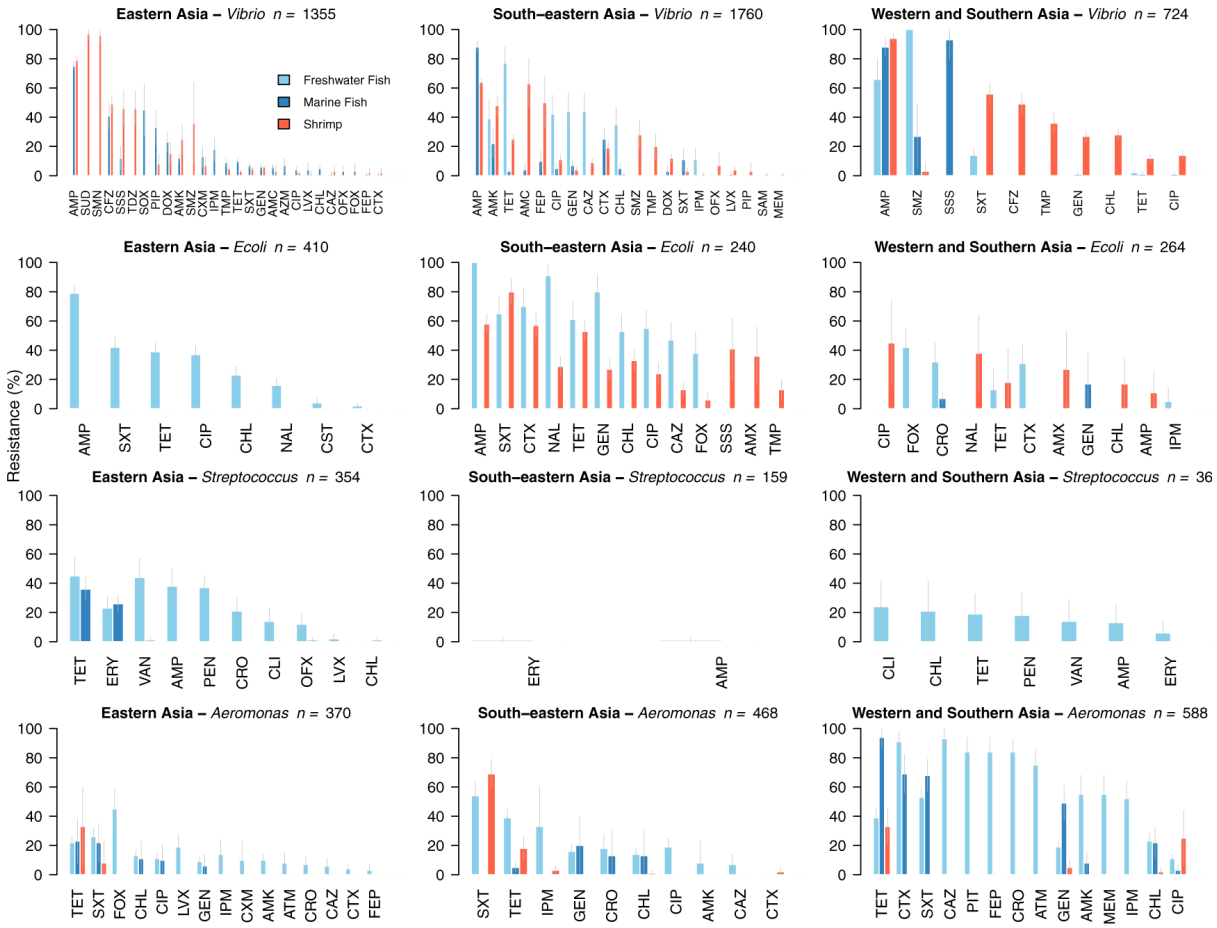


Fig. 2: Antimicrobial resistance in foodborne pathogens isolated from aquatic animals in Asia. Gray bars represent 95% proportion confidence intervals. Resistance is shown for pathogen-drug combinations recommended for susceptibility testing (Supplementary Table S1) and with 10 or more isolates tested. (For drug acronyms, see Supplementary Note 2).

Mean foodborne pathogen resistance to carbapenems was low (2.8%, 95% CI 2 to 4%), with the exception of *Aeromonas* spp. In *Aeromonas* spp. across all regions, carbapenem resistance increased from 5.1% (95% CI 2 to 8%) before 2010 to 51.1% (95% CI 43 to 60%) after 2010 ($p < 0.0001$). *Aeromonas* spp. in Western and Southern Asia exhibited elevated resistance compared with other subregions, particularly in samples originating from freshwater and marine fish. In these aquatic food animals across all years, mean carbapenem resistance in *Aeromonas* spp. was 40.3% (95% CI 33 to 48%), aztreonam resistance was 56.6% (95% CI 46 to 67%), and mean resistance to third-generation and fourth-generation cephalosporins was 69.6% (95% CI 65 to 75%).

Geography of resistance

Predicted hotspots (P50 > 0.5) of multi-drug resistance in freshwater environments included eastern Turkey; southern India—particularly the wetlands of coastal Kerala, Tamil Nadu, and Andhra Pradesh; the Yangtze River in China, both along its upper reaches and at Poyang Lake; and the lower reaches of the Mekong River and its delta in southern Cambodia and Vietnam.

Low P50 (<0.1) was predicted in peri-urban Guangzhou in southern China and in South Korea and Japan (Fig. 3). The interpolation of resistance in this study is associated with uncertainty. Variability in the geolocation of surveys, in covariates, and in estimates of resistance contribute to this uncertainty, which is captured with a 95% confidence interval map on P50 predictions (Supplementary Fig. S18). Uncertainty in P50 predictions was high (95% CI > 0.5) in South Korea and Japan.

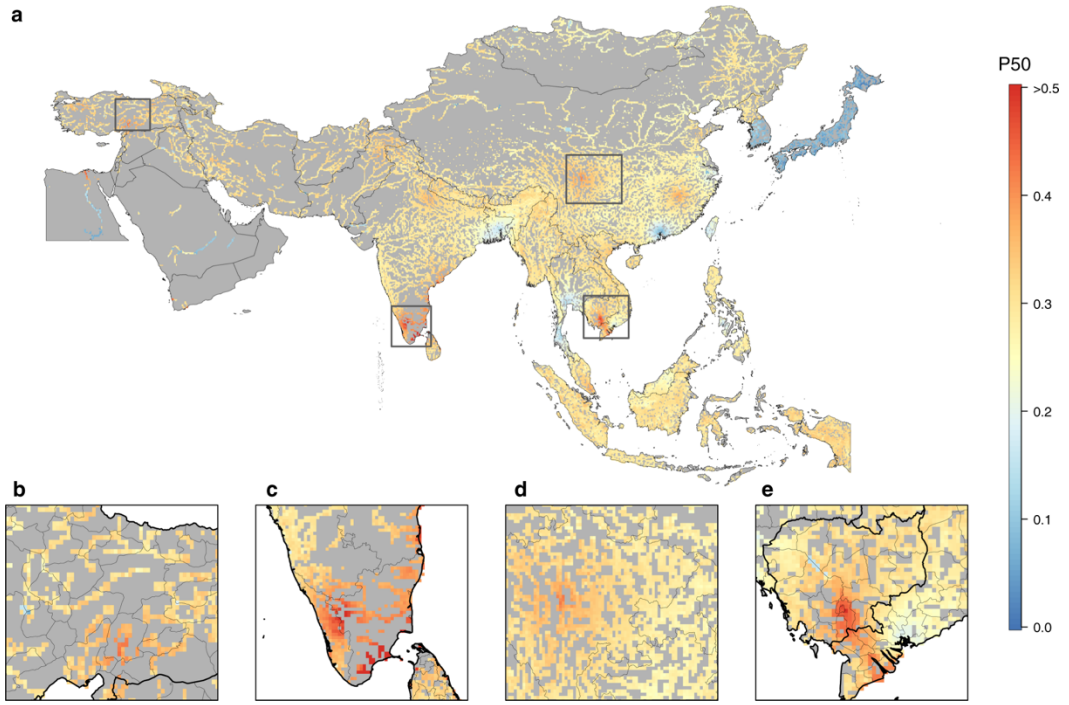


Fig. 3: Geographic distribution of antimicrobial resistance in freshwater environments in Asia. The proportion of antimicrobial compounds in each survey with resistance higher than 50% (P50) at continental scale (a); eastern Turkey (b); southern India (c); Yangtze River drainage basin in China (d); and the Mekong River delta (e).

We identified the locations of 50 hypothetical surveys to be conducted in Asia that would maximize information gained over the current map of AMR in freshwater environments. Future survey locations were optimized using a "need for surveillance" index, calculated as the uncertainty in AMR weighted to areas where resistance is likely to have the greatest impact on human health and the aquaculture industry (see "Methods" section). The majority of future surveys were projected in China (56%) and India (16%), with the highest tier of prioritization distributed predominantly in central and eastern China, western and central India, and along the Indo-Gangetic Plain (Fig. 4). Indonesia accounted for the third-highest number of surveys to be conducted in the future (12%).

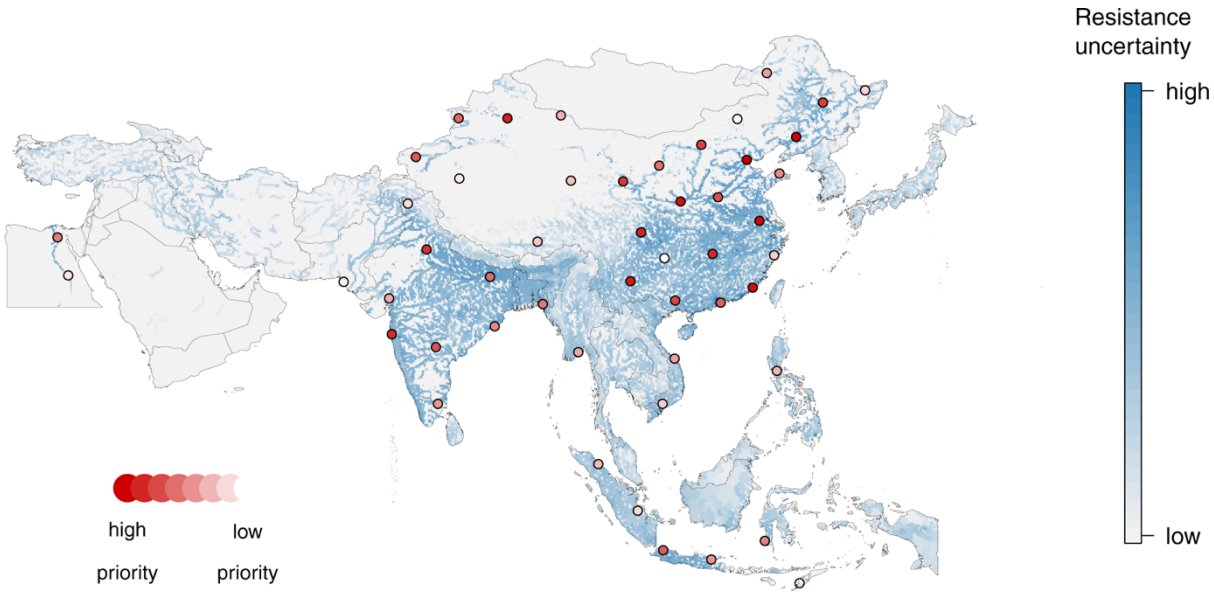


Fig. 4: Future survey locations prioritized to reduce uncertainty in antimicrobial resistance in freshwater environments in Asia. The background color gradient (blue) represents weighted uncertainty in multi-drug resistance (see "Methods" section). An initial set of 50 future surveys optimized to reduce uncertainty in multi-drug resistance is displayed (red).

In marine environments, predicted multi-drug resistance was highest along northeastern China on the Yellow and East China Seas; eastern Hainan Island waters in southern China; the coastal waters of central Vietnam on the South China Sea; southern India coastal waters on the Arabian Sea and the Bay of Bengal between southern India and northern Sri Lanka; and the eastern Mediterranean Sea on the coast of Lebanon. The coastal waters of Thailand and Malaysia on the Gulf of Thailand and Thailand's coastal waters on the Andaman Sea carried lower P50 predicted values (<0.3) (Fig. 5).

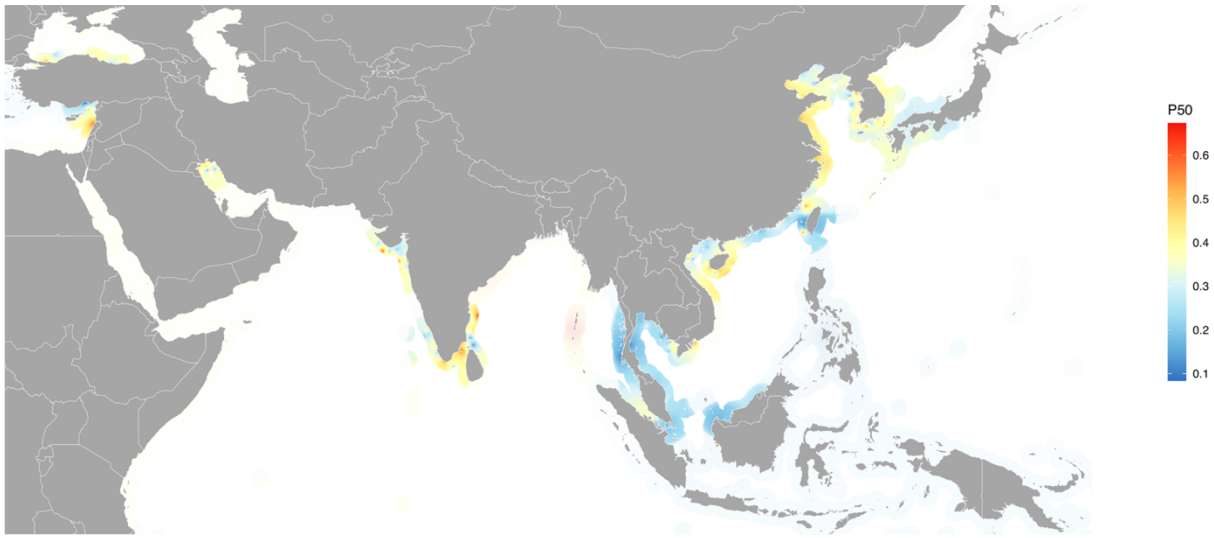


Fig. 5: Antimicrobial resistance in marine environments in Asia. Transparency reflects low survey density; areas of higher relative survey density are represented by increased opacity.

Discussion

We reviewed and mapped antimicrobial resistance in aquatic food animals in Asia during a period of substantial industry growth. Our findings indicate that between 2000 and 2018, antimicrobial resistance in bacteria from cultured aquatic food animals was stable (33%) while the resistance from wild-caught aquatic food animals decreased sharply (52% to 22%). These trends represent currently available evidence from point prevalence surveys, which serve as a surrogate in the absence of systematic surveillance and should be interpreted cautiously. Structured, systematic surveillance will be imperative to document trends in multi-drug resistance at the sub-national level in the future.

Our results are consistent with an analysis of antimicrobial resistance in aquaculture-derived bacteria from forty countries, nearly half of which in Asia, which identified a global mean multi-antibiotic resistance index of .25, and a higher index (>.35) in low-income and middle-income countries in Asia[27]. Although antimicrobial use in surveys from cultured animals was most frequently unspecified, in the limited surveys that recorded whether on-farm antimicrobials were either used or not used ($n = 63$; 11%), use was associated with higher multi-drug resistance than the absence of use ($p < .001$) (Supplementary Fig. S8). The combined effect of multi-source antimicrobial introduction[15] and persistence[30,31] in aquaculture waters and sediments may present a consistent selective pressure facilitating the maintenance of an elevated level of antimicrobial resistance.

Declining resistance observed in bacteria from wild-caught aquatic animals could be associated with reduced human and livestock fecal pollution exposure in wild-caught animals over time. Anthropogenic influences such as human and livestock fecal pollution have been positively associated with antimicrobial resistance gene frequency[32,33] , and in these surveys, we observed a negative correlation between multi-drug resistance and the World Bank basic sanitation index[34] (Supplementary Fig. S9). In addition, surveys from the first half of the study period held a significantly higher proportion of diseased animals than those from the latter half, and disease was positively associated with resistance (Supplementary Fig. S10). Importantly, these findings must be considered in the context of relatively scarce point prevalence surveys (11%, $n = 81$) attributable to capture fisheries. Other contributions—including improved wastewater management and regulatory action—cannot be ruled out and future studies will be valuable in documenting resistance trends and their influencing factors in capture fisheries.

Temporal trends in P50 rates reflect resistance across multiple bacteria, aquatic animal species, production contexts, and subregions in Asia. Production dynamics, antimicrobial exposure intensity, and regulatory standards, amongst other factors, may be highly variable by location and at timescales shorter than the duration of our analysis. Consequently, these trends may not capture more granular country-level resistance profiles in select pathogen species from specific aquatic animal production settings (Supplementary Fig. S11).

Amongst foodborne pathogens, we observed a concerning dual-threat profile, marked by high rates of resistance to first-line antimicrobial classes and, for *Vibrio* and *Aeromonas* spp., moderate to high rates of resistance to antimicrobial classes of last-resort reserved for treatment of multi-drug resistant pathogens.

In Western and Southern Asia, *Aeromonas* spp. isolates exhibit high rates of resistance to tetracyclines, sulfonamides, aminoglycosides, monobactams, carbapenems, and cephalosporins, which, together with moderate resistance to fluoroquinolones, suggests therapeutic options for management of invasive or non-self-limiting *Aeromonas* infections may already be restricted in this subregion as compared with lower resistance profiles reported outside of Asia[35,36]. High rates of carbapenem resistance from *Aeromonas* spp. were also identified in this subregion. *Aeromonas* spp. produce inducible metallo-beta-lactamases, are inherently resistant to ampicillin, amoxicillin-clavulanate, and cefazolin, and are known to express a carbapenemase encoding gene (*cphA*) the significance of which remains unknown[37]. Recent work identified the fish pathogen *Aeromonas allosaccharophila* as the origin of mobile cephalosporinase genes, which confer resistance to beta-lactams, and may be associated with aquaculture antimicrobial use[38]. While beta-lactam antimicrobials are utilized in aquaculture[5,39] , to the best of our knowledge, carbapenems are not used. Elevated carbapenem resistance identified in Western and Southern Asia may reflect inducible resistance under high environmental exposure to beta-lactam antimicrobials in this subregion, and, taken together with increasing carbapenem resistance in *Aeromonas* spp. recently reported from other regions[40,41] , may warrant further investigation.

Although variable by subregion and foodborne pathogen, high rates of resistance to third-generation and fourth-generation cephalosporins were also observed. These findings are consistent with trends identified in terrestrial animal production[26]. Ceftriaxone, Ceftazidime, and Cefotaxime resistance in *E.coli* across subregions was 27.1% (95% CI 25 to 29%), but approached 40% in Western and Southern Asia and exceeded 50% in South-eastern Asia, suggesting that resistance to third-generation cephalosporins is established in at least some members of Enterobacteriaceae in Asia. Notably, resistance in *E.coli* must be interpreted in the context of human and terrestrial animal fecal pollution and may represent the serial accumulation of antimicrobial resistance genes as has been observed in some human pathogens such as *Shigella*[42]. Extended-spectrum beta-lactamase (ESBL)-producing Enterobacteriaceae capable of hydrolyzing third-generation cephalosporins are recognized as a serious threat to human health[43]. Our finding confirms reports[44,45] of an existing pool of ESBL genes that may be plasmid encoded[46] and raises the prospect of horizontal gene transfer conferring broader resistance within and across diverse bacterial genera with the consequent loss of this critically important collection of cephalosporins.

Elevated colistin resistance in Gram-negative bacteria from aquatic animals observed in this study mirrored high rates identified in terrestrial animals in Asia over a similar study period[26]. Colistin resistance across a breadth of bacterial genera in Asia may reflect both intrinsic and acquired resistance mechanisms. Intrinsic resistance to polymyxins in Gram-negative bacteria

has been documented[47]. The possibility that a proportion of the colistin resistance identified is acquired cannot be ruled out, particularly through plasmid-mediated transfer amongst members of the Enterobacteriaceae family[48]. Such resistance may be attributable to the prolific use of colistin and other polymyxins in the region, which is only recently evolving through national regulatory action, and the possible recruitment of *mcr*-family mobile resistance determinants, their broad dissemination through horizontal gene transfer, and the wide distribution potential of aquatic environments[14,48,49]. Indeed, the distribution mechanics of AMR in aquatic systems under differing hydrological conditions is an expanding area of research that holds the promise of contributing to an enhanced understanding of AMR risk across compartments[50].
Previous work has shown that bacterial foodborne disease attributable to aquatic animal consumption contributes a non-negligible burden to human morbidity[51,52]. In this context, the high rates of resistance to multiple classes of critically important antimicrobials in aquatic animal foodborne pathogens in our study raise urgent concerns regarding both therapeutic efficacy of first-line antimicrobials and the further erosion of last-resort therapeutic options for multi-drug resistant infections resulting in severe disease.

Hotspots of AMR in freshwater environments were predicted along several of the region's great river systems, including the Yangtze River, and the estuaries of the Mekong and Nile Rivers. Low rates of AMR were predicted in Japan and South Korea. In these countries, the predictions were associated with high uncertainty (Supplementary Fig. S18). In our freshwater data set, South Korea contributed 2.4% of surveys and there were no surveys from Japan. Consequently, predictions in these two countries should be interpreted cautiously, requiring validation through additional surveillance. However, low rates of AMR could indicate less influence from human and livestock fecal pollution. Low AMR may also reflect comparatively lower rates of integrated livestock-aquaculture farming, enhanced environmental pollution regulatory capacities, and a heightened awareness of antimicrobial stewardship principles in these countries. Low AMR in the peri-urban environment of Guangzhou in southern China was unexpected considering the comparatively dense human population, prevalence of integrated livestock-aquaculture farming systems, and broad agricultural and human connectivity with the aquatic environment in Guangdong Province[14]. Future work will be important to corroborate these findings and identify factors that have the potential to moderate multi-drug resistance.

We identify 50 hypothetical future survey locations prioritized to maximize the knowledge gained over the current map of resistance from freshwater in Asia. Weighting the uncertainty in predicted resistance by the product of human population density and inland aquaculture production prioritizes future surveillance effort to locations where potential multi-drug resistance exposure risk and impact on human health and the aquaculture industry may be most significant (Supplementary Note 4). China and India together account for 72% of future surveys, reflecting their comparatively high population density and predominant role globally—first and second, respectively—in inland aquaculture production[2]. Our approach is scalable to individual country contexts and presents an opportunity to fill national or sub-national knowledge gaps and inform interventions and stewardship under resource-limited settings by maximizing information gained through targeted surveillance.

In marine coastal environments, the highest rates of AMR were identified in northeastern China on the Yellow and East China Seas; southern China and central Vietnam on the South China Sea; southern India on the Arabian Sea and the Bay of Bengal between southern India and northern Sri Lanka; and the eastern Mediterranean Sea on the coast of Lebanon. Lower AMR rates were observed along Thailand and Malaysian coastal waters both on the Gulf of Thailand and the Andaman Sea. Aquaculture production has been correlated with aminoglycoside resistance gene abundance along estuarine waters of coastal China[53]. High rates of AMR in waters also identified as carrying the highest potential for marine aquaculture productivity[54], such as the coastal waters of Hainan Island on the South China Sea, may help prioritize surveillance in an industry poised for substantial growth. However, marine predicted values should be interpreted with caution, given the comparatively limited availability of marine surveys overall ($n = 322$) and an absence of surveys from countries such as Indonesia with sizeable marine aquaculture and nearshore marine fisheries industries.

As with any modeling study, our analyses come with limitations. First, point prevalence surveys, although numerous in the study ($n = 749$), present an inherent challenge to standardization of data from susceptibility testing that are subject to variability in methodologies and protocols resulting in uncertainties[55]. Second, the resistance across species within bacterial genera can vary markedly. Our analyses of drug resistance were conducted at the bacterial genera level—the most granular level at which complete data were available. Although analysis of trends at the bacterial species level may be valuable in context-specific settings, the aim of our study was rather to document temporal trends over twenty years in Asia across a broad range of bacteria. Third, our study includes samples originating from both healthy and diseased animals. An over-representation of samples from diseased animals associated with treatment failure could present a bias toward higher resistance rates[56]. Indeed, samples from diseased animals had higher median P50 values than samples from healthy animals. However, any influence on resistance across the study is likely moderated by a balanced mix between healthy (42%) and diseased (58%) animal samples from the surveys and our treatment of intermediately susceptible isolates as susceptible in our analysis. As drug tolerance precedes clinical resistance[57], it is possible that our results underestimate non-susceptibility. Fourth, the volume of data is currently insufficient to map AMR across different study periods using spatio-temporal interpolation methods. Overall, there was an increasing volume of surveys conducted in recent years (Supplementary Fig. S4). However, when evaluated by time series that most closely partitions each of the cultured and wild-caught data sets in half (cultured: 2000 to 2010 vs. 2011 to 2018; and wild-caught: 2000 to 2009 vs. 2010 to 2016), the directionality of the regressions remained unchanged, suggesting the increasing volume of surveys does not influence the temporal trends in P50. Fifth, due to the contiguous relationship between aquatic animals and their environmental waters, the interpretation of resistance in commensal isolates such as *E.coli* must be contextualized as indicative of human and livestock fecal contamination[32], either in origin waters or through post-harvest contamination. While characterizing the origin of selective pressures driving resistance is critical to risk mitigation at source, our study documents elevated resistance in bacteria from

aquatic animals intended for human consumption—irrespective of origin—and can thus be interpreted as a risk to human health from contact with or consumption of aquatic animals and their products. Finally, it is notable that no surveys were available from Indonesia, despite contributing nearly 10% of global aquatic animal production. Vietnam—the fourth largest global producer of cultured fish[1]—represented less than 3% of all surveys. An absence of surveys from Myanmar and Laos also illustrates gaps in understanding antimicrobial resistance in aquatic food animals in Asia. Despite our broad search parameters, surveys conducted in these countries and which were either not identified in our search or not available for review could potentially influence our findings. Similarly, in the absence of systematic surveillance, reliance on passive surveillance data presents variability in survey coverage, which could introduce bias, as well as uncertainty in survey geolocation, adding potential uncertainty to predictive models (Supplementary Fig. S18. Our study addresses this uneven geographic representation of point prevalence surveys by identifying—on an objective basis—sub-regions in Asia that would benefit most from further surveillance efforts. Such future surveillance will be essential to an enhanced and refined understanding of resistance trends in Asia.

This study identified elevated rates of antimicrobial resistance in bacteria isolated from aquatic animals intended for human consumption in Asia. A growing aquatic food animal production industry may serve as an important pathway for transmission of resistance along the food chain with potential consequences for human health.

A scale-up of an antimicrobial resistance surveillance architecture for aquatic food animals is urgently needed to fill gaps in AMR trends at national and sub-national levels. Our findings help direct the prioritization of this future surveillance effort and provide a foundation for establishing time-bound, measurable targets for reducing antimicrobial resistance[58,59]. The spatial profile of antimicrobial resistance presented here should inform planning for sustainable development of a high-growth aquaculture industry, critical to feeding an expanding global population[54,60] while balancing the imperative for healthy freshwater and marine environments, and the preservation of antimicrobial efficacy for future generations.

Methods

Literature review

We searched PubMed, Web of Science, Scopus, the China National Knowledge Infrastructure database, and grey literature repositories (AGRIS, CGIAR FISH, IFPRI, WorldFish) for point prevalence surveys (PPS) of phenotypic antimicrobial resistance in bacteria isolated from aquatic animals for human consumption in Asia. The search identified surveys published between January 1, 2000, and September 30, 2019, with samples originating from cultured or wild-caught aquatic animals or their products for human consumption. Reviews and meta-analyses were excluded, as were studies with samples originated from bivalve molluscs or ornamental fish (Supplementary Note 1; Supplementary Data 2). The literature search and

systematic review were guided by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement and research synthesis norms[61] (Supplementary Table S4).

Data extraction

Collected data included study details (record identifier, author, year of publication, country), sampling details (latitude and longitude of sample collection, sampling dates, animal species, and whether cultured or wild-caught, number and origin of samples [skin; tissue; intestinal contents; or lesion] collected, health status of animal, history of antimicrobial use), and susceptibility testing details (bacterial genera, species, and strain, number of isolates subjected to antimicrobial susceptibility testing, susceptibility testing method, breakpoints and guidelines used, drug class, compound, resistance rate) (Supplementary Note 2). Where data was missing or required clarifications, the corresponding author was contacted by email. A total of 44 emails were sent requesting clarification, and 15 responses were received (34% response rate). Records were excluded when no response could be provided by the authors, and unclear data precluded further analysis.

There were 104 unique species or groups of species represented in our dataset. To facilitate analysis, species were aggregated into six species groups (Supplementary Fig. S12) reflective of aquatic animal and type of aquatic environment: marine fish, freshwater fish, brackish water fish, shrimp, and a mixed group of aquatic animals spanning these groups and for which resistance rates could not be disaggregated. The remaining species were pooled into a sixth group that included other crustaceans (crab), cephalopods (squid), gastropod molluscs (abalone), amphibians (frogs and salamanders), echinoderms (sea cucumbers and sea urchins), and reptiles (turtles).

We analyzed antimicrobial resistance in surveys at the bacterial genera level. This level of taxonomy was completely available (no missing entries) in our database, whereas the more granular bacterial species and strain level data were either not consistently provided or could not be disaggregated (143 surveys; 19%).

Temporal trends

We used the percentage of antimicrobial compounds in each survey with resistance rates exceeding 50% (P50) as a summary metric of multi-drug resistance. The P50 metric was used in the analysis of temporal trends and for geospatial modeling. We compared P50 with two additional metrics: P30 (calculated as the percentage of antimicrobial compounds in each survey with resistance exceeding 30%) and mean resistance (calculated as the total number of resistant isolates divided by the number of isolates * the number of antibiotics tested in each survey). Across all surveys, there is a positive correlation between P50 and mean resistance (Pearson's correlation coefficient = 0.9596) (Supplementary Note 3; Supplementary Fig. S19). Temporal trends in the P50 for each survey from cultured and wild-caught aquatic food animals

were analyzed using generalized linear model regressions with quasibinomial error distribution weighted by the log of the number of isolates in each survey subjected to susceptibility testing. Root mean square error (RMSE) was used to evaluate the goodness of fit for the temporal trends regression models. RMSE indicated model fits were moderate ($RMSE_{cultured} = 0.223$; and $RMSE_{wild\ caught} = 0.235$), consistent with both the scattered nature and scarcity of the data. The 95% confidence intervals were generated as the fitted values +/– 1.96 * standard error of the fitted value.

One-way analysis of variance (ANOVA) tests conducted on arcsine transformed P50 values were used to analyze the significance of the difference in mean P50 between survey characteristic groups, including culture or wild-caught, history of antimicrobial use, and health status of aquatic animals sampled (Supplementary Figs. S7, S8, and S10). Fisher's exact test was used to compare prevalence between two-time points.

AMR in foodborne pathogens

We calculated the pooled prevalence of resistance from individual pathogen-drug resistance rates to report resistance in foodborne pathogens. Analysis of resistance in foodborne pathogens of aquatic animal origin (*Vibrio* spp., *Streptococcus* spp. and *Aeromonas* spp.) was guided by antimicrobial compounds of relevance for therapeutic use in human clinical settings. We also analyzed resistance in *E.coli* as a marker of potential human and terrestrial animal influence. Resistance rates were calculated for these pathogens from samples originating from marine fish, freshwater fish, and shrimp groups using The Clinical and Laboratory Standards Institute (CLSI M45 and M100) and WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR)[62] pathogen-drug susceptibility testing guidelines (Supplementary Table S1). The 95% confidence interval was calculated on the resistance proportions.

Geospatial modeling

We mapped antimicrobial resistance in freshwater and marine environments at a resolution of 0.08333 decimal degrees, or approximately 10 km at the equator, by interpolating P50 values between point prevalence surveys. In freshwater environments, a two-step procedure was used, in which we first trained multiple child models, and subsequently stacked predictions from these models for universal kriging. Stacked generalization ensemble approaches have been used to model population-level health metrics[63], and previous work[64] has demonstrated that stacking models improve prediction accuracy compared with individual predictive models. The two-step procedure captures both the relationship between P50 and environmental and anthropogenic covariates, as well as spatial autocorrelation in the distribution of P50. This approach has recently been used to model the distribution of AMR in terrestrial animals in low-income and middle-income countries[26]. Although there is an inevitable trade-off in improved accuracy at the expense of reduced model interpretability, we chose an ensemble approach for

prediction accuracy as the focus of our study was to produce the best possible maps of AMR rather than risk factor identification.

In total, 500 surveys were used to map AMR in freshwater environments. Where surveys did not provide precise sampling coordinates, we assigned coordinates at random within a geographic uncertainty range associated with the given sampling location. The uncertainty range was calculated as the mean of the distance in kilometers in the X and Y directions from the centroid to the boundaries of the smallest available administrative unit or place name provided in the survey. We used a random binarization procedure to transform the P50 values into presence (P50 = 1) and absence (P50 = 0) of resistance. We then generated and distributed pseudo-absence points to provide additional covariate values not associated with presence (P50 = 0) using stratified random sampling proportional to the human population density to account for potential P50 observation bias in more densely populated areas. Pseudo-absence points were sampled within a geographic radius of 20 to 500 km from presence points (Supplementary Note 4).

In the first step, environmental and anthropogenic covariates relevant to the freshwater environment were used to train three child models to quantify the association between P50 and these covariates. The 13 covariates were: accessibility to cities; gross domestic product; irrigated land percentage; minimum monthly temperature; terrestrial livestock P50; terrestrial livestock antimicrobial use; human population density; and population densities of cattle, pigs raised intensively, pigs raised semi-intensively, pigs raised extensively, chickens raised intensively, and chickens raised extensively (Supplementary Fig. S13 and Table S2). Child models included boosted regression trees (BRT)[65]; least absolute shrinkage and selection operator applied to logistic regression (LASSO-GLM)[66]; and overlapped grouped LASSO penalties for General Additive Models selection (LASSO-GAM)[67] (Supplementary Fig. S14). BRT models have demonstrated good predictive performance in handling non-linear relationships and interactions amongst a diverse set of covariates and have been frequently used to model disease distribution[25,65,68,69]. By generating and combining a collection of models (decision trees) in a sequential stepwise fashion, boosting reduces both bias and variance while protecting against model overfitting. In addition, BRT models are insensitive to outliers[65]. LASSO regression models—here applied to GLM and GAM—facilitate efficient covariate selection by shrinking some regression coefficients and setting others with minor contributions to zero. These features enable a robust selection of covariates, reducing model complexity and strengthening predictive performance. Child models were fitted using three-fold spatial cross-validation aligned to the Asia sub-regions in our study (Supplementary Fig. S13). This cross-validation procedure takes observations from the training and validation sets which are geographically independent, guarding against overfitting and selection of models with poor capacity to predict to new areas[68]. Models were bootstrapped 10 times to account for variability introduced in the geographic assignment, the random binarization of P50 values, and the stratified random sampling of pseudo-absence points. The mean value of the area under the receiver operator characteristic curve (AUC) for all bootstraps was used to evaluate model predictive ability.

In the second step, predictions from child models were stacked and used as covariates for universal kriging. We fit a Matern variogram at a maximum distance of 500 km, which is where the semi-variogram attained the range. The kriging procedure was weighted by the number of isolates at each location. The output of the kriging procedure was a map of predicted freshwater resistance levels, as well as a map of kriging variance quantifying the spatial interpolation uncertainty (Supplementary Fig. S17). We also produced a "need for surveillance" index map for use in identifying optimal locations for future surveys, calculated as the kriging variance (uncertainty) weighted by the product of human population density[70] and inland aquaculture production volume[2] (described below). We further quantified uncertainty in the interpolation of P50 values by generating a map of the 95% confidence interval on the predicted P50 values (Supplementary Note 4; Supplementary Fig. S18).

In marine environments, a root mean square error (RMSE)-weighted ensemble model was used to map AMR. In contrast to the freshwater model, no association between marine covariates and P50 was identified, and we interpolate P50 values using survey coordinates in the marine AMR map. Surveys from inland freshwater sites were excluded. Wild-caught marine animals sampled at land-based post-harvest sites were randomly assigned coordinates to open ocean within a radius of .54 to 81 nautical miles (1 to 150 km) from their nearest coastal location (Supplementary Fig. S15). The marine data set consisted of two groups of surveys: (i) surveys from animals sampled at land-based post-harvest sites randomly assigned to open ocean; and (ii) surveys originating from marine, coastal marine, brackish water, and coastal brackish water sampling locations (Supplementary Note 4).

This marine data set ($n = 322$) was used to produce inverse distance weighted, natural neighbor, and ordinary kriging models. These models were subsequently stacked and weighted according to their root mean square error to capture the fit and variance of each model in the final ensemble model. The weights were taken as the inverse of the RMSE of each constituent model divided by the sum of RMSE for all models and expressed as their relative proportion in the final RMSE-weighted marine AMR ensemble model (Supplementary Table S3). A transparency function was added proportionally to the spatial kernel density of surveys at a bandwidth of 8.333 decimal degrees to reflect the density of the geographic distribution of surveys contributing to the final marine P50 map.

Optimizing locations for future surveillance

We used the "need for surveillance" index map for freshwater AMR to identify locations for 50 hypothetical surveys—the rounded mean number of annual surveys between 2010 and 2019—that could be conducted across Asia next year. The "need for surveillance" index was calculated as the product of the uncertainty from the spatial interpolation, human population density[70], and inland aquaculture production volume[2]. Human population and aquaculture production terms were standardized to range between 0 and 1 to give equal importance in determining the locations of future surveys. This function weights the necessity for surveillance to locations where the potential exposure risk and impact of AMR is greatest on the aquaculture

industry and human health—via local consumption71 and the cyclical exchange of resistant bacteria and their determinants across humans, aquaculture, and the aquatic environment.
We followed an approach proposed by Zhao et al.[72] that exploits a key feature associated with each additional survey conducted, reducing the uncertainty of the geospatial model in its surrounding area. The survey locations were optimized to reduce uncertainty as quantified through the "need for surveillance" index, thereby maximizing information gained for each successive survey (Supplementary Note 4).

References

1. *The State of World Fisheries and Aquaculture 2020*. https://doi.org/10.4060/ca9229en (FAO, 2020).
2. *Fisheries and aquaculture software. FishStatJ-Software for Fishery and Aquaculture Statistical Time Series*. (2016).
3. Van Boeckel, T. P. et al. Global trends in antimicrobial use in food animals. *Proc. Natl Acad. Sci. USA* **112**, 5649–5654 (2015).
4. Van Boeckel, T. P. et al. Reducing antimicrobial use in food animals. *Science* **357**, 1350–1352 (2017).
5. Schar, D., Klein, E. Y., Laxminarayan, R., Gilbert, M. & Van Boeckel, T. P. Global trends in antimicrobial use in aquaculture. *Sci. Rep.* **10**, 21878 (2020).
6. Laxminarayan, R. et al. Antibiotic resistance—the need for global solutions. *Lancet Infect. Dis.* **13**, 1057–1098 (2013).
7. Aarestrup, F. M. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos. Trans. R. Soc. B* **370**, 20140085 (2015).
8. Robinson, T. P. et al. Animal production and antimicrobial resistance in the clinic. *Lancet***387**, e1–e3 (2016).
9. Liu, C. M. et al. *Escherichia coli* ST131- *H* 22 as a Foodborne Uropathogen. *mBio* **9**, e00470–18 (2018).
10. Buschmann, A. H. et al. Salmon aquaculture and antimicrobial resistance in the marine environment. *PLoS ONE* **7**, e42724 (2012).
11. Done, H. Y., Venkatesan, A. K. & Halden, R. U. Does the recent growth of aquaculture create antibiotic resistance threats different from those associated with land animal production in agriculture? *AAPS J.* **17**, 513–524 (2015).
12. Heuer, O. E. et al. Human health consequences of use of antimicrobial agents in aquaculture. *Clin. Infect. Dis.* **49**, 1248–1253 (2009).
13. Cabello, F. C. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ. Microbiol.* **8**, 1137–1144 (2006).
14. Shen, Y. et al. Integrated aquaculture contributes to the transfer of mcr-1 between animals and humans via the aquaculture supply chain. *Environ. Int.* **130**, 104708 (2019).

15. Cabello, F. C. et al. Antimicrobial use in aquaculture re-examined: its relevance to antimicrobial resistance and to animal and human health: Aquacultural antimicrobial use and antimicrobial resistance. *Environ. Microbiol.* **15**, 1917–1942 (2013).
16. Ryu, S.-H. et al. Antimicrobial resistance and resistance genes in Escherichia coli strains isolated from commercial fish and seafood. *Int. J. Food Microbiol.* **152**, 14–18 (2012).
17. Wellington, E. M. et al. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. *Lancet Infect. Dis.* **13**, 155–165 (2013).
18. Aedo, S., Ivanova, L., Tomova, A. & Cabello, F. C. Plasmid-Related Quinolone Resistance Determinants in Epidemic Vibrio parahaemolyticus, Uropathogenic Escherichia coli, and Marine Bacteria from an Aquaculture Area in Chile. *Microb. Ecol.* **68**, 324–328 (2014).
19. Cabello, F. C., Tomova, A., Ivanova, L. & Godfrey, H. P. Aquaculture and mcr Colistin resistance determinants. *mBio* **8**, e01229–17 (2017). mBio.01229-17.
20. Cabello, F. C., Godfrey, H. P., Buschmann, A. H. & Dölz, H. J. Aquaculture as yet another environmental gateway to the development and globalisation of antimicrobial resistance. *Lancet Infect. Dis.* **16**, e127–e133 (2016).
21. Thornber, K. et al. Evaluating antimicrobial resistance in the global shrimp industry. *Rev. Aquac.* **12**, 966–986 (2020).
22. World Bank. *Drug-Resistant Infections: A Threat to Our Economic Future*. http://documents1.worldbank.org/curated/en/323311493396993758/pdf/final-report.pdf(2017).
23. Watts, J., Schreier, H., Lanska, L. & Hale, M. The rising tide of antimicrobial resistance in aquaculture: sources, sinks and solutions. *Mar. Drugs* **15**, 158 (2017).
24. Henriksson, P. J. G. et al. Unpacking factors influencing antimicrobial use in global aquaculture and their implication for management: a review from a systems perspective. *Sustain. Sci.* **13**, 1105–1120 (2018).
25. Bhatt, S. et al. The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
26. Van Boeckel, T. P. et al. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, eaaw1944 (2019).
27. Reverter, M. et al. Aquaculture at the crossroads of global warming and antimicrobial resistance. *Nat. Commun.* **11**, 1870 (2020).
28. WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance and World Health Organization. *Critically Important Antimicrobials for Human Medicine: Ranking of Antimicrobial Agents for Risk Management of Antimicrobial Resistance Due to Non-human Use*. https://www.who.int/publications/i/item/9789241515528 (2017).
29. World Health Organization. *The AWaRe Classification of Antibiotics Database*. https://adoptaware.org/ (2019).
30. Sarmah, A. K., Meyer, M. T. & Boxall, A. B. A. A global perspective on the use, sales, exposure pathways, occurrence, fate and effects of veterinary antibiotics (VAs) in the environment. *Chemosphere* **65**, 725–759 (2006).

31. Hektoen, H., Berge, J. A., Hormazabal, V. & Yndestad, M. Persistence of antibacterial agents in marine sediments. *Aquaculture* **133**, 175–184 (1995).
32. Karkman, A., Pärnänen, K. & Larsson, D. G. J. Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. *Nat. Commun.***10**, 80 (2019).
33. Su, J.-Q. et al. Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* **5**, 84 (2017).
34. World Bank. *People Using at Least Basic Sanitation Services (% of population)*. data.worldbank.org. (2017).
35. Janda, J. M. & Abbott, S. L. The Genus Aeromonas: taxonomy, pathogenicity, and infection. *Clin. Microbiol. Rev.* **23**, 35–73 (2010).
36. Vila, J. et al. *Aeromonas* spp. and traveler's diarrhea: clinical features and antimicrobial resistance. *Emerg. Infect. Dis.* **9**, 552–555 (2003).
37. CLSI. *Methods for Antimicrobial Dilution and Disk Susceptibility Testing of Infrequently Isolated or Fastidious Bacteria*. 3rd ed. CLSI guideline M45. Wayne, PA: Clinical and Laboratory Standards Institute; 2016.
38. Ebmeyer, S., Kristiansson, E. & Larsson, D. G. J. The mobile FOX AmpC beta-lactamases originated in Aeromonas allosaccharophila. *Int. J. Antimicrob. Agents* **54**, 798–802 (2019).
39. World Organisation for Animal Health. *OIE Annual Report on Antimicrobial Agents Intended for Use in Animals: Fourth Report*. https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/AMR/A_Fourth_Annual_Report_AMR.pdf (2019).
40. Rosso, F. et al. Emerging carbapenem-resistant Aeromonas spp. infections in Cali, Colombia. *Braz. J. Infect. Dis.* **23**, 336–342 (2019).
41. Sinclair, H. A. et al. Genotypic and phenotypic identification of Aeromonas species and CphA-mediated carbapenem resistance in Queensland, Australia. *Diagn. Microbiol. Infect. Dis.* **85**, 98–101 (2016).
42. Njamkepo, E. et al. Global phylogeography and evolutionary history of Shigella dysenteriae type 1. *Nat. Microbiol.* **1**, 16027 (2016).
43. Centers for Disease Control and Prevention (U.S.). *Antibiotic resistance threats in the United States, 2019*. https://stacks.cdc.gov/view/cdc/82532 (2019).
44. Ng, C. et al. Microbial water quality and the detection of multidrug resistant E. coli and antibiotic resistance genes in aquaculture sites of Singapore. *Mar. Pollut. Bull.* **135**, 475–480 (2018).
45. Lu, S.-Y. et al. High diversity of extended-spectrum beta-lactamase-producing bacteria in an urban river sediment habitat. *Appl. Environ. Microbiol.* **76**, 5972–5976 (2010).
46. Paterson, D. L. & Bonomo, R. A. Extended-spectrum β-lactamases: a clinical update. *Clin. Microbiol. Rev.* **18**, 657–686 (2005).
47. Olaitan, A. O., Morand, S. & Rolain, J.-M. Mechanisms of polymyxin resistance: acquired and intrinsic resistance in bacteria. *Front. Microbiol.* **5**, 643 (2014).

48. Liu, Y.-Y. et al. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).
49. Cabello, F. C., Tomova, A., Ivanova, L. & Godfrey, H. P. Aquaculture and mcr colistin resistance determinants. *mBio* **8**, e01229 (2017).
50. Taylor, N. G. H., Verner-Jeffreys, D. W. & Baker-Austin, C. Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends Ecol. Evol.* **26**, 278–284 (2011).
51. Li, M. et al. Global disease burden of pathogens in animal source foods, 2010. *PLoS ONE* **14**, e0216545 (2019).
52. Wu, Y., Wen, J., Ma, Y., Ma, X. & Chen, Y. Epidemiology of foodborne disease outbreaks caused by Vibrio parahaemolyticus, China, 2003–2008. *Food Control* **46**, 197–202 (2014).
53. Zhu, Y.-G. et al. Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* **2**, 16270 (2017).
54. Gentry, R. R. et al. Mapping the global potential for marine aquaculture. *Nat. Ecol. Evol.* **1**, 1317–1324 (2017).
55. Smith, P. Eight rules for improving the quality of papers on the antimicrobial susceptibility of bacteria isolated from aquatic animals. *Dis. Aquat. Organ.* **139**, 87–92 (2020).
56. International Office of Epizootics. *Aquatic Animal Health Code*. https://www.oie.int/en/what-we-do/standards/codes-and-manuals/aquatic-code-online-access/ (2019).
57. Levin-Reisman, I. et al. Antibiotic tolerance facilitates the evolution of resistance. *Science***355**, 826–830 (2017).
58. Laxminarayan, R., Sridhar, D., Blaser, M., Wang, M. & Woolhouse, M. Achieving global targets for antimicrobial resistance. *Science* **353**, 874–875 (2016).
59. Laxminarayan, R. et al. The Lancet Infectious Diseases Commission on antimicrobial resistance: 6 years later. *Lancet Infect. Dis.* **20**, e51–e60 (2020).
60. Troell, M., Jonell, M. & Crona, B. *The role of seafood in sustainable and healthy diets: The EAT-Lancet Commission report through a blue lens*. 21 https://eatforum.org/content/uploads/2019/11/Seafood_Scoping_Report_EAT-Lancet.pdf(2019).
61. Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *PLoS Med.* **18**, e1003583 (2021).
62. World Health Organization. *Integrated surveillance of antimicrobial resistance in foodborne bacteria: application of a one health approach: guidance from the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR)*. https://apps.who.int/iris/handle/10665/255747 (World Health Organization, 2017).
63. Golding, N. et al. Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the SUstainable Development Goals. *Lancet* **390**, 2171–2182 (2017).

64. Bhatt, S. et al. Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* **14**, 20170520 (2017).
65. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
66. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).
67. Chouldechova, A. & Hastie, T. Generalized additive model selection. *arXiv150603850 Stat*(2015).
68. Gilbert, M. et al. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. *Nat. Commun.* **5**, 4116 (2014).
69. Allen, T. et al. Global hotspots and correlates of emerging zoonotic diseases. *Nat. Commun.* **8**, 1124 (2017).
70. Center For International Earth Science Information Network-CIESIN-Columbia University. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. https://doi.org/10.7927/H49C6VHW (2017).
71. Belton, B., Bush, S. R. & Little, D. C. Not just for the wealthy: rethinking farmed fish consumption in the Global South. *Glob. Food Secur.* **16**, 85–92 (2018).
72. Zhao, C. et al. Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China. *Nat. Food*https://doi.org/10.1038/s43016-021-00320-x (2021).
73. Schar, D. et al. Aquatic animal antimicrobial resistance trends dataset and R code. https://doi.org/10.5281/ZENODO.4615703 (2021).

Chapter 3

Predictive Mapping of Antimicrobial Resistance for *Escherichia coli*, *Salmonella*, and *Campylobacter* in Food-Producing Animals, Europe, 2000–2021.

Authors:
Ranya Mulchandani[1], Cheng Zhao[1], Katie Tiseo[1], João Pires[1], Thomas P. Van Boeckel[1,2,*].

Affiliations:
[1]ETH Zürich, Zurich, Switzerland.
[2]One Health Trust, Washington, DC, USA

*Correspondence to: thomas.van.boeckel@gmail.com

Published in:
Emerging Infectious diseases 30(1), 96 (2024). DOI: 10.3201/eid3001.221450
This chapter contains minor revisions compared to the published article.

Contribution remarks:
R.M., C.Z., and T.P.V.B designed the research; R.M., C.Z., and K.T. collected the data; R.M. analysed the data; C.Z. conducted geospatial modelling; R.M., C.Z., J.P., and T.P.V.B wrote the paper.

Abstract

In Europe, systematic national surveillance of antimicrobial resistance (AMR) in food-producing animals has been conducted for decades; however, geographic distribution within countries remains unknown. To determine distribution within Europe, we combined 33,802 country-level AMR prevalence estimates with 2,849 local AMR prevalence estimates from 209 point prevalence surveys across 31 countries. We produced geospatial models of AMR prevalence in *Escherichia coli,* nontyphoidal *Salmonella*, and *Campylobacter* for cattle, pigs, and poultry. We summarized AMR trends by using the proportion of tested antimicrobial compounds with resistance >50% and generated predictive maps at 10 × 10 km resolution that disaggregated AMR prevalence. For *E. coli,* predicted prevalence rates were highest in southern Romania and southern/eastern Italy; for *Salmonella,* southern Hungary and central Poland; and for *Campylobacter*, throughout Spain. Our findings suggest that AMR distribution is heterogeneous within countries and that surveillance data from below the country level could help with prioritizing resources to reduce AMR.

Introduction

Antimicrobial resistance (AMR) is a substantial threat to the health of humans and animals. Among humans, in 2019 an estimated 1.27 million deaths were associated with bacterial AMR[1]. Among food-producing animals (i.e., animals that are used for or produce food items for human consumption), estimates of global AMR burden are still lacking. However, recent work has suggested that among common indicator bacteria of food-producing animals in low- and middle-income countries, the proportion of antimicrobials with resistance >50% increased from 12%–15% in 2000 to 34%–41% in 2018[2], an increase that may have harmful consequences for humans[3]. Moreover, the loss of treatment effectiveness in animals is a long-term threat for animal production and the millions of persons who rely on raising animals for subsistence[4,5]. Therefore, monitoring AMR in food-producing animals has become a global priority for effective prevention strategies.

Since 2009, the European Food Safety Authority (EFSA) has led a harmonized surveillance system for AMR in food-producing animals and products[6]. The system includes AMR prevalence estimates for *Escherichia coli*, nontyphoidal *Salmonella*, and *Campylobacter* among cattle and pigs (odd years) and chickens and turkeys (even years)[7]. Data collected by EFSA have been instrumental for monitoring AMR and for guiding policy decisions in the European Union (e.g., the 2018 ban on prophylactic use of antimicrobials in animals[8]). The efforts to document AMR have also enabled comparison between countries in Europe by estimating prevalence of AMR at the national level. However, recent works have shown that resistance levels in humans and animals can vary at a fine spatial scale, and accumulation of resistance genes in those areas may create geographic hotspots for AMR[2,9]. Identifying geographic hotspots of AMR within countries could help with targeting interventions against AMR, such as improved farm biosecurity and targeted surveillance, where they might have the greatest benefits[10,11,12].

In that context, point prevalence surveys (PPSs) of AMR among food-producing animals, with data points collected at individual geographic locations, provide an opportunity to supplement the national estimates of AMR assembled by EFSA[2]. The resulting mapped predictions could be used to help design regional antibiotic stewardship campaigns or target local investment in farm biosecurity[12]. However, generating robust predictions of AMR pose at least 3 challenges.

First, comparisons need to be made between the resistance trends inferred from PPSs and EFSA; second, subnational predictions should reflect resistance levels reported by EFSA at the national level; and third, an appropriate geospatial modeling approach must be developed to combine data collected at different spatial scales.

In this study, we disaggregated trends in AMR prevalence of *E. coli,* nontyphoidal *Salmonella*, and *Campylobacter* among cattle, pigs, and poultry. We used stacked geospatial models that supplement data from EFSA with individual PPSs to map predictions of AMR prevalence at a resolution of 10 × 10 km for 31 countries in Europe.

Results

EFSA surveillance

At the country level, EFSA data for 2009–2020 provided 33,802 AMR prevalence estimates (resulting in 2,996 P50s). The data were for *E. coli*, nontyphoidal *Salmonella*, *C. coli*, and *C. jejuni* in cattle, pigs, and poultry across 31 countries in Europe.

PPSs

At the local level, for 2000–2021 we identified 209 PPSs, which provided 2,849 AMR prevalence estimates (resulting in 368 P50s). The data were for *E. coli*, nontyphoidal *Salmonella*, and *Campylobacter* in food-producing animals and derived products from 21 countries in Europe. In terms of AMR prevalence, *E. coli* accounted for 44.4%, *Salmonella* for 34.2%, and *Campylobacter* for 21.4%. Poultry accounted for approximately half of the AMR prevalence (n = 1,429, 50.2%), followed by pigs (28.1%) and cattle (21.8%). One third of the sample types tested were meat (34.7%, n = 988), followed by fecal samples (23.4%). Across the countries included in the analysis, geographic coverage was on average 4.21 PPSs (interquartile range 0–11.7)/100,000 km^2. Half of the PPSs identified were from the combination of Spain (20.5%), Italy (18.7%), and Germany (10.5%) (Figure 1). The average number of PPSs published by year increased from 3 during 2000–2005 to 14 during 2015–2021 (Figure 1, panel B).

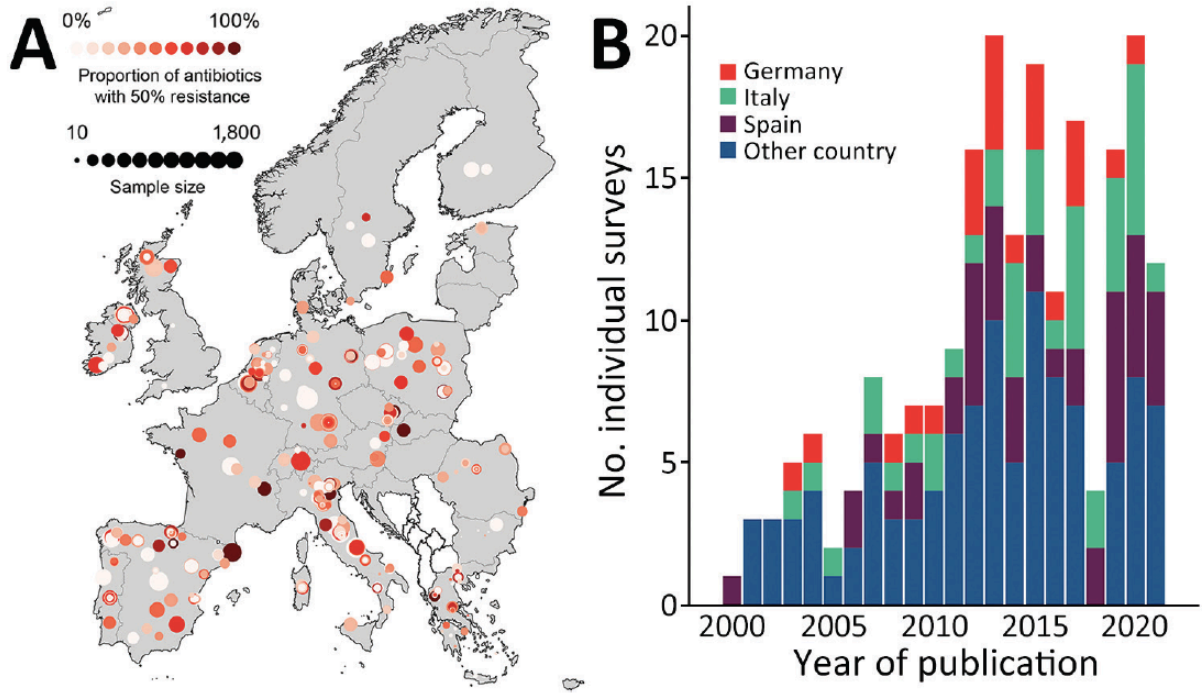


Figure 1. Data from study of predictive mapping for antimicrobial resistance of *Escherichia coli*, *Salmonella*, and *Campylobacter*in food-producing animals, Europe, 2000–2021. A) Geographic distribution of point prevalence surveys (PPSs). B) Number of PPSs published per year. Additional information is provided in the Appendix.

Comparison of PPS and EFSA

AMR prevalence estimates varied considerably between data sources and country. For 2018–2020, Greece, Poland, and Germany accounted for more than double the national average P50 calculated from PPS data compared with P50s calculated from EFSA (Table 1). Conversely, the national average P50 calculated from PPS data from Portugal and Switzerland was <30% lower than that calculated from EFSA.

Country	Mean P50 from PPS	Mean P50 from EFSA	PPS and EFSA P50 ratio
Poland	0.64	0.26	2.47
Germany	0.60	0.25	2.42
Greece	0.39	0.19	2.02
Spain	0.39	0.24	1.67
Belgium	0.29	0.21	1.34
Romania	0.31	0.28	1.10
Italy	0.23	0.25	0.92
Switzerland	0.17	0.22	0.77
Portugal	0.18	0.32	0.57

*EFSA, European Food Safety Authority; PPS, point prevalence surveys; P50, >50% antimicrobial resistance.

Table 1. Three-year mean of proportion of antimicrobial drugs with >50% resistance from PPS and EFSA data and ratios of P50 for countries reporting to both data sources, Europe, 2018–2020*

The highest resistance prevalence estimates were for tetracycline (57.9%–36.4%), ampicillin (58.6%–34.9%), ciprofloxacin (64.6%–13.1%), and nalidixic acid (60.9%–25.5%). The difference in mean P50 between PPSs and EFSA data ranged from 15.2% to –17.4% for *Salmonella* and from 19.1% to –7.96% for *E. coli*. For *Campylobacter*, systematically higher prevalence estimates were obtained from PPSs; differences ranged from 12.1% to 0.78% (Figure 2).

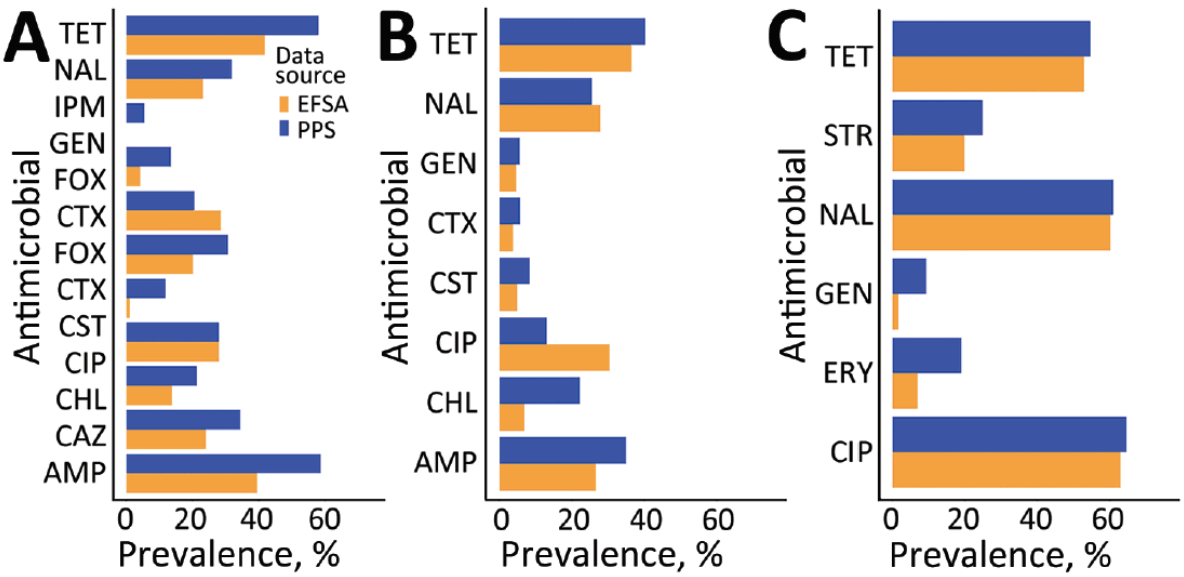


Figure 2. Mean prevalence for antimicrobial class and bacteria combinations, split by data source, Europe, 2009–2020. A) *Escherichia coli*; B) *Salmonella*; C) *Campylobacter*. AMP, ampicillin; CAZ, ceftazidime; CHL, chloramphenicol; CIP, ciprofloxacin; CST, colistin; CTX, clavulanic acid; EFAS, European Food Safety Authority; FOX, cefoxitin; GEN, gentamicin; IPM, imipenem; NAL, nalidixic acid; PPS, point prevalence survey; STR, streptomycin; TET, tetracycline.

Geospatial modelling

We mapped predicted P50s at 10 × 10 km resolution for each of the 3 bacteria across Europe (Figure 3). In the final models, the predicted P50 values ranged from 0 to 79% for *E. coli*, 0 to 40% for *Salmonella*, and 0 to 100% for *Campylobacter* (Figure 3, panel A; prediction uncertainty, Appendix Figure 3, panel A). P50 cutoffs for hotspots of AMR (calculated as the top 95% of the values on the map) were 0.43 for *E. coli*, 0.23 for *Salmonella*, and 0.60 for *Campylobacter*. AMR hotspots for *E. coli* were predicted to be located in southern Romania (Muntenia, Dobrogea) and southern and eastern Italy (Sicily, Emilia-Romagna, Apulia); and for *Salmonella*, predicted hotspots were in southern Hungary, northern Italy, and central Poland. More than 90% of hotspot areas for *Campylobacter* were predicted to be throughout mainland Spain (Figure 3, panel B).

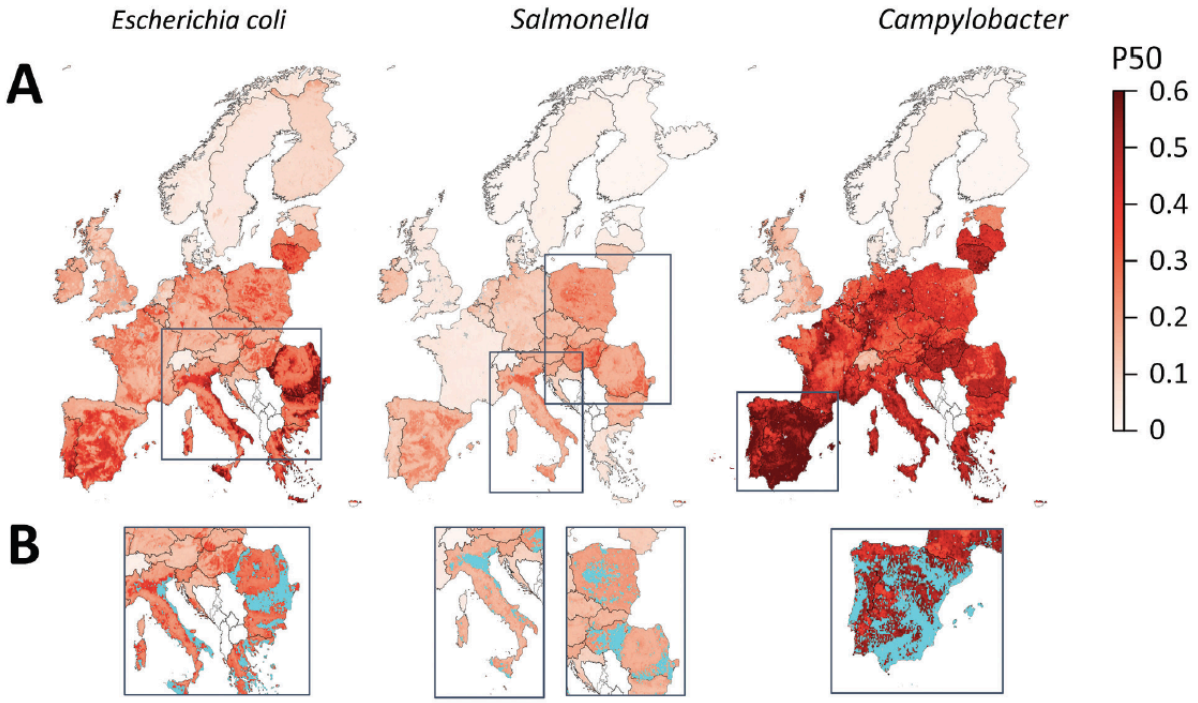


Figure 3. Mapping of predicted P50s and hotspot areas for antimicrobial resistance of *Escherichia coli*, *Salmonella*, and *Campylobacter*, Europe. A) Predicted proportions of antimicrobials with P50 at 10 × 10 km resolution per bacteria. B) Antimicrobial resistance hotspots (light blue) in eastern Europe, Italy, and Spain. Cutoffs: *E. coli*, 0.43; *Salmonella*, 0.23; *Campylobacter*, 0.6 (95% percentile). P50, >50% antimicrobial resistance.

For *E. coli*, the highest geographic variations in predicted P50 levels were in Romania (13% pixel-level SDs), Bulgaria (11%), Greece 1(2%), and Italy (11%). For *Campylobacter*, the highest geographic variations in P50 were in France (10%) and Germany (10%; Appendix Figure 4, panel C). No countries had high spatial variations in predicted P50s for *Salmonella*. Cold spots for all 3 bacteria were identified in Sweden, Norway, Finland, and Iceland (data not shown). Spatial variations of P50 for countries containing coldspots were small, with pixel-level standard deviations of 3.2% (*E. coli*), 0.9% (*Salmonella*), and 1.0% (*Campylobacter*). Restricting PPS by year and antimicrobial bacteria combinations resulted in little difference (mean Pearson correlation coefficient 0.992; mean absolute error 0.932%) to the overall model predictions (Appendix Table 4). In addition, we found little difference when P50 was calculated by antimicrobial class rather than individual compound (Pearson correlation coefficient 0.995, mean absolute error 0.66%) (Appendix Table 4, Figure 4). Importance of environmental covariates to the models varied by organism (Appendix Table 5). For *E. coli* and *Salmonella*, the covariate with highest importance was the percentage of tree coverage (ΔAUC 0.106 for *E. coli* and 0.078 for *Salmonella*). For *Campylobacter*, the covariate with highest importance was antimicrobial use in animals (ΔAUC 0.037), closely followed by yearly average of minimum monthly temperature (ΔAUC 0.034).

In 9 of the 31 countries in Europe, >50% of cattle, pigs, or poultry are estimated to be raised in the predicted AMR hotspot areas (Table 2). For instance, 93% of poultry in Spain, 90% of poultry in Greece, and 97% of poultry and 92% of pigs in Cyprus are raised in AMR hotspots.

Organism, country	Cattle (%)	Pigs (%)	Poultry (%)
E. coli			
France	0	0	0
Germany	0	0	0
Spain	2.1	2.3	1.8
Bulgaria	34.4	51.5	57.8
Cyprus	33.8	68.9	68.5
Greece	39.4	57.9	35.5
Romania	34.8	77.5	57.8
Salmonella			
France	0	0	0
Germany	0	0	0
Spain	8.8	28.2	24.8
Cyprus	51.8	91.6	96.6
Hungary	63.5	64.7	80.6
Italy	52.0	70.2	64.0
Poland	21.6	66.0	74.3
Romania	17.0	65.2	45.0
Campylobacter			
France	0.5	4.6	6.2
Germany	1.8	14.9	23.2
Spain	32.3	87.9	93.0
Cyprus	26.0	44.9	66.3
Greece	10.9	58.4	90.3
Portugal	22.1	74.9	88.0

*Antimicrobial resistance for *Escherichia coli*, nontyphoidal *Salmonella*, and *Campylobacter*.

Table 2. Percentages of food-producing animals raised in each country that fall within an antimicrobial resistance hotspot area (95th percentile per organism) for France, Germany, Spain, and countries in which organism percentage >50% for >1 animal species*

Discussion

In this study, we geographically disaggregated AMR prevalence for *E. coli*, nontyphoidal *Salmonella*, and *Campylobacter* reported among food-producing animals across Europe by supplementing national EFSA data with subnational PPS data to produce maps of estimated AMR prevalence. For multiple countries, such as Italy, Romania, and Poland, rather than consistently high countrywide AMR levels, in our final model we predicted specific geographic hotspots of high AMR prevalence that may coexist within regions of lower AMR prevalence in the same countries. In specific regions, countries in which AMR seems to be consistently high may have made more progress against AMR than previously thought (with only some, rather than all, areas containing high levels) by interpretation of EFSA data or nationally published reports. Further improvements could be made in those countries by targeting interventions (e.g., improved farm biosecurity and targeted surveillance in hotspots where AMR levels remain high). In contrast, largely diffuse and geographically uniform (low) countrywide AMR prevalence was found in countries with low AMR levels (e.g., Sweden, Norway, and Iceland); uncertainty in these predictions were higher for *Campylobacter* than for *E. coli* and *Salmonella*.

For all 3 bacteria studied, AMR prevalence was substantially lower in Norway, Sweden, Denmark, and Switzerland than the average for Europe. Those countries were among the first to establish animal AMR surveillance (i.e., DANMAP in Denmark in 1995[25]) and have now integrated surveillance of zoonotic bacteria in humans and animals. For several decades, they have been guiding national and international control strategies. For instance, in the 1990s, increased prevalence of vancomycin-resistant enterococci reported by DANMAP was instrumental to banning use of antimicrobial drugs for growth promotion in livestock[25].

In contrast, countries in which a high proportion of food-producing animals are raised in areas predicted as hotspots of resistance by our study are Cyprus, Portugal, and Spain. In 2018, one fifth (20.8%) of the pigs in the European Union were reared in Spain[26], where 88% of its pigs were predicted to be raised in geographic hotspots of *Campylobacter* resistance, primarily in Aragon and Catalonia. However, that finding was not the case for other high-density pig regions such as Brittany (France), northwest Germany (Lower Saxony and North Rhine-Westphalia), and Denmark[27]. Those findings suggest that high AMR is not necessarily associated with high animal densities but possibly with other drivers such as farming practices, biosecurity measures, and antimicrobial use[28].

Across Europe, the highest prevalence of resistance in our models was reported for antimicrobial drugs commonly used in animal production: tetracyclines, quinolones, penicillins, and aminoglycosides (gentamicin and streptomycin). Of particular concern were the compounds considered critically important antimicrobials for human medicine[29] and for which AMR prevalence was predicted to be >50% (ampicillin in *E. coli* [58.6%] and ciprofloxacin in *Campylobacter* [64.6%]).

In our study, estimates of P50 for *Salmonella* were much lower than those for *E. coli* and *Campylobacter*, which could potentially be attributed to the success of targets imposed by the European Union (e.g., reducing *Salmonella* prevalence in poultry over the past decade[30]). In addition, several countries had already implemented *Salmonella* control strategies before European Union–wide initiatives. For instance, in the 1970s, the United Kingdom set up national AMR surveillance for *Salmonella*, and in 1969, France had similar initiatives for *Salmonella* and *E. coli*[25]. Switzerland also implemented a stringent control program for *Salmonella* Enteritidis in 1993[31], more than a decade earlier than the first European Union–wide initiative[30].

When we compared estimates of resistance (P50) derived from PPS and EFSA data, the average P50 from PPSs seemed to more closely match national EFSA prevalence values in some countries more than in others. For instance, in Spain and Italy, the ratios of P50 inferred from PPS and EFSA data were close to 1 over the past 3 years. One reason may be the higher number of PPSs from these countries (17 in Spain and 13 in Italy), which average out closer to the EFSA values. In contrast, in countries with P50 ratios >2 or <0.8 (Poland, Germany, Greece, Portugal) inferred from PPS and EFSA data, only 1–4 studies have been conducted in the past 3 years. Therefore, although smaller sample sizes may be insufficient for comparing national averages (PPS vs. EFSA) they may still represent subnational heterogeneity in AMR not observed in the national average from EFSA. A higher coverage of PPSs may further improve the confidence in subnational model predictions.

Among the limitations of our modeling study, the first is that our literature search for PPSs published in Europe during 2000–2021 resulted in a mere 209 PPSs that were associated with

geographic information. In contrast, for the same period, 446 PPSs with geographic information were published in China[12]. Torres et al. also assembled AMR studies of food-producing animals during 1957–2018; however, of the 510 papers from Europe identified, the breakdown of their surveys corresponding to our search criteria was not available in open access[32]. Thus, the limited number of surveys that satisfied our inclusion criteria, particularly the reporting of geographic information, precluded mapping AMR prevalence for individual drug/bacteria combinations or animal species.

Second, with regard to using PPSs for regional estimations, differences in sampling strategy and sample sizes may affect the comparability of surveys and potentially explain why prevalence calculated from PPSs was in some instances higher than the prevalence estimates reported by EFSA. In particular, targeted sampling for bacteria that probably have high-resistance profiles, such as extended-spectrum beta lactamase–producing *E. coli*[33], could lead to comparatively higher AMR in PPS data than in the general population, which are more likely to be observed with the EFSA sampling scheme. In terms of microbiology, the set of tested antimicrobials differed between PPSs, which necessitated use of a composite metric. In addition, there were some transparency issues in terms of which methods or breakpoints were used (i.e., assumptions had to be made in the case of missing data [such as guideline year] and in the harmonization approach used for PPSs that used different guidelines, which may have led to some unintended bias), as well as a diversity of breakpoints used. Despite attempts to reduce variability between surveys, some variability may still exist and therefore efforts should be made to develop standardized protocols in the future, such as for all PPSs to shift to using ECOFF values and to release raw data. The creation of a consensus breakpoint table that could be used by all would also greatly assist with the comparability of those data and reduce the need for such adjustments. Because most studies reported only sampling location or region by name rather than specific coordinates, coordinates and size of region were estimated (and may not always represent the location of the farms where the animals were raised), which may have led to further uncertainty in our models.

Third, because of the limited number of PPSs, as well as their heterogenous distribution across the study period, incorporating the temporal dimension into the modeling framework remains challenging at this stage. Therefore, countries that have had considerably reduced AMR levels since 2009, such as the Netherlands[34], may be associated with higher AMR prevalence in our maps than that in the latest reports. However, as the number of surveys grows in the future, other spatio-temporal approaches, such as the Integrated Nested Laplace Approximation[35], could be used to account for not only spatial but also temporal variations in AMR prevalence extracted from PPSs.

Last, because of the static framework of geospatial modeling, it was not possible to incorporate all relevant data. That limitation may have a dynamic effect on AMR prevalence estimates, notably animal movement.

In conclusion, high-resolution maps that predict subnational hotspots can help support targeted resource allocation and control strategies for reducing AMR burden. Such strategies could include improving farm biosecurity and targeted surveillance. The accuracy of these maps could be gradually improved in the future should countries routinely report geographic location data along with microbiological sampling results.

Methods

EFSA data collection

We reviewed annual EFSA reports published during 2011–2022[13]. We extracted country-level data on AMR prevalence (2009–2020), focusing on the percentage resistance to antimicrobials against *E. coli*, *Salmonella*, *Campylobacter coli*, and *Campylobacter jejuni*. We extracted information on country, year of isolation, animal type (cattle, pigs, chickens, turkeys), sample origin (slaughtered animal, living animal, or meat), bacteria, species, number of samples, antimicrobial tested, and resistance prevalence. We followed European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines to assess microbiological resistance and used microdilution methods and epidemiologic cutoff (ECOFF) values[14]. We retained only antimicrobial/bacteria combinations recommended by the World Health Organization Advisory Group on Integrated Surveillance of Antimicrobial Resistance[15] for antimicrobial susceptibility testing (Appendix Table 1).

PPS data collection

We systematically reviewed PPSs (Appendix) reporting AMR prevalence at individual locations in Europe (Appendix Figure 1). We searched PubMed, Web of Science, and Scopus for PPSs reporting AMR prevalence for *E. coli*, nontyphoidal *Salmonella*, and *Campylobacter* in healthy cattle, pigs, and poultry (combined data for chickens, turkeys, or other poultry), as well as their products (meat and dairy) in Europe during 2000–2021. Environmental samples (e.g., water, soil) were not included. We also extracted information on the geographic location of the PPS (Appendix), the year the PPS was conducted, the year the bacteria was isolated (but not species identification methods used), sample types collected (cecal, cloacal, lymph, or fecal samples taken from living animals, slaughtered animals, dairy products, or meat), animal species, number of samples collected and tested, susceptibility testing guidelines used, and susceptibility guidelines used for resistance interpretation. Duplicated surveys, i.e. surveys analysing data that were previously published, were removed from the collection.

We assessed microbiological resistance across PPSs by using different methods (disk diffusion vs. broth dilution), guidelines (Clinical and Laboratory Standards Institute [https://www.clsi.org] 52%, EUCAST 29%, other 14.6%) and cutoffs (clinical break points vs. ECOFFs[15]). We attempted to account for these differences by using a harmonization approach developed by Van Boeckel et al.[2] (Appendix). We calibrated data from PPSs by using antimicrobial susceptibility testing, guidelines, and breakpoints reported in each study to match those of EUCAST guidelines each year, to enable comparison between those data and data reported by EFSA. As with EFSA data, we retained only antimicrobial/bacteria combinations recommended by the World Health Organization Advisory Group on Integrated Surveillance of Antimicrobial Resistance[15]. In addition, for our analysis we retained only countries that reported to EFSA and that had reported >50 samples during the study period. All prevalence estimates extracted from PPS are available at resistancebank.org (https://resistancebank.org)[16].

Comparative analysis of data sources

We used the proportion of antimicrobials with >50% resistance (P50s) to summarize trends in resistance across each drug/bacteria combination, as in previous works[2,12,17]; all P50s can be

recalculated by using the data available at resistancebank.org. To assess the difference in AMR prevalence between PPS and EFSA data, as well as the implications that that could have for geospatial modeling, we compared the average P50 in countries reporting at >1 PPS and to EFSA during 2018–2020 (Appendix Table 3). A ratio <1 indicated a lower 3-year mean P50 using PPS data, and a ratio >2 meant a more than double 3-year mean P50 from PPS data compared with EFSA data.

Geospatial modelling of P50

We mapped predicted subnational antimicrobial resistance in food-producing animals at a resolution of 0.08333 decimal degrees, corresponding to ≈10 km at the equator. To create the map, we used a 3-step procedure (Appendix Figure 2).

In the first step, we trained 3 child models (one of the individual models that are combined to form the final model) to quantify the relationship between P50 and a set of 9 environmental and anthropogenic covariates (Appendix Table 2). We selected those covariates because of their suspected association with AMR in animals[2,12,17,18,19]. The models used for the first step were boosted regression trees[20]; LASSO (least absolute shrinkage and selection operator) applied to logistic regression[21]; and overlapped grouped LASSO penalties for General Additive Models selection (A. Chouldechova, unpub. data, https://arxiv.org/abs/1506.03850). We calculated the importance of each covariate by comparing the areas under the receiver operator curve (AUCs) between a full model that contained all covariates and a model without each covariate. To evaluate the relative importance of each covariate to the full model, we repeated the procedure sequentially (Appendix Table 5).

We weighted all models by the number of isolates tested in each survey and conducted 10 Monto Carlo simulations on the models to account for the variation introduced by transformation of prevalence estimates into binary variables. The models were trained by using 4-fold spatial cross-validation to prevent overfitting and ensure generalization in geographic regions poorly represented in the training dataset. We defined the 4 spatial folds by using a k-means clustering algorithm[22]. The algorithm clustered the surveys according to their spatial distances and partitioned them into 4 spatially disjointed sets with equal sizes (Appendix). No predictions were made in urban settlements; there were areas defined as artificial surfaces in GlobCover 2009[23]. We conducted sensitivity analyses by restricting PPSs to 2009–2020 only (to match EFSA reporting period), to 6 or 7 of the most common antimicrobial/bacteria combinations only, and to P50 calculated by class (rather than compound) (Appendix).

In the second step, we ensembled predictions from the 3 models according to the models' predictive ability, assessed by using the AUC. We calculated the resulting map of P50 as the mean of the 3 model predictions weighted by their AUC values. We calculated the associated map of prediction uncertainty as the SD of predicted P50 values from the 10 Monte Carlo simulations (Appendix Figure 4, panel A).

In the third step, we adjusted the P50 predictions in each country, using P50 values calculated from EFSA reports. Concretely, we multiplied P50 values in each pixel by the ratio of country-level P50 as reported by EFSA and the mean P50 of all pixels across each country as predicted by the geospatial model. That step ensured that the country-level mean of P50 values corresponded to reports from EFSA while preserving geographic variations in AMR levels within each country. To assess the variations in P50 values within each country, we calculated country-level SDs of P50s (Appendix Figure 4, panel C).

Last, we created the predictive maps of AMR hotspots for each organism. The threshold value for a pixel to be classified as a hotspot corresponded to the 95th percentile of all P50 values across the map and varied for each organism (Appendix Figure 4, panel B). We obtained estimated animal densities associated with those areas from Gilbert et al.[24]. Using those estimates, for each country we calculated the percentage of each animal species living in the hotspot areas.

References

1. Antimicrobial Resistance Collaborators, "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis," *The Lancet,* vol. 399, no. 10325, pp. 629-655, 2022.
2. T. P. Van Boeckel, J. Pires, R. Silvester , C. Zhao and J. Song, "Global trends in antimicrobial resistance in animals in low- and middle-income countries," *Science* , vol. 365, no. 6459, 2019.
3. D. Muloi, M. J. Ward, A. B. Pedersen, E. M. Fèvre, M. E. Woolhouse and B. A. van Bunnik, "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,* vol. 15, no. 8, pp. 467-474, 2018.
4. Food and Agriculture Organisation of the United Nations, "The economic lives of smallholder farmers: an analysis based on household data from nine countries," Rome, 2015.
5. B. Bengtsson and C. Greko, "Antibiotic resistance - consequences for animal health, welfare, and food production," *Upsala Journal of Medical Sciences,* vol. 119, no. 2, 2014.
6. T. Deruelle, "A tribute to the foot soldiers: European health agencies in the fight against antimicrobial resistance," *Health Economics, Policy and Law,* vol. 16, no. 1, 2020.
7. ECDC, "The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019," *The EFSA Journal,* vol. 19, no. 4, pp. 1-179, 2021.
8. European Union, "Regulation (EU) 2019/6 of the European Parliament and of the Council of 11 December 2018 on veterinary medicinal products and repealing Directive 2001/82/EC (Text with EEA relevance)," 2019. [Online]. Available: Available at http://data.europa.eu/eli/reg/2019/6/oj/eng. [Accessed 27 04 2022].
9. D. A. Ramos, J. A. H. Pulgarín, G. A. M. Gómez, J. A. Alzate, J. C. O. Gómez, I. C. Bonilla and C. V. Mosquera, "Geographic mapping of Enterobacteriaceae with extended-spectrum B-lactamase (ESBL) phenotype in Pereira, Colombia," *BMC Infectious Diseases,* vol. 20, no. 540, 2020.
10. Y. Agersø and F. Aarestrup, "Voluntary ban on cephalosporin use in Danish pig production has effectively reduced extended-spectrum cephalosporinase-producing Escherichia coli in slaughter pigs," *Journal of Antimicrobial Chemotherapy,* vol. 68, no. 3, pp. 569-572, 2013.
11. S. Raasch, L. Collineau and M. Postma, "Effectiveness of alternative measures to reduce antimicrobial usage in pig production in four European countries," *Porcine Health Management,* vol. 6, no. 6, 2020.
12. C. Zhao, Y. Wang, K. Tiseo, J. Pires, N. G. Criscuolo and T. P. Van Boeckel, "Geographically targeted surveillance of livestock could help prioritise intervention against antimicrobial resistance in China," *Nature Food,* no. 2, pp. 596-602, 2021.
13. EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control). The European Union Summary Report on Antimicrobial Resistance

in zoonotic and indicator bacteria from humans, animals and food in 2019-2020. EFSA Journal 2022;20(3):7209, 197 pp.
14. European Society of Clinical Microbiology and Infectious Diseases, "Standard Operating Procedure. MIC distributions and the setting of epidemiological cut-off (ECOFF) values.," 2019.
15. World Health Organization, "Integrated surveillance of antimicrobial resistance in foodborne bacteria: application of a one health approach: guidance from the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR)," World Health Organization, Geneva, 2017.
16. N. G. Criscuolo, J. Pires, C. Zhao and T. P. Van Boeckel, "resistancebank.org, an open-access repository for surveys of antimicrobial resistance in animals," *Nature Scientific Data,* vol. 8, no. 189, 2021.
17. D. Shar, C. Zhao, Y. Wang, D. G. Joakim Larsson, M. Gilbert and T. P. Van Boeckel, "Twenty-year trends in antimicrobial resistance from aquaculture and fisheries in Asia," *Nature Communications,* vol. 12, no. 1, p. 5384, 2021.
18. D. R. MacFadden, S. F. McGough, D. Fisman, M. Santillana and J. S. Brownstein, "Antibiotic resistance increases with local temperature," *Nature Climate Change,* vol. 8, pp. 510-514, 2018.
19. K. Tang, N. Caffrey, D. Nóbrega, S. Cork, P. Ronksley and H. Barkema, "Restricting the use of antibiotics in food-producing animals and its association with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis," *The Lancet Planetary Health,* vol. 1, no. 8, pp. E316-E327, 2017.
20. J. Elith, J. R. Leathwick and T. Hastie, "A working guide to boosted regression trees," *J. Anim. Ecol.,* vol. 77, pp. 802-813, 2008.
21. R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *J. R. Stat. Soc. Ser. B-Stat. Methodol.,* vol. 73, pp. 273-282, 2011.
22. A. Chouldechova and T. Hastie, "Generalized additive model selection," *ArXiv Prepr.,* vol. ArXiv150603850, 2015.
23. J. A. Hartigan and M. Wong, "Algorithm AS 136: A k-means clustering algorithm," *J. R. Stat. Soc. Ser. C Appl. Stat,* vol. 28, pp. 100-108, 1979.
24. S. Bontemps, P. Defourny, E. Van Bogaert, O. Arino, V. Kalogirou and J. Ramos Perez, "GLOBCOVER 2009".
25. M. Gilbert, G. Conchedda, T. P. Van Boeckel, G. Cinardi, C. Linard, G. Nicolas, W. Thanapontharm, L. D'Aietti, W. Wint, S. H. Newman and T. P. Robinson, "Income disparities and the global distribution of intensively farmed chicken and pigs," *PLoS ONE,* vol. 10, no. 7, p. e0133381, 2015.
26. S. Simjee, P. McDermott, D. J. Trott and R. Chuanchuen, "Present and future surveillance of antimicrobial resistance in animals: principles and practices," *Clinical Microbiology,* vol. 6, no. 4, 2018.
27. European Union, "The EU pig meat sector: European Parliament Briefing," 09 2020. Online]. Available: https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/652044/EPRS_BRI(2020)652044_EN.pdf. [Accessed 27 04 2022].
28. "Pig farming sector - statistical portrait 2014: density of sows.," Eurostat, 05 10 2018. Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Pig_farming_sector_-_statistical_portrait_2014. Accessed 27 04 2022].
29. L. Huber, A. Agunos, S. P. Gow, C. A. Carson and T. P. Van Boeckel, "Reduction in antimicrobial use and resistance to Salmonella, Campylobacter, and Escherichia coli in

broiler chickens, Canada, 2013-2019," *Emerging Infectious Diseases,* vol. 27, no. 9, pp. 2434-2444, 2021.
30. World Health Organization, "Critically important antimicrobials for human medicine: 6th revision 2018," 2019. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/312266/9789241515528-eng.pdf. [Accessed 29 04 2022].
31. EUR-Lex, "COMMISSION REGULATION (EC) No 1177/2006: requirements for the use of specific control methods in the framework of the national programmes for the control of salmonella in poultry," Official Journal of the European Union, 02 08 2006. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32006R1177. [Accessed 19 04 2022].
32. R. K. Koop, "The swiss control programme for Salmonella enteritis in laying hens: experience and problems," *Rev Sci Tech,* vol. 16, no. 3, pp. 885-90, 1997.
33. R. T. Torres, J. Carvalho, J. Fernandes, J. D. Palmeira, M. V. Cunha and C. Fonseca, "Mapping the scientific knowledge of antimicrobial resistance in food-producing animals," *One Health,* vol. 13, p. 100324, 2021.
34. N. Geser, R. Stephan, P. Kuhnert, R. Zbinden, U. Kaeppeli, N. Cernela and H. Haechler, "Fecal carriage of extended-spectrum β-lactamase-producing Enterobacteriaceae in swine and cattle at slaughter in Switzerland," *Journal of Food Protection,* vol. 74, no. 3, pp. 446-449, 2011.
35. A. Hesp, K. Veldman, J. van der Goot, D. Mevius and G. van Schaik, "Monitoring antimicrobial resistance trends in commensal Escherichia coli from livestock, the Netherlands, 1998 to 2016," *Eurosurveillance,* vol. 24, no. 25, 2019.
36. M. Blangiardo, M. Cameletti, G. Baio and H. Rue, "Spatial and spatio-temporal models with R-INLA," *Spatial and Spatio-temporal Epidemiology,* vol. 4, no. 2013, pp. 33-49, 2013.

Chapter 4

Global surveillance of antimicrobial resistance in food animals using priority drugs maps.

Authors:
Cheng Zhao[1], Yu Wang[1], Ranya Mulchandani[1], Thomas P. Van Boeckel[1,2,3,*]

Affiliations:
[1]Health Geography and Policy Group, ETH Zürich, Zürich, Switzerland.
[2]One Health Trust, Washington DC, USA.
[3]Spatial Epidemiology Lab, Université Libre de Bruxelles, Brussels, Belgium.

*Correspondence to: thomas.van.boeckel@gmail.com

Published in:
Nature Communications 15, 763 (2024). DOI: 10.1038/s41467-024-45111-7
This chapter contains minor revisions compared to the published article.

Contribution remarks:
C.Z. and T.P.V.B. designed the research; C.Z. and Y.W. collected the data; C.Z. analysed data; C.Z., R.M., and T.P.V.B wrote the paper.

Abstract

Antimicrobial resistance (AMR) in food animals is a growing threat to animal health and potentially to human health. In resource-limited settings, allocating resources to address AMR can be guided with maps. Here, we mapped AMR prevalence in 7 antimicrobials in *Escherichia coli* and nontyphoidal *Salmonella* species across low- and middle-income countries (LIMCs), using 1088 point-prevalence surveys in combination with a geospatial model. Hotspots of AMR were predicted in China, India, Brazil, Chile, and part of central Asia and southeastern Africa. The highest resistance prevalence was for tetracycline (59% for *E. coli* and 54% for nontyphoidal *Salmonella*, average across LMICs) and lowest for cefotaxime (33% and 19%). We also identified the antimicrobial with the highest probability of resistance exceeding critical levels (50%) in the future (1.7–12.4 years) for each 10 x 10 km pixel on the map. In Africa and South America, 78% locations were associated with extended-spectrum penicillins or tetracyclines crossing 50% resistance in the future. In contrast, in Asia, 77% locations were associated with extended-spectrum penicillins or sulphonamides. Our maps highlight diverging geographic trends of AMR prevalence across antimicrobial classes, and can be used to target AMR surveillance in AMR hotspots for priority antimicrobial classes.

Introduction

Antimicrobials are life-saving drugs used to treat infections in humans. However, the majority (73%) of antimicrobials sold globally are used in animals raised for food[1]. In animals, antimicrobials are used for treatment but also as surrogates for good hygiene practices and to increase productivity on farm in some regions of the world[2]. Antimicrobials have facilitated the intensification of animal farming, and enabled meeting a growing demand for animal proteins worldwide. From 2000 to 2020, In Brazil and China – the largest exporter and importer of meat among low- and middle- income countries (LMICs), meat production has grown by 89% and 23% respectively[3]. However, in LMICs, during the same period, the percentage of antimicrobials with prevalence of resistance higher than 50%[4] rose from 15% to 41% in chicken, and from 13% to 34% in pigs, with important consequences for animal health, and potentially for human health[5,6].

In high-income settings such as the US[7], Canada[8], and the EU countries[9], animal AMR has been the focus of systematic surveillance for decades. Surveillance data have supported policies that helped limiting the use of certain classes of antimicrobials in animals (e.g. third-generation cephalosporins such as ceftiofur[8]). However, in LMICs, systematic surveillance remains at best nascent, and point prevalence surveys (PPS) have been used as surrogates to systematic surveillance to infer regional trends in AMR in animals[10]. Thus far, these attempts at documenting trends in AMR using PPS relied on summary metrics such as the fraction of antimicrobials tested in a survey with prevalence of resistance higher than 50% (P50). For LMICs, trends in AMR have not yet been disaggregated for individual antimicrobial-bacteria combinations. This is a major limitation for potentially taking targeted actions on individual antimicrobial classes. One such action was the 2005 ban of fluoroquinolones in poultry in the United States that was supported by surveillance data of fluoroquinolone-resistant *Campylobacter*[11]. For humans, systematic reviews – including in LMICs – helped estimate the global burden of AMR for 88 individual antimicrobial-bacteria combinations[12]. Conducting a symmetrical exercise for animals would enable a more targeted approach to the management of AMR in animals, and also comparison with patterns of AMR in humans[13].

The World Health Organization's list of Medically Important Antimicrobials (MIA)[14] is a natural starting point for developing drug-specific guidelines for surveillance of AMR in animals, and define priorities for actions. However, the MIA list does not explicitly account for considerable geographic variations of AMR levels within countries such as Kenya, China, Thailand where subsistence farming and industrial farming co-exist, and where access to veterinary services varies considerably between regions. In regions with high AMR levels, first line antimicrobials for disease treatment may have lost efficacy. Having the ability to predict which antimicrobials will cross critical resistance levels in the future could help assess the risk of antimicrobial resistant infections acquired from animal sources, as well as strengthening local surveillance effort. To the best of our knowledge, there are currently no maps that prioritize antimicrobial classes for surveillance in animals based on local epidemiological patterns of AMR. This is largely due to the lack of fine-grained geographic information on AMR prevalence for individual antimicrobial-bacteria combinations in systematic surveillance systems. PPS conducted at individual locations provide a unique opportunity for supplementing these efforts, and mapping priority antimicrobials for AMR surveillance. However, several challenges must be addressed to transform data extracted from event-based surveillance (PPS) into actionable epidemiological information. Firstly, combinations of antimicrobial-bacteria vary between PPS, and a panel of combinations that are abundantly represented across PPS must be selected to ensure comparability. Secondly, not only local environmental and anthropogenic covariates but also patterns of co-resistance between antimicrobials observed in PPS[15,16] can be informative of the future resistance profiles, but an appropriate computational framework must be developed to transform these statistical associations into predictions of classes of antimicrobials that will reach critical resistance levels.

In this study, we used 1,088 PPS to map, at 10x10 kilometer resolution, the prevalence of resistance to 7 antimicrobials in *E. coli* and nontyphoidal *Salmonella* species in food animals. We combined the maps of resistance prevalence with environmental and anthropogenic covariates as well as patterns of co-resistance to predict, in each location, which antimicrobials had the highest probability of exceeding critical levels of resistance (10%, 25% or 50%) in the near future. Our output is a global map displaying fine-scale variations of these drugs that will reach critical resistance levels, and could serve a basis to refine AMR surveillance efforts across regions.

Results

Trends of AMR

The mean prevalence of resistance weighted by the number of samples in each PPS, in *E. coli* and nontyphoidal *Salmonella*, was respectively 59% (n = 745) and 54% (n = 597) for tetracycline (TET), 57% (n = 779) and 46% (n = 632) for ampicillin (AMP), 45% (n = 649) and 36% (n = 501) for sulfamethoxazole-trimethoprim (SXT), 35% (n = 656) and 26% (n = 553) for chloramphenicol (CHL), 30% (n = 796) and 26% (n = 624) for ciprofloxacin (CIP), 28% (n = 882) and 23% (n = 650) for gentamicin (GEN), and 33% (n = 446) and 19% (n = 334) for cefotaxime (CTX). Between 2000 and 2019, changes in the prevalence of resistance were +12% (TET), +33% (AMP), +19% (SXT), +20% (CHL), +16% (CIP), +11% (GEN),

and +37% (CTX) (Figure 1). The temporal increases of resistance were significant ($p < 0.05$) for all antimicrobials apart from TET.

Prevalence of resistance was investigated in poultry in 52% (n = 570) of PPS, in cattle in 38% (n = 409) of PPS, and in pigs in 28% (n = 303) of PPS. Prevalence of resistance increased significantly for AMP, CHL, CIP, and CTX for poultry, and for AMP, SXT, CHL, CIP, GEN, and CTX for pigs (Supplementary Figures 1 to 2). However, temporal trends of resistance were not significant for any antimicrobial classes for cattle (Supplementary Figures 3).

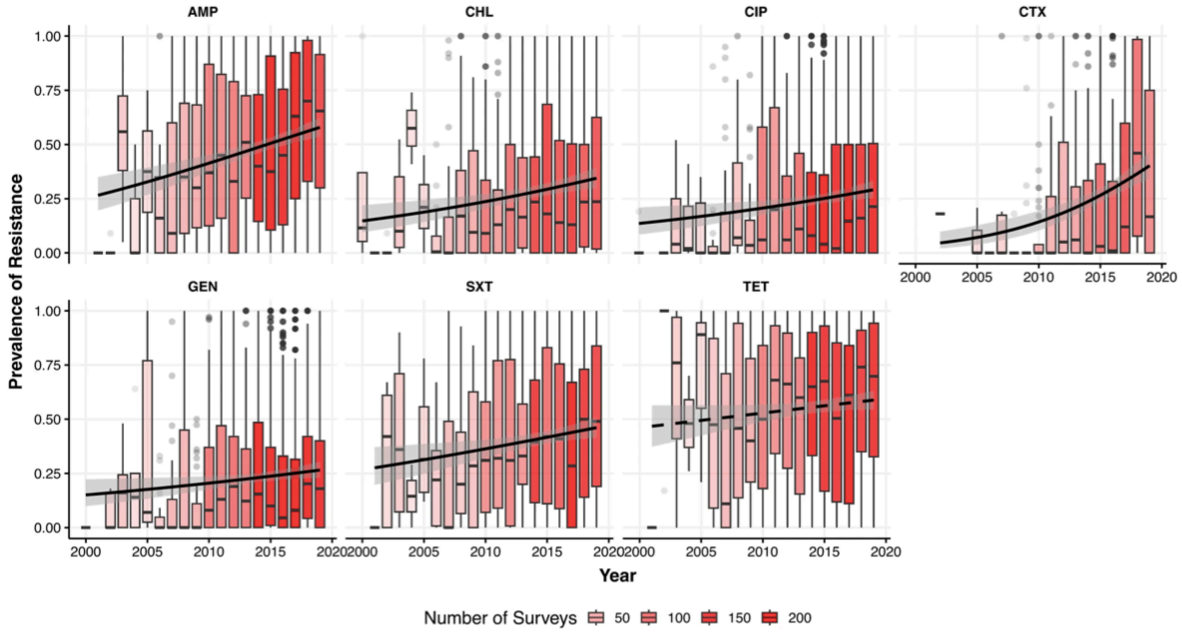


Figure 1. Temporal trends of the prevalence of resistance in low- and middle-income countries for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Transparency levels of the red colors are proportional to the number of surveys published each year. The 25th and 75th percentiles are represented by the lower and upper limits of each box, and the median value is marked with a horizontal line. Lengths of whiskers are 1.5 times the inter-quartile ranges, and values outside of this range are shown as individual points. Logistic regression is used to fit temporal trends of resistance prevalence. Solid lines represent significant temporal trends ($p < 0.05$; AMP: $p = 0.00000043$, CHL: $p = 0.00057$, CIP: $p = 0.0043$, CTX: $p = 0.00000023$, GEN: $p = 0.023$, SXT: $p = 0.0050$), and dashed lines represent nonsignificant trends (TET: $p = 0.061$). No adjustments are made for multiple comparisons. The 95% confidence intervals of the estimated temporal trends are shown in the grey areas.

We used ensemble geospatial modelling to map the prevalence of resistance to 7 antimicrobials in *E. coli* and nontyphoidal *Salmonella* in animals (Methods). In *E. coli*, resistance hotspots defined as N50 ≥ 3 were predicted in southern and eastern China, central Asia, northern India, northern Brazil, and Chile (Figure 2h). In nontyphoidal *Salmonella*, resistance hotspots were predicted in northeastern China (Figure 3d). Maps of resistance using other cutoff values (N20 and N35) were shown in Supplementary Figure 7. Northern and eastern Brazil was also resistance hotspots for AMP resistance in *E. coli* (Figure 2b).

Northeastern China was resistance hotspots for SXT and GEN resistance in *E. coli*, as well as CHL, CIP, and GEN resistance in nontyphoidal *Salmonella* (Figure 2c, 2f, 3d, 3e, 3f). Uncertainty of the predictions was the highest for CTX resistance in nontyphoidal *Salmonella* and the lowest for TET resistance in nontyphoidal *Salmonella* –standard deviation of predictions across all pixels on the map was on average 19.9% and 4.5%, respectively (Supplementary Figures 8 and 9).

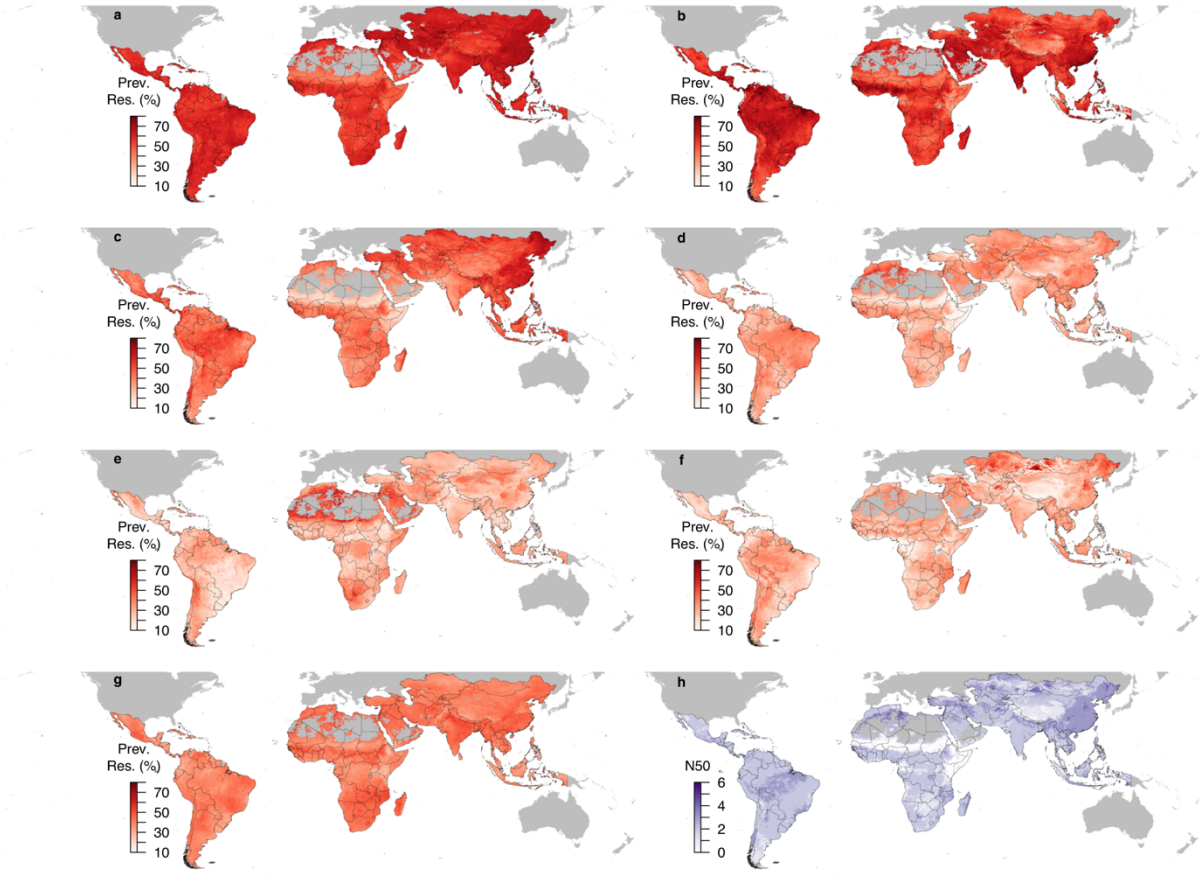


Figure 2. Geographic distribution of antimicrobial resistance in *E. coli* in low- and middle-income countries between 2000 and 2019 (median year 2015). Prevalence of resistance (Prev. Res.) for tetracycline (a), ampicillin (b), sulfamethoxazole-trimethoprim (c), chloramphenicol (d), ciprofloxacin (e), gentamicin (f), cefotaxime (g). Overall resistance level across antimicrobials measured using the number of antimicrobials (out of 7) with resistance higher than 50% (N50; h) (See Supplementary Figure 7 for maps generated using cutoff values other than 50%). Maps of resistance prevalence for the 7 antimicrobials are available on resistancebank.org.

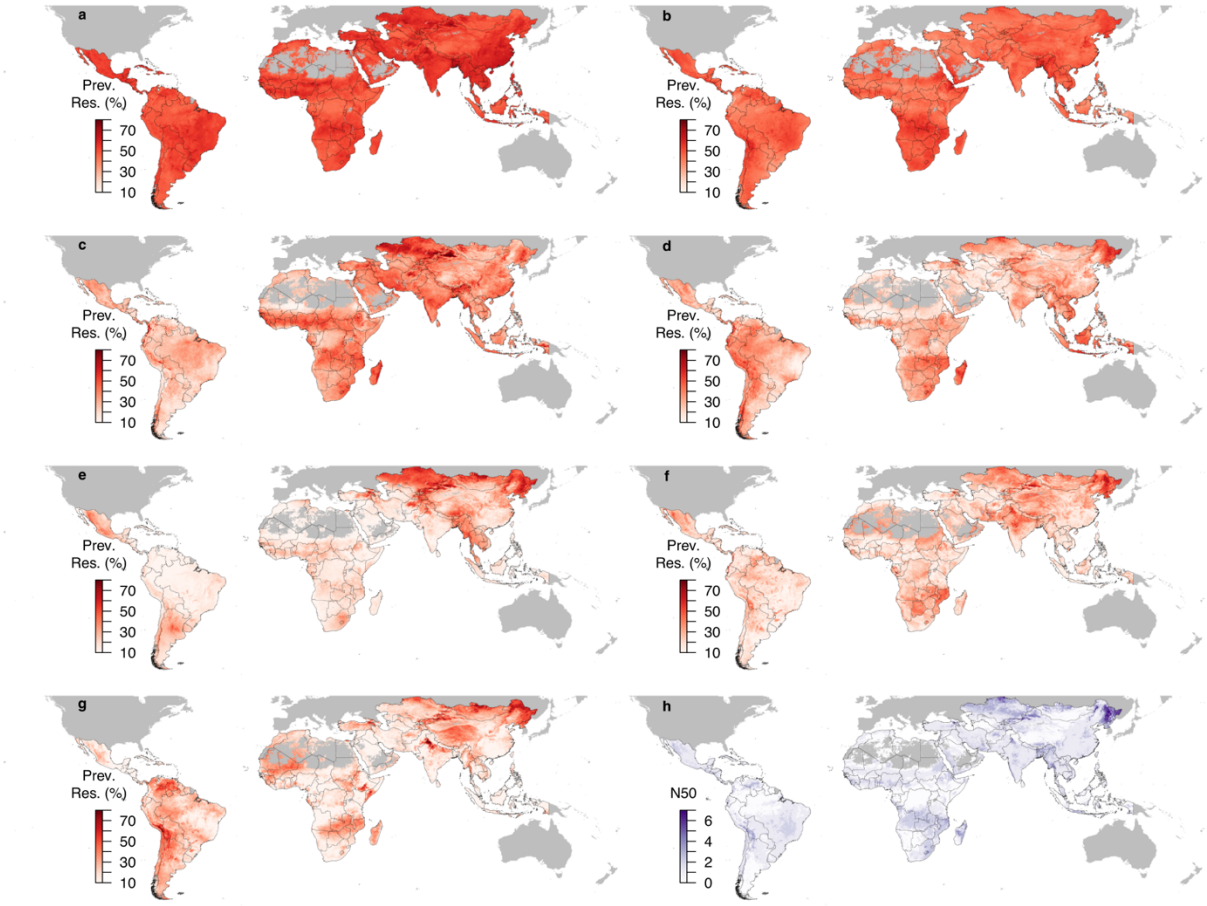


Figure 3. Geographic distribution of antimicrobial resistance in nontyphoidal *Salmonella* in low- and middle-income countries between 2000 and 2019 (median year 2015). Prevalence of resistance (Prev. Res.) for tetracycline (a), ampicillin (b), sulfamethoxazole-trimethoprim (c), chloramphenicol (d), ciprofloxacin (e), gentamicin (f), cefotaxime (g). Overall resistance level across antimicrobials measured using the number of antimicrobials (out of 7) with resistance higher than 50% (N50; h) (See Supplementary Figure 7 for maps generated using cutoff values other than 50%). Maps of resistance prevalence for the 7 antimicrobials are available on resistancebank.org.

Mapping priority antimicrobials for AMR surveillance – where and for which antimicrobial class will resistance prevalence exceed critical levels in the future?

We used risk factors (Supplementary Table 3) associated with the locations of each resistance profile reported in PPS, in combination with histories of acquisition of resistance phenotypes, to map which antimicrobials had the highest probability of its resistance prevalence exceeding critical levels (10%, 25% or 50%) in the future (Methods). This resulted in global maps of 'priority antimicrobials' for AMR surveillance. Using 50% as the critical level of resistance prevalence, the predicted priority antimicrobials were TET or AMP in 78% locations In Africa and South America (Figure 4a). In contrast, in Asia, 77% locations were associated with AMP or SXT, because resistance to TET has already exceeded 50% in the vast majority of locations (83%). Concretely, SXT was the priority antimicrobial in northeastern India, southern and northeastern China, southern Brazil, Turkey, and Iran; AMP

was the priority antimicrobial in northern and western China, Mongolia, and western India. In southern and eastern China, CHL was the predicted priority antimicrobial. Predictions of GEN and CTX having the highest probability of exceeding 50% resistance were not frequent (0.02%) and scattered in Asia and South America. The uncertainty associated with the predicted priority antimicrobials was on average 12% across all pixels (Figure 4b), and was high (> 40%) in parts of western Brazil, South Sudan and North Korea. The percentage of pixels with high uncertainty (> 40%) for each country was calculated in Supplementary Table 11. We estimated the time for resistance prevalence to exceed 50%, for the predicted priority antimicrobial in each 10 x 10 km pixel (Supplementary Figure 10). Across locations where AMP was the predicted priority antimicrobial (Supplementary Figure 10b), the average time weighted by animals' biomass was 1.7 years, while for CIP the average time was 12.4 years (see Supplementary Table 4 for the average estimated time for each antimicrobial class). Maps of priority antimicrobials using 10% and 25% as critical levels of resistance prevalence are shown in Supplementary Figure 7.

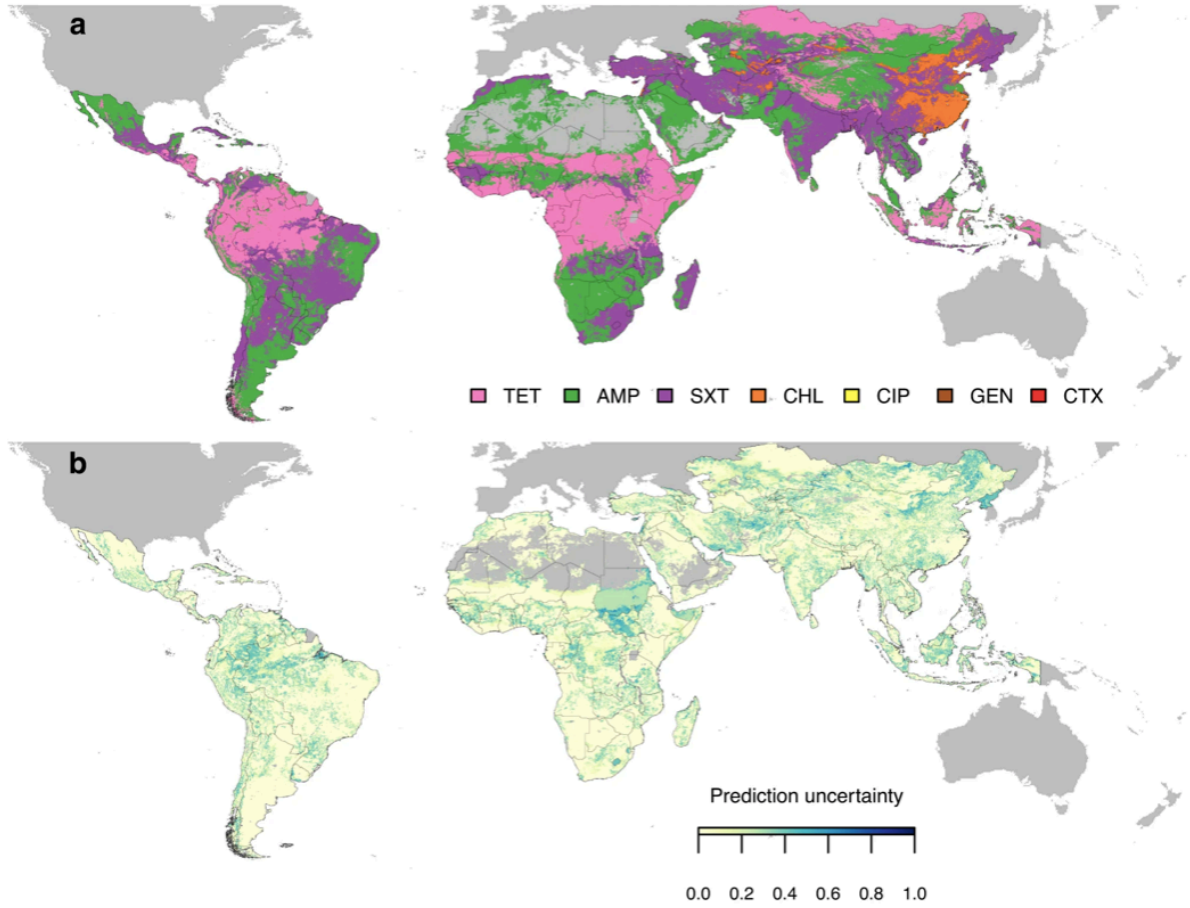


Figure 4. Geographic distribution of priority antimicrobials. a Geographic distribution of antimicrobials with the highest probability of their resistance prevalence exceeding 50% in the future in low- and middle-income countries. TET: tetracycline; AMP: ampicillin; SXT: sulfamethoxazole-trimethoprim; CHL: chloramphenicol; CIP: ciprofloxacin; GEN: gentamicin; CTX: cefotaxime. b Estimated uncertainty of the predictions shown in panel a, introduced by the imputation of missing resistance prevalence in the input dataset. Blues shades indicate the proportion of Monte Carlo simulations of imputed datasets, which generated different predictions compared with panel a (Methods).

We assessed the accuracy of the predicted priority antimicrobials using spatial cross-validation (Methods) and calculated the area under the receiver operating characteristic curve (AUC) for predicting the probability of resistance prevalence exceeding 50% for each antimicrobial. The AUCs ranged from 0.880 to 0.994 between antimicrobials. We calculated the influence of each covariate for explaining the divergence in prediction accuracy by sequentially excluding these covariates from the models, and calculated the loss in AUC. Co-resistance patterns had the highest influence on predicting resistance to all antimicrobials, with ΔAUC ranging from 0.224 to 0.494. In contrast, environmental and anthropogenic covariates had limited added value for predicting whether resistance exceeds 50% in TET and AMP (ΔAUC 0.002 and -0.003), yet they increased prediction accuracy for other antimicrobials (ΔAUC ranging from 0.109 to 0.416). Covariates that were most frequently associated with the probability of resistance prevalence exceeding 50% were antimicrobial use, pesticide application rate, tri-annual cycles of precipitation, and amplitudes of night land surface temperature (Supplementary Table 5).

Discussion

In this study, we mapped the distribution of resistance prevalence for 7 antimicrobials in *E. coli* and nontyphoidal *Salmonella* in food animals in low- and middle-income countries. We mapped the antimicrobials with the highest probability of their resistance prevalence exceeding critical levels (10%, 25% or 50%) in the future.

Geographic distribution of AMR

The predicted maps of AMR based on the number of antimicrobials with resistance higher than 50% (N50) were consistent with previous global estimates of AMR in Van Boeckel & Pires *et al*. 2019. This consistency can be partly attributed to the incorporation of a subset of PPS used in Van Boeckel & Pires *et al*. 2019 into the present analysis (Supplementary Information). In both analyses, China, Turkey, Iran, India and Brazil were identified as hotspots of AMR. However, the previous authors estimated trends of AMR for four pathogens combined. Our analysis was conducted for *E. coli* and nontyphoidal *Salmonella* separately, and ensured comparability for monitoring AMR trends by including data on 7 drugs each representing a medically important class of antimicrobials. Our choice of proxies was also in line with other global surveillance initiatives, such as the Global Tricycle Surveillance that uses ESBL-producing *E. coli* as the proxy[17].

In this analysis, we showed that the geographic distribution of AMR varied depending on the bacteria considered. For example, Iran was resistance hotspots of extended-spectrum penicillins and amphenicols resistance in *E. coli* but not in nontyphoidal *Salmonella*. On average, *E. coli* had higher prevalence of resistance compared with nontyphoidal *Salmonella* for all antimicrobials. We also showed that the geographic distribution of AMR varied depending on the classes of antimicrobials considered. For example, in either *E. coli* or nontyphoidal *Salmonella*, northeastern China was identified as resistance hotspots for all antimicrobials except tetracyclines and extended-spectrum penicillins. These two classes of antimicrobials have already reached high levels of resistance globally, leaving the preservation of the other antimicrobials of particular importance. Therefore, this region may need intensified policy intervention to contain AMR. Despite variations of AMR trends between antimicrobials, there were also consistencies on their geographic distributions. For example, Africa had consistently lower AMR prevalence compared the rest of the world for all antimicrobials, possibly because it consumes the least amount of veterinary antimicrobials compared with the rest of the world [18].

The 7 antimicrobial classes included in the analysis are the most frequently cited classes across 1,088 point prevalence survey, and are important for treating infectious diseases in food animals. For example, tetracycline is widely used for treating Mycoplasma in chicken[19], gentamicin is used for treating *Pseudomonas aeruginosa* infections[20], and third- and fourth-generation cephalosporins are used for treating cattle mastitis[20]. Therefore, rising resistance levels in these drugs may lead to therapy failure, and thereby negatively impact animal health and the agricultural economy. Measures to contain AMR in the identified hotspot regions will need to be focused on reducing antimicrobial use as well as strengthening biosecurity in farms. Enforcing a regulation with a cap of 50 milligram antimicrobial used per kilogram of food animal products was estimated to reduce global antimicrobial consumption by 64%[1]. However, major investment on the surveillance of antimicrobial use is needed for such regulations to be effective. Improving biosecurity in farms may reduce the reliance on

antimicrobials for keeping the animals healthy. Measures to improve biosecurity include stricter hygienic control on farm entry and better separation between compartments in the farm, and can be facilitated by risk-based quantitative tools[21].

Priority antimicrobials for AMR surveillance

We developed a computational approach (Methods), to map priority antimicrobials for surveillance that incorporates dependencies on local risk factors such as antimicrobial use and animal density, as well as history of acquisition of resistance phenotypes in one location. This approach uses spatial variations of resistance profiles of multiple drugs to infer which antimicrobial has the highest probability of its resistance prevalence exceeding critical levels (10%, 25% or 50%) in the future. If 50% was considered the critical level of resistance prevalence, in regions with currently low resistance levels (N50 = 0 or 1), tetracyclines and extended-spectrum penicillins were the most frequently predicted antimicrobials with their resistance prevalence exceeding 50% in the future. Predictions of these two antimicrobial classes were based primarily on patterns of co-resistance between antimicrobials, with little influence from environmental and anthropogenic covariates. This suggested that such patterns were universal across regions with low AMR, with the following possible reasons. Firstly, tetracyclines are among the cheapest and most accessible antimicrobials globally[22]. Secondly, tetracycline and ampicillin were discovered the earliest among the 7 antimicrobials included in the analyses. Their routine application for growth promotion in farms started as early as in the 1950s[23]. These factors may make them drugs of choice for application in food animals in regions with limited budgets and where their resistance has not yet been established[24].

In contrast, in regions with high AMR levels (N50 ≥ 2), sulfonamides and amphenicols were the antimicrobials with the highest probability that their prevalence of resistance will exceed 50% in the future. For amphenicols, the predictions were in eastern and southern China, where resistance to tetracyclines, extended-spectrum penicillins, and sulfonamides were already above 50%. In China, despite chloramphenicol being banned for use in food animals since 2002 and other amphenicols being banned as growth promoters in 2020, increases in the prevalence of resistance to chloramphenicol[25] and florfenicol[26] continued to be observed years after the restrictions took place. The increases may be caused by the continued use of the drugs despite changes in regulation, or by co-selection of their resistance (e.g., associated with class 1 integrons) due to the use of other drugs such as dihydrostreptomycin and trimethoprim[27]. Our predictions suggested that future surveillance on use of amphenicols and its resistance could be intensified in these regions to better understand mechanisms underlying these trends.

Environmental and anthropogenic covariates were predictive of the priority antimicrobials for AMR surveillance, particularly in regions associated with high levels of AMR. Population densities of animals were influential covariates, possibly because commonly applied antimicrobials differ between animal species[28]. Therefore, the difference in antimicrobial use across animal species may lead to difference in AMR. Temperature may affect the prevalence of animal injuries and therefore the frequency of (preventive) drug application[29].

We estimated the time it may take for resistance prevalence of the predicted priority antimicrobials to exceed a critical level. For locations where tetracyclines, extended-spectrum penicillins or cephalosporins were the predicted priority antimicrobials, the average time for resistance to reach 50% across locations was below 7 years. Given that the median year of

publication of the PPS was 2015, this implies that resistance may have already exceeded 50% now at these locations. For locations where amphenicols or quinolones were the predicted priority antimicrobials, their prevalence of resistance was estimated to exceed 50% in 2026 and 2027 on average across locations. However, the temporal trends of AMR used for estimating the time was based on data across low- and middle-income countries, and may differ depending on the geographic region. Estimating separate trends for each region is challenged by the limited amount of data in America and Africa countries, with large uncertainty associated with the estimated coefficients (Supplementary Table 2). Future work may be able to make region-specific projections of AMR as the number of PPS published each year steadily increases and more data becomes available.

Our prediction of priority antimicrobials was based on surveys conducted exclusively on commensal *E. coli* and nontyphoidal *Salmonella* from healthy animals, and the majority of surveys used human clinical breakpoints to determine resistance phenotype. However, our approach could also be adapted to databases of AMR of other animal pathogens using veterinary clinical breakpoints, to help inform veterinarians on possible treatment options in regions of high AMR levels.

Co-resistance between antimicrobials

Across surveys (n=1,088), resistance prevalence was significantly correlated between antimicrobials. All correlations were positive, a finding consistent with studies that interpreted collateral resistance using Markov network[30]. However, these observations were based on resistance profiles at the population level, rather than at the strain level where a diversity of both collateral resistance and sensitivity have been shown *in silico*[31] and *in vitro*[32]. Our results based on the amalgamation of PPS suggested that, at the population level, higher resistance in one drug is consistently associated with higher resistance in other drugs.

The highest correlations of resistance between antimicrobials were observed for sulfamethoxazole-trimethoprim and chloramphenicol, and for sulfamethoxazole-trimethoprim and tetracycline. One reason could be the co-location of several resistance genes on the same genetic element. For example, in *E.coli* isolated from pigs, chloramphenicol resistance gene *cmlA* was found on large plasmids that were linked to sulphonamide resistance genes *sul1* or *sul3*[33]. In addition, animals may often be exposed simultaneously to tetracyclines and sulfonamides, as these are antimicrobials the most frequently used in food animals[18].

Limitations

As with any modelling study, our analysis comes with limitations. Firstly, predictive maps, as well as the imputation of missing resistance prevalence for modelling priority antimicrobials introduces uncertainty. The number of imputations was highest for cefotaxime – its resistance prevalence was missing in half (51%) of the surveys. However, the uncertainty of the missing values was captured by the high standard deviation (24%) of the multiple imputed values for cefotaxime. We attempted to quantify the uncertainty by combining Monte Carlo simulations of the imputed input datasets, and the variance of the Bayesian posterior predictive distribution for each simulation (Methods). Secondly, due to the limited number of surveys reporting resistance prevalence for individual antimicrobial-bacteria combinations, mapped predictions of AMR were restricted to 7 drugs and 2 bacteria. These drugs were amongst the most frequently used antimicrobial classes and the most frequently cited classes across 1,088 point prevalence surveys. Additionally, predictions of nontyphoidal *Salmonella* were not

disaggregated for individual serovars. However, this is in consistency with Murrey et al. 2022 who mapped AMR in humans[12]. The limited number of surveys available also made it challenging to conduct spatio-temporal modelling, and we pooled together surveys from all years for AMR mapping. As the number of point prevalence surveys[34] published each year is growing, future efforts to map AMR may incorporate more antimicrobial-bacteria combinations and investigate both spatial and temporal effects on AMR maps, while insuring statistical robustness in the extrapolations. Thirdly, the maps of priority antimicrobials were built under the assumption that resistance prevalence will increase at the same rate as in the past 20 years, implying that the drivers behind AMR - including policies regulating antimicrobial use (AMU) - will remain unchanged in the near future. However, due to temporal changes in these policies – e.g. a 30% decline in antimicrobial use in Thailand from 2017 to 2019[18], the drivers behind AMR patterns may change in the future. Our predictions were intended to show how resistance may evolve without interventions on AMU policies, for the purpose of guiding such interventions. Fourthly, due to the lack of a systematic inventory of country-specific regulations on antimicrobial use, we did not explicitly include these regulations as covariates. For example, ciprofloxacin is banned in poultry in the US[11], but not in China[35]. However, the regulations were implicitly considered in the modelling process, with the inclusion of AMU in 2013 and 2020 for each antimicrobial class as covariates. In 2013, Maron and colleagues reviewed restrictions on antimicrobial use in food animals[36]. However, to the best of our knowledge, an up-to-date global database on antimicrobial use policies has not been conducted. Fifthly, we dichotomized resistance prevalence using 50% threshold to define priority antimicrobials for AMR surveillance. We conducted sensitivity analysis by mapping priority antimicrobials using other thresholds (10% and 25%) as well. However, the choice of thresholds is dependent on multiple factors and in its nature subjective.

The maps of AMR produced in this study helps outline priorities for action. Firstly, in AMR hotspots – including China, Iran, India, Brazil and Chile, measures should be taken to further ongoing efforts to reduce antimicrobial use in food animals. Secondly, our analyses showed how AMR for 7 medically important antimicrobials may evolve in the future, without policy interventions. This could provide a baseline scenario where revisions of AMR policy could be based. Thirdly, the 3rd Global High-Level Ministerial Conference on AMR has set out a global target to reduce antimicrobials used in agrifood systems by 30-50% by 2030. Our maps could serve as a reference for more targeted measures aimed at specific antimicrobial classes in their corresponding hotspot regions of resistance. Possible measures include stricter regulations and on-farm monitoring on antimicrobial use, targeted awareness campaigns among veterinarians and farmers, as well as investments on improving farm hygiene to reduce dependence on antimicrobials.

Methods

This analysis is structured in five steps (Figure 5a-e): a) collection and extraction of epidemiological information from point prevalence surveys (PPS); b) mapping distribution of resistance prevalence using three machine learning models; c) ensembling predictions using Gaussian process stacked generalization; d) mapping priority antimicrobials for surveillance; and e) estimating prediction uncertainty of maps generated in steps c and d. The literature review was conducted using Zotero (version 5.0.96.2) and Microsoft Excel (version 16.53), and all data analysis was conducted using R (version 4.1.1).

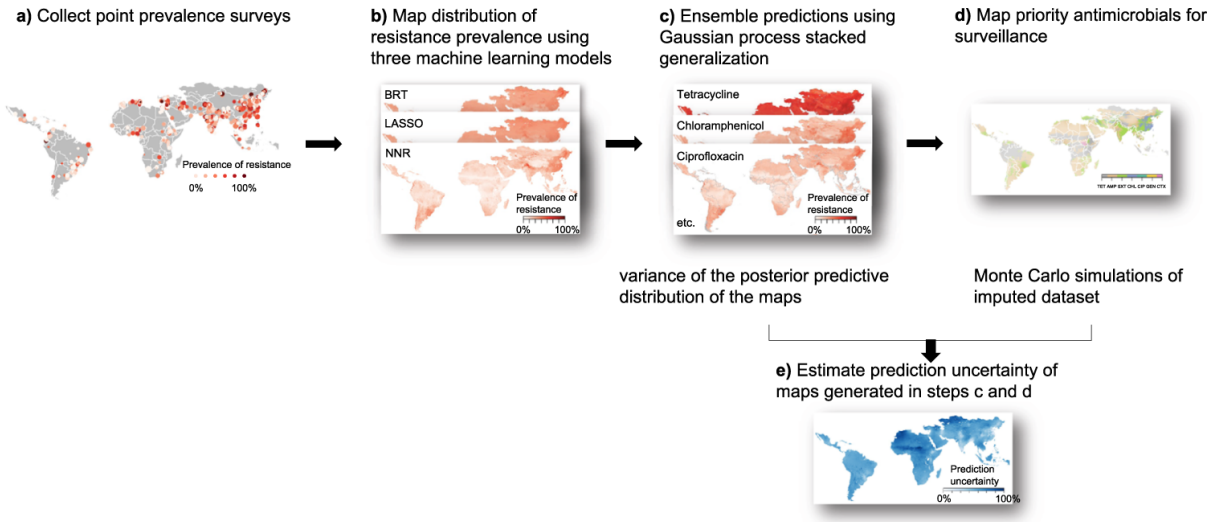


Figure 5. Modelling framework. a Collect point prevalence surveys. b Map distribution of resistance prevalence using three machine learning models: boosted regression trees (BRT), LASSO logistic regression (LASSO), feed-forward neural network (NNR). c Ensemble predictions using Gaussian process stacked generalization. d Map priority antimicrobials for surveillance. e Estimate prediction uncertainty of maps generated in steps c and d.

Data collection and imputation

We extracted 1,088 point prevalence surveys on AMR of *E. coli* and nontyphoidal *Salmonella* in healthy food animals across low- and middle-income countries (LMICs) across two decades between 2000 and 2019 (Supplementary Table 3). These surveys were collected through three rounds of literature review of four databases (PubMed, Scopus, ISI Web of Science, and China National Knowledge Infrastructure). The process of data extraction is explained in detail in the Supplementary Information section "Literature review and data extraction". These surveys were conducted on major food animal species including cattle (n=409), pigs (n=303), poultry (n=570), sheep (n=89), horse (n=2), and goat (n=2). The animal samples used to determine resistance prevalence were taken from their meat (34% of total resistance prevalence), swabs from living animals on farm or in wet markets (32%), food products such as milk and eggs (16%), swabs from slaughtered animals (9%), and fecal samples on farm (7%).

In each survey, we extracted information on resistance prevalence, method used for antibiotic susceptibility testing (AST), guideline document used for performing AST, breakpoints used for assessing AST results, sample origin, number of animal samples and bacterial isolates, as well as the geographic location and time of the survey. The majority (91%) of the studies used the performance standards for antimicrobial susceptibility testing developed by the Clinical and Laboratory Standards Institute (CLSI) or the European Committee for Antimicrobial Susceptibility Testing (EUCAST). Each performance standards set breakpoints to classify resistance phenotypes, which are updated annually. These variations in breakpoints were adjusted using methods developed by Van Boeckel and colleagues[4], to maximize comparability between surveys.

For this analysis, we focused on 7 antimicrobial drugs: tetracycline (TET), ampicillin (AMP), sulfamethoxazole-trimethoprim (SXT), chloramphenicol (CHL), ciprofloxacin (CIP), gentamicin (GEN), and cefotaxime (CTX). The resistance prevalences of these drugs were

the most frequently reported for their individual antimicrobial classes in the collected surveys (Supplementary Table 6), and therefore ensured robustness in comparisons made between surveys. These antimicrobial classes were all classified as critically important in veterinary medicine[20], and were also classified as either critically important or highly important for human medicine[14]. For each of the 7 drugs, we used all PPS that reported its resistance prevalence individually to map its distribution, with methods explained in the next sector. However, the subsequent prediction of priority antimicrobials requires complete resistance profiles with resistance prevalence of all 7 drugs. Therefore, 806 PPS that reported resistance prevalence of at least 4 out of these 7 drugs were included for this part of analysis. For the unreported antimicrobials, we imputed their resistance prevalence based on correlations between antimicrobials, using multivariate imputation by chained equations[37] (MICE; Supplementary Methods; Figure 5a). The MICE algorithm imputed plausible values for 21% out of 9,877 antimicrobial resistance prevalence estimates in these surveys, while also providing a mechanism for integrating the uncertainty of imputation in the following analysis, as explained in section "Uncertainty".

Trends of AMR for each antimicrobial class

We used logistic regression models to estimate temporal trends of resistance prevalence between 2000 and 2019 for each antimicrobial. For TET and AMP, we removed one outlier (DOI of PPS: 10.1264/jsme2.2000.173) out of 758 PPS reporting resistance for TET and 797 PPS reporting resistance for AMP, to ensure that the assumption of linearity between the logit of dependent variable and the independent variable was met based on results of Box-Tidwell test.

We mapped the distribution of the prevalence of resistance for each antimicrobial at 10x10 kilometer resolution using Gaussian process stacked generalization, an ensemble approach of multiple models. This approach has been shown to increase prediction accuracy for disease mapping compared with other methods such as Gaussian process regression[38]. This mapping procedure comprised two steps (Figure 5b, 5c). In the first step, we trained three 'child models' to predict resistance prevalence based on a set of environmental and anthropogenic covariates, such as total antimicrobial use in 2013 and 2020, animal population density, and temperature (Supplementary Table 3; Supplementary Method). For each antimicrobial class, we also included the quantities (kg) used in 2020[18] disaggregated at 10x10 kilometer resolution as a covariate. This was calculated by disaggregating the total antimicrobial use per country proportionally to the distribution of animals' biomass in 2020[18]. Animals' biomass was calculated as the population correction units of food animals in 2020, using methods described in Van Boeckel et al. 2015[39]. In the second step, the child model predictions were stacked using Gaussian process regression, fitted using the integrated nested Laplace approximations (INLA)[40] (Supplementary Methods). This second step allowed to simultaneously capture the influence of environmental and anthropogenic covariates, as well as the residual spatial correlation.

For each antimicrobial, we defined resistance hotspots as regions with resistance prevalence higher than the 95% percentile of all pixels on the map. We combined the drug-level resistance maps using summary metrics for the overall AMR level – N10, N25, or N50: the number of antimicrobials (out of 7) with resistance prevalence higher than 10%, 25% or 50% in each pixel. For the summary AMR level across antimicrobial classes, resistance hotspots were defined as regions with $N50 \geq 3$.

Mapping priority antimicrobials for AMR surveillance

Priority antimicrobials for AMR surveillance were defined as antimicrobials that have the highest probability of their resistance prevalence exceeding a critical level (defined as 10%, 25% or 50%) in the near future. Here, we assumed that prevalence of resistance will continue to increase in the future, based on temporal trends of AMR between 2000 and 2019. We developed an approach to predict priority antimicrobials at each 10x10 kilometer pixel, based on local risk factors as well as patterns of co-resistance in PPS. In the following, we explain the modelling process using 50% as the critical resistance level, while similar procedures were followed for the other cutoff values of resistance prevalence (10% or 25%). We illustrate the model formulation, with the following example of a pixel with N50=4 (Figure 6).

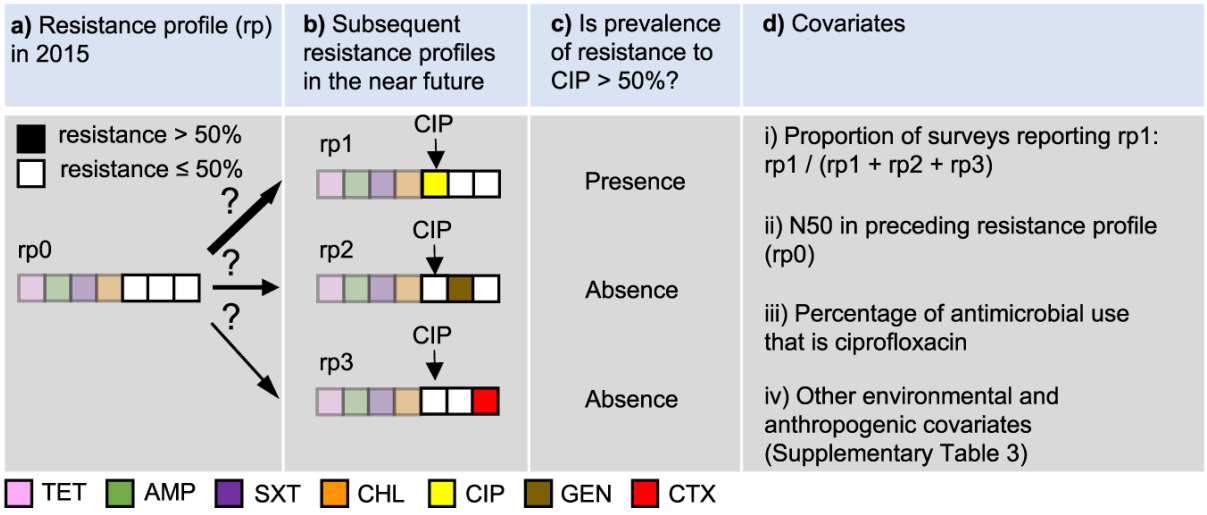


Figure 6. LASSO logistic regression model to predict the probability that resistance prevalence of ciprofloxacin (CIP) will exceed 50% in the future, in pixels with predicted resistance profile (rp) of [1,1,1,1,0,0,0] (rp0) in 2015. a Resistance profile in 2015. b Subsequent resistance profiles in the near future. c Determine whether prevalence of resistance to CIP is above 50%. d Covariates, including (i) the proportion of point prevalence surveys reporting the resistance profile in which resistance prevalence of ciprofloxacin exceeds 50% (rp1), out of all alternative antimicrobials (gentamicin in rp2, and cefotaxime in rp3), (ii) the number of antimicrobials with resistance above 50% (N50) in the predicted resistance profile in 2015 (rp0), (iii) the percentage of antimicrobial use (kg) of ciprofloxacin, and (iv) a set of environmental and anthropogenic covariates. TET: tetracycline; AMP: ampicillin; SXT: sulfamethoxazole-trimethoprim; CHL: chloramphenicol; CIP: ciprofloxacin; GEN: gentamicin; CTX: cefotaxime.

Firstly, we binarized the resistance profile in 2015 for a given pixel (e.g. TET 70%, AMP 75%, SXT 60%, CHL 55%, CIP 40%, GEN 30%, and CTX 30%) by reclassifying the antimicrobials with resistance higher than 50% as 1, and the opposite as 0, such that the resistance profile for the 7 drugs considered in this analysis was: [1,1,1,1,0,0,0] (Figure 6a). Secondly, for each of the three antimicrobials classified as 0 (e.g. CIP, GEN, CTX), we predicted whether their resistance prevalence will exceed 50% as a binary response variable (Figure 6c), using covariates extracted from the collected surveys (Figure 6d). The model

considers future scenarios where only one additional antimicrobial will exceed 50% resistance (Figure 6b). The model was constructed using least absolute shrinkage and selection operator (LASSO) applied to logistic regression. Using CIP as an example, its resistance prevalence exceeds 50% in resistance profile rp1, while it is absent in resistance profiles rp2 and rp3 (Figure 6b, 6c). The covariates used to predict its presence and absence included two components. The first component considers patterns of co-resistance between antimicrobials, implying that probabilities of occurrence vary between resistance profiles. This variation is captured by using the proportion of surveys recording rp1 out of all surveys recording rp1, rp2 or rp3 as a covariate (Figure 6d.i). Patterns of co-resistance also implies that the development of resistance of CIP is dependent on resistance of other antimicrobials. This dependence is captured by using the number of antimicrobials with resistance higher than 50% in the resistance profile in 2015 as a covariate (Figure 6d.ii). The second component of covariates includes risk factors for predicting the development of resistance. This includes the percentage of CIP use (kg) out of all three antimicrobials at the location of the survey (Figure 6d.iii), as well as a set of environmental and anthropogenic covariates associated with the locations of the surveys, such as total antimicrobial use in 2013 and 2020, temperature, and animal density (Figure 6d.iv; Supplementary Table 3).

The above example was based on the current resistance profile rp0 (Figure 6a). For CIP, there were in total 64 permutations of current resistance profiles – all six antimicrobials apart from CIP could have resistance of 0 or 1. A complete model for CIP was trained by including all permutations in the procedure described in Figure 6. This model was then applied to each pixel on the map where resistance to CIP has not yet exceeded 50%, to generate the probability that it will exceed 50% in the future. Similarly, the probabilities for the other antimicrobials were generated. Finally, at each pixel, we mapped the antimicrobial with the highest probability of its resistance prevalence exceeding 50% in the future.

The accuracy of the models for each antimicrobial was quantified by calculating the area under the receiver operating characteristic curve (AUC) using four-fold spatial cross validation[4]. The predictive capacity of the model came from two components of the covariates. The first component was based on co-resistance between drugs (Figure 6d.i and Figure 6d.ii). The second component was environmental and anthropogenic covariates associated with resistance to individual drugs (Figure 6d.iii and Figure 6d.iv). We quantified the relative contribution of these two covariate components to the model prediction accuracy, by calculating the drop in AUC following the withdrawal of each covariates compared with a full model including all covariates.

Furthermore, based on predictions of the priority antimicrobial for AMR surveillance at each 10x10 km pixel (Figure 6), we estimated the time it takes for resistance prevalence of this antimicrobial to reach 50% in the future (Supplementary Figure 16). Concretely, we extracted the current resistance prevalence estimated at each pixel, and calculated the time difference from the current resistance prevalence (Supplementary Figure 16, time point a) until it reaches 50% (Supplementary Figure 16, time point b), using the corresponding regression models fitted in section “Trends of AMR for each antimicrobial class”.

Uncertainty

The uncertainty of the mapped predictions of resistance prevalence (Figure 5c) was calculated as the variance of the posterior predictive distribution for each map. The uncertainty of the mapped priority antimicrobials was calculated in two steps. Firstly, we

generated 15 Monte Carlo simulations of imputed datasets of resistance prevalence, to incorporate the uncertainty introduced by imputation in the following analyses. Secondly, using the imputed datasets, we generated 15 maps of priority antimicrobials. We quantified its uncertainty by calculating – at each pixel - the proportion of maps that generated different predictions of antimicrobials as compared with the final map:

$$\text{Uncertainty} = \frac{\text{N}_{\text{maps with different predictions}}}{m} \quad (1)$$

References

1. Van Boeckel, T. P. *et al*. Reducing antimicrobial use in food animals. *Science* **357**, 1350–1352 (2017).
2. Laxminarayan, R., Van Boeckel, T. & Teillant, A. The economic costs of withdrawing antimicrobial growth promoters from the livestock sector. (2015).
3. Food and Agriculture Organization of the United Nations. FAOSTAT Statistical Database. (2020).
4. Van Boeckel, T. P. *et al*. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, 1266-+ (2019).
5. Laxminarayan, R. *et al*. Antibiotic resistance—the need for global solutions. *Lancet Infect. Dis.* **13**, 1057–1098 (2013).
6. Aarestrup, F. M. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140085 (2015).
7. Karp, B. E. *et al*. National antimicrobial resistance monitoring system: two decades of advancing public health through integrated surveillance of antimicrobial resistance. *Foodborne Pathog. Dis.* **14**, 545–557 (2017).
8. Dutil, L. *et al*. Ceftiofur resistance in Salmonella enterica serovar Heidelberg from chicken meat and humans, Canada. *Emerg. Infect. Dis.* **16**, 48 (2010).
9. European Food Safety Authority & European Centre for Disease Prevention and Control. The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019. *EFSA J.* **19**, e06490 (2021).
10. Zhao, C. *et al*. Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China. *Nat. Food* **2**, 596–602 (2021).
11. FDA, U. Animal drugs, feeds, and related products; enrofloxacin for poultry; withdrawal of approval of new animal drug application. *Fed Regist* **70**, 44048–44049 (2005).
12. Murray, C. J. *et al*. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).
13. Iwamoto, M. *et al*. Ceftriaxone-resistant nontyphoidal Salmonella from humans, retail meats, and food animals in the United States, 1996–2013. *Foodborne Pathog. Dis.* **14**, 74–83 (2017).
14. World Health Organization. WHO Medically Important Antimicrobial List.
15. Wang, X. *et al*. Antibiotic resistance in Salmonella Typhimurium isolates recovered from the food chain through national antimicrobial resistance monitoring system between 1996 and 2016. *Front. Microbiol.* **10**, 985 (2019).

16. Lehtinen, S., Blanquart, F., Lipsitch, M., Fraser, C., & with the Maela Pneumococcal Collaboration. On the evolutionary ecology of multidrug resistance in bacteria. *PLoS Pathog*. **15**, e1007763 (2019).
17. World Health Organization. Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2020. (2020).
18. Mulchandani, R., Wang, Y., Gilbert, M. & Van Boeckel, T. P. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. *PLOS Glob. Public Health* **3**, e0001305 (2023).
19. Puvača, N. *et al*. Influence of different tetracycline antimicrobial therapy of mycoplasma (Mycoplasma synoviae) in laying hens compared to tea tree essential oil on table egg quality and antibiotic residues. *Foods* **9**, 612 (2020).
20. OIE, W. OIE list of antimicrobial agents of veterinary importance. *J OIE Int Commit* **33**, 1–9 (2015).
21. Gelaude, P., Schlepers, M., Verlinden, M., Laanen, M. & Dewulf, J. Biocheck. UGent: a quantitative tool to measure biosecurity at broiler farms and the relationship with technical performances and antimicrobial use. *Poult. Sci*. **93**, 2740–2751 (2014).
22. Chopra, I. & Roberts, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol. Mol. Biol. Rev*. **65**, 232–260 (2001).
23. Heth, D. & Bird, H. Growth response of chicks to antibiotics from 1950 to 1961. *Poult. Sci*. **41**, 755–760 (1962).
24. Eliopoulos, G. M., Eliopoulos, G. M. & Roberts, M. C. Tetracycline therapy: update. *Clin. Infect. Dis*. **36**, 462–467 (2003).
25. Du, Z. *et al*. The prevalence of amphenicol resistance in Escherichia coli isolated from pigs in mainland China from 2000 to 2018: A systematic review and meta-analysis. *PloS One* **15**, e0228388 (2020).
26. Wen, R., Li, C., Zhao, M., Wang, H. & Tang, Y. Withdrawal of antibiotic growth promoters in China and its impact on the foodborne pathogen Campylobacter coli of swine origin. *Front. Microbiol*. 3505 (2022).
27. Harada, K., Asai, T., Kojima, A., Ishihara, K. & Takahashi, T. Role of coresistance in the development of resistance to chloramphenicol in Escherichia coli isolated from sick cattle and pigs. *Am. J. Vet. Res*. **67**, 230–235 (2006).
28. Economou, V. & Gousia, P. Agriculture and food animals as a source of antimicrobial-resistant bacteria. *Infect. Drug Resist*. **8**, 49 (2015).
29. Diana, A., Manzanilla, E. G., Calderón Díaz, J. A., Leonard, F. C. & Boyle, L. A. Do weaner pigs need in-feed antibiotics to ensure good health and welfare? *PloS One* **12**, e0185622 (2017).
30. Love, W. J., Zawack, K. A., Booth, J. G., Gröhn, Y. T. & Lanzas, C. Markov networks of collateral resistance: national antimicrobial resistance monitoring system surveillance results from Escherichia coli isolates, 2004-2012. *PLoS Comput. Biol*. **12**, e1005160 (2016).
31. Nichol, D. *et al*. Antibiotic collateral sensitivity is contingent on the repeatability of evolution. *Nat. Commun*. **10**, 1–10 (2019).
32. Munck, C., Gumpert, H. K., Wallin, A. I. N., Wang, H. H. & Sommer, M. O. Prediction of resistance development against drug combinations by collateral responses to component drugs. *Sci. Transl. Med*. **6**, 262ra156-262ra156 (2014).
33. Bischoff, K. M., White, D. G., Hume, M. E., Poole, T. L. & Nisbet, D. J. The chloramphenicol resistance gene cmlA is disseminated on transferable plasmids that confer multiple-drug resistance in swine Escherichia coli. *FEMS Microbiol. Lett*. **243**, 285–291 (2005).

34. Criscuolo, N. G., Pires, J., Zhao, C. & Van Boeckel, T. P. resistancebank.org, an open-access repository for surveys of antimicrobial resistance in animals. *Sci. Data* **8**, 189 (2021).
35. Ministry of Agriculture and Rural Affairs of China. National food safety standard - Maximum residue limits for veterinary drugs in foods. (2019).
36. Maron, D. F., Smith, T. J. & Nachman, K. E. Restrictions on antimicrobial use in food animal production: an international regulatory and economic survey. *Glob. Health* **9**, 1–11 (2013).
37. Van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**, 1–67 (2011).
38. Bhatt, S. *et al*. Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* **14**, 20170520 (2017).
39. Van Boeckel, T. P. *et al*. Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci.* **112**, 5649–5654 (2015).
40. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71**, 319–392 (2009).

Chapter 5

Forecasting intensive care occupancy in Switzerland during the COVID-19 pandemic

Chapter 5.1 icumonitoring.ch: a platform for short-term forecasting of intensive care units occupancy during the COVID-19 epidemic in Switzerland.

Authors:
Cheng Zhao[1,†], Burcu Tepekule[2,†], Nicola G. Criscuolo[1,†], Pedro D. Wendel Garcia[3], Matthias P. Hilty[3], RISC-19-ICU Investigators for Switzerland, Thierry Fumeaux[4,5], and Thomas P. Van Boeckel[1,6,*].

Affiliations:
[1] Health Geography and Policy Group, ETH Zürich, Switzerland.
[2] University Hospital Zürich, Switzerland.
[3] Institute of Intensive Care Medicine, University Hospital of Zurich
[4] Service de médecine et des soins intensifs, Hôpital de Nyon.
[5] President, Swiss Society of Intensive Care Medicine
[6] Center for Disease Dynamics Economics and Policy, New Delhi, India.

*Correspondence to: thomas.vanboeckel@env.ethz.ch

[†]these authors contributed equally.

Published in:
Swiss Medical Weekly. 150:w20277. DOI: 10.4414/smw.2020.20277.
This chapter contains minor revisions compared to the published article.

Contribution remarks:
T.P.V.B conceived the work, cleaned and analysed the data on ICU occupancy, and wrote the first version of the manuscript. C.Z. and B.T. produced epidemic projections. C.Z. coordinated the biweekly update of content displayed on *icumonitoring.ch*. N.G.C developed the Shiny web application. P.D.W.G, and M.H. facilitated access to and pre-treated data from the RISC-19-ICU registry. T.F, facilitated access to data from SSIM, and informed the model development. All authors critically revised and edited the manuscript.

Abstract

In Switzerland, the COVID-19 epidemic is progressively slowing down owing to 'social distancing' measures introduced by the Federal Council on March 16th. However, the gradual ease of these measures may initiate a second epidemic wave, for which length and intensity are difficult to anticipate. In this context, hospitals must prepare for a potential increase of patient admissions with acute respiratory distress syndrome in intensive care units. Here, we introduce icumonitoring.ch, a platform providing hospital-level projections for intensive care unit occupancy. We combined current data on the number of beds and ventilators with canton-level projections of COVID-19 cases from two S-E-I-R models. We disaggregated epidemic projection in each hospital in Switzerland for the number of COVID-19 cases, hospitalizations, hospitalizations in ICU, and ventilators in use. The platform is updated every 3-4 days and can incorporate projections from other modelling teams to inform decision makers with a range of epidemic scenarios for future hospital occupancy.

Introduction

The COVID-19 epidemic currently affecting Switzerland seems to progressively slow down. The inflection point of the epidemic curve of deaths[1] was passed on 06.04.2020, and thus far, the number of COVID-19 cases with acute respiratory distress syndrome (ARDS) needing an intensive care unit (ICU) admission or mechanical ventilation has not exceeded the ad hoc increase in ICU bed capacity and ventilator availability. However, the gradual ease of the lockdown measures that have been in place since 17.03.2020[2] may initiate a second epidemic wave. As in other countries, there is currently considerable uncertainty about the true prevalence[3,4] of COVID-19 in the Swiss population, and thus also about whether the country might achieve herd-immunity and if so, when. The absence of specific therapies against the SARS-CoV-2 virus responsible for COVID-19 and the difficulty to anticipate the effect of lifting lockdown measures on movement intensity[5], and future infection rates[6] further contribute to this uncertainty. In this context, hospitals must prepare for a potential secondary increase in ICU admissions of unknown magnitude and duration.

Since the onset of the COVID-19 outbreak, disease modelers have tried to anticipate the trajectory of the COVID-19 epidemic in Switzerland. Some have focused on long-term policies at the national scale[7,8,9] while others focused on capturing and forecasting the dynamics of COVID-19 in individual Cantons[6]. For forecasting COVID-19 cases, disease dynamics modelling is a particularly useful approach. Its model structures capture the intricate transmission dynamics of the disease between groups of individuals, categorized by their infection status such as being susceptible, infected, or recovered. This enables real-time adaptation of model parameters to reflect changes in the traits of the virus, such as its contagiousness and severity, as well as changes in public health measures impacting transmission dynamics. It also allows for scenario analysis, enabling researchers to explore the potential impact of different interventions, such as social distancing measures or vaccination

campaigns. However, thus far, little attention has been placed on generating forecasts at the spatial level where most intervention can take place: hospitals, and specifically ICUs. During the same period, hospitals, health care facilities, government agencies, and the Swiss Armed Forces have reacted to the COVID-19 outbreak on a day-to-day basis. For example, their actions consisted in expanding bed capacities[10], building stocks of personal protection equipment[11], or dispatching medical troops in support of hospitals[12]. Their actions have been guided by multiple surveillance efforts conducted in parallel by federal and cantonal authorities and professional health care societies. Amongst these is the IES system managed by the Coordinated Sanitary Service (CCS). This system should provide bi-daily reports of the occupancy of emergency departments and hospital beds across the country. However, its use as a monitoring platform during the COVID-19 outbreak has proven difficult, due to slow, incomplete, and uneven reporting across hospitals. On 14.03.2020, just 15 hospitals did effectively report their bed occupancy, as compared with 156 hospitals on the 03.05.2020, after several measures were taken by the CCS in collaboration with the Swiss Society of Intensive Care Medicine (SSICM). In the coming weeks, resources available to attend to COVID-19 patients with ARDS will need to be optimally deployed (within and between ICU) to minimize the risk of overflow.

At least three challenges must be addressed to generate reliable hospital-level projections that could help ICU managers to anticipate the need for additional resources. First, the IES system needs to be accurately and regularly documented. Second, projections from national and cantonal epidemic models must be downscaled at the hospital-level by making reasonable assumptions that reflect the situation experienced by clinicians. Third, the output of epidemic models must be summarized and rapidly transferred to clinicians in a format that is straightforward to inform management decisions in hospitals.

Addressing the first challenge belongs to individual hospitals, which should ensure and control accuracy of IES collected data: models can inform decisions, but only reliable data can help modelers helping hospitals. For the second challenge, tools from the field of spatial analysis can be used to disaggregate information generated in polygons (Cantons) to individual hospitals (latitude/longitude coordinates) while accounting for the respective 'catchment areas' of these hospitals. These approaches have been used extensively in spatial epidemiology to study the treatment-seeking behavior of HIV patients on antiretroviral therapy[13], the allocation of bed nets against malaria[14], and access to emergency maternal care[15]. For the third challenge, the recent development of web-based applications enables a rapid display and update of model outputs using a simple web-browser. In particular, Shiny apps give users the possibility to query regions/hospitals interactively, and thus represent an improvement from static maps in 'one-off' publications.

Here, we introduce icumonitoring.ch a platform of ICU bed occupancy forecasting for individual hospitals in Switzerland based on projections from two Canton-level epidemic models. Our framework is flexible, and projections from other modeling groups can be integrated using a 'forecast template'. Due to confidentiality reasons, this article only presents

aggregated results at the Canton-level. Access to projections for individual hospitals in icumonitoring.ch are available upon request to the communicating author; the password will be automatically issued for ICU healthcare workers.

Methods

Data

Time series of hospitalization in intensive care units (ICU) in Switzerland as reported in the IES system were provided by the Swiss Armed Forces. This dataset consists in reports of ICU bed occupancy for COVID19 and non-COVID19 patients in adults, and children. This database, which is updated twice a day, reports the number of patients in ICU beds, and the number of these beds equipped with ventilators. The number of Extracorporeal Membrane Oxygenation (ECMO) is comprised in the number of beds with ventilators. Importantly, this database contains an estimate of the number of SSMIC-certified ICU beds, as well as the number of ad hoc beds since the start of the COVDI19 outbreak. In some hospitals, the number of COVID-19 patients entered in the IES system was higher than the number total number of patients. Similarly, in some hospitals the number of ventilated COVID-19 patients was higher than the total numbers of COVID-19 patients. As these situations are de facto impossible, we assumed that the person who entered the data points considered COVID-19 not to be part of the pool of all patients, which it should. These inputs were corrected such as if the number of COVID-19 patients was higher than the total number of patients then the total number of patients was calculated as the sum the number of COVID-19 patients reported and the number of patients reported. If the number of ventilated COVID-19 patients was higher than the number of COVID-19 patients then the number or COVID-19 patients was calculated as the sum of the number of COVID-19 patients reported, and the number of ventilated COVID-19 patients reported. For hospitals that did not report on 03.05.2020, we used numbers (beds, patients, ventilators) provided for the last date of reporting available.

Near real-time data on ICU length of stay, mortality has been collected using the RIsk Stratification in COVID-19 patients in the Intensive Care Unit (RISC-19-ICU) registry, a collaborative effort with the participation of a majority of the Swiss ICUs to provide a basis for decision support during the ongoing public health crisis. The registry was deemed exempt from the need for additional ethics approval and patient informed consent by the ethics committee of the University of Zurich (KEK 2020-00322, ClinicalTrials.gov Identifier: NCT04357275). The data were collected using a secure REDCap infrastructure provided by the Swiss Society of Intensive Care Medicine. As of 03.05.2020, 68.5% of critically ill patients with COVID-19 admitted to an ICU in Switzerland have entered the registry had already been dismissed from the ICU or have died.

This analysis accounts for ‘patient disappearance’ from the IES system in Ticino (131 patients on 28.03.2020) at Clinica Luganese Moncucco and Ospedale Regionale di Lugano (42 patients on 01.04.2020), as well as in Vaud (148 patients on 25.03.2020) at Hôpital Riviera-Chablais,

Centre hospitalier de Rennaz. These institutions seem to have either stopped reporting or transferred all of their patients on the dates mentioned above. In these hospitals, patients were removed from the IES system and did not reappear in other hospitals in the Canton. Media sources referred to only a very small number of patients from Ticino that were hosted in the German-speaking part of the country. We have attempted to gather information from 'Clinica Luganese Moncucco', but they declined to answer our questions regarding the number of patients in their ICU. We accounted for 'patient disappearance' by creating a 'Hospital X' in the two Cantons concerned. This adjustment is meant as a way to account for all active acute COVID-19 cases.

Epidemic forecasting

CZ Model (adapted from Althaus et al.). We used an S-E-I-R model developed by Althaus and colleagues[8] (accessed April 24th, 2020) to model epidemics of COVID-19 in Swiss Cantons. The model assumed constant uncontrolled transmission until the soft lockdown measures on 17.03.2020[16]. The basic reproduction number and the reduction in transmission after the soft lockdown were estimated using a maximum likelihood framework. Following the announcement from the Federal Council to ease lockdown measures from the 27th of April[17], the model assumes that contact would resume to 50% of their pre-lock level form that date. The model was fitted to the reported numbers of deaths in 18 Cantons, where enough data on times series of death was available for parameter inference. The inference was done using Maximum Likelihood with the Nelder & Mead algorithm implemented in the function optim in the R statistical Software.The number of deaths until 28.04.2020 per Canton was retrieved at 21:00, 02.05.2020 from OpenZH[1]. The number of deaths on 03.05.2020 was incomplete and subject to future modifications and was therefore not used for the epidemic modeling. For each Canton, the model predictions included five categories: infected cases (IF), hospitalized cases (HS), ICU cases (IC), immune cases (IM), and death cases (DE). Infected cases were calculated as the sum of exposed cases, infectious cases, hospitalized cases, and ICU cases. In the remaining eight cantons (AI, GL, JU, NW, OW, SH, UR, ZG), models could not be fitted due to the lower number of deaths. There we calculated the model predictions in proportion to the number of COVID-19 cases reported in each of these Canton, in the last eight days. The model predictions in each Canton were adjusted, such that they summed up to the model predictions at the national level. The final outcome of the epidemic model prediction was an estimation of the number of IF, HS, IC, IM, and DE for each day in each Canton, as well as the 95% confidence intervals of the predictions.

The equations used in the CZ model are listed below, with the descriptions and values of the notations in Table 1, and the descriptions of compartments listed in Table 2:

$$\beta = \frac{R_0}{N} \times \kappa \times \gamma$$

$$\frac{dS(t)}{dt} = -S\beta I,$$

$$\frac{dE(t)}{dt} = +S\beta I - \sigma E\,,$$

$$\frac{dI(t)}{dt} = +\sigma E - \gamma I\,,$$

$$\frac{dP(t)}{dt} = +\varepsilon_1\gamma I - \omega_1 P\,,$$

$$\frac{dH_1(t)}{dt} = +\omega_1 P - \omega_2 H_1\,,$$

$$\frac{dH_2(t)}{dt} = +(1-\varepsilon_2)\omega_2 H_1 - \omega_3 H_2\,,$$

$$\frac{dU(t)}{dt} = \varepsilon_2\omega_2 H_1 - (1-\varepsilon_4)\omega_4 U - \varepsilon_4\omega_5 U\,,$$

$$\frac{dR(t)}{dt} = +(1-\varepsilon_1)\gamma I + (1-\varepsilon_3)\omega_3 H_2 + (1-\varepsilon_4)\omega_4 U\,,$$

$$\frac{dD(t)}{dt} = +\varepsilon_3\omega_3 H_2 + \varepsilon_4\omega_5 U\,,$$

$$\frac{dC(t)}{dt} = +\gamma I\,.$$

Table 1. **Parameters descriptions, and values in CZ model**.

*Obtained for patients (n=382) included in the RISC-19-ICU registry supported by Swiss Society of Intensive Care Medicine (https://www.sgi-ssmi.ch).

Parameters	Description	Value
R_0	Basic reproduction number	fitted
κ	Percentage of R_0 after lockdown	fitted
σ, γ	Serial interval	1/2.6 days [18]
ω_1	Duration from onset of symptoms to hospitalization	1/5 days [19]
ω_2	Initial hospitalization	1/6 days [19]
ω_3	Additional days of hospitalization until recovery/death	1/10 days [19]
ω_4	Additional days in ICU until recovery	1/11.2 days*
ω_5	Additional days in ICU until death	1/10.5 days*
ε_1	Rate of H admission of infected	3.5%[4,5] [20, 21]
ε_2	Hospitalized cases requiring critical care in ICU	30% [19]
ε_3	Death outside of ICU	35%[6] [8]
ε_4	Death rate from ICU	23%*

Table 2. **States variables in model CZ, and model BT (epidemic compartments)**.

Variable (model CZ)	Variable (BT model)	Description
S	S	Susceptible
E	E	Exposed
I	I	Infected
H	$H_1 + H_2$	Hospitalized patients
-	H_1	Initial hospitalization until transfer to ICU
-	H_2	Addit. hospitalization until recovery/death
U	U	ICU patients

D	D	deaths
R	R	Recovered
C	C	cumulative number of infected
-	P	infected but not yet hospitalized

BT model. The second model used was an extended SEIR model which additionally includes the hospitalized and ICU patients. In the BT model, people who are infected by the virus are assumed to develop symptoms in 2 to 3 days but may be infectious in the community for another 2 to 3 days, adding up to a generation time of between 4 to 6 days (Ganyani et al. assumes 5.2 days). People who become sick are hospitalized at a proportion ranging between 1% to 15%, where they are isolated, and thus are considered as non-infectious to the community. People who are admitted to the hospital are assumed to stay at the ward for 6-18 days, and an additional 2-11 days if they need to stay in the ICU, which is the case for 30-80% of hospitalized patients. The death rate in the ICU is assumed to be between 30-80%. The effect of the lockdown is assumed to vary between 60-80%. We assume a combined probability of positive diagnosis and detection for the infected patients to make use of the reported case data, and this probability varies between 0.05 to 0.35 (5%-35% of the total infected). Four different time series (number of daily deaths, number of daily reported cases, number of people at the hospital ward, and number of people at the ICU, obtained from OpenZH[1]) are used simultaneously to do the model fitting for each Canton separately. We used Hamiltonian Markov Chain Monte Carlo (MCMC) for model inference, as implemented in RStan[22]. Hundred chains with random initial parameter vectors are used with 10,000 jumps in total. The first 5000 jumps were considered for the 'burn in period', we used Geweke statistics on each chain to assess convergence, and chains were thinned with a sampling rate of 100 samples. As for the CZ model, the change in contact patterns resulting from the lockdown measures (16.03.2020), and their subsequent release (27.04.2020) are accounted for through a parameter reducing infection rates r_{lock}. This parameter varied across Cantons from 0.57 (SZ) to 0.76 (BL) during the lockdown period (17.03.2020 – 27.04.2020).

To calculate the time series output of the fitting, we include the mean values of the posterior distributions of 50% of the chains with the best likelihood output among the ones which have converged. Chains that haven't converged are omitted and not used in the analysis. Due to the high dimensionality of the parameter space, we used a mixed sampling approach: first, we determine the likelihood of each chain among the chains that have converged. Second, we sample from the posterior distributions of these chains proportional to the mean likelihood value they have converged to. Confidence intervals of the results are calculated for each time point over the population outputs. By allowing model parameters to vary in between these ranges, we have more freedom to fit our model to the number of daily deaths, number of people in the hospital ward, and the number of ICU patients simultaneously, in a Canton specific manner. This is especially important due to the differences in the treatment and ICU transfer policies of different Cantons and hospitals. As an example, restricting the ICU length of stay to a value that is necessarily smaller or larger than the length of hospital ward stay might not

apply for all Cantons in question. For both models, the number of hospitalizations on 03.05.2020 was estimated by back-casting from the model's output.

The equations used in BT model are listed as following, with the descriptions and values of the notations in Table 3, and the descriptions of compartments listed in Table 2:

$$\beta = \frac{R_0}{N} \times (1 - r_{lock}) \times \gamma$$

$$\frac{dS(t)}{dt} = -S\beta I$$

$$\frac{dE(t)}{dt} = S\beta I - \tau E$$

$$\frac{dI(t)}{dt} = \tau E - \gamma I$$

$$\frac{dH(t)}{dt} = \epsilon_H \gamma I - \gamma_H H$$

$$\frac{dU(t)}{dt} = \gamma_H \epsilon_{H2I} H - \gamma_U U$$

$$\frac{dR(t)}{dt} = \gamma_H (1 - \epsilon_{H2I}) H + \gamma_U (1 - \epsilon_D) U$$

$$\frac{dD(t)}{dt} = \gamma_U \epsilon_D U$$

$$\frac{dC(t)}{dt} = \gamma_U \epsilon_D U$$

Table 3. Parameters descriptions, and values in BT model. All parameters fitted (except N).

Parameters	**Description**	**Value**
R_0	Basic reproduction number	2-3
$r_{lock}(t)$	Time-dependent reduction in infectiousness	60-80%
τ	Incubation period	1/2-3 days
γ	Duration of infection of I	1/2-3 days
γ_H	Duration of hospital ward stay	1/6-18 days
γ_U	Duration of ICU stay	1/2-11 days
ϵ_H	Rate of direct H admission of infected	1-15%
ϵ_{H2I}	Transfer rate from H to ICU	30-80%
ϵ_D	Death rate from ICU	30-80%
r_d	Diagnosis rate	5-35%
N	Population size by canton	fixed

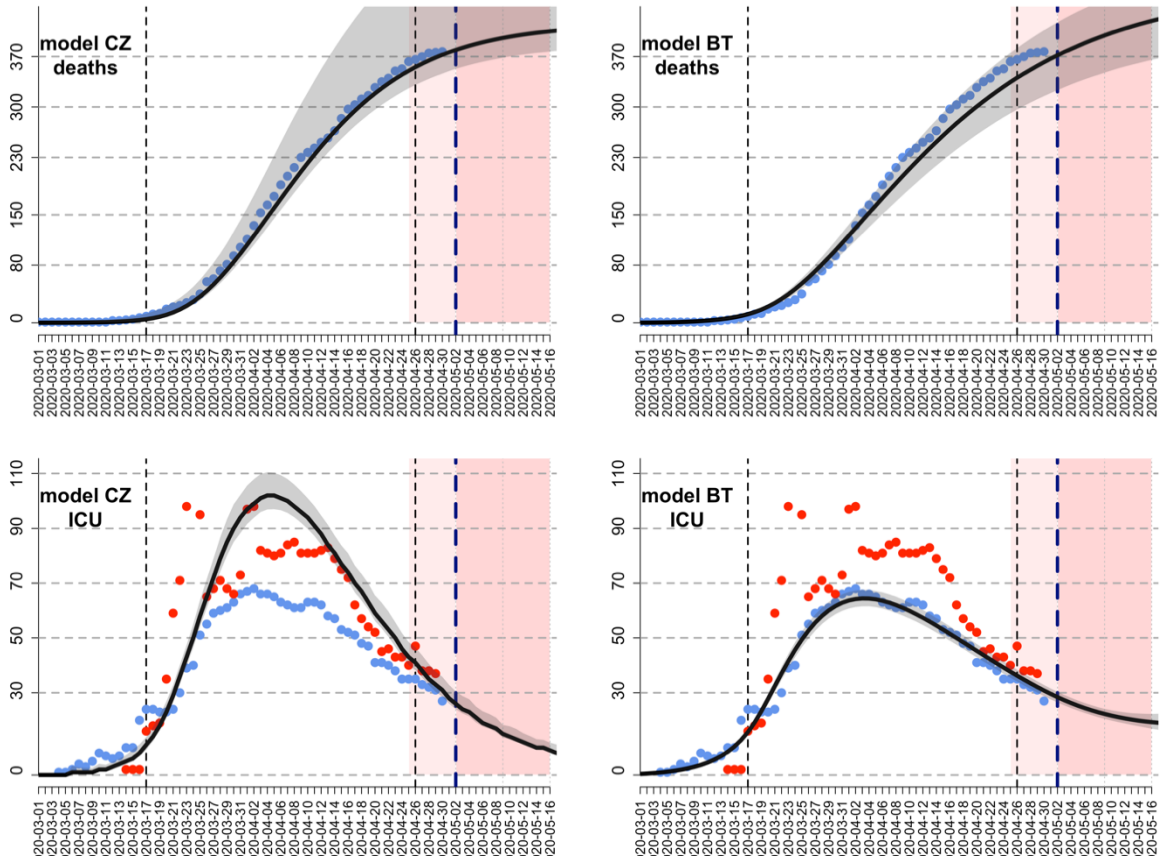


Fig. 1. Canton-level fit of epidemic model CZ and BT in Vaud on 03.05.2020. Black shades describe 95% confidence interval, red dots are ICU occupancy as reported in the IES systems, and blue dots are ICU occupancy as reported by OpenZH. The BT model is fitted to deaths and ICU occupancy times series from OpenZH, while the CZ models is fitted only to deaths.

Uncertainty and forecasting capacities

For the CZ model, we simulated 10,000 R_0 and kappa values, from a multivariate normal distribution with mean values equal to the fitted values of R_0 and kappa, and a covariance matrix estimated from the maximum likelihood estimation. The confidence interval for predictions in each canton was estimated, with the 2.5% and 97.5% quantiles of the 10,000 predictions. The mean values of the cantonal predictions in the CZ model were adjusted, by multiplying with an adjustment ratio, such that they sum up to the model predictions at the national level. The upper and lower bounds of confidence intervals for each canton were adjusted, by multiplying with the same adjustment ratio that were applied to the mean values of model predictions in each canton. For the BT model, similarly, confidence intervals were calculated using the converged MCMC parameter estimates (posteriors) and extracting the 2.5% and 97.5% quantiles of the predictions generated with these parameters' values.

Confidence intervals at the hospital-level were estimate from the cantonal-level 95% CI in a two-step procedure. First, we calculate the percentage of deviation between the upper/lower bound of the 95% CI and the mean number of cases, hospitalization, ICU hospitalizations, and

deaths. Second these percentages of deviation were applied to the estimates of the same outputs downscaled at the hospital level based on population density and travel times (next section). For example, a hospital where 10 ICU beds were projected to be occupied and that is located in a Canton where the total number of ICU bed was 100 [95% CI 90-120] would have a 95% CI between 9 and 12 beds.

The ability of our epidemic models to temporal projections 4 days ahead – the update frequency of icumonitoring.ch- was evaluated at the cantonal level by comparing projections and observations for the number of deaths and COVID-19 patients in ICU on the 03.04.2020 using a model calibrated on the 29.05.2020. The metrics used to evaluate the accuracy of projections the number of deaths, and the number of ICU cases were the spearman correlation coefficient between projections, and observations, as well as the average percentage deviation between projections and observation across Cantons.

Geographic downscaling of epidemic projections in hospital

In each hospital, we estimated the bed occupancy on 02.05.2020 as the sum of: (i) the number of beds in use by non-COVID-19 patients, which was assumed to be stable since 29.04.2020, (ii) the number of ICU beds in use by COVID-19 patients that were admitted before 03.05.2020 that remained in the ICU by 06.05.2020, and (iii) the number of new COVID-19 patients who required a bed in an ICU between 03.04.2020 and 06.05.2020. For (ii), we assumed an exponential survival function with a discharge rate equal to: α x LOS_{deaths} + (1-α) x $LOS_{recovered}$, where α is the mortality rate of COVID-19 patients in ICU (0.23), LOS_{deaths} is the length of stay for deceased COVID-19 patients (10.5 days), and $LOS_{recovered}$ is the length of stay of patients that recover (11.2 days). For (iii), we spatially disaggregated the total number of ICU cases projected by Canton (see epidemic forecasting) minus (ii). Each future COVID-19 patient requiring care in an ICU (iii) was assigned a latitude and longitude in each Canton via stratified random sampling inside the corresponding Canton. The weighting factor for the stratification was population density. Each case was assigned to the 'nearest' hospital, measured in travel time (minutes). We used a friction surface[23] to estimate travel times to hospitals. Based on each location, each case was assigned to a nearby hospital using a gravity model. The probability of having attended a hospital from any given pixel was given by: $\log10(ICU_{beds})/(travel\ time_{pixel->H})$, for hospital 'H'. Each patient was assigned to the hospital with the higher probability of attendance. This process was repeated 10 times through Monte Carlo simulations. The hospital that was selected with the highest frequency across the 10 simulations was designated as the hospital likely attended by a patient in the future. The number of patients on ventilators on 03.05.2020 was estimated by applying the current rate of ventilation of COVID-19 patients in an ICU (76%) to the future number of COVID-19 patients admitted in an ICU.

Online platform

All epidemic model outputs at the cantonal- and hospital-level are uploaded to an online platform *icumonitoring.ch* twice per week Sunday and Thursday evening. The platform is a

'Shiny'[24] interactive application developed in the open-access R programming language[25] inside the RStudio Environment. In addition, we used JavaScript actions and CSS code to adjust aesthetic features of the platform into a dashboard. In its current version, icumonitoring.ch relies on the following dependencies: aws.s3 (0.3.21), grDevices (3.6.3), htmlwidgets (1.5.1), leaflet (2.0.3), RColorBrewer (1.1-2), readr (1.3.1), rgdal (1.4-4), shiny (1.4.0.2), shinydashboard (0.7.1), shinyjs (1.1), and tidyverse (1.3.0). *icumonitoring.ch* is hosted on a password-protected shinyapps.io server. The databases and model outputs displayed on the platform are stored on an encrypted storage service of ETH Zürich (polybox).

Results

As of 03.04.2020, the number of patients requiring an ICU bed in Switzerland was 713. This estimate is below the effective *ad hoc* ICU bed capacity in Switzerland (1,275). The number of patients requiring ICU beds (for COVID-19 and non-COVID-19 causes) is decreasing and unlikely to exceed the effective capacity in the next week. By 06.05.2020, we project that the need for ICU beds could range between 739 [CI 95% 669 - 871] (model CZ) and 761 [CI 95% 541 - 1164] (model BT). As of 03.05.2020, 367 patients were ventilated in ICU, out of 1,064 ventilators available. Assuming a ventilation rate of 76% for ICU COVID-19 patients, as reported on 03.05.2020, the number of ventilators required by 06.05.2020 could be 398 [CI 95% 365 - 466] (CZ model) or 416 [CI 95% 295 - 658] (BT model).

For the number of COVID-19 ICU cases, on 03.05.2020, the correlation between projections (4 days ahead) and observations by canton was 0.62, and 0.86 for the CZ and BT model respectively. The percentage deviation between projection and observation for the number COVID-19 ICU cases was --16.7% for the CZ and -14.4% for the BT model. For the number of deaths, on 03.05.2020, the correlation between projections (4 days ahead) and observations by canton was 0.99 for the CZ and the BT. The percentage deviation between projection and observation for the number of deaths was +2.14% for the CZ model and -0.4% for the BT model.

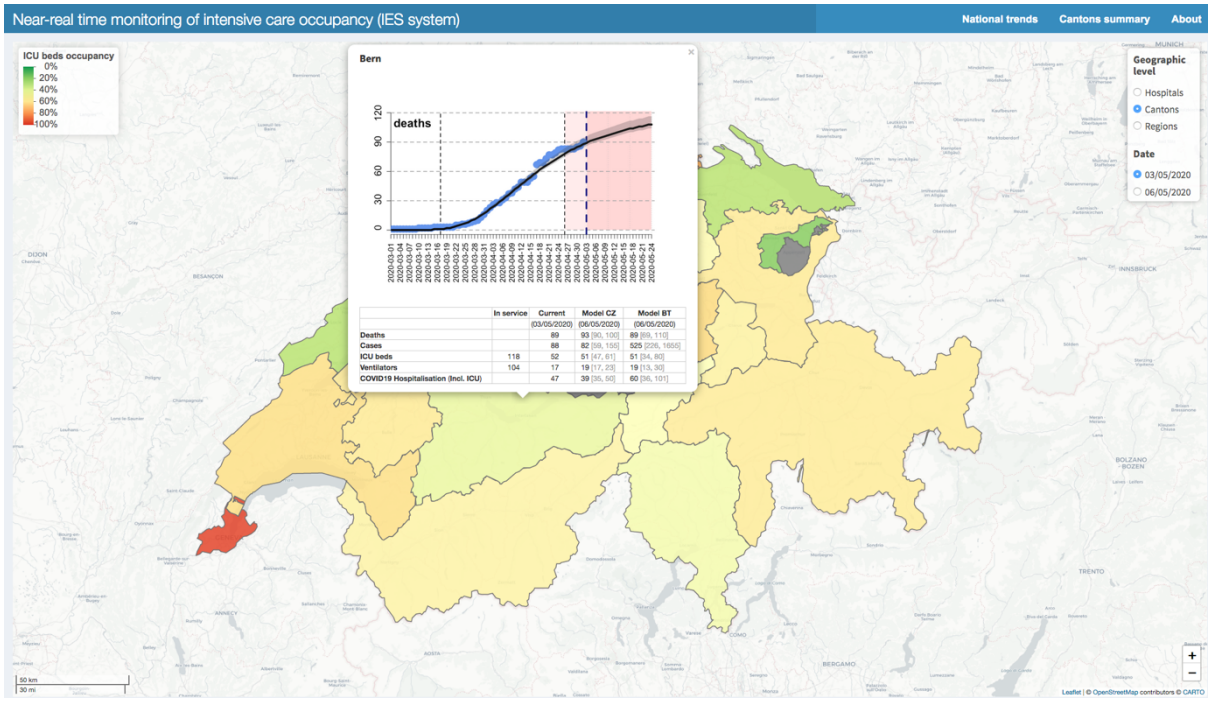


Fig 2. Canton-level ICU occupancy, colors in other Cantons indicate ICU bed occupancy compared to the number of beds in service. Pop-up windows indicate the situation in the Canton of Bern as reported in *icumonitoring.ch* for 03.05.2020 and projected for 06.05.2020.

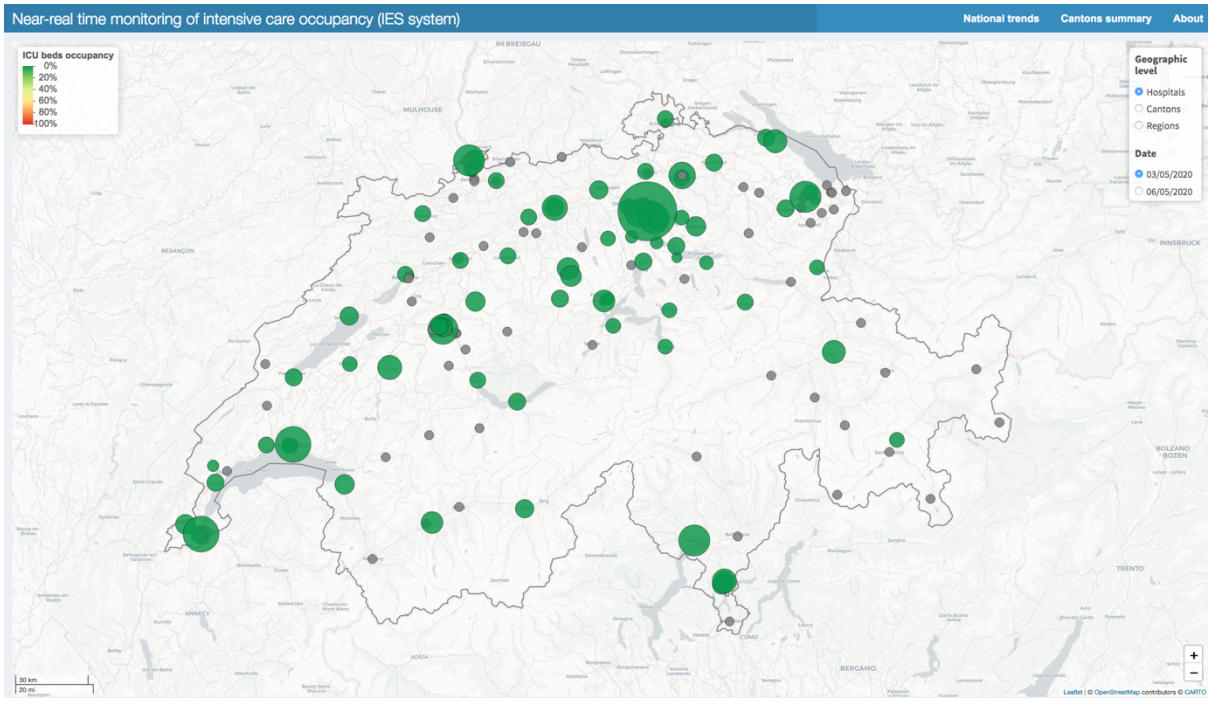


Fig 3. Hospitals with ICU (yellow), **and without ICU** (grey). bed occupancy in individual hospitals set to 0 artificially but available upon request to the corresponding author (bottom).

Icumonitoring.ch disaggregates these findings by Canton, and by hospital. It is an interactive web application that displays Intensive Care Units (ICU) occupancy. Each geographic element (i. e. hospitals or cantons) can be interrogated via a pop-up window. The pop-up window shows the number of deaths attributed to COVID-19; the number of estimated COVID-19 infections

(Cases); the number of ventilators available and used; the number of COVID-19 patients in ICU, and beds available; the total number of COVID-19 patients hospitalized (including ICU). Projections for these quantities are available 3-4 days in advance (02.05.2020) and re-calibrated every 3-4 days based on epidemic data. At the cantonal level, the pop-up window also provides a visual of the epidemic model fit to the time series of deaths in each Canton. The data presented in this article are aggregated at the Canton-level, but access to hospital-level information is available to healthcare workers upon request to the communicating author. *icumonitoring.ch* also provides a comparative summary of current and future bed occupancy, ventilators occupancy, and hospitalizations in each canton.

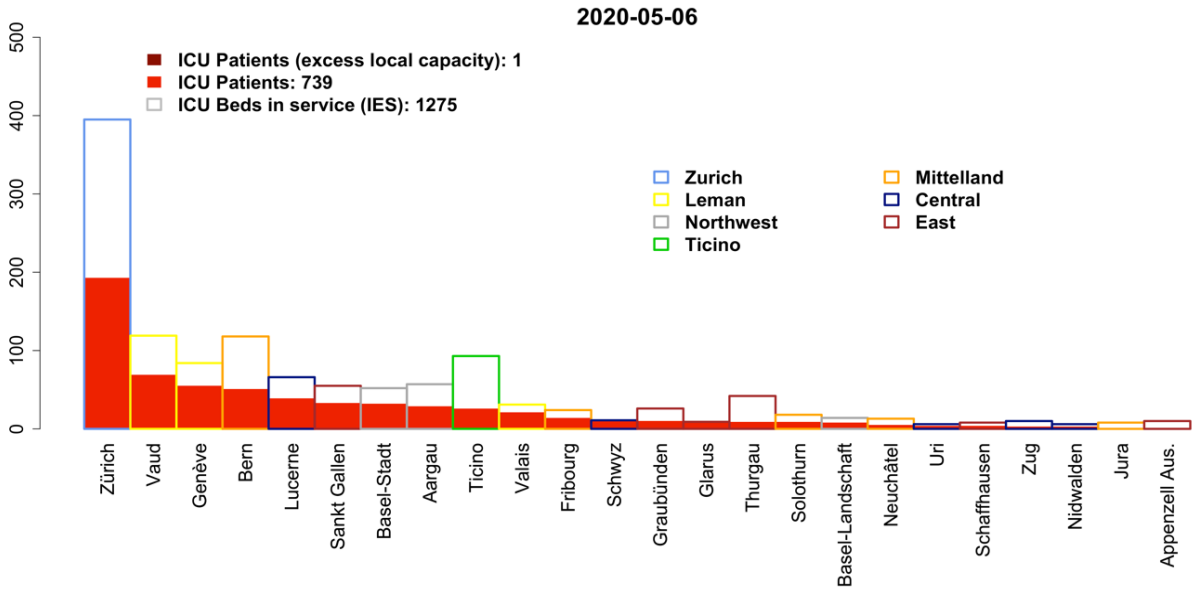


Fig 4. Projection for ICU bed occupancy on 06.05.2020 ranked by Regions and Cantons (model CZ).

Table 4. On 06.05.2020, the number of hospitalizations, and ICU hospitalizations for COVID-19 was projected to grow in 5 cantons and decline in 21 Cantons compared to 03.05.2020. On 03.05.2020, the number of ICU hospitalizations was projected to grow in 7 cantons and decline in 19 Cantons compared to 03.05.2020. On 03.05.2020 the number of ventilators needed was projected to grow in 6 cantons and decline in 20 Cantons compared to 03.05.2020.

Canton	COVID-19 Hospitalizations 03.05.2020	COVID-19 Hospitalizations 06.05.2020	ICU Beds In service 03.05.2020	ICU Beds In use 03.05.2020	ICU Beds projected 06.05.2020	Ventilators In service 03.05.2020	Ventilators In use 03.05.2020	Ventilators projected 06.05.2020
AG	39	36	57	28	29	56	15	16
AI	0	0	0	0	0	0	0	0
AR	2	0	10	2	0	6	2	1
BE	47	39	118	52	51	104	17	19
BL	5	5	14	7	8	6	1	2
BS	13	9	52	34	32	28	12	10
FR	12	10	24	14	14	24	8	7
GE	62	28	84	80	55	82	33	21
GL	31	35	9	5	10	4	2	5
GR	9	3	26	15	10	23	9	4
JU	7	3	8	5	1	8	3	1
LU	14	3	66	49	39	55	19	13
NE	9	8	13	4	5	13	3	3
NW	1	1	6	3	3	6	0	0
OW	1	1	0	0	0	0	0	0
SG	11	10	55	32	33	49	9	9
SH	3	1	8	6	4	5	4	2
SO	14	11	18	11	9	14	7	4
SZ	14	15	11	7	10	6	1	3
TG	15	10	42	13	9	42	7	4
TI	57	30	93	44	26	90	28	9
UR	3	4	6	3	4	2	0	1
VD	74	53	119	74	69	112	34	31
VS	33	30	31	17	21	31	13	15
ZG	3	3	10	3	3	8	1	1
ZH	35	18	395	205	193	290	139	131

Discussion

icumonitoring.ch is a tool to support decision-makers anticipate ICU occupancy during the COVID-19 outbreak. Ultimately, its goal is to prevent hospital overflow[26,27], such as in Italy or Spain by projecting when capacities may need to be expanded, or the transfer of patients has to be considered such as in France[28]. Conversely, this tool can also be used to progressively reduce costly expansion of nominal hospital capacities[29] and request for additional medical staff[30].

The accuracy of the projections available in icumoniroting.ch relies on epidemic models, but also on the completeness of the data reported in the IES system. In normal times, the IES system is seldom maintained, without apparent consequences. However, these are not normal times. While acknowledging that healthcare workers face unprecedented demand in the clinic, we urge them to maintain the IES system up-to-date as to help us helping them. This may require additional personnel/training. The 'epidemiology community' would welcome a display of leadership from the competent authorities in empowering hospital managers with the

appropriate resources to maintain the IES. What cannot be measured cannot be managed, let alone projected in the future.

As with any modeling study, the projection presented in icumonitoring.ch comes with a series of limitations. The development of icumonitoring.ch started on 10.04.2020 and remains a 'work-in-progress'. In particular, the following adjustments will be considered in the future. First, the number of non-ICU beds in hospitals reported in the IES system remains inaccurate (personal communication) and is therefore not displayed on the platform at this stage. Second, the two epidemic models used for projection do not yet account for the age structure of the population. Given the strong dependency of the case- fatality rate of COVID-19 on age[31], this would need to be included in a future iteration of our models. Third, the epidemic model used here show deviations between projection (4 days ahead) and observations for a given day. For the projections of the number of deaths – on which both models are fitted – these deviations are minimal (~2%). However, for the number of the number of COVID-19 case requiring ICU beds – 16.7% for the CZ model, and -14.4% for the BT model, respectively. The BT thus model slightly outperform the CZ model. From a hospital management perspective, the underestimation of the number of ICU cases by both models could make our projections seen as a conservative minimal threshold for hospital managers to consider before downscaling the ad-hoc capacities currently in place. The reasons for the underestimation of the capacity may be associated with current estimate of the length of stay in ICU. Here we attempted to include the most up-to-date estimates of LOS from the RISC-19-ICU registry to which >30 Swiss ICU units are contributing. However, it should be acknowledged that 31.5% of patients with acute COVID-19 are still currently in ICU and may have longer than average LOS due to the severity of their infections. This may artificially decrease the LOS used in this analysis which are calculated from patients that have already been discharged or are deceased. Similarly, another potential source of bias for LOS in the ICU in is the limitation of therapy due to a patient's wishes. These decisions do not necessarily correlate to disease severity but could be motivated by a patient's previous health status, advance directive, or substitute directives. Fourth, the geographic downscaling is based on population density. We thus implicitly make the hypothesis that a future patient is equally likely to have contracted the disease in cities or in rural settings. This may lead to an underestimation of the clustering of cases in cities where contact rates may be higher than in the countryside. Fourth, uncertainty in epidemic model lies in the ability to infer transmission parameters but also the intrinsic formulation of a model in different epidemic compartments[32]. Here we attempt to address this concern by using two epidemic models with different inference methods, and compartmental structures as to capture the uncertainty inherent to the model structure. The objective of our platform is also to include projection from other modelling group with a forecast template such as to allow further comparison between models that may have radically different structure, such as agent-based model[9,33]. The authors would also welcome suggestions from the intensive medicine community for relevant parameters to be added to icumonitoring.ch that can help guide hospital management decisions.

References

1. openZH/covid_19 [Internet]. Specialist Unit for Open Government Data Canton of Zurich; 2020 [cited 2020 Apr 18]. Available from: https://github.com/openZH/covid_19
2. Federal Council to gradually ease measures against the new coronavirus [Internet]. [cited 2020 Apr 18]. Available from: https://www.admin.ch/gov/en/start/documentation/media-releases/media-releases-federal-council.msg-id-78818.html
3. Streeck H, Hartmann G, Exner M, Schmid M. Vorläufiges Ergebnis und Schlussfolgerungen der COVID-19 Case-Cluster-Study (Gemeinde Gangelt). University of Bonn.
4. COVID-19 Antibody Seroprevalence in Santa Clara County, California | medRxiv [Internet]. [cited 2020 Apr 18]. Available from: https://www.medrxiv.org/content/10.1101/2020.04.14.20062463v1
5. Entwicklung des Mobilitätsverhaltens während der COVID-19-Krise [Internet]. [cited 2020 Apr 18]. Available from: https://statistik.zh.ch/internet/justiz_inneres/statistik/de/aktuell/mitteilungen/2020/covid_mobilitaetsverhalten.html
6. Monitoring COVID-19 spread in Switzerland [Internet]. [cited 2020 Apr 18]. Available from: https://bsse.ethz.ch/cevo/research/sars-cov-2/real-time-monitoring-in-switzerland.html
7. Lemaitre J, Perez-Saez J, Azman A, Rinaldo A, Fellay J. Switzerland COVID-19 Scenario Report. École Polytechnique Fédérale de Lausanne; 2020.
8. Real-time modeling and projections of the COVID-19 epidemic in Switzerland [Internet]. [cited 2020 Apr 18]. Available from: https://ispmbern.github.io/covid-19/swiss-epidemic-model/
9. COVID-19 Epidemic in Switzerland: Growth Prediction and Containment Strategy Using Artificial Intelligence and Big Data | medRxiv [Internet]. [cited 2020 Apr 18]. Available from: https://www.medrxiv.org/content/10.1101/2020.03.30.20047472v2
10. Suisse romande: Les soins intensifs ne manquent pas de lits - News Suisse: Suisse romande - 24heures.ch [Internet]. [cited 2020 Apr 18]. Available from: https://www.24heures.ch/suisse/suisse-romande/soins-intensifs-manquent-lits/story/20067350
11. Pelda K. L'armée mène une opération secrète pour importer des masques par millions. TDG [Internet]. 2020 Apr 15 [cited 2020 Apr 18]; Available from: https://www.tdg.ch/suisse/armee-mene-operation-secrete-importer-masques-millions/story/12405340
12. Coronavirus: la Suisse mobilise l'armée. Le Temps [Internet]. 2020 Mar 16 [cited 2020 Apr 18]; Available from: https://www.letemps.ch/suisse/coronavirus-suisse-mobilise-larmee
13. Houben RM, Van Boeckel TP, Mwinuka V, Mzumara P, Branson K, Linard C, et al. Monitoring the impact of decentralised chronic care services on patient travel time in

rural Africa-methods and results in Northern Malawi. International journal of health geographics. 2012;11(1):1.
14. Macharia PM, Odera PA, Snow RW, Noor AM. Spatial models for the rational allocation of routinely distributed bed nets to public health facilities in Western Kenya. Malaria Journal. 2017 Sep 12;16(1):367.
15. Myers BA, Fisher RP, Nelson N, Belton S. Defining Remoteness from Health Care: Integrated Research on Accessing Emergency Maternal Care in Indonesia. AIMS Public Health. 2015 Jul 1;2(3):257–73.
16. Coronavirus: Federal Council declares 'extraordinary situation' and introduces more stringent measures [Internet]. [cited 2020 Apr 18]. Available from: https://www.admin.ch/gov/en/start/documentation/media-releases.msg-id-78454.html
17. FOPH FO of PH. New coronavirus: Measures, ordinance and explanations [Internet]. [cited 2020 Apr 30]. Available from: https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/massnahmen-des-bundes.html
18. Estimating the generation interval for COVID-19 based on symptom onset data | medRxiv [Internet]. [cited 2020 Apr 29]. Available from: https://www.medrxiv.org/content/10.1101/2020.03.05.20031815v1
19. Report 9 - Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand [Internet]. Imperial College London. [cited 2020 Apr 29]. Available from: http://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/
20. Estimates of the severity of coronavirus disease 2019: a model-based analysis - The Lancet Infectious Diseases [Internet]. [cited 2020 Apr 29]. Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext
21. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China | Nature Medicine [Internet]. [cited 2020 Apr 29]. Available from: https://www.nature.com/articles/s41591-020-0822-7
22. R Interface to Stan [Internet]. [cited 2020 May 3]. Available from: https://mc-stan.org/rstan/
23. Weiss DJ, Nelson A, Gibson HS, Temperley W, Peedell S, Lieber A, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature. 2018;553(7688):333.
24. Shiny [Internet]. [cited 2020 Apr 18]. Available from: https://shiny.rstudio.com/
25. R: The R Project for Statistical Computing [Internet]. [cited 2020 Apr 20]. Available from: https://www.r-project.org/
26. Wynne A. Horrifying images show coronavirus patients in Madrid hospital [Internet]. Mail Online. 2020 [cited 2020 Apr 19]. Available from: https://www.dailymail.co.uk/news/article-8142013/Horrifying-images-coronavirus-patients-lying-floor-packed-Madrid-hospital.html
27. Hume T. Coronavirus Has Northern Italy's Hospitals on the Brink of Collapse [Internet]. Vice. 2020 [cited 2020 Apr 19]. Available from:

https://www.vice.com/en_us/article/k7ex4a/coronavirus-has-northern-italys-hospitals-on-the-brink-of-collapse
28. Keohane D. France's TGV speeds Covid-19 patients to spare hospital beds [Internet]. 2020 [cited 2020 Apr 19]. Available from: https://www.ft.com/content/619bd7b0-7424-11ea-95fe-fcd274e920ca
29. Coût de mesures d'urgence estimé à 35 milliards. TDG [Internet]. 2020 Jul 4 [cited 2020 Apr 19]; Available from: https://www.tdg.ch/economie/cout-mesures-urgence-estime-35-milliards/story/13977524
30. Service d'appui : licenciement partiel du personnel sanitaire assorti de contraintes de disponibilité [Internet]. Swiss Armed Forces; 2020. Available from: https://www.vtg.admin.ch/fr/actualite/coronavirus.detail.news.html/vtg-internet/verwaltung/2020/20-04/20-04-17-arbeitgeberbrief_nr2.html
31. Adjusted age-specific case fatality ratio during the COVID-19 epidemic in Hubei, China, January and February 2020 | medRxiv [Internet]. [cited 2020 Mar 21]. Available from: https://www.medrxiv.org/content/10.1101/2020.03.04.20031104v1
32. Foss AM, Vickerman PT, Chalabi Z, Mayaud P, Alary M, Watts CH. Dynamic Modeling of Herpes Simplex Virus Type-2 (HSV-2) Transmission: Issues in Structural Uncertainty. Bull Math Biol. 2009 Apr 1;71(3):720–49.
33. Müller SA, Balmer M, Neumann A, Nagel K. Mobility traces and spreading of COVID-19. 2020 Mar 20 [cited 2020 Apr 30]; Available from: https://depositonce.tu-berlin.de/handle/11303/10945

Chapter 5.2 A hybrid Neural Network-SEIR model for forecasting intensive care occupancy in Switzerland during COVID-19 epidemics

Authors:
Riccardo Delli Compagni[1], Cheng Zhao[1], Stefania Russo[2], Thomas P. Van Boeckel[1,3*]

Affiliations:
[1]Health Geography and Policy Group, ETH Zürich, Switzerland
[2]Ecovision Lab, Photogrammetry and Remote Sensing, ETH Zürich, Switzerland
[3]Center for Diseases Dynamics Economics and Policy, Washington DC, United States

*Correspondence to: thomas.vanboeckel@env.ethz.ch

Published in:
PLoS ONE 17(3): e0263789. DOI: 10.1371/journal.pone.0263789

Contribution remarks:
R.D.C., C.Z. and T.P.V.B conceptualized the research and curated the data; R.D.C., C.Z. and S.R. developed the methodology; R.D.C. analysed the data; R.D.C., C.Z., S.R. and T.P.V.B wrote the paper.

Abstract

Anticipating intensive care unit (ICU) occupancy is critical in supporting decision makers to impose (or relax) measures that mitigate COVID-19 transmission. Mechanistic approaches such as Susceptible-Infected-Recovered (SIR) models have traditionally been used to achieve this objective. However, formulating such models is challenged by the necessity to formulate equations for plausible causal mechanisms between the intensity of COVID-19 transmission and external epidemic drivers such as temperature, and the stringency of non-pharmaceutical interventions. Here, we combined a neural network model (NN) with a Susceptible-Exposed-Infected-Recovered model (SEIR) in a hybrid model and attempted to increase the prediction accuracy of existing models used to forecast ICU occupancy. Between 1st of October, 2020 - 1st of July, 2021, the hybrid model improved performances of the SEIR model at different geographical levels. At a national level, the hybrid model improved, prediction accuracy (i.e., mean absolute error) by 74%. At the cantonal and hospital levels, the reduction on the forecast's mean absolute error were 46% and 50%, respectively. Our findings illustrate those predictions from hybrid model can be used to anticipate occupancy in ICU, and support the decision-making for lifesaving actions such as the transfer of patients and dispatching of medical personnel and ventilators.

Introduction

On March 11th, 2020, the World Health Organization (WHO) declared the COVID-19 pandemic an international health emergency[1]. Since then, COVID-19 has caused infections in millions of people[2], with a substantial proportion of infections (e.g. 9-11%[3]) requiring hospitalization in intensive care units (ICU). In multiple countries, demand of ICU beds exceeded bed availability[4,5,6], leading to excess mortality of COVID-19 patients as well as backlogs of patients for other pathologies that require hospitalization in ICU[7,8,9]. Monitoring and anticipating ICU occupancy has become critical to support decision-makers to impose (or relax) non-pharmaceutical interventions that can help mitigate the transmission of COVID-19, and thereby reduce its impact on healthcare systems.

Mathematical models have been used extensively to anticipate the evolution of epidemic indicators, including the occupancy of ICU[10,11,12,13]. In particular, two families of mathematical models have been predominantly used: 1) mechanistic models (MMs), including Susceptible-Infected-Recovered (SIR) models[14] and their extensions into agent-based models[15], as well as, 2) statistical approaches[16], including machine learning (ML) models[17]. Each family of model present advantages and disadvantages: MMs typically consist of differential equation systems that reflect biological mechanisms that govern the dynamic of infections. The parameters of these equations usually have a biological meaning (i.e., an infectious period) and therefore can be used for predictions outside of their calibration space (i.e., scenario analysis). However, for MM, accounting for the causal mechanisms between ICU occupancy and environmental covariates (e.g. changes in environmental conditions)[18,19] comes at the cost of additional parameters to be estimated in a differential equations system. In contrast, ML models seek to establish statistical associations between response variables and potential covariates without making assumptions about potential biological mechanisms[20]; however, because ML models are based on statistical associations and not causation, their validity is bound to their calibration space, and every prediction outside such a space can lead to inconsistent results[21,22]. Therefore, the combination of MMs and ML models in "hybrid models" has been explored in a variety of fields[21,22,23] (e.g., earth systems, climate science, biology, hydrology, etc.), and have showed promising results for improving

prediction accuracy[20] from MM models. One of the most common configurations for a hybrid model is known as "residuals modelling", and is of particular interest when the MM formulation may be too limited to capture complex associations between a response variable and its covariates[23]. Concretely, this configuration consists of using a MM to capture the overall temporal trend of a temporally autocorrelated process while letting the ML model compensate for any residual error that is potentially associated with external drivers of the process of interest. Neural Networks (NN) are one of the most commonly used ML models in this framework due to their ability to implicitly capture nonlinearities and interactions[24]. MMs have been coupled with NNs in different fields, thereby improving performances of the corresponding MMs: for example, Chu et al.[25] improved prediction accuracy of a MM to simulates performances of a centrifugal compressor; Lee et al.[26] also improved prediction accuracy of a MM to simulate the operations of a waste-water treatment plant; Thompson and Kramer[22] used a NN to model a fed-batch penicillin fermentation reaction. Thus far, multiple works have shown how hybrid models can be used to predict the evolution of the COVID-19 epidemic[27,28,29]; however, to the best of our knowledge, these works did not implement the configuration of residual modelling using a NNs as a ML model.

In this study, we developed a hybrid model based on the residual modelling configuration aimed at increasing the prediction accuracy of an SEIR model (Susceptible-Exposed-Infected-Recovered) across spatial scales for producing short-term (3- and 7-days ahead) predictions of ICU occupancy. The accuracy of the modelling framework was tested in Switzerland, where data on ICU occupancy were available at different geographical levels (i.e., national, cantonal, hospital). Finally, we also downscale predictions of the hybrid model at the hospital-level to support hospital management actions.

Methods

2.1. Mechanistic model (MM)

We used the SEIR model described in Zhao et al.[30]. to simulate the dynamics of occupancy of ICU from the 6th of November 2020 until 1st of July 2021. This model was expanded to include the impact of vaccination campaigns[31]. This period included three epidemic phases: phase 1, from the lockdown (19th of October, 2020) until the start of second-dose vaccinations (15th of January, 2021); phase 2, from the start of second-dose vaccinations until the relaxation (14th of April, 2021); and phase 3, from the relaxation until 1st of July, 2021. The phases are reported in Fig S1 in the Supplementary Information (SI).

$$\beta = \frac{R_0}{N} \times k \times \gamma,$$

$$\frac{dS(t)}{dt} = -S\beta I - c,$$

$$\frac{dE(t)}{dt} = +S\beta I - \sigma E\ ,$$

$$\frac{dI(t)}{dt} = +\sigma E - \gamma I\ ,$$

$$\frac{dP(t)}{dt} = +\varepsilon_1 \gamma I - \omega_1 P$$

$$\frac{dH_1(t)}{dt} = +\omega_1 P - \omega_2 H_1\ ,$$

$$\frac{dH_2(t)}{dt} = +(1-\varepsilon_2)\omega_2 H_1 - \omega_3 ICU\ ,$$

$$\frac{dICU(t)}{dt} = \varepsilon_2\omega_2 H_1 - (1-\varepsilon_4)\omega_4 ICU - \varepsilon_4\omega_5 ICU\ ,$$
$$\frac{dR(t)}{dt} = +(1-\varepsilon_1)\gamma I + (1-\varepsilon_3)\omega_3 H_2 + (1-\varepsilon_4)\omega_4 ICU\ ,$$
$$\frac{dD(t)}{dt} = +\varepsilon_3\omega_3 H_2 + \varepsilon_4\omega_5 ICU\ ,$$
$$\frac{dC(t)}{dt} = +\gamma I\ .$$

Where *S* (Susceptible), *E* (Exposed), *I* (Infected), *P* (infected but not yet hospitalized), *H* (= $H_1 + H_2$, Hospitalized), *ICU*, *D* (Death), *R* (Recovered), and *C* (Cumulative Infected) are the model variables; R_0 the basic reproduction number, *c* the vaccination rate, and *k* the reduction/increase in transmission rate after a non-pharmaceutical intervention is introduced/relaxed. Parameter values and their meanings are reported in Table 1. Daily vaccination rates were obtained from the public dashboard of the Swiss Federal Office of Public Health[32].

Table 1. Model parameters of the SEIR model (adapted from Zhao et al.[30]).

Parameter	Description	Value
R_0	Basic reproduction number	Estimated
κ	Percentage of R_0	Estimated
c	Vaccination rate	Estimated
σ, γ	Serial interval	1/2.6 days[33]
ω_1	Duration from onset of symptoms to hospitalization	1/5 days[34]
ω_2	Initial hospitalization	1/6 days[34]
ω_3	Additional days of hospitalization until recovery/death	1/10 days[34]
ω_4	Additional days in ICU until recovery	1/13.1 days *
ω_5	Additional days in ICU until death	1/12.7 days *
ε_1	Rate of *H* admission of infected	0.0161[35]
ε_2	Hospitalized cases requiring critical care in ICU	30%[34]
ε_3	Death outside of ICU	35%[36]
ε_4	Death rate from ICU	22%*

* Obtained for patients (n = 382) included in the RISC-19-ICU registry supported by Swiss Society of Intensive Care Medicine (https://www.sgi-ssmi.ch).

2.2. Machine learning model

2.2.1. Model structure

We used a feed-forward NN with a single hidden layer[37,38] to predict the residuals ($\hat{\varepsilon}_{t+\Delta t)}$) of the SEIR model (Fig. 1). This choice was based on two properties that make this type of NN suitable for our purpose: first the ability to account for nonlinearities and interactions between response variables and covariates[24,39]. Second, ability of NN models to be trained on relatively small datasets that is comparatively higher than for other ML models such as deep neural networks[40].

Covariates (section 2.2.2.) were introduced in the NN with a lag corresponding to the maximum correlation with the response variable via cross-correlation. In particular, the lag of

t - Δt, ..., t – *42* was explored among the possibilities, with 42 days as the minimum value. An additional covariate, ε_t, was also added to account for the autoregressive nature of process.

The NN formulation was:

$$\hat{\varepsilon}_{t+\Delta t} = \mathcal{T}\left(\omega_0 + \sum_{j=1}^{J} \omega_j \cdot \mathcal{T}\left(\omega_{0j} + w_j^T x\right)\right) \quad (1)$$

Where Δt was set to 3 or 7 days; ω_0 is the intercept of the output layer, and ω_{0j} the intercept of j^{th} hidden node; ω_j is the weight (also known as parameter) associated with the connection from the j^{th} hidden node to the output layer, and w_j^T is the vector of weights associated with the connection to the j^{th} hidden node; Γ is the Rectified Linear Unit (ReLU) activation function; x is the vector of covariates. The size of the of the hidden layer, determined by the number of hidden nodes, was optimized together with other hyperparameters. Specifically, hyperparameters are different from parameters: parameters are learned during model training, while hyperparameters need to be optimized externally to model training (see section 2.3.).

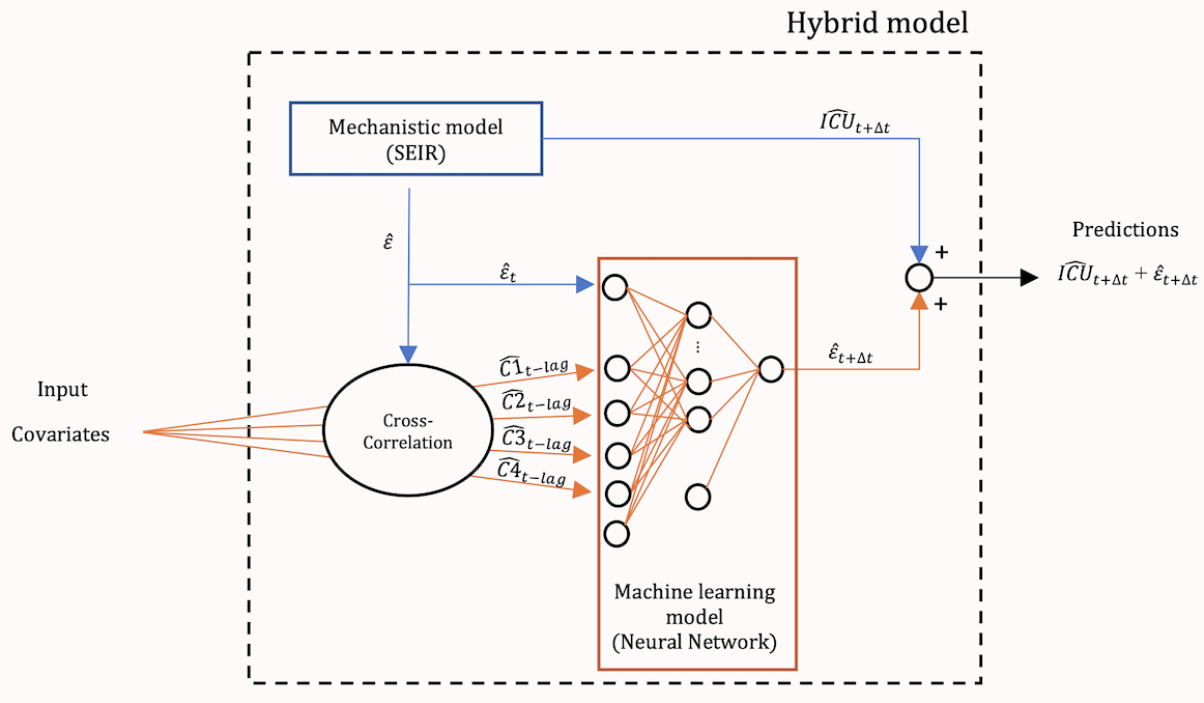


Fig 1. Configuration of the hybrid model. The hybrid model combines a mechanistic model (SEIR) with a machine learning model (Neural Network).

2.2.2. Covariates

These covariates were used to predict ICU occupancy : i) the number of COVID-19 cases; ii) the number of COVID-19 cases associated with the Alpha variant (better known as UK variant); iii) the level of non-pharmaceutical interventions (e.g., school closures, workplace closures, and travel bans) as identified by the Containment and Health Index (i.e., a subindex of the Stringency Index[41]); and iv) the mean daily air temperature.

Table 2. List of covariates.

Name	Source [Reference]
COVID-19 cases	Open Swiss Government data set[42]
Proportion of COVID-19 cases associated to the Alpha variant	Github repository[43]
Index of Containment and Health	Github repository[44]
Mean environmental temperature	opendata.swiss[45]

2.3. Model training and performance evaluation

We adopted a temporal cross-validation scheme[46] similar to the one used by Vollmer et al[29]. This scheme (Fig 2) allowed us to train and evaluate the performance of the hybrid model multiple times (n = 85 for the prediction at 3-days and n = 36 for the prediction at 7-days) over the simulated period.

The scheme works as follows:
First, the time series is divided in three successive time windows: the training set, validation set, and test set (Fig 2). The training and the validation sets are used for the optimization of the parameters of the SEIR model (one κ, the reduction applied to R_0, estimated separately for of the 3 phases), as well as the optimization of hyperparameters of the NN (i.e., the number of nodes in the hidden layer, learning rate, and dropout rate). A training-validation split of 90%/10% is adopted. The initial training and validation set included data from the 8th of October, 2020 until the 6th of November, 2020, in order to meet a minimum amount of data for model training.

Second, the performance of the trained model is evaluated on the test set using the mean absolute error. The test set consisted of n = Δt values, with Δt equals to 3 and 7 for the predictions at 3- and 7-days ahead, respectively.

Third, the training-validation set is expanded to include the test set of the previous iteration.

At the end of the iterative validation scheme, the overall performance of each model is estimated using the average MAE across iterations on the successive test sets (average MAE on red block for iteration 1 to n Fig. 2).

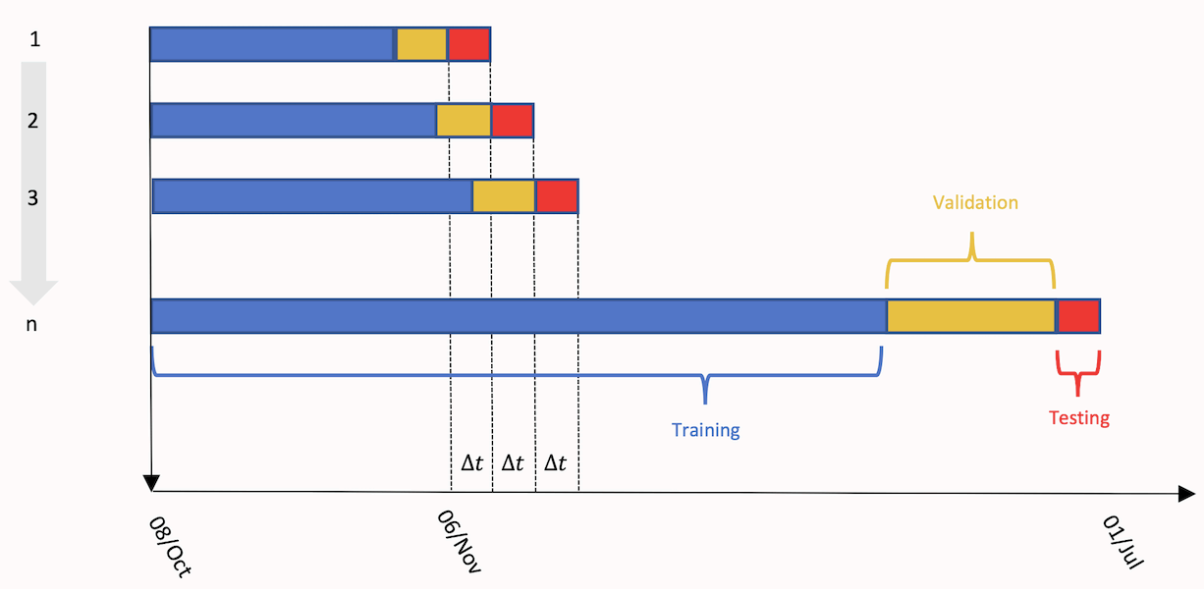


Fig 2. Temporal cross-validation scheme. Δt was set to 3 and 7 days for the predictions at 3 and 7 days ahead, respectively.

In step 2, the optimization of the SEIR model was performed using maximum likelihood (Nelder and Mead algorithm[47]) on the complete training-validation set; residuals of the SEIR model are then calculated. The optimization of the hyperparameters of the NN was done as follows: a sampling space of 100 combinations of hyperparameters was generated using a Latin Hypercube[48]. A back-propagation algorithm (based on gradient descent) was used as a learning algorithm to modify the values of the weights and obtain the best matches possible between the true and estimated values of the residuals of the SEIR model in the training set. The mean absolute error (MAE) was used as fitting criteria on the validation set, and an early stopping mechanism was applied to stop the learning algorithm if the MAE did not achieve a decrease of 5 units within 100 epochs (i.e., number of iterations that the learning algorithm worked through the training set). The largest number of possible epochs was set to 4,000. We accounted for the stochastic nature of the optimization by repeating the simulation 10 times for each combination of hyperparameters. For each set of 10 simulations, we calculated the mean and standard deviation of the MAE in the validation set. The combination of hyperparameters that generated the minimum mean MAE in the validation set was selected as optimal for the NN.

Performance evaluation was evaluated on the test set as follows: i) predictions of the SEIR model ($\widehat{ICU}_{t+\Delta t)}$) were based on the parameters inferred in the training-validation set (extrapolation); ii) predictions of the NN ($\hat{\varepsilon}_{t+\Delta t)}$) were obtained after training a NN with the optimal combination of hyperparameters on both the training and validation set; iii) the sum of the two contributions ($\widehat{ICU}_{t+\Delta t)} + \hat{\varepsilon}_{t+\Delta t)}$) was compared with the observed ICU occupancy. Confidence intervals for the NN were generated using the standard deviation calculated on the validation set, while they were obtained as described in Zhao et al. for the SEIR model[30], with the 2.5% and 97.5% quantiles of the 10,000 predictions.
Furthermore, the predictions (and accuracy evaluated via the MAE) of the hybrid model were compared to that of the SEIR and NN model independently.

2.4. Downscaling at hospital level

Model predictions were obtained at the national- and cantonal-level, and from the cantonal-level downscaled to the hospital-level. Particularly, cantonal-level predictions were downscaled based on the percentage of occupancy of ICU beds in each hospital, calculated as moving average of the past Δt days. The ICU occupancy data of Swiss hospitals were provided daily by the Coordinated Sanitary Service of the Swiss Armed Forces, and refer to the number of ICU beds occupied by COVID-19 patients. The canton of Zurich was selected for model testing, since it is the most populated Canton in Switzerland with 15 hospitals with ICU.

2.5. Importance of covariates

We determined the relative importance of each covariate in making predictions for the hybrid model. Concretely, we computed a Deviation metric (2) between the MAE of the full model including n=5 covariates with that of a reduced model with n=1 covariates[49]. The Deviation was calculated on the test set at the end of each epidemic phase. The procedure was repeated n=5 times excluding one covariate at a time. The Deviation (%) was calculated as follows:

$$Deviation = \frac{MAE_{Reduced\ Model} - MAE_{Full\ Model}}{MAE_{Full\ Model}} x\ 100 \quad (2)$$

A positive Deviation signifies that the excluded covariate was important for the model. Specifically, a positive Deviation corresponds to a decreased accuracy of the reduced model compared to the accuracy of the full model that included all covariates.

Results

3.1. Model comparison

We compared 3 types of epidemic models (i.e., SEIR, NN and hybrid,) to predict short-term (3 and 7 days ahead) ICU occupancy at the national- and cantonal-level. Fig 3 shows the 3-day predictions at the national-level from the three models (a), and with its associated MAE calculated on the test set (b).

During phase 1 (19th of October, 2020 - 15th of January, 2021) the hybrid model (average MAE = 19 beds) outperformed both the NN (average MAE = 27 beds) and the SEIR model (average MAE = 78 beds). During phase 2 (15th of January, 2021 - 14th of April, 2021), the hybrid model remained the most accurate model (average MAE = 16 beds), although the performance of the NN (average MAE = 21 beds) and the SEIR model (average MAE = 59 beds) improved in comparison with phase 1.

During phase 3 (14th of April, 2021 - 1st of July, 2021), the hybrid model (average MAE = 19 beds) was slightly outcompeted by the NN (average MAE = 13 beds); while the SEIR model was associated with the worse performances (average MAE = 125 beds). Fig S1 in Supporting Information (SI) showed predictions at 7-days ahead and its corresponding MAE. Predictions 3-days ahead (Fig S2) were then downscaled from the cantonal-level to the hospital-level. Results for a medium-sized hospital, as well as the biggest hospital in the canton of Zurich are shown in Fig 4. At the hospital-level, the hybrid model outperformed the SEIR model for both the medium-sized hospital (average MAE_{hybrid} = 1.2 beds, average

MAE_{SEIR}= 2.2 beds) and largest hospital (average MAE_{hybrid} = 3.1 beds, average MAE_{SEIR}= 6.2 beds) in the canton of Zurich. In comparison, the NN model performed on average as good as the hybrid model for both hospitals. Similar to the national-level scenario, the highest average MAE for the SEIR model was observed during phase 3, during which the SEIR model was not capable of capturing the occupancy increase of ICU that occurred two months after the start of vaccination (15th of January, 2021).

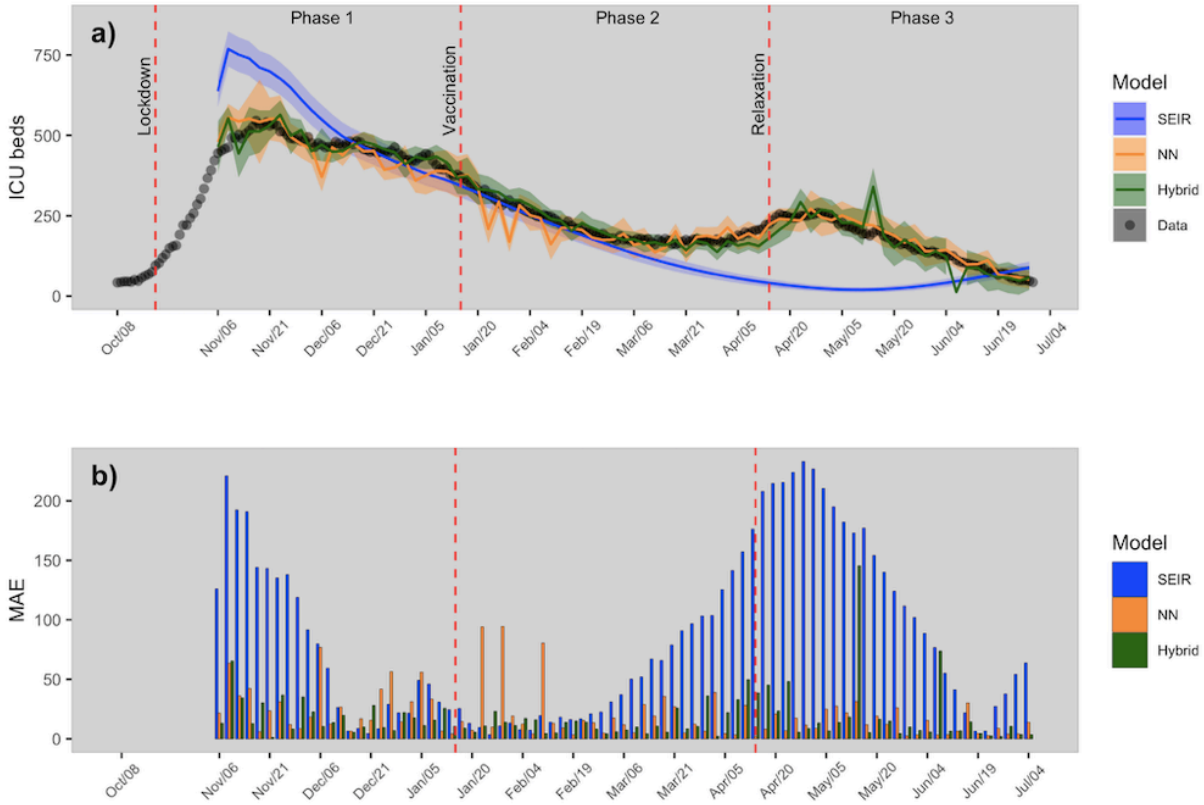


Fig 3. Model predictions of intensive care bed occupancy at the national-level. a) Predictions 3-days ahead of intensive occupancy at the national-level for the three models (shaded areas represent 95% confident intervals); b) corresponding Mean Absolute Error (MAE) calculated on test data.

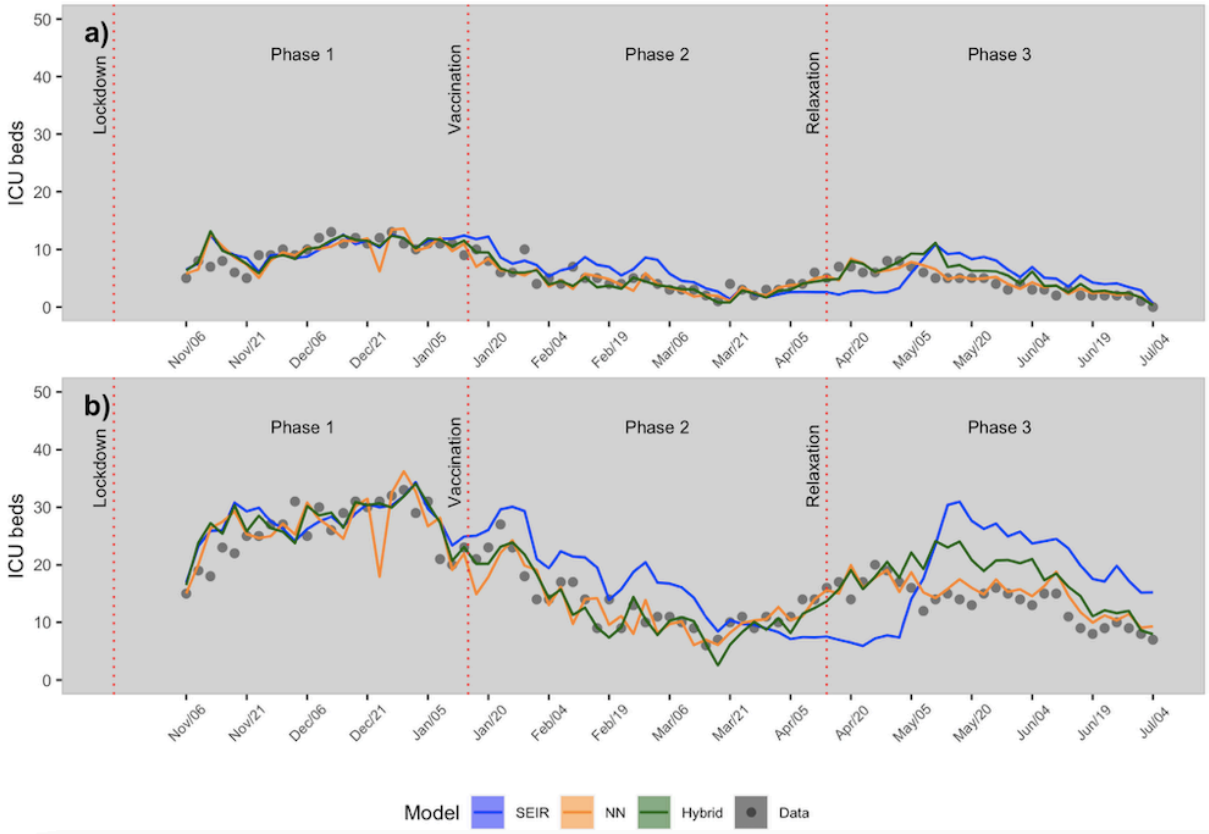


Fig 4. Model predictions of intensive care bed occupancy at the hospital-level. Prediction at the hospital-level for a medium-sized hospital (a) and the largest hospital (b) in the canton of Zurich.

3.2. Relative importance of covariates

The relative importance of covariates during each of the three phases is reported in Fig 5. In phase 1, a negative Deviation (marked with an asterisk in the Figure) was observed for a majority of the covariates (i.e., COVID-19 cases, proportion of COVID-19 cases associated to the alpha variant, Index of Containment and Health, and mean environmental temperature), meaning that their exclusion from the full model improved prediction accuracy. Conversely, the autoregressive covariate was important for making predictions, with Deviations equal to 92% and 66% for the hybrid and NN model, respectively. In phase 3, all of the covariates were informative; in this last phase, the NN predictions were more affected by the exclusion of covariates in comparison to the hybrid model. The average Deviation was 230% and 53%, for the NN and hybrid model, respectively.

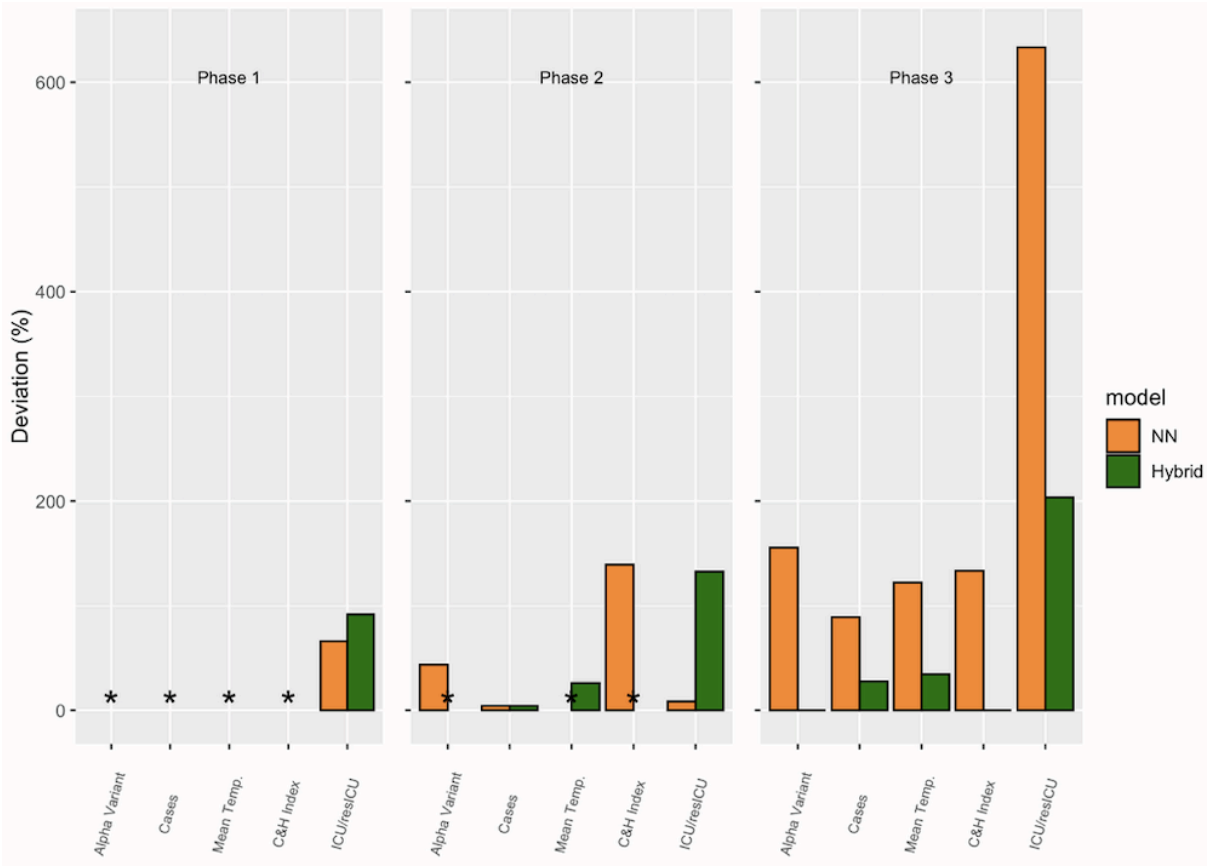


Fig 5. Covariance importance for each phase. An asterisk represents a negative Deviation.

Discussion

4.1. Prediction accuracy

In this study, we showed increased prediction accuracy of ICU occupancy using a hybrid model combining a SEIR and a NN model. The model developed here could help guide interventions against future COVID-19 epidemics. At a national-level, during phase 1 (19th of October, 2020 - 15th of January, 2021) the overestimation of ICU occupancy by the SEIR model could be associated with its intrinsic nature to predict exponential growth at the beginning of a new wave. In contrast, the SEIR model underestimated the ICU occupancy during phase 3 (14th of April, 2021 - 1st of July, 2021). This could be explained by the fact that the model lacks important covariates such as temperature, which may have been responsible of an increase of cases during the winter and thus for an increased ICU occupancy. For the NN model, its worst performance was observed during phase 1, where abrupt oscillations occurred. These oscillations could be attributed to the short time series available for model training at that stage, thereby compromising model training and limiting predictive performance. This interpretation is supported by the fact that the prediction accuracy of the NN model improved during phase 2 (15th of January, 2021 - 14th of April, 2021) and 3 (14th of April, 2021 - 1st of July, 2021), when longer time series became available for model training.

At the national- and cantonal-levels, the SEIR model was unable to capture the increase in ICU occupancy that occurred two months after the beginning of the second-dose vaccination campaign. The causal mechanisms behind this trend remains unclear, but may be associated with other drivers such new variants (e.g., Delta variant) that are not incorporated in the SEIR model. In contrast, the hybrid and NN model could capture this trend, suggesting that both models succeeded in learning potential non-linear relationships between covariates and occupancy of ICU.

4.2. Relative importance of covariates

The fact that the relative importance of each covariate for our models changed between phases has multiple possible interpretations. The first is that a covariate is important for making predictions during one phase, while it is not important for another phase. For example, the proportion of the Alpha variant was not informative during phase 1 when its prevalence was < 10% of the total confirmed COVID-19 cases, while it was informative during phase 3, when its prevalence was > 50% of the total confirmed Covid-19 cases. The second reason could be associated to the length of the time series. For example, the model had limited data for training during phase 1, while the amount of data tripled for phase 3. This could have caused the full model (i.e., with 5 covariates) to perform worse than the reduced model (i.e., with 1-5 covariates)[50,51], leading to negative Deviation.

During phase 1, the autoregressive term was the only informative covariate, meaning that the models behaved similarly to an Automatic Regressive Integrated Moving Average (ARIMA) model. Furthermore, on average, the Deviation associated with the hybrid model was always lower than the one associated to the NN. This means that the hybrid model was more robust in the exclusion of a specific covariate compared to the NN.

4.3. Possible applications and limitations

In Switzerland, a number of studies have focussed on providing long-term (>2 weeks)[52,53,54] and short-term (<2 weeks)[36] predictions using MMs. Predictions have predominantly on the national scale; while many of the lifesaving actions (e.g., transfer of patients) need to be planned at the cantonal- (provincial-level) or hospital-level. In this study, we attempted to increase the prediction accuracy of a SEIR model by coupling it with a NN, generating a so-called hybrid model. Among all the possible ways to combine a MM with a ML model, we opted for a configuration called residual modelling. In particular, we used a SEIR model for predicting occupancy of ICU beds under future scenarios at different geographical levels (national, cantonal, and hospital) in Switzerland; we trained a NN to supplement these predictions using the information embedded in covariates (temperature, etc.). This modelling framework could be applied in other geographic regions for which a MM (e.g., of the SIR family), and spatially explicit covariates are available. Specifically, different extension of the SIR model[14] can be used, from simple examples (like the SEIR used in this study), to increasingly complex frameworks such as SIDARTHE[12]. As for the ML model, we used a feed-forward NN with a single hidden (see section 2.2.1.). However, alternative formulations could have been chosen. For example, Maher Ala'raj et al.[27] coupled an ARIMA model, a very popular ML model for time series forecasting with a SEIRD model; Watson et al.[17] embedded a Bayesian time series model and a random forest algorithm within a SIRD model; Rahmadani and Lee[28] combined a deep-learning algorithm with a SEIR model.

As with any modelling study, our analysis also comes with limitations. For example, training of the NN is often computationally intensive and the selection of optimal hyperparameters is based on empirical rules such as try-and-error approaches[46]. In this study, the optimization of hyperparameters required a significant effort in terms of computational cost. Specifically, all simulations were run in parallel on ETH High Performance Computing facilities (Euler cluster)[55], requiring, on average, 1 minute per simulation on one CPU, and thus 15 CPU minutes for each iteration running simultaneously on 15 CPU cores. We optimized three hyperparameters, namely the number of nodes in the hidden layer, learning rate, and dropout rate; however, other hyperparameters such as the type of activation function, the number of batches, the number of epochs, etc., could also have been subjected to optimization. Furthermore, other type of search algorithms such as the sequential model-based optimization (SMBO, also known as Bayesian optimization)[56] could have been explored. Another drawback of the residual modelling configuration is the inability to enforce real-world constraints (e.g., ICU beds ≥ 0), since the residuals are modelled instead of based on the actual ICU occupancy. One possible alternative could be to combine the SEIR model and the NN in series. In this case, the NN estimates intermediate variables to be used in the SEIR model, although it would impose structural changes on the SEIR model based on the variables selected, which may be challenging to implement.

As for the downscaling at the hospital-level, we used a simple method based on the moving average to downscale predictions at the cantonal-level (see section 2.4.), demonstrating a satisfactory degree of accuracy in hospitals in the Canton of Zurich. However, this method requires the availability of ICU beds at the hospital-level, which is not always the case. Consequently, more complex methods could be tested. In particular, Zhao et al.[30] presented a method to distribute ICU patients based on travel time from the location of the patient to the hospital. In the future, our modelling framework can be updated as growing knowledge is gained on the covariates associated with the spread of COVID-19. For example, new covariates such as other virus variants and mobility patterns in different regions (e.g., people coming in and out of Switzerland) could be included to improve predictions. Lastly, the framework could be applied to improve predictions of other infectious diseases, for which a MM already exists.

References

1. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. Acta Biomed Atenei Parm. 2020;91: 157–160. doi:10.23750/abm.v91i1.9397
2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20: 533–534. doi:10.1016/S1473-3099(20)30120-1
3. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? The Lancet. 2020;395: 1225–1228. doi:10.1016/S0140-6736(20)30627-9
4. Fagiuoli S, Lorini FL, Remuzzi G. Adaptations and Lessons in the Province of Bergamo. N Engl J Med. 2020;382: e71. doi:10.1056/NEJMc2011599
5. Li R, Rivers C, Tan Q, Murray MB, Toner E, Lipsitch M. The demand for inpatient and ICU beds for COVID-19 in the US: lessons from Chinese cities. medRxiv. 2020 [cited 31 May 2021]. doi:10.1101/2020.03.09.20033241
6. Covid-19 in India: Why second coronavirus wave is devastating. BBC News. 20 Apr 2021. Available: https://www.bbc.com/news/world-asia-india-56811315. Accessed 31 May 2021.

7. Phua J, Hashmi M, Haniffa R. ICU beds: less is more? Not sure. Intensive Care Med. 2020;46: 1600–1602. doi:10.1007/s00134-020-06162-8
8. Harris S, Singer M, Sanderson C, Grieve R, Harrison D, Rowan K. Impact on mortality of prompt admission to critical care for deteriorating ward patients: an instrumental variable analysis using critical care bed strain. Intensive Care Med. 2018;44: 606–615. doi:10.1007/s00134-018-5148-2
9. Metzler B, Siostrzonek P, Binder RK, Bauer A, Reinstadler SJ. Decline of acute coronary syndrome admissions in Austria since the outbreak of COVID-19: the pandemic response causes cardiac collateral damage. Eur Heart J. 2020;41: 1852–1853. doi:10.1093/eurheartj/ehaa314
10. Aleta A, Martín-Corral D, Pastore y Piontti A, Ajelli M, Litvinova M, Chinazzi M, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. Nat Hum Behav. 2020;4: 964–971. doi:10.1038/s41562-020-0931-9
11. Balabdaoui F, Mohr D. Age-stratified discrete compartment model of the COVID-19 epidemic with application to Switzerland. Sci Rep. 2020;10: 21306. doi:10.1038/s41598-020-77420-4
12. Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nat Med. 2020;26: 855–860. doi:10.1038/s41591-020-0883-7
13. Supino M, d'Onofrio A, Luongo F, Occhipinti G, Dal Co A. The effects of containment measures in the Italian outbreak of COVID-19. BMC Public Health. 2020;20: 1806. doi:10.1186/s12889-020-09913-w
14. Kermack WO, McKendrick AG, Walker GT. A contribution to the mathematical theory of epidemics. Proc R Soc Lond Ser Contain Pap Math Phys Character. 1927;115: 700–721. doi:10.1098/rspa.1927.0118
15. DeAngelis DL, Mooij WM. Individual-Based Modeling of Ecological and Evolutionary Processes. Annu Rev Ecol Evol Syst. 2005;36: 147–168. doi:10.1146/annurev.ecolsys.36.102003.152644
16. Gitto S, Mauro CD, Ancarani A, Mancuso P. Forecasting national and regional level intensive care unit bed demand during COVID-19: The case of Italy. PLOS ONE. 2021;16: e0247726. doi:10.1371/journal.pone.0247726
17. Watson GL, Xiong D, Zhang L, Zoller JA, Shamshoian J, Sundin P, et al. Pandemic velocity: Forecasting COVID-19 in the US with a machine learning & Bayesian time series compartmental model. PLOS Comput Biol. 2021;17: e1008837. doi:10.1371/journal.pcbi.1008837
18. Rohrer M, Flahault A, Stoffel M. Peaks of Fine Particulate Matter May Modulate the Spreading and Virulence of COVID-19. Earth Syst Environ. 2020;4: 789–796. doi:10.1007/s41748-020-00184-4
19. Poirier C, Luo W, Majumder MS, Liu D, Mandl KD, Mooring TA, et al. The role of environmental factors on transmission rates of the COVID-19 outbreak: an initial assessment in two spatial scales. Sci Rep. 2020;10: 17002. doi:10.1038/s41598-020-74089-7
20. Baker RE, Peña J-M, Jayamohan J, Jérusalem A. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol Lett. 2018;14: 20170660. doi:10.1098/rsbl.2017.0660
21. Karpatne A, Atluri G, Faghmous J, Steinbach M, Banerjee A, Ganguly A, et al. Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data. IEEE Trans Knowl Data Eng. 2017;29: 2318–2331. doi:10.1109/TKDE.2017.2720168
22. Thompson ML, Kramer MA. Modeling chemical processes using prior knowledge and neural networks. AIChE J. 1994;40: 1328–1340. doi:https://doi.org/10.1002/aic.690400806

23. Willard J, Jia X, Xu S, Steinbach M, Kumar V. Integrating Physics-Based Modeling with Machine Learning: A Survey. ArXiv200304919 Phys Stat. 2020 [cited 17 May 2021]. Available: http://arxiv.org/abs/2003.04919
24. Hornik K, Stinchcombe M, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Netw. 1990;3: 551–560. doi:10.1016/0893-6080(90)90005-6
25. Chu F, Wang F, Wang X, Zhang S. A hybrid artificial neural network—mechanistic model for centrifugal compressor. Neural Comput Appl. 2014;24: 1259–1268. doi:10.1007/s00521-013-1347-5
26. Lee DS, Vanrolleghem PA, Park JM. Parallel hybrid modeling methods for a full-scale cokes wastewater treatment plant. J Biotechnol. 2005;115: 317–328. doi:10.1016/j.jbiotec.2004.09.001
27. Ala'raj M, Majdalawieh M, Nizamuddin N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. Infect Dis Model. 2021;6: 98–111. doi:10.1016/j.idm.2020.11.007
28. Rahmadani F, Lee H. Hybrid Deep Learning-Based Epidemic Prediction Framework of COVID-19: South Korea Case. Appl Sci. 2020;10: 8539. doi:10.3390/app10238539
29. Vollmer MAC, Glampson B, Mellan T, Mishra S, Mercuri L, Costello C, et al. A unified machine learning approach to time series forecasting applied to demand at emergency departments. BMC Emerg Med. 2021;21: 9. doi:10.1186/s12873-020-00395-y
30. Zhao C, Tepekule B, Criscuolo NG, Wendel Garcia PD, Hilty MP, RISC-19-ICU consortium Investigators in Switzerland, et al. icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. Swiss Med Wkly. 2020;150. doi:10.4414/smw.2020.20277
31. Yıldırım A, Cherruault Y. Analytical approximate solution of a SIR epidemic model with constant vaccination strategy by homotopy perturbation method. Kybernetes. 2009;38: 1566–1575. doi:10.1108/03684920910991540
32. Federal Office of Public Health FOPH. Vaccinated People. Available: https://www.covid19.admin.ch/en/epidemiologic/vacc-persons
33. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance. 2020;25: 2000257. doi:10.2807/1560-7917.ES.2020.25.17.2000257
34. Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College London; 2020 Mar. Available: http://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/
35. Hauser A, Counotte MJ, Margossian CC, Konstantinoudis G, Low N, Althaus CL, et al. Estimation of SARS-CoV-2 mortality during the early stages of an epidemic: A modeling study in Hubei, China, and six regions in Europe. PLOS Med. 2020;17: e1003189. doi:10.1371/journal.pmed.1003189
36. Althaus CL. Real-time modeling and projections of the COVID-19 epidemic in Switzerland. 24 Apr 2020. Available: https://ispmbern.github.io/covid-19/swiss-epidemic-model/

37. Günther F, Fritsch S. neuralnet: Training of Neural Networks. R J. 2010;2: 30. doi:10.32614/RJ-2010-006
38. Bishop CM. Neural networks for pattern recognition. New York: Oxford University Press; 1995.
39. Belue LM, Bauer KW. Determining input features for multilayer perceptrons. Neurocomputing. 1995;7: 111–121. doi:10.1016/0925-2312(94)E0053-T
40. San O. Neural network closures for nonlinear model order reduction. : 34.
41. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav. 2021;5: 529–538. doi:10.1038/s41562-021-01079-8
42. OpenDataZH. SARS-CoV-2 open government data reported by the Swiss Cantons and the Principality of Liechtenstein. Available: https://github.com/openZH/covid_19/tree/master#community-contributions
43. Chen C, Nadeau S, Topolsky I, Manceau M, Huisman JS, Jablonski KP, et al. Quantification of the spread of SARS-CoV-2 variant B.1.1.7 in Switzerland. medRxiv. 2021; 2021.03.05.21252520. doi:10.1101/2021.03.05.21252520
44. Dünner RP, Penny MA. COVID_measures_CH. Available: https://github.com/SwissTPH/COVID_measures_by_canton
45. Opendata.swiss. Available: https://opendata.swiss/en
46. Raschka S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. ArXiv181112808 Cs Stat. 2020 [cited 7 May 2021]. Available: http://arxiv.org/abs/1811.12808
47. Nelder JA, Mead R. A Simplex Method for Function Minimization. Comput J. 1965;7: 308–313. doi:10.1093/comjnl/7.4.308
48. McKay MD, Beckman RJ, Conover WJ. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics. 1979;21: 239–245. doi:10.2307/1268522
49. Fisher A, Rudin C, Dominici F. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. J Mach Learn Res. 20: 1–81.
50. Cibas T, Soulié FF, Gallinari P, Raudys S. Variable selection with neural networks. Neurocomputing. 1996;12: 223–248. doi:10.1016/0925-2312(95)00121-2
51. Deng X, Li Y, Weng J, Zhang J. Feature selection for text classification: A review. Multimed Tools Appl. 2019;78: 3797–3816. doi:10.1007/s11042-018-6083-5
52. Marini M, Chokani N, Abhari RS. COVID-19 Epidemic in Switzerland: Growth Prediction and Containment Strategy Using Artificial Intelligence and Big Data. medRxiv. 2020; 2020.03.30.20047472. doi:10.1101/2020.03.30.20047472
53. Shattock AJ, Rutte EAL, Dünner RP, Sen S, Kelly SL, Chitnis N, et al. Impact of vaccination and non-pharmaceutical interventions on SARS-CoV-2 dynamics in Switzerland. medRxiv. 2021; 2021.04.14.21255503. doi:10.1101/2021.04.14.21255503
54. Neher RA, Dyrdak R, Druelle V, Hodcroft EB, Albert J. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. Swiss Med Wkly. 2020;150. doi:10.4414/smw.2020.20224
55. ETH Euler Cluster. Available: https://scicomp.ethz.ch/wiki/Euler
56. Bergstra J, Yamins D, Cox DD. Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms.

Chapter 6

Discussion

6.1 Summary of findings

In this thesis, we map trends in AMR in food animals, and develop novel approaches to guide future surveillance efforts. We also present our contributions to inform decision-making in near real-time during the COVID-19 pandemic in Switzerland.

In Chapter 2.1, we focus on China, the world's largest consumer of veterinary antimicrobials. Our work builds up on efforts to review PPS to map AMR in LMICs in 2019[1], which we considerably expanded to include publications in the Chinese language. This results in a three-fold increase in our database of PPS that was used to create an improved map of AMR. We show that between 2000 and 2019, in China, AMR in pigs underwent the biggest increase by +59% in *E. coli*, +148% in *Salmonella* and +85% in *S. aureus*. In comparison, AMR in chicken has plateaued, albeit at high levels: the proportion of antimicrobials with resistance higher than 50% was 60% in *E. coli*, 42% in *Salmonella*, and 37% in *S. aureus*. Furthermore, we show that hotspots of AMR are currently concentrated in eastern China (e.g., eastern Jiangsu Province), and part of central and northwestern China (e.g., southeastern Sichuan, and Xinjiang Uyghur Autonomous Region).

In Chapter 3, we focus on Europe where national-level systematic surveillance of AMR has existed for decades. We use PPS for supplementing AMR trends at finer spatial scale within countries. For multiple countries with relatively high country-level AMR prevalence, such as Italy and Romania, we show that rather than having consistently high AMR throughout the country, specific regions in southern Romania and southern/eastern Italy are AMR hotspots. Interventions to reduce AMR could be targeted in these areas as a priority.

In Chapter 4, we refine and disaggregate the index of AMR introduced by Van Boeckel and Pires et al.[1] that are used to summarize trends in AMR in Chapter 2 and Chapter 3. Instead, we focus on characterizing resistance prevalence for individual drug-bacteria combinations. We select 7 drugs each from a different antimicrobial class of medical importance, and map their AMR prevalence in *E. coli* and *Salmonella* at global scale. We show that AMR levels are higher in *E. coli* than in *Salmonella* for all antimicrobial classes. We also show that patterns of AMR vary between antimicrobials. For example, northeastern China is predicted as a resistance hotspot for sulphonamides and aminoglycoside in *E. coli*, while northern and eastern Brazil are resistance hotpots for penicillins in *E. coli*.

In Chapter 2 and Chapter 4, we also investigate how resources can be optimally deployed for future AMR surveillance. In Chapter 2.1, we show that future AMR surveillance could be targeted at southwestern and northeastern China, using a computational approach aimed at minimizing the uncertainty of mapped AMR predictions. We show that instead of distributing surveys evenly across administrative divisions, our geographically targeted approach of surveillance could reduce AMR prediction uncertainty by 2-fold. This approach is used in Chapter 2.2 on PPS of AMR in aquaculture in Asia, and suggests that eastern China and India could benefit the most from a scale-up of surveillance efforts in aquaculture in the future. In Chapter 4, we use patterns of co-resistance in combination with local environmental and anthropogenic covariates, to predict which antimicrobial has the highest probability of its resistance exceeding critical levels (e.g. 50%) in the future. In the majority of locations in Africa and South America, future surveillance could be targeted at penicillins or tetracyclines; while in Asia, the target antimicrobial classes are penicillins or sulphonamides.

In Chapter 5, we move from large-scale risk mapping to local-scale disease dynamics modelling for the operational management of epidemics. During the COVID-19 pandemic, we built an online platform *icumonitoring.ch* to anticipate hospital occupancy in the intensive care units 3- and 7-days ahead. These near real-time predictions were provided to authorities (Federal Office of Public Health, Armed Forces, hospitals) to help optimally allocate medical resources, including doctors and hospital beds, in preparation for potential hospital overload.

6.2 Global initiatives on AMR

AMR is a global issue with far-reaching impacts across sectors including human health, terrestrial food animal production, aquaculture, and the environment. Therefore, the effective containment of AMR can only be achieved through collective and multisectoral endeavours. Globally, several organizations have launched initiatives dedicated to addressing AMR in their respective sectors.

In the human sector, the World Health Organization launched the Global Antimicrobial Resistance and Use Surveillance System (GLASS) in 2015[2]. The programme conducts surveillance on AMR in bacteria causing infections in humans, and provides annual estimates of AMR prevalence per country, categorised by infectious syndromes, bacterial pathogens, and antimicrobial classes. Since 2020, GLASS also started surveillance on human antimicrobial consumption. By the end of 2022, 127 countries, territories and areas are involved in GLASS. In 2022, Murray et al. moved beyond surveillance of AMR prevalence, and combined it with disease incidence data to estimate the burden of AMR to human health[3].

In the environment sector, Hendriksen et al. used untreated sewage samples across countries to monitor the abundance and diversity of AMR genes[4]. Metagenomic analyses were conducted to characterise the resistome of bacteria that were mainly of faecal and environmental origin. Although this approach suffers from the limitation that the presence of AMR genes is not necessarily linked to actual expression of a resistant phenotype, it represents a flexible and affordable approach to survey resistance genes LMICs. However, to the best of our knowledge, there is an absence of systematic reporting of AMR in the environment sector with a dedicated system akin to GLASS for the human sector. One significant challenge is the considerable variability in the methods and targets adopted in different environmental monitoring programs of AMR. Therefore, ongoing collaborative efforts[5] need to be strengthened to establish a standardized framework for methods and quality control in this sector.

In the food animal sector, since 2016, the World Organisation for Animal Health has been collecting country-level antimicrobial use data on a voluntary basis. However, these data are aggregated at the regional level upon public release[6]. In 2018, the World Health Organization launched a One Health Module, incorporating AMR surveillance in the food animal sector. However, the scope of this surveillance is limited on extended-spectrum beta-lactamase (ESBL)-*E. coli* in poultry. In 2019, we produced the first global map of AMR prevalence at 10x10 km resolution across LMICs. The map was based on an amalgamation of PPS between 2000 and 2019[1]. We further improved our estimations of AMR in this thesis. In Chapter 2.1, we improved the mapped predictions of AMR in China; In Chapter 3, we disaggregated the overall trends of AMR across LMICs into trends for 7 antimicrobial classes of medical importance; In Chapter 2.2, we extended the mapping exercise to the aquaculture sector. These maps contribute to a more comprehensive overview of AMR across sectors. However, our maps did not quantify the burden of animal disease attributable to AMR, compared with the

work of Murray et al. in the human sector. With the rollout of the Global Burden of Animal Diseases project in 2018, a comprehensive overview of animal disease incidences will hopefully be made available in the near future[7]. By combining these disease data with our estimates of AMR prevalence, we may be able to map the burden of AMR for animal health in the future.

6.3 Policy implications

6.3.1 Priorities for intervention

The use of antimicrobials is a double-edged sword. Antimicrobials can be used to treat infections in animals, which not only safeguards their well-being but also has a positive effect on neighbouring animals and the livelihood of those who rely on livestock for subsistence. However, this practice also fuels the rise of AMR, making future treatments less effective. Adding to this problem is the shrinking global supply of new antimicrobial drugs, with only 4 new classes of antimicrobials developed between 1970 and 2010[8]. Consequently, the problem of AMR can be seen as a common goods challenge, where individual consumption generates externalities affecting the community as a whole. In light of these insights, and drawing from the findings presented in this thesis, two key priorities for policy interventions are outlined below.

Global challenges need global coordination

First, as with other externalities of global importance such as climate change, international collaboration is key for containing AMR. Woolhouse et al. called for the establishment of an intergovernmental panel on antimicrobial resistance (IPAMR), akin to the existing Intergovernmental Panel on Climate Change (IPCC)[9]. They argued that such a panel should identify key knowledge gaps and assess short- and long-term feasible solutions to reduce AMR. A well-defined road map can then be designed with internationally agreed and evidence-based targets of reduced antimicrobial use and subsequently contained AMR levels. As an analogy to how IPCC bases its target of reaching net zero carbon emission in 2050 on past temperature measurement and models of future climate, our maps of AMR in food animals could serve as a first set of evidence base for designing targets and intervention strategies for an IPAMR. We aim to revise our AMR maps on a regular basis, to incorporate the most up-to-date evidence on AMR trends. This can help policymakers to assess where we stand as compared to our targets, in a similar way that climate change models are adjusted regularly to update the likelihood of meeting previously set targets.

An IPAMR could use our maps of AMR to identify where future surveillance could be prioritized. Currently, there are no PPS available in countries such as Mongolia and Kazakhstan. AMR prevalence in these regions were estimated based on geospatial models trained using PPS data from other countries such as China and India. An IPAMR could coordinate efforts across countries to conduct surveillance in regions with a lack of AMR data. In addition, an IPAMR could help address a significant limitation of the current amalgamation of PPS – the differences in sampling and antimicrobial susceptibility testing methodologies between surveys. It could help improve the quality of future surveillance data, by encouraging countries to adopt a harmonized and systematic surveillance approach such as the one we developed with resistancebank.org.

In addition, we provided a concrete and computationally feasible method for an IPAMR to optimize its allocation of AMR surveys. This method was described in detail in Chapter 2.1, where we used the map of prediction uncertainty to identify the locations where conducting new surveys could be most valuable to improve the confidence level of a prevalence map. We used AMR in China as a case study, and showed that compared with distributing new surveys evenly across administrative divisions, our method could reduce AMR prediction uncertainty by two-fold. However, for the effective rollout of harmonized surveillance across LMICs in the future, international assistance from high-income countries will also be necessary, as many locations without data are also locations with limited resources available.

A targeted approach for policy intervention

Second, we identified hotspots of AMR across LMICs, which may carry the highest negative externalities and call for targeted interventions to reduce both antimicrobial use and AMR. A previous literature review identified three main methods[10] for taking externalities of AMR into account. These methods include regulatory measures, levies or taxes on antimicrobial usage, and the allocation of trade permits or licenses for antimicrobial utilization. Currently, only regulatory measures are adopted to reduce antimicrobial use. In several high-income countries, such regulations have been implemented for decades. For example, Scandinavian countries were the first to initiate restrictions on agricultural antimicrobial use[11]. In Sweden, farmers petitioned to ban all antimicrobial growth promoters, which took effect in 1986. In Denmark, farmers voluntarily stopped using avoparcin and it was officially banned by the government in 1995. The European Union member states also banned four commonly used antimicrobial growth promoters in 1998, and eventually banned all antimicrobial growth promoters in 2006. Other high-income countries also followed. For example, the US banned fluoroquinolones in poultry production in 2005, and banned all antimicrobial growth promoters in 2017. These regulations have proven successful in decreasing AMR in animals for some antimicrobial classes. For example, in Denmark, the ban of cephalosporin use in pig production in 2010 led to a decrease in the prevalence of cephalosporinase-producing *E. coli* from 11.8% to 3.6% within a year[12]. In Canada, the withdrawal of ceftiofur in chicken in 2005 led to a decrease in its resistance in *Salmonella* for both chicken and humans[13]. Similarly, another study in Canada found that AMR in chicken declined by 6% - 38% after decreasing antimicrobial use[14].

In comparison, in LMICs, changes in regulations to limit antimicrobial use and contain AMR in animals have been slow. First, while meat consumption in many high-income countries have plateaued, it continues to grow in LMICs[15], driving expansions of antimicrobial-dependent intensive farming systems. Second, the transformation of food animal systems towards less antimicrobial dependency requires significant investment on improving biosecurity measures, hygiene, and animal welfare to prevent loss of productivity. LMICs lack resources to support such transformation. Currently, global funders such as the Fleming Fund spend the most of their resources on strengthening surveillance and AMR workforces in LMICs[16]. Our predictions of AMR hotspots can provide funders with targets for investing in the transformation of agricultural systems. Engagement with potential funders can be established through a direct communication of our results with them, clearly summarizing the research findings and potential benefits of their adoption in practice.

In addition to restricting antimicrobial use, policy interventions can be implemented to incentivize the development of novel antimicrobials. These interventions may include both “push” and “pull” economic incentives for producers. On the "push" side, financial support, grants, and subsidies can be provided to research institutions and pharmaceutical companies

engaged in antimicrobial development. This investment can help offset the high costs and risks associated with early-stage research and development, encouraging entities to invest in ground-breaking projects. On the "pull" side, governments or international bodies can offer financial rewards, market exclusivity, or advanced purchase commitments for successful antimicrobial products. These incentives aim to attract producers by ensuring a return on investment and a viable market for their developed antimicrobials. By combining both push and pull mechanisms, policymakers can establish a supportive system that encourages innovation, research, and production of urgently needed antimicrobial solutions. This dual-pronged strategy seeks to address economic barriers and stimulate sustained efforts in the fight against AMR.

6.3.2 The utility of science and modelling in assisting decision-making

"Following the science" vs. "Choosing the science that fits"

In this thesis, we combined scientific data with models to assist decision-making during pandemics, and to anticipate actions that could reduce AMR on a global level. Throughout history, science has played a pivotal role in shaping policy decisions. A notable example was the formation of Montreal Protocol, based on scientific evidence linking the depletion of the ozone layer to the use of chemicals such as chlorofluorocarbons. The Montreal Protocol led to the phase-out of nearly 99% of ozone-depleting substances, and subsequently notable recovery of the protective ozone layer[17]. However, when science meets politics, it is not always science that is being followed as other factors intervene in decision-making. For example, in the fight against infectious diseases, lockdowns are clearly one of the most effective measures to reduce disease transmission. However, the cost of lockdowns is not merely a reduction in economic activities. In LMICs such as India with great occupational vulnerability, it could mean lives lost due to starvation[18].

The complex interaction between science and politics was highlighted during the COVID-19 pandemic. The acute and sudden outbreak of COVID-19 drew immediate attention from all sectors, including the scientific community. Within days of the pandemic onset, there was a dramatic increase of scientific publications regarding the sources, transmission routes and interventions of COVID-19, with sometimes contradictory conclusions. This was somehow anticipated by the scientific community because science, by nature, is an iterative process that evolves through correcting itself, often slowly. This iterative nature of science – particularly during early phases of a pandemic – comes with large uncertainties, and can sometimes be used by politicians to cherry-pick evidence that supports an agenda. As Stevens pointed out, in the case of disease dynamics modelling, "small changes in the assumptions made by the modellers can have large effects on their estimates and implications."[19] With multiple models present, a government may cherry-pick results that support their narrative. This raises an important question: what determines the degree to which science is followed by policymakers?

Utility of science and epidemic models in policy making

First, it should be emphasized that the robustness of the scientific evidence itself is one of the bases for its acceptance in policymaking. An example is the contribution of human activities to climate change, where its evidence has been so thoroughly investigated that a dispute would today be beyond reasonable. AMR is also an example where a growing body of evidence has shown a link between its development and the use of antimicrobials. Therefore, based on scientific evidence, interventions to cut antimicrobial use have been implemented in several

high-income countries. For example, Denmark banned the use of cephalosporin in pig production in 2010. This intervention successfully led to a decrease in the prevalence of cephalosporinase-producing *E. coli* from 11.8% to 3.6% within a year[12].

Second and perhaps of greater importance, the timescale at which decisions need to be made could affect the importance given to scientific evidence in the decision-making process. In this context, my PhD work on both COVID-19 and AMR was a privileged opportunity to observe contrasting situations in how scientific evidence is used (or not) by decision makers. For acute crisis like the COVID-19 pandemic, countries have taken largely different approaches towards pandemic control, from strict measures in New Zealand to voluntary social distancing in Sweden. In Sweden's case, the central role played by the Public Health Agency in guiding the country's response to the COVID-19 pandemic contributed largely to its comparatively relaxed approach. The variations in approaches across countries seem to have been primarily political or cultural, rather than scientific since all countries had access to the same scientific evidence but nevertheless followed different policies. Furthermore, in the case of Switzerland, the scientific evidence that my colleagues and I contributed to generate – upon request from authorities – was not always followed. Sometimes this came with negative consequences such as the second wave of the COVID-19 epidemic in September and October 2020, which was characterized by high excess mortality[20]. In contrast, AMR is a health crisis that spans decades. Our experience suggests that scientific evidence for such longer-term challenges tends to draw greater attention from decision makers. In particular, the output of our work on AMR has been requested by international funders, such as the Fleming Fund, to inform their efforts. Two factors may contribute to the differences in the utility of scientific evidence during acute *vs* non-acute crises. First, non-acute crises often lack highly visible and immediate consequences that demand immediate action. This relative absence of immediate pressure creates an environment for open discussion, allowing science to play a more influential role in shaping decisions and policies. Second, in the context of non-acute crises, a temporal disconnect exists between decision-making and accountability. In countries with regular elections, decision makers who base their actions on scientific evidence for long-term challenges like AMR might no longer be in office when the consequences of their decisions become visible. This situation may make them more willing to embrace the risk associated with uncertain outcomes guided by scientific recommendations.

Although science may play a more prominent role in shaping policy interventions for enduring crises like AMR, the lack of data remains an important obstacle for drawing clear conclusions. For example, only 42 countries publicly report their antimicrobial use in food animals; national surveillance data of AMR in food animals are mostly only publicly available in high-income countries. While such limitations may be perceived as preventing modelling exercises to even be pursued, we argue that modelling in such situations can play a dual role in guiding policymaking. First, it provides a summary of available evidence to raise public awareness on the health crisis of AMR. For example, our maps of AMR provide a visual representation that is easier for the public to grasp the extent of its spread and impact. It can also raise awareness among the scientific community to push this research forward. Second, the uncertainties associated with AMR estimates might incentivize governments to disclose previously withheld surveillance data to challenge output from models such as those presented in this thesis. This is especially relevant when predicted AMR levels exceed the recorded surveillance data, as governments would be motivated to make the data public.

6.4 Future directions and limitations

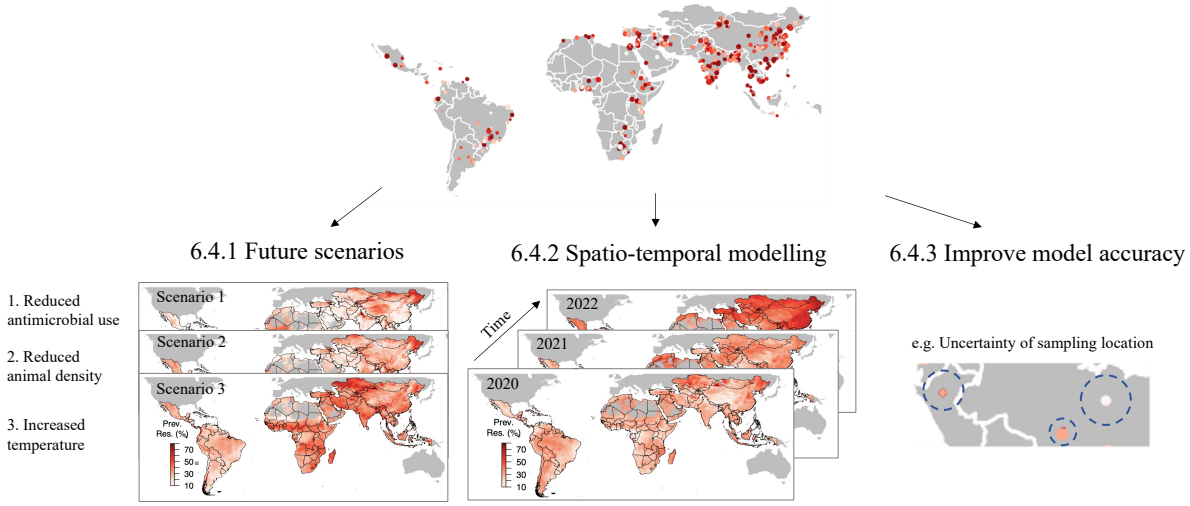


Figure 1. Directions for future research of AMR.

6.4.1 Trends of AMR under different future scenarios

In this thesis, we used species distribution modelling and Gaussian process stacked generalization to map AMR distribution. The models quantified the relationship between AMR and environmental covariates as well as anthropogenetic covariates to inform patterns of AMR prevalence. Covariates that are shown to contribute the most to the distribution of AMR include travel time to cities, population density of animals, and temperature. The predictive power of these covariates on AMR may be due to their indirect influence on antimicrobial use: locations closer to cities may have easier access to antimicrobials, thereby encouraging the overuse of antimicrobials. For example, in Thailand, Huber et al. showed that pig farms with AMR *E. coli* were located closer to drugstores, compared with farms without AMR *E. coli*[21]. High densities of animals can result in more antimicrobial use if per-animal usage is stable; higher temperature can lead to more animal injuries that require antimicrobial treatment[22]. In addition, the influence of animal density on AMR can be related to farm management practices such as poor sanitation and overcrowding among intensive farms, which contribute to AMR transmission.

Based on the relationship between covariates and AMR that were trained by our geospatial models, future scenarios with changes in the values of one or more of these covariates can be modelled to produce future trajectories of AMR prevalence. For example, the Third Global High-Level Ministerial Conference on Antimicrobial Resistance set the goal to reduce antimicrobial use in agriculture by 30% to 50% by 2030. We can analyse the potential impact of achieving this goal by projecting the corresponding reduction in AMR. Furthermore, a previous study explored the effect of rising temperature on AMR in humans[23], and such analysis can be extended to animals. By combining the reduction in antimicrobial use with climate change scenarios, we can evaluate their collective effects on the future trajectories of AMR.

In addition, trajectories of AMR may differ between high-income countries and LMICs, under the same future scenarios. This is due to differences in the effect of environmental and anthropogenic covariates on AMR. For example, the practice of intensive farming may have a greater influence on antimicrobial use in LMICs than in high-income countries. The reason

behind is that intensive farming may require higher standards of biosafety measures and veterinary services to keep animals healthy due to comparatively higher stocking densities. LMICs have less resources to meet such standards, leading to more dependence on antimicrobial use. A comparative analysis of AMR trajectories between high-income countries and LMICs can assist policymakers in fine-tuning their strategies for managing and mitigating AMR. With more data available in the future, it may also be possible to run the scenario analyses by country or antimicrobial classes, potentially generating more detailed insights on AMR containment strategies.

6.4.2 Integrating the temporal dimension into AMR prevalence mapping

The analyses presented throughout Chapter 2 to Chapter 4 were focused on the spatial dimension of AMR. Our current model formula included effects of covariates and a spatial random effect, and was trained using data collected between 2000 and 2019. Therefore, an important limitation of our approach is that by collapsing all years of data together to guarantee statistical power of the spatial interpolation, we did not consider the temporal variations of AMR. This may potentially lead to an overestimation of AMR in regions where data have been collected in more recent years, and an underestimation of AMR in regions where data were older. However, incorporating a temporal random effect into the modelling framework will significantly increase model complexity, and is currently challenged by the limited amount of data available (n=1,088). With the steady increase of PPS conducted each year[1], future work may be able to develop a spatio-temporal model using Integrated Nested Laplace Approximation with Stochastic Partial Differential Equations. Such models will not only improve the accuracy of AMR prevalence maps, but also capture temporal trends of AMR at a fine spatial resolution.

6.4.3 Capturing spatial uncertainties

The accuracy of our modelled AMR distribution is dependent, amongst other factors, upon the amount and spatial coverage of PPS. Several improvements can be made in these regards. First, as the number of PPS steadily increases each year[1], the species distribution models will be better able to learn the non-linear and complex relations between the covariates and risk of AMR. Second, a majority of PPS do not report exact coordinates where animal samples are collected, but rather the administrative unit such as cities[24], with varying degrees of accuracy in the administrative unit reported (state, district, city, village, farm, etc.). This leads to varying degrees of uncertainty of the assigned locations of the data points, which further influences the extracted covariate values at the corresponding locations. In the methods used throughout Chapters 2 and 4, we used the geographic centroid of the reported administrative unit as the location. Future efforts will need to incorporate the uncertainty of assigned locations to better assess the robustness of the resulting maps. For example, Monte Carlo simulations can be conducted over areas of the administrative units, to generate multiple maps from each simulation. The uncertainty of locations can then be estimated through variations of the mapped values across the simulations.

Covariates used to train the model also play a pivotal role for the accuracy of the results. Among the covariates used for predicting AMR, antimicrobial use is expected to have the highest importance. However, results throughout Chapters 2 and 4 showed that this covariate was not identified as contributing significantly to the modelled outputs. One important underlying reason is the lack of reported data on antimicrobial use in food animals. In 2020, only 42 countries reported antimicrobial use in food animals[25]. Mulchandani et al. extrapolated these

data to other countries and produced maps of antimicrobial use, which we used in this thesis to predict AMR[25]. However, this limited amount of data inevitably results in uncertainty in the mapped antimicrobial use. Data from more countries are needed to better characterize the relation between antimicrobial use and AMR at a global scale. In addition, the accuracy of our species distribution models for mapping AMR is moderate, compared with higher prediction accuracy for vector-borne diseases that are more explicitly linked to the environmental conditions: the mean area under the curve (AUC) statistic calculated using cross validation was 0.66 in Chapter 2.1, as compared with 0.82 for a map of Zika virus risk[26] and 0.72 for a map of dengue risk[26]. This indicates that future work may need to include additional socio-economic covariates with potentially greater impact on AMR, and at a finer spatial resolution.

Apart from covariates, our models also include spatial random effects in a Bayesian framework[27]. The spatial effect results in increased similarity of AMR levels in locations that are closed to each other. Consequently, the bigger the spatial effect, the smoother the map tends to be. It is therefore essential that this term does not over-smooth the map and mask finer-scale patterns, and that the priors of the covariance function are properly set based on existing evidence. Future work will need to conduct sensitivity analysis on how prior knowledge of the spatial random effect changes the resulting spatial patterns of AMR.

Conclusions

This thesis lays the foundation for estimating trends of AMR in food animals worldwide, serving as a starting point for its ongoing and iterative refinements in the future. A successful precedent for such iterative improvements on global disease mapping was set by the Malaria Atlas Project (MAP). Since its reception nearly two decades ago, the project has produced multiple updated global maps of malaria risk. In just seven years, MAP achieved collaboration with the World Health Organization, and it continues to expand its reach with the addition of new partner countries with endemic malaria. Our AMR mapping project started in 2019, with the production of the first global map of AMR in food animals in LMICs. Given the ongoing rise in meat consumption in LMICs and the looming threat of AMR to both animal and human health, we hope to see our AMR mapping projects grow and assist decision makers to safeguard the effectiveness of life-saving antimicrobials.

6.5 References

1. Van Boeckel, T. P. *et al*. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, eaaw1944 (2019).
2. World Health Organization. Global antimicrobial resistance and use surveillance system (GLASS) report: 2021. (2021).
3. Murray, C. J. *et al*. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).
4. Hendriksen, R. S. *et al*. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
5. Liguori, K. *et al*. Antimicrobial resistance monitoring of water environments: a framework for standardized methods and quality control. *Environ. Sci. Technol.* **56**, 9149–9160 (2022).
6. Góchez, D. *et al*. OIE annual report on antimicrobial agents intended for use in animals: methods used. *Front. Vet. Sci*. 317 (2019).
7. Rushton, J. *et al*. Roll-out of the Global Burden of Animal Diseases programme. *The Lancet* **397**, 1045–1046 (2021).
8. Cooper, M. A. & Shlaes, D. Fix the antibiotics pipeline. *Nature* **472**, 32–32 (2011).
9. Woolhouse, M. & Farrar, J. Policy: an intergovernmental panel on antimicrobial resistance. *Nature* **509**, 555–557 (2014).
10. Leal, J. R., Conly, J., Henderson, E. A. & Manns, B. J. How externalities impact an evaluation of strategies to prevent antimicrobial resistance in health care organizations. *Antimicrob. Resist. Infect. Control* **6**, 1–11 (2017).
11. Kirchhelle, C. Pharming animals: a global history of antibiotics in food production (1935–2017). *Palgrave Commun.* **4**, (2018).
12. Agersø, Y. & Aarestrup, F. M. Voluntary ban on cephalosporin use in Danish pig production has effectively reduced extended-spectrum cephalosporinase-producing Escherichia coli in slaughter pigs. *J. Antimicrob. Chemother.* **68**, 569–572 (2013).
13. Dutil, L. *et al*. Ceftiofur resistance in Salmonella enterica serovar Heidelberg from chicken meat and humans, Canada. *Emerg. Infect. Dis.* **16**, 48 (2010).
14. Huber, L., Agunos, A., Gow, S. P., Carson, C. A. & Van Boeckel, T. P. Reduction in antimicrobial use and resistance to Salmonella, Campylobacter, and Escherichia coli in broiler chickens, Canada, 2013–2019. *Emerg. Infect. Dis.* **27**, 2434 (2021).
15. Food and Agriculture Organization. Crops and livestock products (Production).
16. The Fleming Fund. Taking action against drug resistance for a healthier world. (2019).
17. Bornman, J. F., Barnes, P. W. & Pandey, K. Environmental effects of stratospheric ozone depletion, UV radiation, and interactions with climate change: 2022 Quadrennial Assessment. *Photochem. Photobiol. Sci*. 1–2 (2023).
18. Ray, D. & Subramanian, S. India's lockdown: An interim report. in *The Impact of COVID-19 on India and the Global Order: A Multidisciplinary Approach* 11–61 (Springer, 2022).
19. Stevens, A. Governments cannot just 'follow the science'on COVID-19. *Nat. Hum. Behav.* **4**, 560–560 (2020).
20. Riou, J. *et al*. Direct and indirect effects of the COVID-19 pandemic on mortality in Switzerland. *Nat. Commun.* **14**, 90 (2023).
21. Huber, L. *et al*. Geographic drivers of antimicrobial use and resistance in pigs in Khon Kaen Province, Thailand. *Front. Vet. Sci.* **8**, 659051 (2021).

22. Diana, A., Manzanilla, E. G., Calderón Díaz, J. A., Leonard, F. C. & Boyle, L. A. Do weaner pigs need in-feed antibiotics to ensure good health and welfare? *PLoS One* **12**, e0185622 (2017).
23. MacFadden, D. R., McGough, S. F., Fisman, D., Santillana, M. & Brownstein, J. S. Antibiotic resistance increases with local temperature. *Nat. Clim. Change* **8**, 510–514 (2018).
24. Criscuolo, N. G., Pires, J., Zhao, C. & Van Boeckel, T. P. resistancebank. org, an open-access repository for surveys of antimicrobial resistance in animals. *Sci. Data* **8**, 189 (2021).
25. Mulchandani, R., Wang, Y., Gilbert, M. & Van Boeckel, T. P. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. *PLOS Glob. Public Health* **3**, e0001305 (2023).
26. Messina, J. P. *et al*. The current and future global distribution and population at risk of dengue. *Nat. Microbiol*. **4**, 1508–1515 (2019).
27. Lindgren, F. & Rue, H. Bayesian spatial modelling with R-INLA. *J. Stat. Softw*. **63**, (2015).

Supplementary Information

nature portfolio

https://doi.org/10.1038/s43016-021-00320-x

Supplementary information

Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China

In the format provided by the authors and unedited

This PDF file includes:

Supplementary Text S1-S5

Supplementary Figures 1-16

Supplementary Tables 1-3

Supplementary Text

S1. Literature Review and Data Harmonization

We searched for veterinary literature reporting antimicrobial resistance (AMR) rates in China on three international databases (PubMed, Scopus, and ISI Web of Science), as well as the leading Chinese-language academic search engine – China's National Knowledge Infrastructure (CNKI). We focused on four common indicator pathogens: *Escherichia coli*, nontyphoidal *Salmonella* spp., *Staphylococcus aureus*, and *Campylobacter* spp. The search was conducted on March 11, 2020, and included all studies published between 2000 and 2019. The search query used the generic structure:

(Resistance) AND (Bacterial Species) AND (Animals and Sample types)

The key words used on CNKI was as follows, with all possible combinations of drugs, pathogens, and animals: ('抗生素' + '抗菌' + '兽药' + '兽用药' + '兽用抗生素' + '用药' + '抗微生物') AND TI = (escherichia + (E*coli) + coliform + salmonella + enterococcus + enterococci + VRE + (E*faecalis) + (E*faecium) + (S*aureus) + staphylococcus + MRSA + MSSA + campylobacter + (C*jejuni) + (C*coli) + '埃希菌' + '大肠杆菌' + '大肠菌' + '沙门氏菌' + '肠球菌' + '葡萄球菌' + '弯曲杆菌' + '弯曲菌' + '曲状杆菌') AND TI = ('动物' + '食物' + '牛' + '鸡' + '猪' + '肉' + '奶' + '蛋' + '食用'). The full search query used in the advanced search functionality was: (TI = key

words) OR (KY = key words) OR (AB = key words), where TI, KY, and AB stand for title, keywords, and abstract, respectively.

The key words used on PubMed, Scopus, and ISI Web of Science was as follows: (antibiotic resistance OR antimicrobial resistance OR resistance OR susceptibility OR antibiogram OR antibiotic susceptibility testing OR antibiotic OR antimicrobial OR antibacterial) AND (Escherichia OR E. coli OR coliform OR salmonella OR salmonella spp. OR enterococcus OR enterococcus spp. OR enterococci OR VRE OR E. faecalis OR E. faecium OR S. aureus OR staphylococcus OR Staphylococcus spp. OR MRSA OR MSSA OR campylobacter OR campylobacter spp. OR C. jejuni OR C. coli) AND (animal OR food OR food producing OR farm OR farm animal OR meat OR cow OR cattle OR beef OR bovine OR buffalo OR pig OR piggeries OR pork OR chicken OR flock OR broiler OR layer OR egg OR poultry OR avian OR milk OR dairy OR cheese) AND (China). On PubMed, the key words were directly used as the full search query. On Scopus, the full search query was specified as TS = (key words), where TS stands for search topic. On ISI Web of Science, the full search query was specified as TITLE-ABS-KEY = (key words), TITLE-ABS-KEY stands for title, abstract and key words.

From CNKI, the literature search resulted in 8,481 documents in Chinese being identified. Abstracts were screened manually. We removed reviews, meta-analysis, and publications that did not report resistance rates such as experiments on antimicrobial effect of traditional Chinese medicine, and retained 1,080 publications as potentially relevant point prevalence surveys (PPS) to read in full. Among these, we excluded:

manuscripts characterizing a defined set of strains not derived from PPS (strain surveys), surveys on diseased animals, surveys in which samples were pooled between host species, or resistance rates were pooled between pathogens, and studies without information on sampling locations. We further excluded PPS focusing on animals with small sample sizes, such as sheep and camel. Finally, 4,023 antimicrobial resistance rates were extracted from 318 PPS for drug-pathogens combinations recommended for susceptibility testing by the WHO AGISAR consortium [1]. Amongst the publications published in Chinese, we listed the institutions most frequently associated with the publications, as well as their geographic location (Table S2).

From PubMed, Scopus, and ISI Web of Science, 101 PPS conducted in China between 2000 and 2018 were extracted by Van Boeckel & Pires *et al* [2], including 1,817 antimicrobial resistance rates. An additional search for PPS in China published in 2019 was conducted on March 11, 2020. The literature search identified 1,136 documents published in 2019. Abstracts were screened manually. We retained 91 publications as potentially relevant point prevalence surveys (PPSs), and read in full. Finally, 455 antimicrobial resistance rates were extracted from 27 PPS published in 2019. In total, 2,272 resistance rates were extracted from 128 PPS published in English from 2000 to 2019, and were combined with 4,023 resistance rates extracted from 318 PPS published in Chinese, to build a consolidated database of 6,295 resistance rates that was used for geospatial mapping. The average prevalence of *E. coli*, *Salmonella*, *S. aureus* and *Campylobacter* in samples of food animals across all PPS were 60%, 23%, 22% and 27%, respectively.

We associated clinical breakpoint values with each resistance rates reported in the surveys. Among the 446 PPS, only 126 reported breakpoint values directly in the publication. However, 382 PPS mentioned guidelines used for susceptibility testing. The overwhelming majority (369; 83%) of PPSs conducted in China used guidelines from Clinical & Laboratory Standards Institute (CLSI) for susceptibility testing. Therefore, for surveys that did not mention the guidelines that were used, it was assumed to be CLSI. Among the PPS that reported which guidelines were used, 291 also mentioned the year of the guideline. The average lag between the year of guideline used, and the publication date of a survey was 4 years. Therefore, for surveys that did not report the year of the guidelines, we used the CLSI guidelines 4 years prior to the publication date of the surveys. After a set of guidelines were associated to each surveys, we used a methodology [2] to harmonize resistance rates. This approach accounts for differences in the types of susceptibility testing method used, as well as potential temporal variation in breakpoints values.

Following the harmonization of resistance rates, we summarized resistance levels in each PPS by calculating the proportion of antimicrobial compounds with resistance higher than 50% (P50). The mean P50 across all PPSs was 39.8%, and the standard deviation of the P50 distribution was ±29.5%. We then conducted a robustness check to ensure that the number of surveys for which data were extracted (n=446) would allow to make inference on the mean value of P50 across China. Concretely, we calculated the

minimum number of PPSs needed to obtain a confidence interval within ±5% of the estimated mean P50. We used a non-parametric Monte Carlo sampling scheme:

Step 1. We generated values of P50 for 78,181,840 hypothetical farms, the total number of farms in 2018 for chicken (30,170,356), pigs (37,746,624) and cattle (10,264,860) in China. The number of farms was obtained from the China Animal Husbandry and Veterinary Yearbook 2018[3]. In the next steps, this distribution of P50 values was called the farm distribution.

Step 2. We drew n values from the farm distribution, and calculated the mean P50 value. We repeated the operation 10,000 times. These distributions of P50 values were called the survey distribution.

Step 3. We calculated the 95% confidence interval on the survey distribution of n P50 values.

Step 2 and step 3 were repeated iteratively, and n increased by +1 at each iteration until the confidence internal on the mean P50 of the survey distribution was within ±5% of the estimated mean P50 (39.8%). The resulting minimum number of surveys that satisfied this condition was n=133 PPS. The dataset assembled amounted to 446 surveys and thus exceed the minimal number of surveys to make inferences of mean P50 value within the ±5% confidence interval.

S2. Geospatial Modelling

We interpolated P50 values from the survey locations to create a map of P50 at a resolution of 10 x 10 Km across China. The approach followed a two-step procedure [4]. In the first step, three "child models" were trained to quantify the relation between P50 and environmental and anthropogenic covariates (Table S1; Supplementary Figure 13). In the second step, we used universal kriging to stack predictions from the three child models. Stacking has been shown to improve accuracy of disease maps [5]. The second step simultaneously account for the influence of covariates on P50 and the spatial autocorrelation of P50.

Step 1. Training child models

We identified 10 covariates such as antimicrobial use in animals and travel time to cities (Table S1; Supplementary Figure 13) to train child models. All covariates were log10-transformed and resampled from their original resolution to 0.0833 decimal degrees (10 kilometers at the equator). Three child models used were: boosted regression trees [6] (BRT), least absolute shrinkage and selection operator applied to logistic regression [7] (LASSO-GLM), and feed-forward neural network implemented in Keras [8] (FFNN).

For the BRT model, we used a tree complexity of three, learning rate of 0.0005 and 50 initial trees. For feed-forward neural network, we used the adaptive moment estimation (adam) optimizer, and rectified linear unit (relu) activation for the hidden layer, and a sigmoid activation function for the output layer. We used random search to optimize the hyperparameters for the neural network. Concretely, we sampled 1,000 combinations of

dropout rate, learning rate, and the number of neurons. The combination of hyperparameters that generated the highest area under the ROC curve (AUC) was retained to train the feed-forward neural network. The final model had one hidden layer of 13 neurons, a dropout rate of 0.5 and a learning rate of 0.07.

P50 values were transformed into presence/absence as input into all child models, using a random binarization procedure: All P50 values were replicated five times, and compared with one random number between zero and one. If the P50 value was higher than the random number, then it was categorized as presence; alternatively, it was categorized as absence. Additionally, pseudo-absences (P50 = 0) were sampled within 20 to 1,000 km from presences, using stratified random sampling proportional to the log10 of the population density outside urban areas. This provided the models additional covariate values that were not associated with presences. Four-fold spatial cross validation was used to train all child models. Concretely, we defined four folds using k-means clustering on the latitudes and longitudes of sampling locations (Supplementary Figure 14). Spatial sorting bias (SSB) was negligible (mean SSB = 0.9941). Each model was run four times, each time using data from one fold as testing set, and data from the other three folds as training set. Spatial cross-validation repented local overfitting, and ensured extrapolation accuracy outside training regions. Stratified sampling of pseudoabsences and random binarization of P50 were bootstrapped 10 times using Monte Carlo simulations. The average over these 10 simulations was the final predictions of each child model (Supplementary Figure 15). The accuracies of all child models were: $AUC_{BRT} = 0.65$, $AUC_{LASSO\text{-}GLM} = 0.68$, $AUC_{FFNN} = 0.66$.

163

164 *Step 2. Universal kriging*

165 In the second step, we used universal kriging to stack predictions from three child

166 models:

167

168 $$\boldsymbol{Z}(\boldsymbol{x}) = \boldsymbol{\mu}(P_{BRT}(x), P_{LASSO-GLM}(x), P_{FFNN}(x)) + \boldsymbol{Y}(\boldsymbol{x})$$

169

170 Where $\boldsymbol{Z}(\boldsymbol{x})$ is the predicted P50 value at location x;

171 $\boldsymbol{\mu}(P_{BRT}(x), P_{LASSO-GLM}(x), P_{FFNN}(x))$ are a linear combination of the predictions of the

172 child models; $\boldsymbol{Y}(\boldsymbol{x})$ is the model residual, which is spatially autocorrelated. A Matern

173 semi-variogram model with a maximum range of 800 km was fitted. The spatial weight

174 of each survey in the kriging prediction was proportional to the number of biological

175 samples recovered in each survey. Duplicated coordinates, due to multiple surveys

176 conducted at the same location, were randomly redistributed within a radius of 1km of

177 the sampling sites, multiplied by log10 transformed sample size to reflect greater spatial

178 range of large surveys. Predictions below 0 or above 1 were replaced with 0 and 1,

179 respectively, such that all kriging predictions were bounded within that interval.

180

181 *Uncertainty of P50 predictions*

182 We produced a map of uncertainty associated with the spatial predictions of P50. This

183 map consisted of two sources of variability and was given by:

184

185 $$Var_{total} = Var(P_{BRT}, P_{LASSO-GLM}, P_{FFNN}) + Var_K$$

Where, $Var(P_{BRT}, P_{LASSO-GLM}, P_{FFNN})$ was the variance of the predictions of each child model across 10 Monte Carlo simulations, and Var_K was the kriging variance of the interpolation procedure. Var_K was calculated as follows: first, the kriging variance was obtained from an ordinary kriging on the residuals of the stacked child model predictions, in order to subtract the part of variance attributed to estimating the trend in universal kriging, and to obtain the variance attributable to the spatial interpolation procedure. Second, the kriging variance was standardized such that it equaled zero at the locations of observation. The percentage of Var_{total} attributable to Var_K was 80.3% across all pixels.

S3. Economic factors associated with antimicrobial resistance rates

We estimated the association between resistance rates and the ease of obtaining antimicrobials from the market. The latter is estimated using the price and market availability of individual antimicrobials. We used price estimates from online stores as a general proxy for antimicrobial prices., because drug prices from offline stores were challenging to estimate. Besides, the co-existence of these two sales channels in the same market suggests that their prices could not fundamentally differ. We searched for prices of products of each antimicrobial class on the Alibaba platform (1688.com; the major e-wholesale market in China) on 10 December 2019. For each antimicrobial, we extracted the prices (CNY/kg) from the three online retailers with the highest sales volume in the previous month (Yuan/kg), and the highest number of online retailers (Table S3). These two variables were used as proxies for price, and market availability in China. We fitted a

linear mixed-effects model to resistance rates recorded in all PPS, using price and market availability as covariates, and PPS as random effect.

Antimicrobials with lower price and higher market availability were associated with higher resistance rates ($p < 0.05$). A decrease of 112 CHY/Kg (approximately 16 USD/Kg) in price, and an increase of 19 online retailers selling the antimicrobials were associated with an increase of 1% resistance.

S4. Existence of convenience sampling for PPS?

We tested whether veterinary institutes conducted surveys on AMR following convenience sampling, i.e. the tendency to collect samples in nearby location instead of following a stratified random sampling proportional to animal population densities. Concretely, we extracted locations of the institutes associated with the last authors of the 437 surveys. First, we calculated the distance between each sampling location and the institute conducting the survey. Second, we tested a scenario of sampling proportional to animal densities, where 437 locations were sampled weighted by the population corrected units (PCU) of food animals. We assigned each of the 437 locations to the closest institute, and calculated the distances between each sampling location and institute. We then compared the distances between the locations and the institutes under the assumed scenario of sampling proportional to animal density, with the distances between the actual sampling locations and the institutes, using a t-test. The sampling procedure for the stratified random sampling was repeated 50 times using Monte Carlo simulations for locations.

The distances identified by stratified random sampling were significantly larger (+84 km; +78%; t-test < 0.05) than the actual distances (107 km) between the actual sampling locations and veterinary institutes who led the surveys (Supplementary Figure 16). This suggests that a large number of surveys have been conducted by convenience sampling.

S5. Identifying (optimal) locations for future surveys on AMR

We predicted the locations of 50 hypothetical new surveys, such that the sum of uncertainty in trends in AMR for all pixels across China was minimized. Uncertainty levels were represented by a map of "necessity for additional surveillance" (NS), defined as:

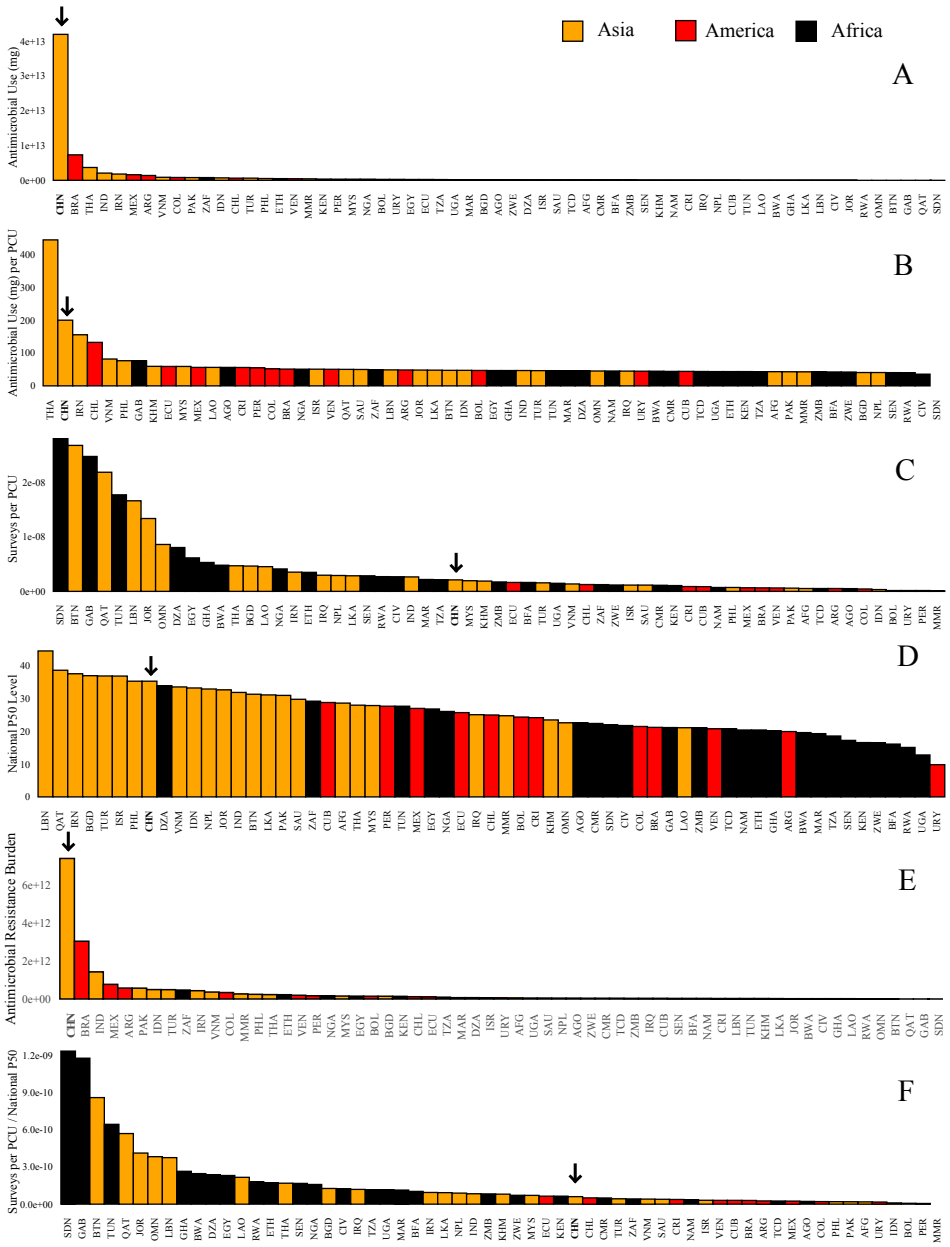
$$NS = Var_K \cdot W$$

Where Var_K was the kriging variance of the geospatial model which reflected the uncertainty of the spatial interpolation, and W is log10 transformed population density of animals [9] or humans [10] which reflected exposure. The population density of animals used here was the sum of population corrected units of chicken, pigs, and cattle, calculated using the method described in Van Boeckel and colleagues [11].

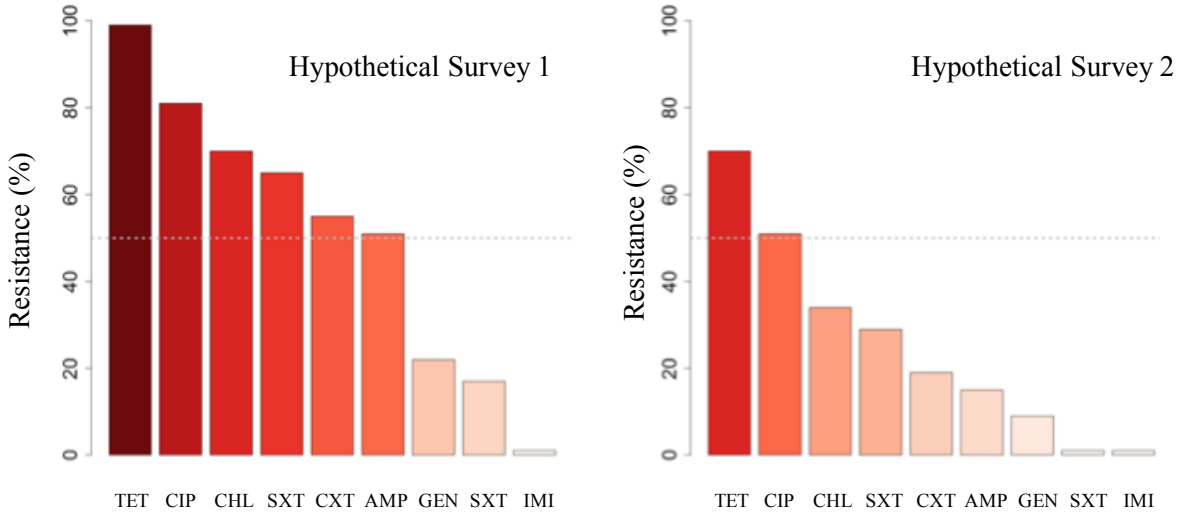
We used four approaches to identify optimal locations for future surveys across China (Materials and Methods). For the efficient implementations of these approaches, the resolution of the map of NS was increased to 40 x 40 km. This resulted in a map of 20,976 pixels (114 x 184), including 7,857 pixels with positive NS values.

For the "overlap approach" (Materials and Methods; Supplementary Figure 12), the size of a new survey's neighborhood area was optimized using approximate Bayesian computation based on sequential Monte Carlo [12] (ABC-SMC). Concretely, the neighborhood area was considered a circle with radius *r* (km). Priors of *r* were set as a uniform distribution between 50 km and 150 km. *r* was sampled 50 times per step. The average NS across 7,857 pixels was calculated for each sampled *r* value, and its medium in each step was set as the upper limit for the acceptance criteria in the following step. The optimization procedure proceeded until the convergence tolerance was met, i.e., the decrease in average NS in the last step was less than 0.5% of the total decrease in average NS after all previous steps, or until a maximum of 10 steps were run. The radius size (*r*) of 114.1 km yielded an optimal reduction in average NS, if human population was considered to determine exposure; The radius size (*r*) of 161.4 km yielded an optimal reduction in average NS, if animal population was considered to determine exposure.

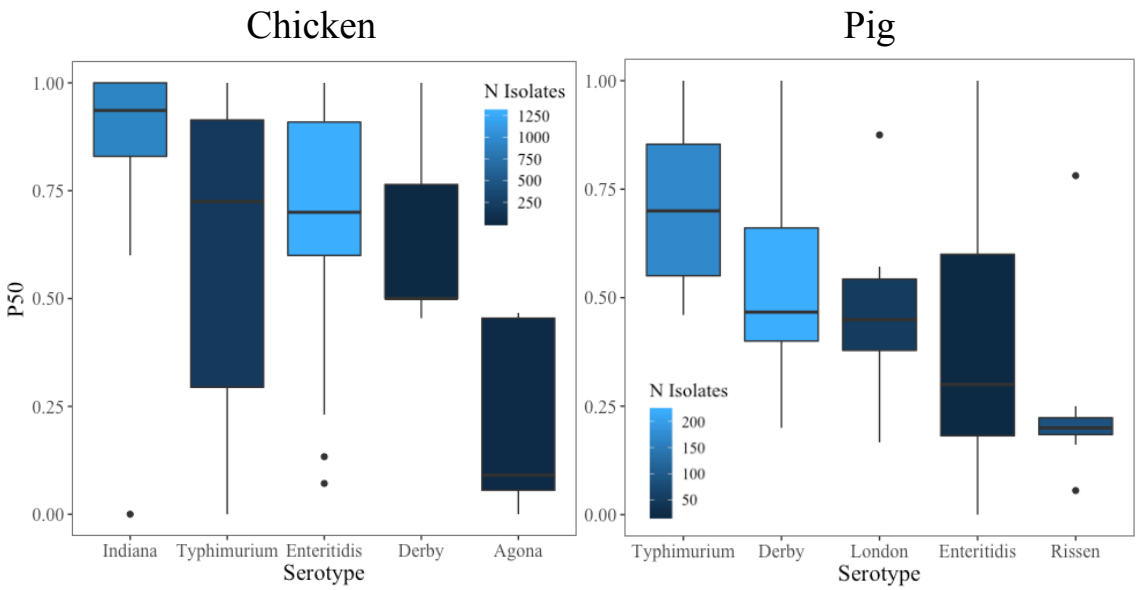
Supplementary Figures



Supplementary Figure 1. Country-level antimicrobial use (A), antimicrobial use per kilogram of food animal, number of available surveys per kilogram of food animal (C), antimicrobial resistance level measured using metric "P50" (D), antimicrobial resistance burden calculated as the product of P50 and total weight of food animals (E), and number of available surveys per kilogram of food animal per P50 (F). Data was extracted from Van Boeckel et al 2017[11], Tiseo et al 2020[13], and Van Boeckel & Pires et al 2019[2].



Supplementary Figure 2. Example of the calculation of P50, the proportion of antimicrobial compounds with resistance higher than 50%. The two panels show resistance rates of 9 antimicrobial compounds in two hypothetical surveys. The dashed line indicates the resistance rate of 50%. In "hypothetical survey 1", 6 out of 9 antimicrobial compounds have resistance rates higher than 50%, resulting in P50 = 6/9 = 67%; in "hypothetical survey 2", 2 out of 9 antimicrobial compounds have resistance rates higher than 50%, resulting in P50 = 22%.



285

286 **Supplementary Figure 3.** Proportion of antimicrobials with resistance rates higher than

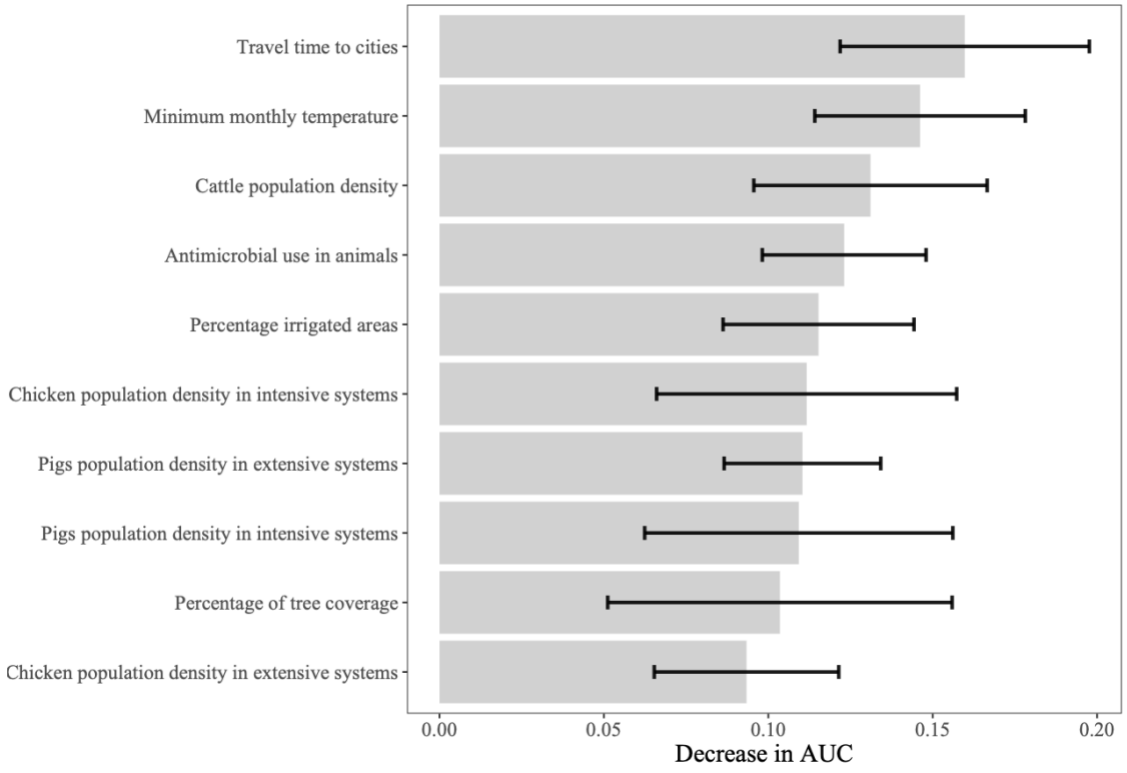
287 50% (P50) in *Salmonella* serotypes. For P50 in chicken and pigs separately, 5 serotypes

288 with the highest number of resistance rates recorded from all available surveys were

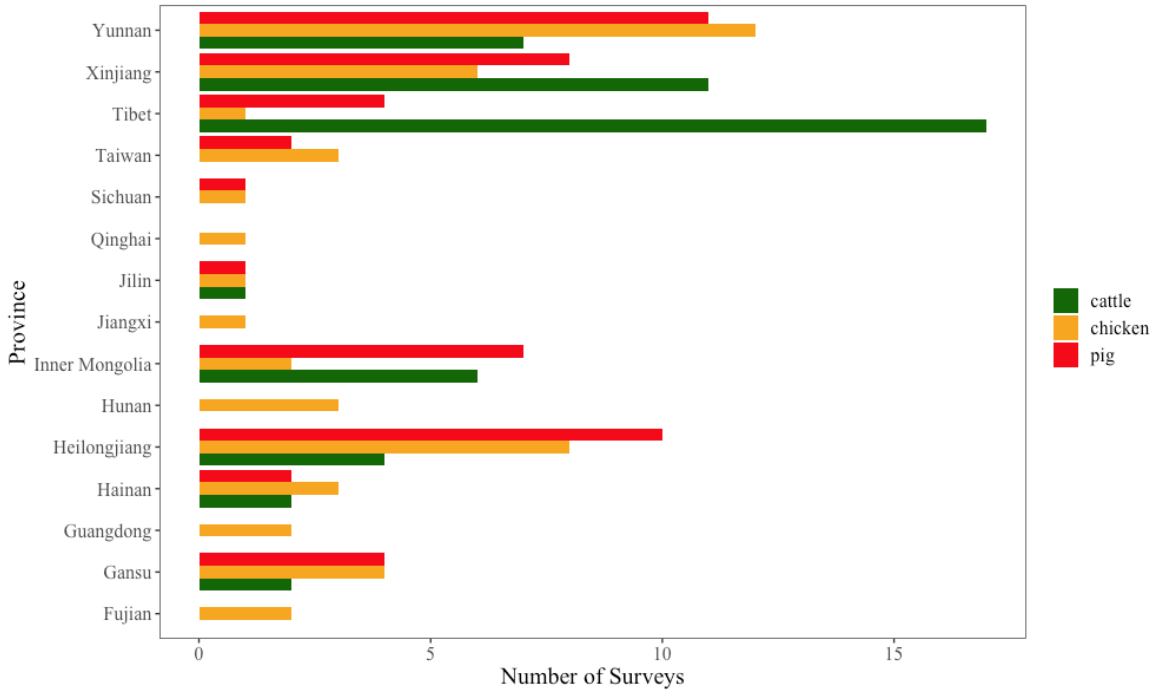
289 shown. Colors represent the number of isolates used to calculate the P50.

290

291



Supplementary Figure 4. Importance of covariates in the prediction of P50 was assessed using the decrease in AUC by sequential permutation of each covariate. The covariate with the largest predictive power, and thus associated with the largest decrease in AUC was the travel time to cities of more than 50,000 people. Minimum monthly temperature and antimicrobial use also showed moderated importance.



300

301 **Supplementary Figure 5.** Number of surveys assigned in each province in China, using

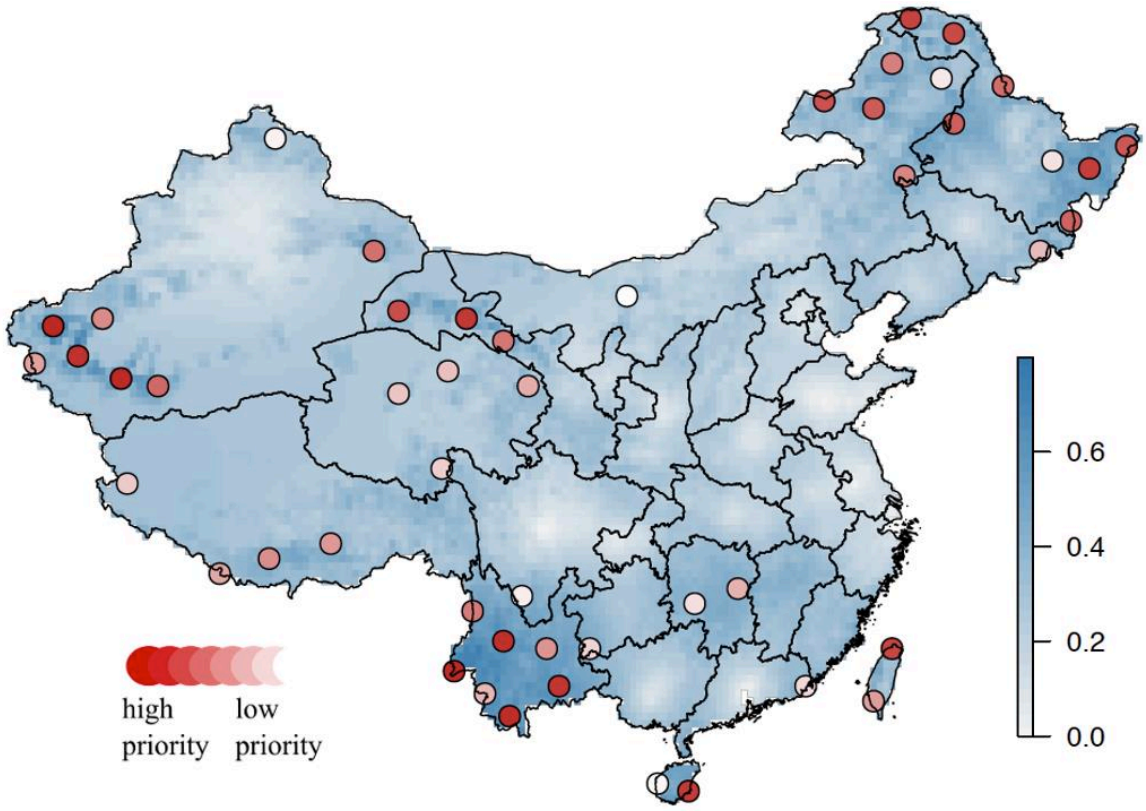
302 the overlap approach weighted by the population corrected units of cattle, chicken, and

303 pigs, respectively.

304

305

306

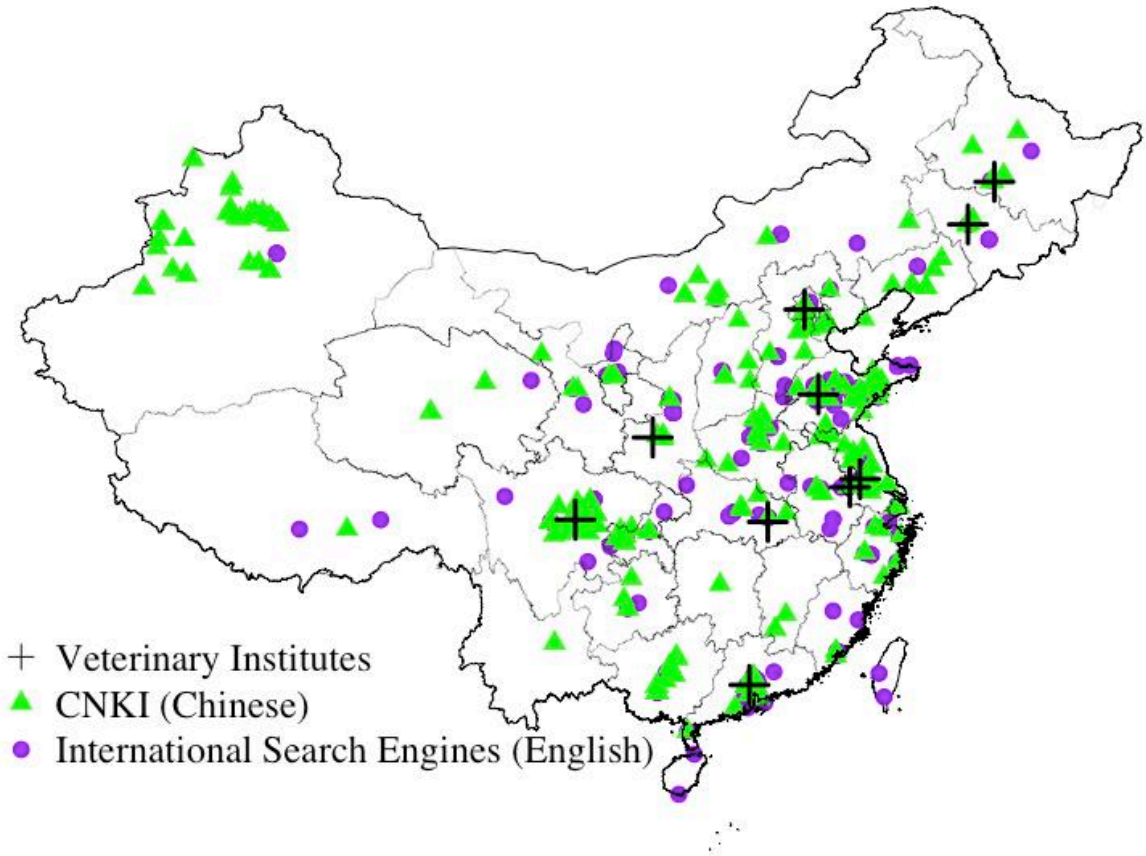


307

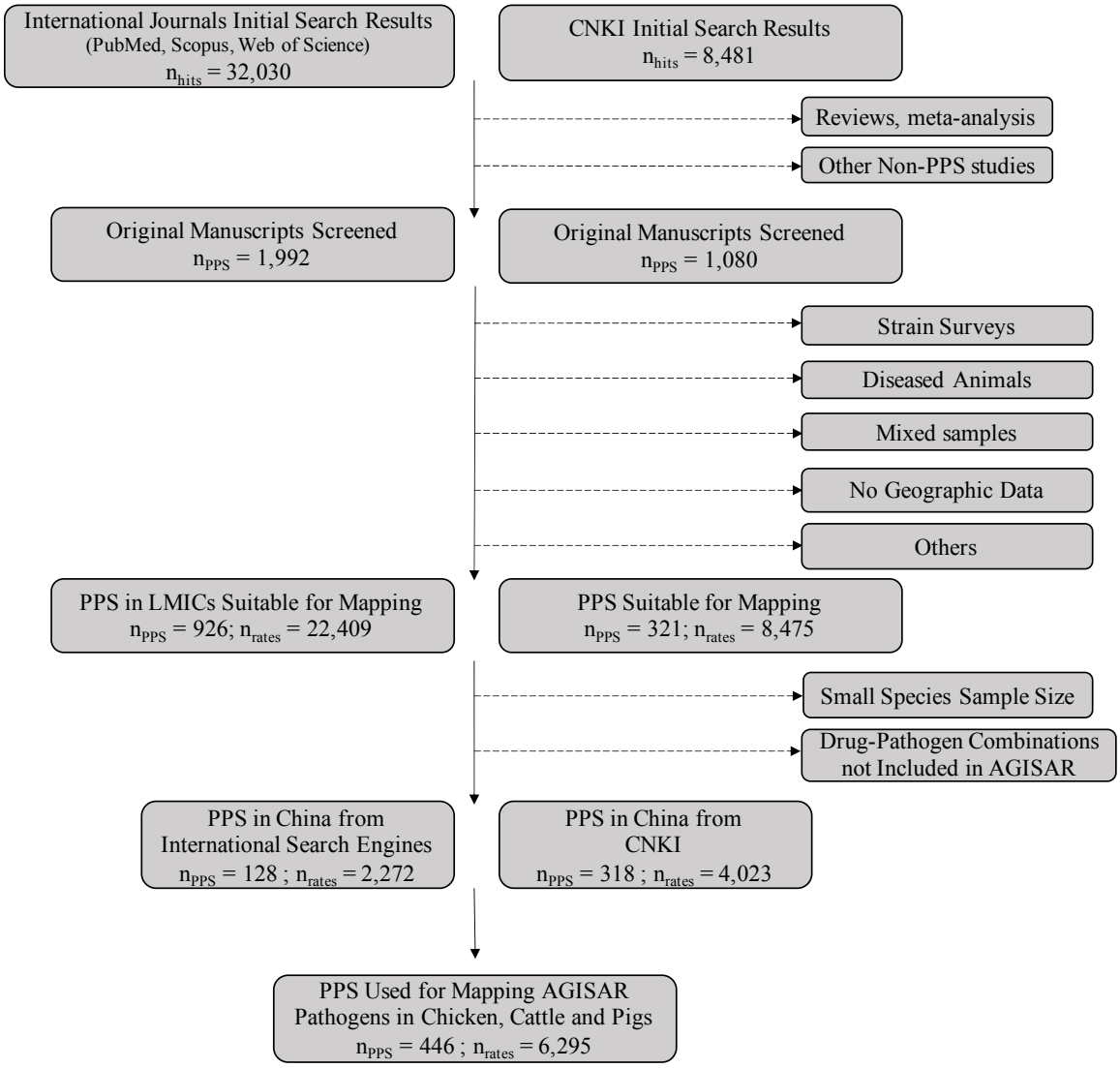
308

309 **Supplementary Figure 6.** Predicted optimal locations for future surveys using the
310 "overlap approach" weighted by human population density. The background color
311 represents the "necessity for additional surveillance" (NS): the product of the kriging
312 variance and animal population density (standardized from 0 to 1).

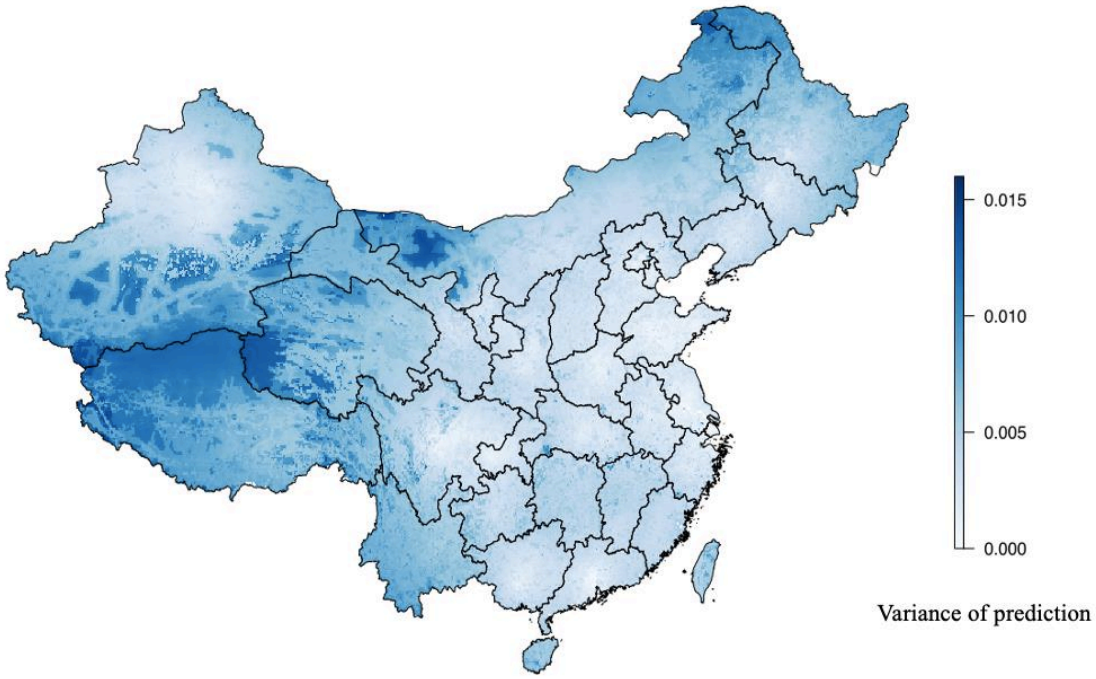
313



Supplementary Figure 7. Locations of surveys. PPS extracted from China National Knowledge Infrastructure (CNKI; red); Point prevalence surveys (PPS) obtained from [2] (blue), with extended literature review to include all PPS published in 2019 (*SI Appendix S1*). **Locations of 10 veterinary institutes** (green) most frequently associated with studies retrieved from CNKI search engine with the keywords "animal + antimicrobial + resistance" ("动物 抗生素 耐药性") (*SI Appendix S1*; Table S2).



Supplementary Figure 8. Procedure of Literature Review. Number of resistance rates (n_{rates}), and point-prevalence surveys (n_{PPS}) identified, exclusion criteria, and records used for mapping antimicrobial resistance. PPS in China extracted from international search engines were a subset of the collection in Van Boeckel & Pires et al 2019[2]. AGISAR = Advisory Group on Integrated Surveillance of Antimicrobial Resistance.

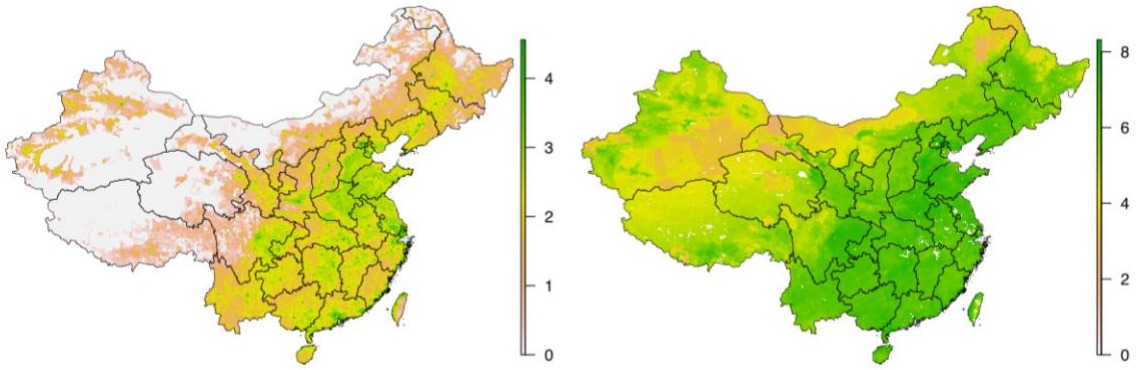


330

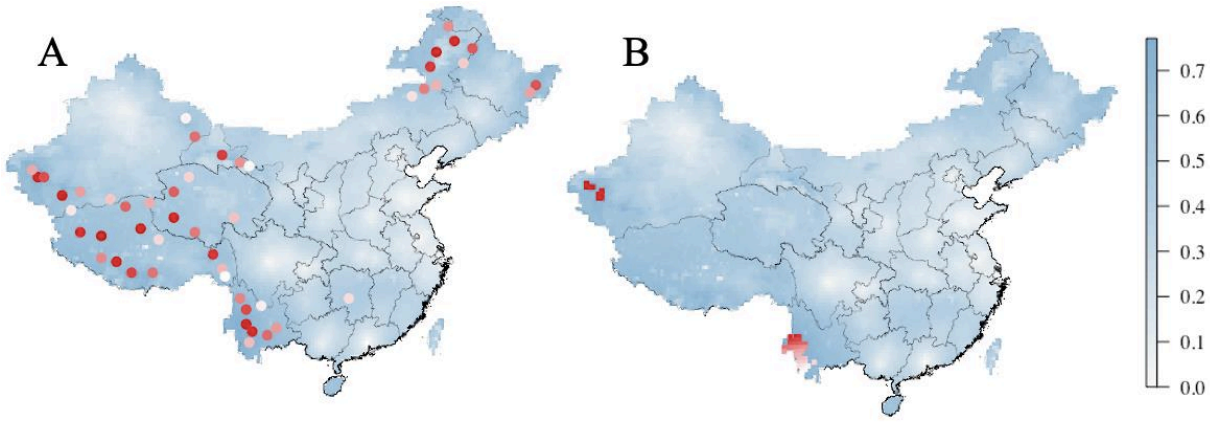
331 **Supplementary Figure 9.** Uncertainty in P50 predictions. Variance in the predicted P50

332 values across China.

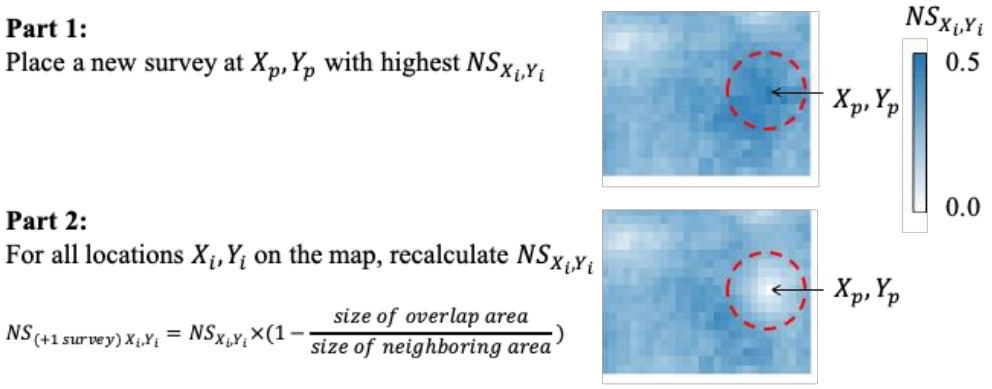
333



Supplementary Figure 10. Geographic distribution of population density of humans (left) and animals (right), log10 transformed.



Supplementary Figure 11. (A) Locations of 50 future surveys as predicted by the greedy approach (Materials and Methods), if animal population was considered to determine exposure. (B) The 50 pixels with the 50 highest values of NS (necessity for surveillance; Materials and Methods).



344

345

346 **Supplementary Figure 12.** Procedure of the overlap approach to identify optimal

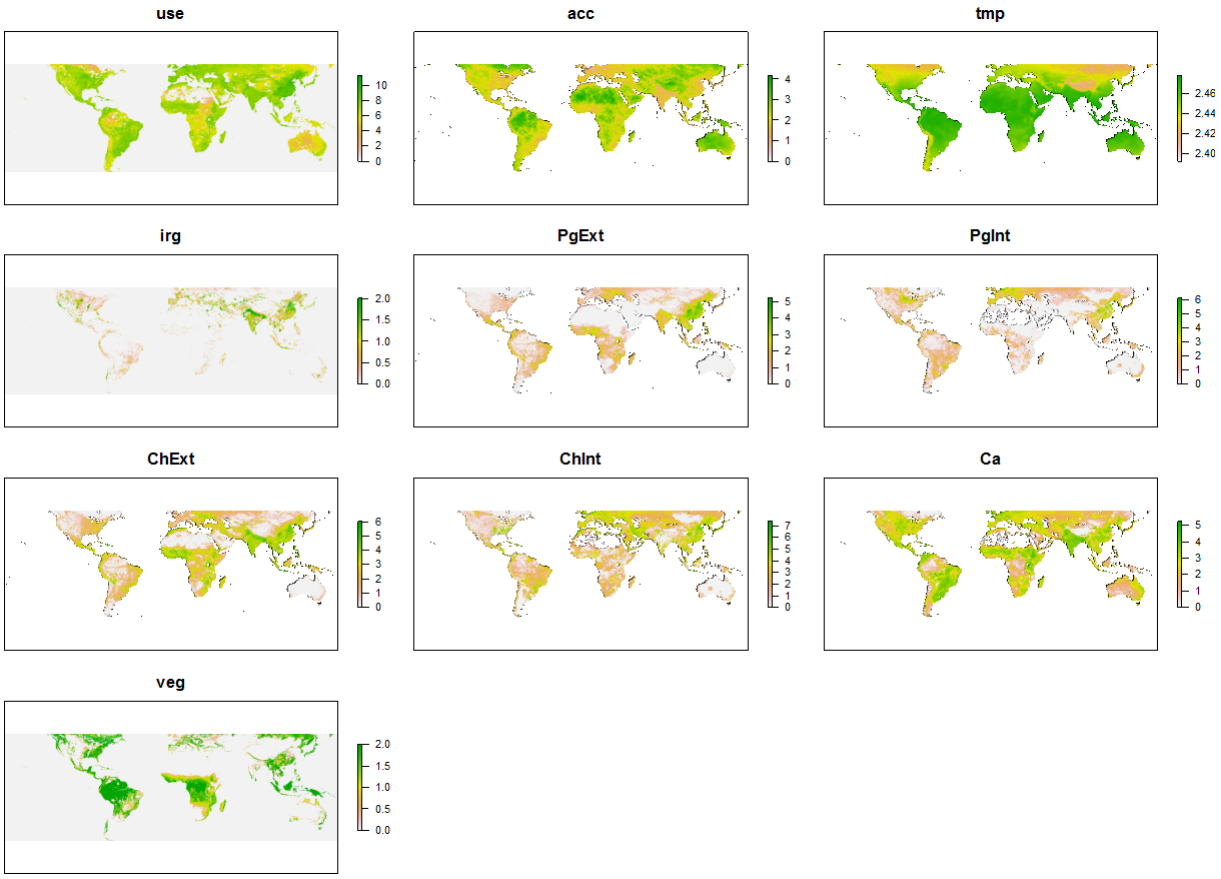
347 locations for 50 hypothetical additional surveys. Red circle indicates the neighborhood

348 area of a new survey located at X_p, Y_p. NS ("necessity for additional surveillance") is

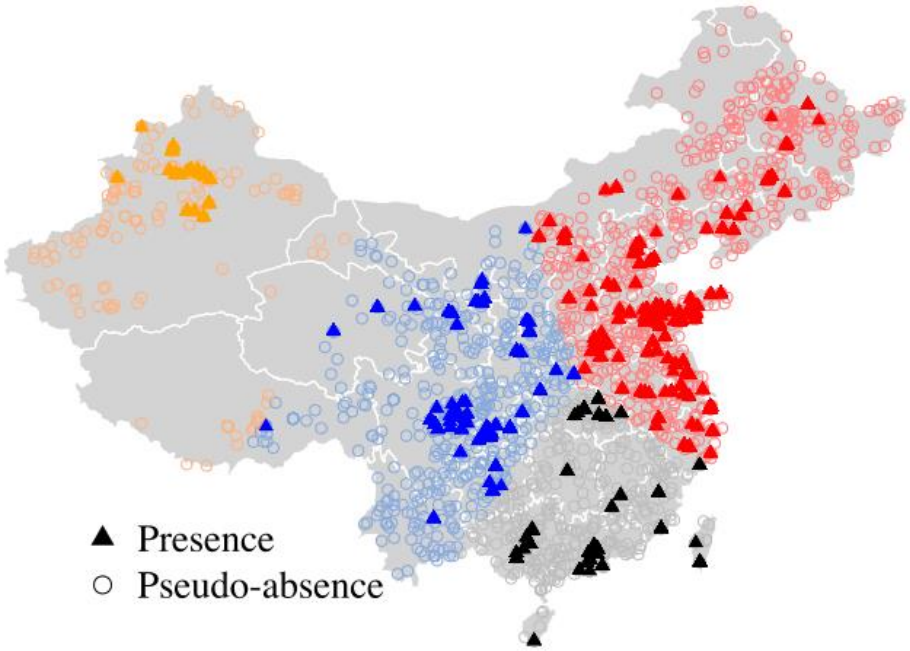
349 calculated as the product of the kriging variance and human population density

350 (standardized from 0 to 1).

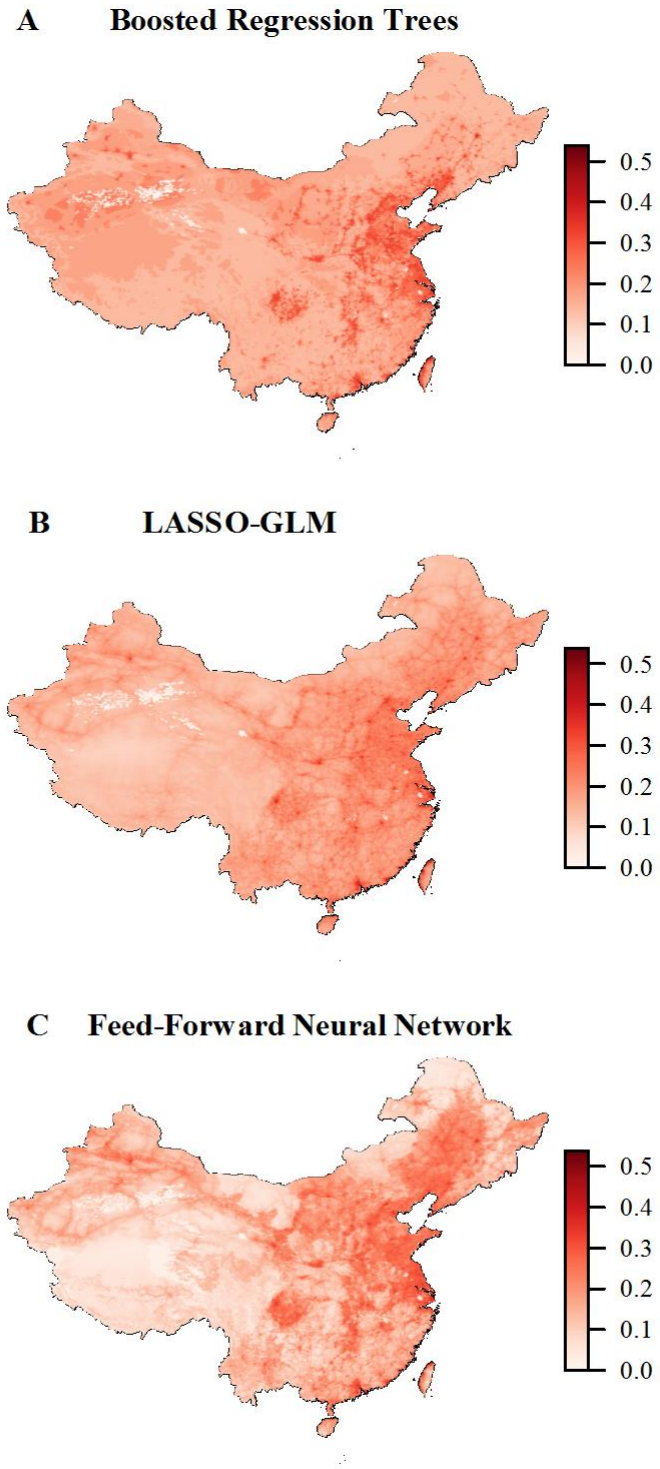
351



Supplementary Figure 13. Environmental and anthropogenic covariates used for training the child models (log10 scaled). Predicted antimicrobial use in animals (use), travel time to cities of more than 50,000 people (acc), yearly average of minimum monthly temperature (tmp), percentage of pixel area irrigated (irg), population densities of extensively raised pigs (PgExt), intensively raised pigs (PgInt), extensively raised chicken (ChExt), intensively raised chicken (ChInt), Cattle (Ca), and percentage are covered in vegetation (veg).



Supplementary Figure 14. Geographic distribution of presence and pseudo-absence. Points in four regions (represented in four colors) were used for the K-fold spatial cross-validation procedure of the child models.

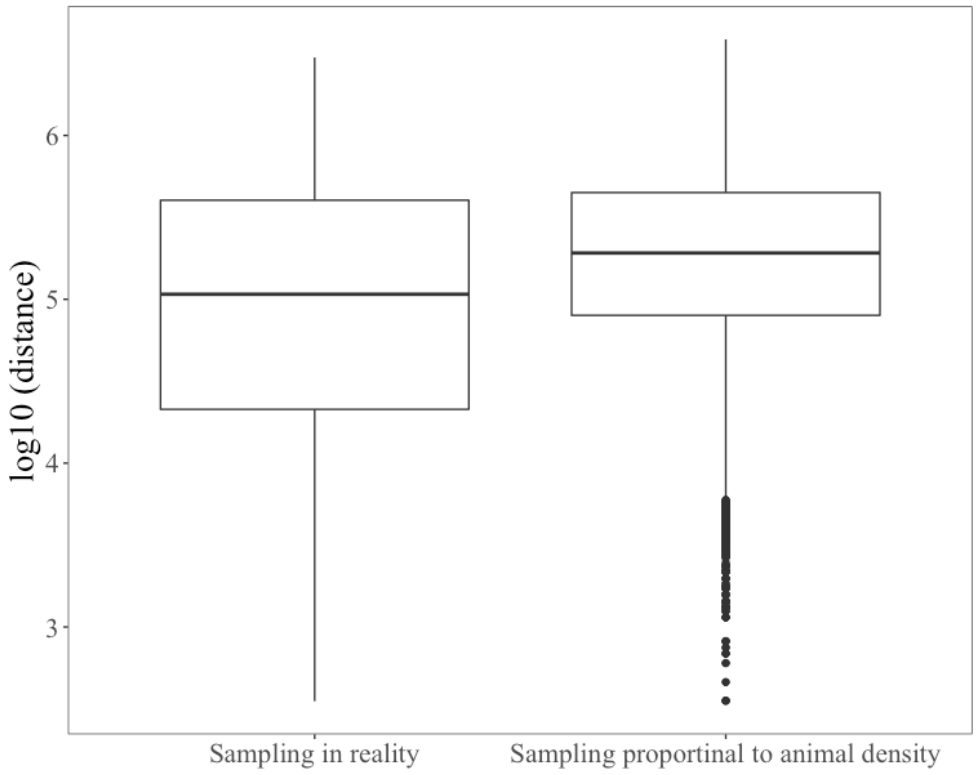


366

367 **Supplementary Figure 15.** Distribution of P50 obtained from three child models using

368 environmental covariates. A: boosted regression trees. B: Logistic regression with

369 LASSO regularization. C: Feed-forward neural network.



Supplementary Figure 16. The distribution of log10 transformed distance (meter) between institutes and their sampling locations, under a scenario that sampling is conducted proportional to animal density, as compared with sampling locations in reality.

Supplementary Tables

Supplementary Table 1. Environmental and anthropogenic covariates used for training the child models.

Acronym	Name	Year	Original Resolution	Source
use	Antimicrobial use in animals	2013	0.083333 decimal degrees	Van Boeckel et al 2017 [11] http://science.sciencemag.org/content/357/6358/1350.full
acc	Travel time to cities	2015	30-arcsec resolution	Weiss et al 2018 [14] https://www.map.ox.ac.uk/accessibility_to_cities/.
tmp	Yearly average of minimum monthly temperature	1970-2000	2.5 minutes	Worldclim [15] http://worldclim.org/version2
irg	Percentage irrigated areas	2005	0.083333 decimal degrees	Global Map of Irrigation Areas (Siebert et al., 2013) http://www.fao.org/nr/water/aquastat/irrigationmap/index10.stm
ChExt ChInt PgExt PgInt Ca	Population density pigs, chicken, and cattle, (extensive vs intensive systems)	2013	0.083333 decimal degrees	Gridded Livestock of the World v3 [9] https://dataverse.harvard.edu/dataverse/glw
veg	Percentage of Tree Coverage	2013	0.008333 decimal degrees	https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.2.html [17]

Supplementary Table 2. Research institutes most frequently associated with studies retrieved from CNKI search engine (www.cnki.net; accessed 11 March 2020) with the keywords "animal+antimicrobial+resistance" ("动物 抗生素 耐药性").

Institution	**Province**	**Number of studies**
Sichuan Agricultural University	Sichuan	75
Huazhong Agricultural University	Hubei	71
Nanjing Agricultural University	Jiangsu	57
Shandong Agricultural University	Shandong	55
Jilin University	Jilin	54
Northwest A&F University	Shaanxi	52
Yangzhou University	Jiangsu	49
China Agricultural University	Beijing	47
Northeast Agricultural University	Heilongjiang	44
South China Agricultural University	Guangdong	37

Supplementary Table 3. Prices of antimicrobial from the three online retailers with the highest sales volume, and the number of online retailers selling each antimicrobial on Alibaba (1688.com; accessed 10 December 2019).

Antimicrobial Searched on Alibaba	Chinese Name	Price_1	Price_2	Price_3	Average Price	Number of Online Retailers
Ampicillin	氨苄西林	380	260	180	273	68
Amoxicillin	阿莫西林	205	100	200	168	295
Sulfamonomethoxine	磺胺间甲氧嘧啶	315	210	295	273	261
Sulfamethoxazole	磺胺甲恶唑	180	128	100	136	177
Trimethoprim	甲氧苄啶	118	245	190	184	333
Sulfadiazine	磺胺嘧啶	140	160	200	167	542
Trimethoprim	甲氧苄啶	118	205	200	174	216
Oxacillin	苯唑西林	600	580	600	593	10
Sulfamethoxazole-Trimethoprim	新诺明	145	220	210	192	237
Nalidixic acid	萘啶酸	NA	NA	NA	NA	20
Chloramphenicol	氯霉素	185	230	270	228	330
Tetracycline	四环素	170	180	180	177	313
Sulfafurazole/ Sulfisoxazole	磺胺异恶唑	NA	NA	NA	NA	26
Penicillin	青霉素	130	110	120	120	886
Cefotaxime	头孢噻肟	600	520	500	540	122
Ciprofloxacin	环丙沙星	190	160	180	177	81
Doxycycline	强力霉素	530	380	520	477	425
Gentamicin	庆大霉素	360	850	680	630	323
Gentamicin	庆大霉素	360	850	680	630	184
Clindamycin	克林霉素	490	298	330	373	93
Ceftazidime	头孢他啶	900	1100	1000	1000	38
Erythromycin	红霉素	480	550	440	490	361
Nitrofurantoin	呋喃妥因	600	NA	NA	600	20
Ceftriaxone	头孢曲松钠	500	420	380	433	69

Cefepime	头孢吡肟	NA	NA	NA	NA	12
Cefoxitin	头孢西丁	500	NA	NA	500	14
Azithromycin	阿奇霉素	709	950	700	786	25
eptomycin	链霉素	220	180	190	197	348
Rifampicin	利福平	360	800	960	707	22
Colistin	粘菌素	420	140	100	220	216
Meropenem	美罗培南	NA	NA	NA	NA	16
Imipenem	亚胺培南	NA	NA	NA	NA	0
Tigecycline	替加环素	400	300	NA	350	7
Vancomycin	万古霉素	6000	6000	5500	5833	117

SI References

1. World Health Organization. Integrated surveillance of antimicrobial resistance in foodborne bacteria: Application of a One Health Approach. (2017) doi:https://www.who.int/foodsafety/publications/agisar_guidance2017/en/.
2. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, 1266-+ (2019).
3. Editorial Committee of China Animal Husbandry and Veterinary Yearbook. *China Animal Husbandry and Veterinary Yearbook*. (2018).
4. Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000-15: a baseline analysis for the Sustainable Development Goals. *Lancet* **390**, 2171–2182 (2017).
5. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
6. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802–813 (2008).
7. Tibshirani, R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **73**, 273–282 (2011).
8. Chollet, F. Keras. (2015).
9. Gilbert, M. *et al.* Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data* **5**, 180227 (2018).

10. Doxsey-Whitfield, E. *et al.* Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Papers in Applied Geography* **1**, 226–234 (2015).
11. Van Boeckel, T. P. *et al.* Reducing antimicrobial use in food animals. *Science* **357**, 1350–1352 (2017).
12. Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. H. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* **6**, 187–202 (2009).
13. Tiseo, K., Huber, L., Gilbert, M., Robinson, T. P. & Van Boeckel, T. P. Global trends in antimicrobial use in food animals from 2017 to 2030. *Antibiotics* **9**, 918 (2020).
14. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333-+ (2018).
15. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315 (2017).
16. Siebert, S., Henrich, V., Frenken, K. & Burke, J. Global Map of Irrigation Areas version 5. *Rheinische Friedrich-Wilhelms-University, Bonn, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy* (2013).
17. Hansen, M. C. *et al.* High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).

Supplementary Information for

Twenty-year trends in antimicrobial resistance from aquaculture and fisheries in Asia

Daniel Schar [a,*,1], Cheng Zhao [b], Yu Wang [b], D.G. Joakim Larsson [c,d], Marius Gilbert [a,e,2], Thomas P. Van Boeckel [b,f,*,1,2]

[a]Spatial Epidemiology Laboratory, Université Libre de Bruxelles, B1050 Brussels, Belgium;

[b]Institute for Environmental Decisions, ETH Zurich, Zurich 8006, Switzerland;

[c]Center for Antibiotic Resistance Research, University of Gothenburg, Sweden;

[d]Department of Infectious Diseases, Institute for Biomedicine, University of Gothenburg, Sweden;

[e]Fonds National de la Recherche Scientifique, B1000 Brussels, Belgium;

[f]Center for Diseases Dynamics, Economics, and Policy, New Delhi, India.

[1]Co-corresponding authors

[2]MG and TPVB supervised the work equally.

[*]To whom correspondence may be addressed.

Daniel Schar, Unit 7201 Box 692

DPO AE 09974-0692

Email: dlschar@gmail.com

Thomas P. Van Boeckel

Email: thomas.vanboeckel@env.ethz.ch

This file includes:

Supplementary Note 1: Systematic Review and Literature Search

A systematic review of the literature was conducted to identify point prevalence surveys (PPS) of phenotypic antimicrobial resistance inclusive of the period January 1, 2000 through September 30, 2019. Eligibility criteria included: PPS of antimicrobial resistance in bacterial pathogens of production significance or aquatic animal-associated bacterial zoonoses; PPS with samples originating from cultured or wild caught aquatic animals or their products; PPS conducted in Asia from 2000 to 2019. Target bacterial pathogens were defined according to OIE Aquatic Animal Health Code[1] and the FAO List of Important Bacterial Diseases in Aquaculture, and included: *Aerococcus*, *Aeromonas*, *Chlamydia*, *Clostridium*, *Edwardsiella*, *Enterobacterium*, *Escherichia coli*, *Flavobacterium*, *Francisella*, *Hepatobacter*, *Lactococcus*, *Mycobacterium*, *Nocardia*, *Photobacterium*, *Piscirickettsia*, *Pseudomonas*, *Renibacterium*, *Streptococcus*, *Vibrio*, and *Yersinia*.

Exclusion criteria included: reviews; meta-analyses; strain surveys describing individual strain characteristics not associated with a PPS; data from an experimental protocol not associated with a PPS; PPS with samples originating from bivalve molluscs; PPS with samples originating from ornamental fish; and PPS where no source or methodology for derivation of data was provided. We assessed data quality in our search, excluding records where resistance rates were unclear or missing; no geographic information on survey location was provided; samples originated from imported products; or samples were not clearly identified as originating from an aquatic animal or animal product.

The literature search was conducted across four databases (PubMed, Web of Science, Scopus, China National Knowledge Infrastructure) and grey literature repositories (AGRIS, CGIAR FISH, IFPRI, WorldFish). The Russian Science Citation Index and Korean Journal Database (KCI) were included in the Web of Science search. The search was conducted through September 2019 without restriction as to language, and records in English, Turkish, Japanese, Korean, Thai, and Chinese were identified and screened. The systematic review was not registered. The literature search and systematic review were guided by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement and research synthesis norms[2] (Table S4).

Search strategies were tailored to individual databases to achieve optimal search sensitivity and specificity guided by the following general formula:

(Resistance) AND (Bacteria species) AND (Aquaculture/fisheries/aquatic animals) AND (Country/region)

The full search queries were:
PubMed: (Resistance OR "antibiotic resistance" OR "antimicrobial resistance") AND ("Escherichia coli" OR "E. coli" OR *vibrio* OR Photobacterium OR Aeromonas* OR Edwardsiell* OR Yersinia OR Pseudomonas* OR Flavobacter* OR Piscirickettsia OR Hepatobacter OR Francisella OR Chlamydia OR Mycobacter* OR Nocardia OR Streptococc* OR Lactococc* OR Aerococc* OR Renibacter* OR Clostridium OR Enterobacterium) AND (aquaculture OR aquatic OR *fish* OR shellfish OR marine OR freshwater OR carp OR catfish OR prawn OR salmon OR shrimp OR tilapia OR trout) AND (Asia OR "Southeast Asia" OR "South Asia" OR "East Asia" OR Mekong OR Afghanistan OR "American Samoa"

OR Bahrain OR Bangladesh OR Bhutan OR "Brunei Darussalam" OR Cambodia OR China OR "Chinese Taipei" OR "Cook Islands" OR "Democratic People's Republic of Korea" OR Fiji OR "French Polynesia" OR Guam OR "Hong Kong" OR India OR Indonesia OR Iran OR Iraq OR Jordan OR Kiribati OR Korea OR Kuwait OR "Lao People's Democratic Republic" OR Lao OR Laos OR Lebanon OR Macau OR Malaysia OR Maldives OR "Marshall Islands" OR Micronesia OR Mongolia OR Myanmar OR Nauru OR Nepal OR "New Caledonia" OR Niue OR "Norfolk Island" OR "Northern Mariana Islands" OR Oman OR Pakistan OR Philippines OR Palau OR Palestine OR "Papua New Guinea" OR "Pitcairn Islands" OR Qatar OR Samoa OR "Saudi Arabia" OR Singapore OR "Solomon Islands" OR "Sri Lanka" OR "Syrian Arab Republic" OR Taiwan OR Thailand OR Timor-Leste OR Tokelau OR Tonga OR Turkey OR Tuvalu OR "United Arab Emirates" OR Vanuatu OR "Viet Nam" OR Vietnam OR "Wallis and Futuna Islands" OR Yemen)

Web of Science [All Databases, inclusive of Russian Science Citation Index and Korean Journal Database (KCI)]: TOPIC: (Resistance OR "antibiotic resistance" OR "antimicrobial resistance") AND TOPIC: ("Escherichia coli" OR "E. coli" OR *vibrio* OR Photobacterium OR Aeromonas* OR Edwardsiell* OR Yersinia OR Pseudomonas* OR Flavobacter* OR Piscirickettsia OR Hepatobacter OR Francisella OR Chlamydia OR Mycobacter* OR Nocardia OR Streptococc* OR Lactococc* OR Aerococc* OR Renibacter* OR Clostridium OR Enterobacterium) AND TOPIC: (aquaculture OR aquatic OR *fish* OR shellfish OR marine OR freshwater OR carp OR catfish OR prawn OR salmon OR shrimp OR tilapia OR trout) AND TOPIC: (Asia OR "Southeast Asia" OR "South Asia" OR "East Asia" OR Mekong OR Afghanistan OR "American Samoa" OR Bahrain OR Bangladesh OR Bhutan OR "Brunei Darussalam" OR Cambodia OR China OR "Chinese Taipei" OR "Cook Islands" OR "Democratic People's Republic of Korea" OR Fiji OR "French Polynesia" OR Guam OR "Hong Kong" OR India OR Indonesia OR Iran OR Iraq OR Jordan OR Kiribati OR Korea OR Kuwait OR "Lao People's Democratic Republic" OR Lao OR Laos OR Lebanon OR Macau OR Malaysia OR Maldives OR "Marshall Islands" OR Micronesia OR Mongolia OR Myanmar OR Nauru OR Nepal OR "New Caledonia" OR Niue OR "Norfolk Island" OR "Northern Mariana Islands" OR Oman OR Pakistan OR Philippines OR Palau OR Palestine OR "Papua New Guinea" OR "Pitcairn Islands" OR Qatar OR Samoa OR "Saudi Arabia" OR Singapore OR "Solomon Islands" OR "Sri Lanka" OR "Syrian Arab Republic" OR Taiwan OR Thailand OR Timor-Leste OR Tokelau OR Tonga OR Turkey OR Tuvalu OR "United Arab Emirates" OR Vanuatu OR "Viet Nam" OR Vietnam OR "Wallis and Futuna Islands" OR Yemen)

Scopus: TITLE-ABS-KEY (resistance OR "antibiotic resistance" OR "antimicrobial resistance") AND TITLE-ABS-KEY ("Escherichia coli" OR "E. coli" OR *vibrio* OR photobacterium OR aeromonas* OR edwardsiell* OR yersinia OR pseudomonas* OR flavobacter* OR piscirickettsia OR hepatobacter OR francisella OR chlamydia OR mycobacter* OR nocardia OR streptococc* OR lactococc* OR aerococc* OR renibacter* OR clostridium OR enterobacterium) AND TITLE-ABS-KEY (aquaculture OR aquatic OR *fish* OR shellfish OR marine OR freshwater OR carp OR catfish OR prawn OR salmon OR shrimp OR tilapia OR trout) AND TITLE-ABS-KEY (Asia OR "Southeast Asia" OR "South Asia" OR "East Asia" OR mekong OR afghanistan OR "American Samoa" OR bahrain OR bangladesh OR bhutan OR "Brunei Darussalam" OR cambodia OR china OR "Chinese Taipei" OR "Cook Islands" OR "Democratic People's Republic of Korea" OR fiji OR "French Polynesia" OR guam OR "Hong Kong" OR india OR indonesia OR iran OR iraq OR jordan OR kiribati OR korea OR kuwait OR "Lao People's Democratic Republic" OR lao

OR laos OR lebanon OR macau OR malaysia OR maldives OR "Marshall Islands" OR micronesia OR mongolia OR myanmar OR nauru OR nepal OR "New Caledonia" OR niue OR "Norfolk Island" OR "Northern Mariana Islands" OR oman OR pakistan OR philippines OR palau OR palestine OR "Papua New Guinea" OR "Pitcairn Islands" OR qatar OR samoa OR "Saudi Arabia" OR singapore OR "Solomon Islands" OR "Sri Lanka" OR "Syrian Arab Republic" OR taiwan OR thailand OR timor-leste OR tokelau OR tonga OR turkey OR tuvalu OR "United Arab Emirates" OR vanuatu OR "Viet Nam" OR vietnam OR "Wallis and Futuna Islands" OR yemen)

CNKI: TI = (‘抗生素’ + ‘抗菌’ + ‘兽药’ + ‘兽用药’ + ‘兽用抗生素’ + ‘用药’ + ‘抗微生物’) AND TI = (escherichia + (E*coli) + coliform + vibrio + Photobacterium + Aeromonas + Edwardsiella + Edwardsiellosis + Yersinia + Pseudomonas + Flavobacter + Piscirickettsia + Hepatobacter + Francisella + Chlamydia + Mycobacter + Nocardia + Streptococcus + Lactococcus + Aerococcus + Renibacterium + Clostridium + Enterobacterium + ‘大肠菌’ + ‘埃希菌’ + ‘弧菌’ + ‘光细菌’ + ‘气单胞菌’ + ‘埃德华氏菌’ + ‘耶尔森氏菌’ + ‘假单胞菌’ + ‘黄杆菌’ + ‘立克次氏体’ + ‘肝杆菌’ + ‘弗朗西菌’ + ‘衣原体’ + ‘分枝杆菌’ + ‘诺卡氏菌’ + ‘链球菌’ + ‘乳球菌’ + ‘空气球菌’ + ‘肾菌’ + ‘梭菌’ + ‘肠杆菌’) AND TI = (‘水产’ + ‘鱼’ + ‘渔’ + ‘贝’ + ‘海’ + ‘淡水’ + ‘鲤’ + ‘虾’ + ‘鲑’ + ‘罗非’ + ‘鳟’)) OR (KY = (‘抗生素’ + ‘抗菌’ + ‘兽药’ + ‘兽用药’ + ‘兽用抗生素’ + ‘用药’ + ‘抗微生物’) AND KY = (escherichia + (E*coli) + coliform + vibrio + Photobacterium + Aeromonas + Edwardsiella + Edwardsiellosis + Yersinia + Pseudomonas + Flavobacter + Piscirickettsia + Hepatobacter + Francisella + Chlamydia + Mycobacter + Nocardia + Streptococcus + Lactococcus + Aerococcus + Renibacterium + Clostridium + Enterobacterium + ‘大肠菌’ + ‘埃希菌’ + ‘弧菌’ + ‘光细菌’ + ‘气单胞菌’ + ‘埃德华氏菌’ + ‘耶尔森氏菌’ + ‘假单胞菌’ + ‘黄杆菌’ + ‘立克次氏体’ + ‘肝杆菌’ + ‘弗朗西菌’ + ‘衣原体’ + ‘分枝杆菌’ + ‘诺卡氏菌’ + ‘链球菌’ + ‘乳球菌’ + ‘空气球菌’ + ‘肾菌’ + ‘梭菌’ + ‘肠杆菌’) AND KY = (‘水产’ + ‘鱼’ + ‘渔’ + ‘贝’ + ‘海’ + ‘淡水’ + ‘鲤’ + ‘虾’ + ‘鲑’ + ‘罗非’ + ‘鳟’)) OR (AB = (‘抗生素’ + ‘抗菌’ + ‘兽药’ + ‘兽用药’ + ‘兽用抗生素’ + ‘用药’ + ‘抗微生物’) AND AB = (escherichia + (E*coli) + coliform + vibrio + Photobacterium + Aeromonas + Edwardsiella + Edwardsiellosis + Yersinia + Pseudomonas + Flavobacter + Piscirickettsia + Hepatobacter + Francisella + Chlamydia + Mycobacter + Nocardia + Streptococcus + Lactococcus + Aerococcus + Renibacterium + Clostridium + Enterobacterium + ‘大肠菌’ + ‘埃希菌’ + ‘弧菌’ + ‘光细菌’ + ‘气单胞菌’ + ‘埃德华氏菌’ + ‘耶尔森氏菌’ + ‘假单胞菌’ + ‘黄杆菌’ + ‘立克次氏体’ + ‘肝杆菌’ + ‘弗朗西菌’ + ‘衣原体’ + ‘分枝杆菌’ + ‘诺卡氏菌’ + ‘链球菌’ + ‘乳球菌’ + ‘空气球菌’ + ‘肾菌’ + ‘梭菌’ + ‘肠杆菌’) AND AB = (‘水产’ + ‘鱼’ + ‘渔’ + ‘贝’ + ‘海’ + ‘淡水’ + ‘鲤’ + ‘虾’ + ‘鲑’ + ‘罗非’ + ‘鳟’))

Grey literature searches:

AGRIS: (Resistance OR "antibiotic resistance" OR "antimicrobial resistance") AND (bacteria species) AND (aquaculture OR aquatic OR fish OR shellfish OR marine OR freshwater OR fish spp.) AND (Asia OR "Southeast Asia" OR "South Asia" OR "East Asia" OR Mekong)

CGIAR FISH: Resistance OR "antibiotic resistance" OR "antimicrobial resistance"

IFPRI: Resistance OR "antibiotic resistance" OR "antimicrobial resistance"

WorldFish: (Resistance OR "antibiotic resistance" OR "antimicrobial resistance") AND (Asia OR "Southeast Asia" OR "South Asia" OR "East Asia" OR Mekong)

The literature search identified 5,804 unique records screened by titles and abstracts. Of these records, 1,131 full text articles were assessed for eligibility, and 343 records yielding 749 point prevalence surveys were included in the data analysis (Fig. S1).

Supplementary Note 2: Data Extraction

Data extraction was performed according to the following protocols to generate the database used for analysis. Data was extracted using Google Sheets and Microsoft Excel (version 16.51). The protocols are adapted from *Resistancebank*[3], adjusted for the contexts and characteristics of PPS conducted from aquatic animals and settings. The data extracted from each record included the Direct Object Identifier (DOI), author, year of publication, country, latitude and longitude of sample collection, derivation of lat/long coordinates, location type of sample origin, sampling dates, animal species sampled and whether cultured or wild caught, sample type collected, health status of animal, history of antimicrobial use, pathogen and strain, number of samples collected, number of isolates subjected to antimicrobial susceptibility testing (AST), AST method, breakpoints and guidelines used, drug class, compound, resistance rate, and author contact details.

The complete database legend is available on the Zenodo public repository (doi: 10.5281/zenodo.4609884)[4].

The following broad antibiotic classes were included in the database: PEN (Penicillins), CEP (Cephalosporins), MON (Monobactams), CAR (Carbapenems), AMI (Aminoglycosides), QUI (Quinolones), AMP (Amphenicols), TET (Tetracyclines), SUL (Sulfonamides), MAC (Macrolides), Glycopeptides (GLY), POL (Polymyxins), and OTH (Others).

Antimicrobial compounds used for susceptibility testing were designated by a 3-letter code and the Anatomical Therapeutic Chemical (ATC) Classification code designation (https://www.whocc.no/atc_ddd_index/ or https://www.whocc.no/atcvet/atcvet_index/). ATC-Code starting with J0 stand for antibiotics for human systemic use while QJ01for veterinary use. For antibiotics without attributed ATC codes, a pseudo-code was constructed by using the ATC code of the molecular classification (5 or 6 characters for human and veterinary antibiotics, respectively) and adding the first character of the compound's name separated by a - (e.g. Sarafloxacin – J01MA-S). Some ATC codes are

provided for mixture of compounds (e.g. J01RA01 for penicillins in combination with other antibacterials). Active ingredients' names were reported when commercial drugs were used.

The antibiotics found across all studies are the following (3 letter code, ATC-code): Amoxicillin-Clavulanic Acid (AMC, J01CR02); Ticarcillin-Clavulanic acid (TIM, J01CR03); Piperacillin-Tazobactam (PIT, J01CR05); Ampicillin-Sulbactam (SAM, J01CR01); Ampicillin (AMP, J01CA01); Amoxicillin (AMX, J01CA04); Amoxicillin-Sulbactam (AMS, J01CR02); Azlocillin (AZL, J01CA09); Ticarcillin (TIC, J01CA13); Cloxacillin (CLO, J01CF02); Oxacillin (OXA, J01CF04); Piperacillin (PIP, J01CA12); Flucloxacillin (FLU, J01CF05); Carbenicillin (CAR, J01CA03); Methicillin (MET, J01CF03); Penicillin (PEN, J01CE01); Mezocillin (MEZ, J01CA10); Ceftriaxone (CRO, J01DD04); Ceftazidime (CAZ, J01DD02); Cefalexin (CLX, J01DB01); Cefotaxime (CTX, J01DD01); Cefepime (FEP, J01DE01); Cefoxitin (FOX, J01DC01); Cefalotin (CFL, J01DB03); Ceftiofur (CFU, QJ01DD90); Cefuroxime (CXM, J01DC02); Cefpodoxime (CPD, J01DD13); Cefazolin (CFZ, J01DB04); Cefixime (CFM, J01DD08); Cefamandole (CMD, J01DC03); Cefoperazone (CFP, J01DD12); Moxalactam (MOX, J01DD06); Cefradine (CFR, J01DB09); Sulbactam-CFP (SFP, J01DD62); Ceftizoxime (CZM, J01DD07); Cephaloridine (CLD, J01DB02); CAZ-Clavulanic Acid (CAC, J01DD52); Cefotiam (CFT, J01DC07); Cefpimizole (CPM, J01DC-C); Cefminox (CMX, J01DC12); Cefaclor (CFC, J01DC04); Cefadroxil (CFR, J01DB05); Aztreonam (ATM, J01DF01); Imipenem (IPM, J01DH51); Ertapenem (ERT, J01DH03); Meropenem (MEM, J01DH02); Kanamycin (KAN, J01GB04); Gentamicin (GEN, J01GB03); Neomycin (NEO, J01GB05); Streptomycin (STR, J01GA01); Amikacin (AMK, J01GB06); Tobramycin (TOB, J01GB01); Apramycin (APR, QA07AA92); Netilmicin (NET, J01GB07); Spectinomycin (SPT, J01XX04); Fleroxacin (FLR, J01MA08); Enoxacin (ENO, J01MA04); Ciprofloxacin (CIP, J01MA02); Nalidixic acid (NAL, J01MB02); Pipemidic acid (PIM, J01MB04); Enrofloxacin (ENR, QJ01MA90); Norfloxacin (NOR, J01MA06); Ofloxacin (OFX, J01MA01); Oxolinic Acid (OXO, J01MB05); Flumequine (FLQ, J01MB07); Moxifloxacin (MXF, J01MA14); Levofloxacin (LVX, J01MA12); Pefloxacin (PEF, J01MA03); Marbofloxacin (MRB, QJ01MA93); Gatifloxacin (GAT, S01AE0E); Lomefloxacin (LOM, J01MA07); Danofloxacin (DAN, QJ01MA92); Sarafloxacin (SAR, J01MA-S); Chloramphenicol (CHL, J01BA01); Florfenicol (FFC, QJ01BA90); Thiamphenicol (TFC, J01BA02); Tetracycline (TET, J01AA07); Oxytetracycline (OXT, J01AA06); Doxycycline (DOX, J01AA02); Minocycline (MIN, J01AA08); Chlortetracycline (CTE, J01AA03); Sulfamethoxazole-Trimethoprim (SXT, J01EE01); Sulfamethoxazole (SMZ, J01EC01); Sulfafurazole or Sulfisoxazole (SOX, J01EB05); Sulfadiazine (SUD, J01EE-S); Sulfonamides (SSS, J01E); Trimethoprim-Sulfadiazine (TDZ, QJ01EW10); Trimethoprim (TMP, J01EA01); Sulfamonomethoxine (SMN, QJ01EQ18); Kitasamycin (KIT, QJ01FA93); Erythromycin (ERY, J01FA01); Oleandomycin (OLD, J01FA05); Lincomycin (LIN, J01FF02); Clindamycin (CLI, J01FF01); Clarithromycin (CLR, J01FA09); Tylosin (TYL, QJ01FA90); Azithromycin (AZM, J01FA10); Spiramycin (SPI, J01FA02); Tilmicosin (TIL, QJ01FA91); Roxithromycin (ROX, J01FA06); Midecamycin (MID, J01FA03); Vancomycin (VAN, J01XA01); Teicoplanin (TEC, J01XA02); Polymyxin B (PMB, J01XB02); Colistin (CST, J01XB01); Linezolid (LIZ, J01XX08); Nitrofurantoin (NIT, J01XE01); Nitrofurazone (NFZ, D08AF01); Bleomycin (BLM, L01DC01); Rifampicin (RIF, J04AB02); Bacitracin (BAC, J01XX10); Fosfomycin (FOF, J01XX01); Fusidic acid (FUS, J01XC01); Metronidazole (MTD, J01XD01); Pristinamycin (PRI, J01FG01); Furazolidone (FRZ, QJ01XE90); Novobiocin (NOV, QJ01XX95); Bicyclomycin (BCM, J01-B); Virginiamycin (VRG, D06AX10).

The median of the interval between mid-point of sampling dates and year of publication—set to mid-point of year (e.g. July 2nd)—was calculated for all records where such details were given. In records where sampling dates were not provided, the estimated sampling dates were then calculated by subtracting the median interval in days from the publication date.

For records requiring clarification or with missing data, corresponding authors were contacted. A total of 44 emails were sent requesting clarification, and 15 responses were received. Records were excluded when no response was received and missing or unclear data precluded further analysis.

Asia sub-regions and their corresponding countries and territories were defined according to United Nations classifications (https://unstats.un.org/unsd/methodology/m49/) (Fig. S3), with modification to merge Western Asia and Southern Asia into a single region for analysis.

There were 104 unique species or groups of species represented in our dataset. To facilitate analysis, species were aggregated into six groups reflective of aquatic animal and type of aquatic environment: marine fish, freshwater fish, brackish water fish, shrimp, and a mixed group where aquatic animal species sampled spanned categories and for which resistance rates were not disaggregated. The remaining species were pooled into a sixth group that included other crustaceans (crab), cephalopods (squid), gastropod molluscs (abalone), amphibians (frogs and salamanders), echinoderms (sea cucumbers and sea urchins), and reptiles (turtles).

Supplementary Note 3: Data Analysis

The percentage of antimicrobial compounds in each survey with resistance exceeding 50% was calculated (P50). With the exception of select *Aeromonas* and *Flavobacterium* spp. pathogens of primarily cold water salmonid species[5], the majority of aquatic animal pathogens have neither standard interpretive criteria for susceptibility testing nor clinical breakpoints to guide therapeutic interventions in aquatic animals. In the absence of standard interpretive criteria, surveys frequently used human clinical breakpoints when available, either at the bacterial species, genera or family level. Therefore, the P50 is best used as an index of multi-drug resistance rather than indicative of expected therapeutic outcomes in aquatic animals.

Analysis of Antimicrobial Resistance Trends Across All Bacteria Isolated from Aquatic Food Animals

The P50 metric was used in the analysis of temporal trends and for geospatial modeling. We compared P50 with two additional metrics: P30 (calculated as the percentage of antimicrobial compounds in each survey with resistance exceeding 30%) and mean resistance (calculated as the total number of resistant isolates divided by the number of isolates * the number of antibiotics tested in each survey). Across all surveys, there is a positive correlation between P50 and mean resistance (Pearson's correlation coefficient = 0.9596). In cultured animals, P50 (RMSE =0.223 ; coefficient = 0.004; p = 0.633) and Mean (RMSE =0.207 ; coefficient = 0.002 ; p = 0.758) were comparable with a positive coefficient slope, whereas for P30 (RMSE =0.235 ; coefficient = -0.012; p = 0.158) the slope is negative. In wild caught animals, the Mean model fit improved (RMSE = 0.176; coefficient = -0.065; p = 0.002) when compared with P50 (RMSE = 0.235; coefficient = -0.085; p = 0.003) and P30 metrics (RMSE = 0.262; coefficient = -0.106; p

= 0.001), however the directionality of the trends is consistent and remain statistically significant across all metrics (Fig. S19). We use P50 as an index of multi-drug resistance to document the temporal and geographic trends in resistance in bacteria of aquatic animal origin intended for human consumption.

Temporal trends were analyzed by fitting regression lines using generalized linear models with quasibinomial error distribution weighted by the log of the number of isolates in each survey subjected to susceptibility testing in order to reflect uncertainty in surveys with a limited number of isolates. Root mean square error (RMSE) was used to evaluate goodness of fit for the temporal trends regression models. RMSE indicated model fits were moderate ($RMSE_{cultured}$ = 0.223; and $RMSE_{wild\ caught}$ = 0.235), consistent with both the scattered nature and scarcity of the data. The 95% confidence intervals were generated as the fitted values +/- 1.96 * standard error of the fitted value.

One-way analysis of variance (ANOVA) tests were used to analyze the significance of the difference in mean P50 across all surveys when comparing samples from (*i*) cultured and wild caught aquatic animals; (*ii*) animals with or without history of antimicrobial use prior to sampling; and (*iii*) diseased and healthy aquatic animals. One-way ANOVA tests were conducted on arcsine transformed P50 values to normalize the distributions of these proportions. The distributions of the residuals were checked visually using histograms and q-q plots and with the Shapiro-Wilk normality test. Homogeneity of variance between groups was confirmed using a Bartlett test and by examining residual plots.

Analysis of Antimicrobial Resistance in Foodborne Pathogens Isolated from Aquatic Animals

The pooled prevalence of resistance was calculated from individual pathogen-drug resistance rates to report resistance in foodborne pathogens specifically (*Vibrio* spp., *E.coli*, *Streptococcus* spp. and *Aeromonas* spp.). We analyzed antimicrobial resistance in surveys at the bacterial genera level. This level of taxonomy was completely available (no missing entries) in our database, whereas the more granular bacterial species and strain level data were either not consistently provided or could not be disaggregated (143 surveys; 19%). In 2019, the U.S. FDA National Antimicrobial Resistance Monitoring System (NARMS) initiated a pilot surveillance program for pathogens from seafood[6]. This NARMS pilot study design targets *Vibrio*, *Aeromonas* and *Enterococcus* spp. reported at the genera level. Analysis of resistance in foodborne bacteria of aquatic animal origin was guided by antimicrobial compounds of relevance for therapeutic use in human clinical settings. Resistance rates were calculated for *Vibrio* spp., *E.coli*, *Streptococcus* spp. and *Aeromonas* spp. from samples originating from marine fish, freshwater fish, and shrimp groups using The Clinical & Laboratory Standards Institute (CLSI M45 and M100) and WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR) pathogen-drug susceptibility testing guidelines (Table S1). The 95% confidence interval was calculated for the population proportion.

Supplementary Note 4: Geospatial modeling

P50 values from point prevalence surveys were interpolated to map AMR in freshwater and marine environments at a resolution of 0.0833 decimal degrees, or approximately 10 km at the equator.

Freshwater Protocol

Using a two-step procedure, we first trained multiple child models, and subsequently stacked model predictions for universal kriging. This stacked generalization ensemble modeling approach has been used to model population level health metrics[7] as well as the distribution of AMR in terrestrial animals[3]. Such approaches have been shown to improve overall predictive accuracy when compared with individual models[8].

The freshwater data set was replicated five times to expand values for modeling. Each survey was randomly redistributed within a discrete uncertainty range in kilometers defined as the mean of the uncertainty boundaries around the X and Y coordinates for the smallest available administrative unit or place name provided in the survey. In practice, the uncertainty range for a survey where precise sampling coordinates were provided was zero, and such surveys (n=43) were therefore not redistributed. P50 values from the expanded data set were first transformed into presence or absence of resistance using random binarization, where P50 values were compared with a random number between zero and one and classified as presence if they exceed this number or absence if they fell below this number. Next, pseudo-absence points were generated and distributed to provide additional covariate values that were not associated with presences (P50 = 0). Pseudo-absence points were sampled within a radius of 20 to 500 km from presence points using stratified random sampling proportional to the human population density to account for potential P50 observation bias in more densely populated areas. Child models contained equal numbers of true presence versus absences (true absence + pseudo absences) as balanced data sets have been shown to enhance predictive accuracy of spatial models[9].

In the first step, we trained three classes of child models: boosted regression trees (BRT)[10]; least absolute shrinkage and selection operator applied to logistic regression (LASSO-GLM)[11]; and overlapped grouped LASSO penalties for General Additive Models selection (LASSO-GAM)[12]. Models were trained to quantify the association between P50 and a set of environmental and anthropogenic covariates relevant to the freshwater environment (Fig. S13 and Table S2). Covariates were log10-transformed and resampled from their original resolution to 0.0833 decimal degrees.

Prior to model fitting, spatial sorting bias (SSB) was calculated to determine whether mean distance between sets of points and their nearest reference points for the training and validation data sets for each spatial fold differed[13]. The SSB approached one (SSB = .95), indicating negligible bias. The BRT model was fit using a tree complexity of two, a learning rate of 0.0001, and a step size of 60, controlling interactions between variables, the weights of each individual tree in the final model, and the number of trees added at each cycle, respectively. All child models were fit using three-fold spatial cross validation to prevent local overfitting and bootstrapped 10 times to account for variability introduced in the redistribution of surveys according to their geographic uncertainty range; in the random binarization of P50 values; and in the stratified random sampling of pseudo-absence points. For all child models, the modified UN Asia sub-regions served as the three cross validation regions: Eastern Asia; Western and Southern Asia; and South-eastern Asia. Model predictive accuracy was evaluated by taking the mean value of the area under the receiver operator characteristic curve (AUC) for all bootstrap runs. The BRT AUC = 0.60, LASSO-GLM AUC = 0.57, and LASSO-GAM AUC = 0.56. The distribution of relative influence[10] for each covariate across all bootstraps was used to evaluate the contribution of covariates in the BRT models (Fig. S16).

In the second step, predictions from child models were stacked and used as covariates for universal kriging of P50 values between survey locations. Duplicate coordinates from surveys conducted at the same location were randomly redistributed within 1 km of the survey coordinates, weighted by the log10 of the number of isolates in the survey to reflect broader geographic range from larger surveys. We fit a Matern variogram with a cutoff of 500 km, which is where the semi-variogram attained the range. The kriging procedure was weighted by the number of isolates at each location.

We quantified spatial uncertainty in our P50 map, producing a 95% confidence interval on the predicted values (Fig. S18). We first calculated the standard deviation in predictions in each pixel in each child model. Next, we calculated a standardized kriging variance such that variance was zero at the location of observations (Fig. S17). These were used to produce the 95% confidence interval map on the predictions as follows:

$$95\%\ CI = 1.96 \times (sd(P_{BRT}, P_{LASSO-GLM}, P_{LASSO-GAM}) + \sqrt{Var}_K)$$

Where $P_{BRT}, P_{LASSO-GLM}, P_{LASSO-GAM}$ are the P50 predicted values from each child model and Var_K is the standardized kriging variance after stacking.

Predictions were masked to freshwater using a hybrid of the Global Lakes and Wetlands Database[14], the HydroSHEDS lakes database[15] of freshwater lakes with surface area greater than 10 hectares, and the HydroSHEDS rivers database[16] of rivers with long-term average discharge greater than 10 m^3/s.

Marine Protocol

A root mean square error (RMSE)-weighted ensemble model was used to map AMR in marine environments. In the marine model, P50 values from inland freshwater surveys were excluded. Surveys from wild caught marine animals sampled at land based post-harvest sites were randomly assigned coordinates to open ocean within a radius of .54 to 81 nautical miles (1 to 150 km) from their nearest coastal location. This range captures a distance intended to include the largest distribution of both artisanal and industrial fishing fleets in Asia, and falls within the 200 nautical mile distance from coastline exclusive economic zones for which countries retain rights to explore their marine resources. The marine data set consisted of these surveys combined with surveys originating from marine, coastal marine, brackish water, and coastal brackish water sampling locations, which were randomly distributed within 1 km radius from their original sampling location to avoid duplicate coordinates. Coastal marine surveys with land based coordinates were redistributed to their nearest coastline.

The combined marine data set was then used to interpolate P50 using only the sampling coordinates as covariates. Inverse distance weighted, natural neighbor, and ordinary kriging models were produced. In the natural neighbor model, the maximum number of neighbor locations was set to eight and the inverse distance power set to one. In the ordinary kriging model, we fit a Matern variogram model with a maximum range of 2,000 km. The kriging procedure was weighted by the number of isolates subjected to antimicrobial susceptibility testing in each survey to capture uncertainty associated with surveys with a limited number of isolates.

We then stacked the models and weighted their P50 predictions according to their root mean square error (RMSE) to capture the fit and variance of each model in the final ensemble model. The weights were taken as the inverse of the RMSE of each constituent model divided by the sum of RMSE for all models, and expressed as their relative proportion in the final RMSE-weighted marine AMR ensemble model (Table S3). A transparency function was added proportional to the spatial kernel density of surveys to reflect geographic distribution of surveys contributing to the final marine P50 map. The spatial kernel density was calculated at a distance bandwidth of 8.333 decimal degrees (approximately 1000 km at the equator), which is where the variogram of the surveys leveled or reached range.

Optimizing Locations for Future Surveillance

We identified the locations for 50 hypothetical surveys that could be conducted across Asia in the next year, aimed at maximizing the information gained from future AMR surveillance in freshwater aquaculture. Future survey locations were identified using a map of "need for surveillance" (NS), defined as:

$$NS = Var_K \cdot W_h \cdot W_a$$

Where Var_K represents kriging variance; and W_h and W_a are human population density and inland aquaculture production, respectively. These terms weight the necessity for surveillance by those areas where AMR is likely to have the greatest impact on human health and the aquaculture industry. All three terms were standardized to range [0,1], thus given equal weights in determining the need for surveillance.

We applied the approach proposed by Zhao et al. (submitted) that maximizes information gain in determining each successive survey location. First, we identified the pixel location with the highest value on the map of NS—denoted as X_p, Y_p—as the first survey location. Then the value of NS at each pixel location X_i, Y_i was recalculated as

$$NS_{(+1\ \text{survey})\ X_i,Y_i} = NS_{X_i,Y_i} \times \left(1 - \frac{\text{overlap area}}{\text{neighborhood area}}\right)$$

Where the "neighborhood area" consists of all pixels within the distance of decreased *NS* to a new survey, represented as a circular area with radius *dist*. "Overlap area" is the overlapped neighborhood areas between location X_p, Y_p and X_i, Y_i. Each additional survey location was then placed successively at the pixel location with highest value on the recalculated map of NS until all surveys were assigned.

The radius *dist* of the neighborhood area was optimized using a sensitivity analysis, through approximate Bayesian computation (sequential Monte Carlo)[17], such that the sum of all pixel values on the map of NS was minimized after all survey locations were assigned and added in the kriging model. The priors of *dist* were a uniform distribution between 50 km and 500 km.

Data analysis was conducted in R version 3.6.3.

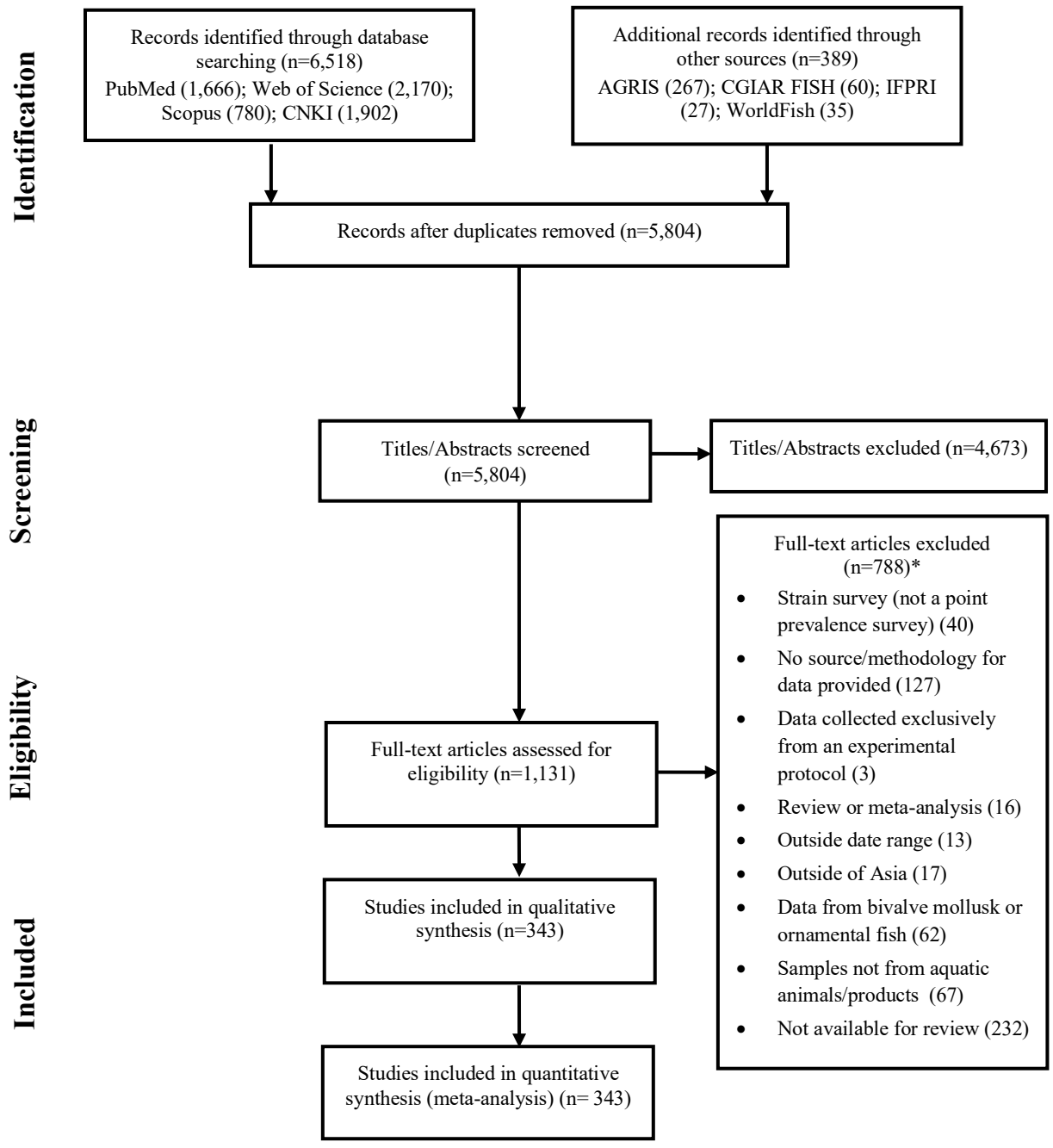


Fig. S1. Systematic review[2] and meta-analysis of antimicrobial resistance point prevalence surveys in aquatic animals. *Exclusion criteria for CNKI search records are not reflected in individual criteria totals.

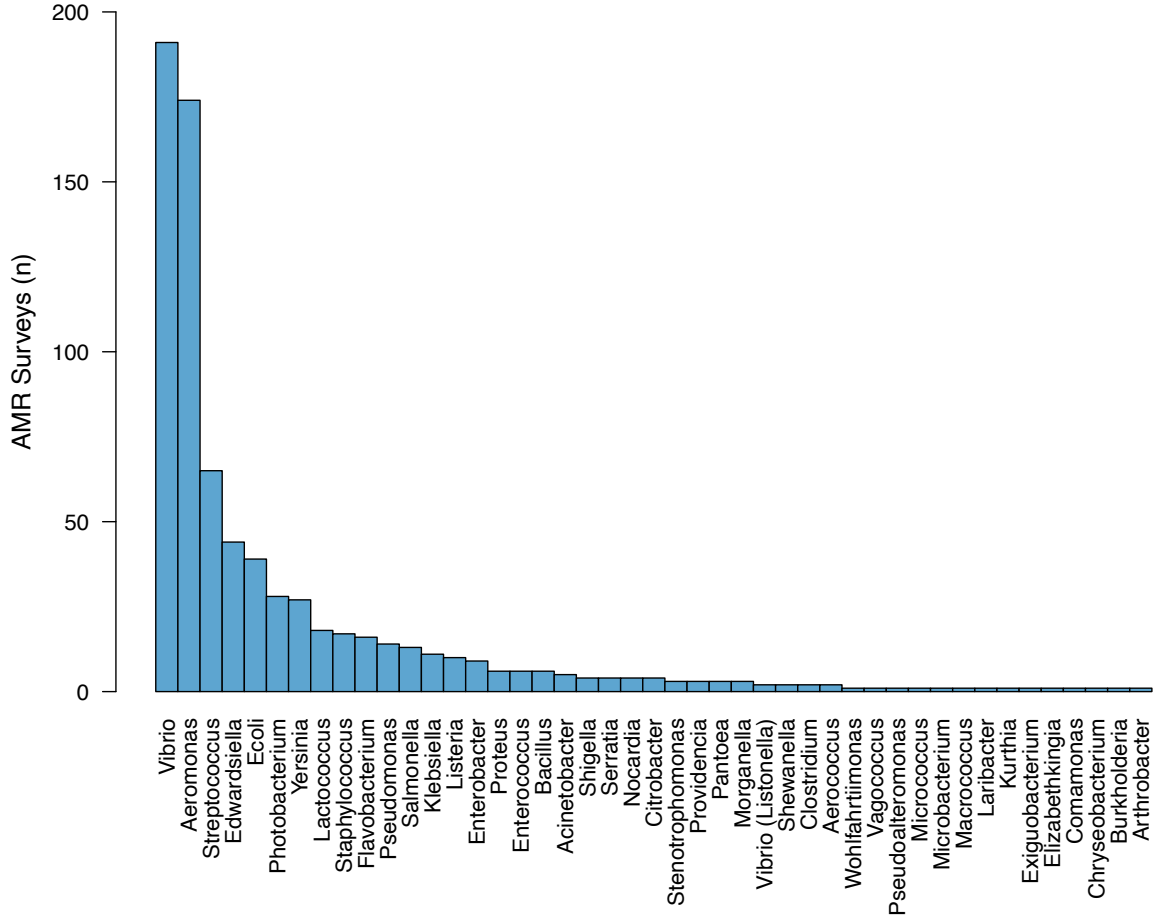


Fig. S2. **Pathogens represented in resistance surveys in aquatic animals.**

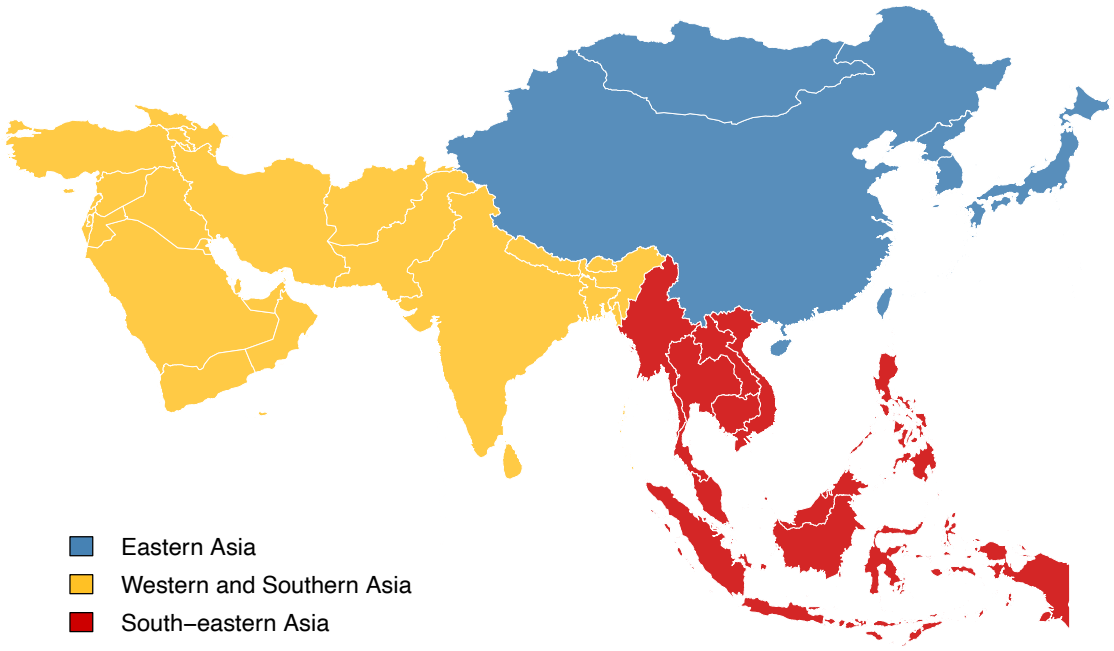


Fig. S3. Asia sub-regions from which the point prevalence surveys originate. Sub-regions are defined according to the United Nations "Standard Country or Area Codes for Statistical Use", with Western Asia and Southern Asia merged.

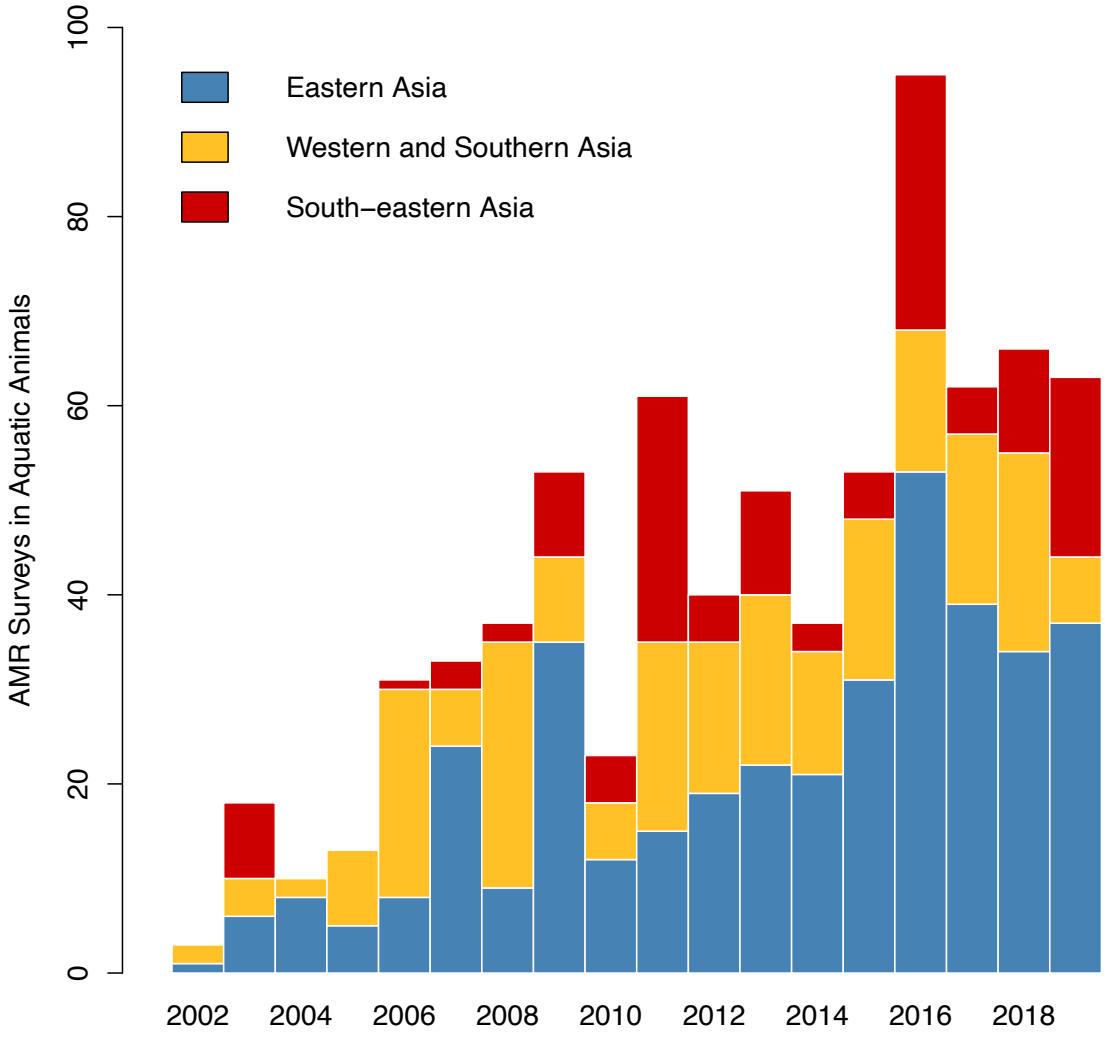


Fig. S4. Resistance surveys conducted in aquatic animals. Countries constituting each sub-region (Eastern Asia: CHN, HKG, JPN, KOR, TWN; Western and Southern Asia: BGD, IND, IRN, KWT, LBN, PAK, SAU, TUR; South-eastern Asia: KHM, MYS, PHL, SGP, THA, VNM) are represented by their International Organization for Standardization country codes. Year of publication is shown.

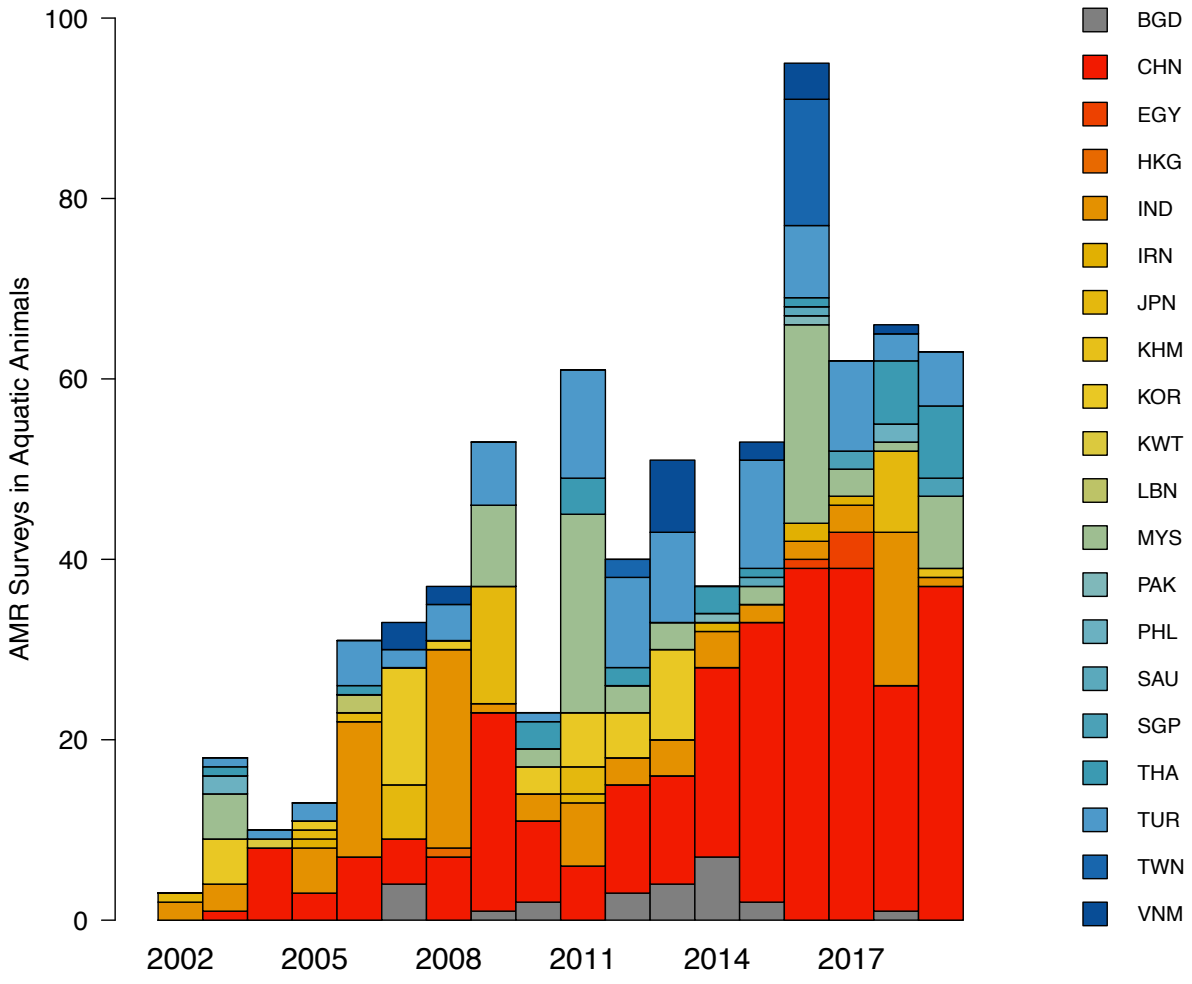


Fig. S5. Countries conducting resistance surveys in aquatic animals. Countries are represented by their International Organization for Standardization country codes. Year of publication is shown.

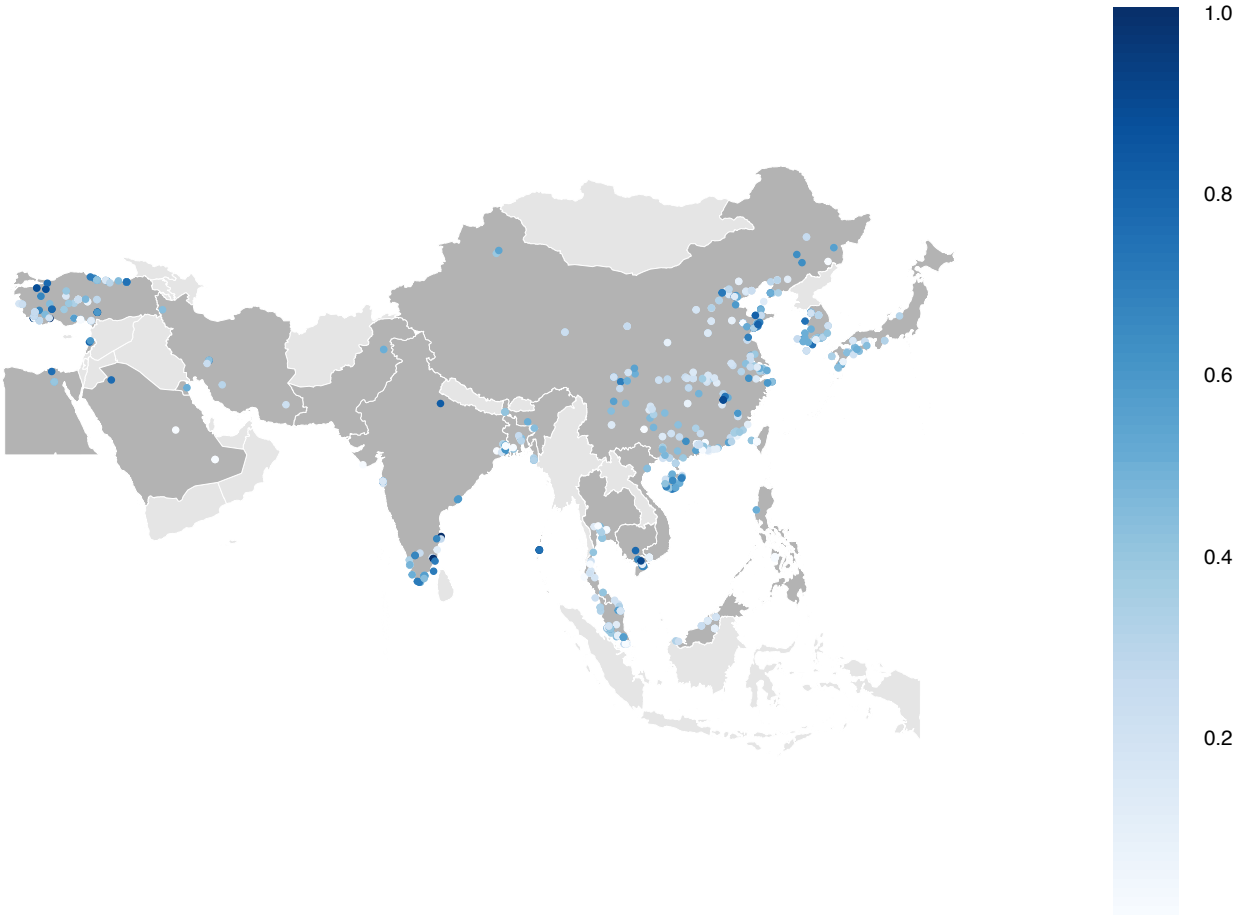


Fig. S6. Proportion of drugs with resistance higher than 50% (P50) in each point prevalence survey in Asia. P50 is shown from all aquatic animal species from all aquatic environments.

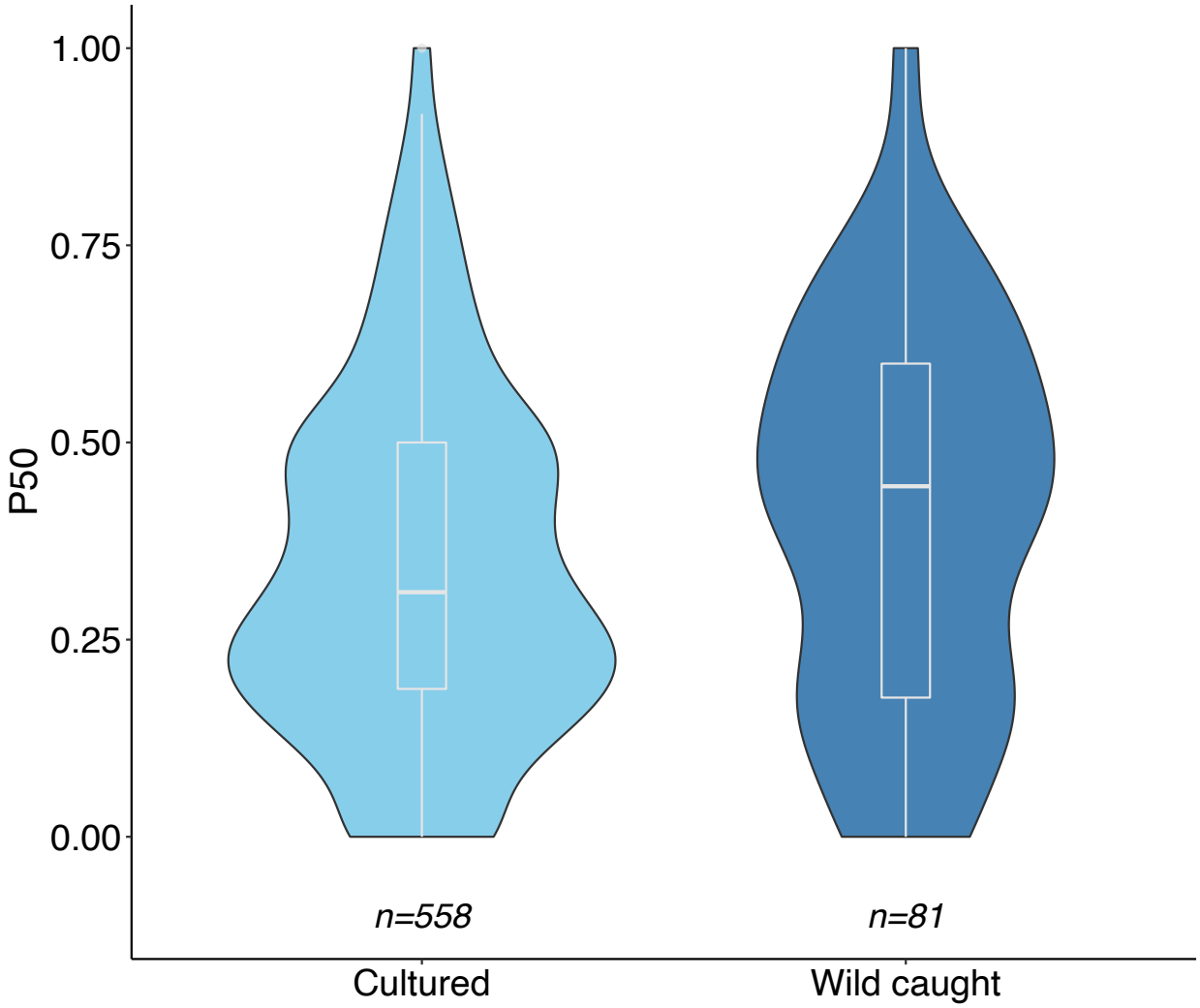


Fig. S7. Distribution of P50 for cultured and wild caught aquatic animals. The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. n = 639 individual point prevalence surveys. One-way ANOVA p=.059

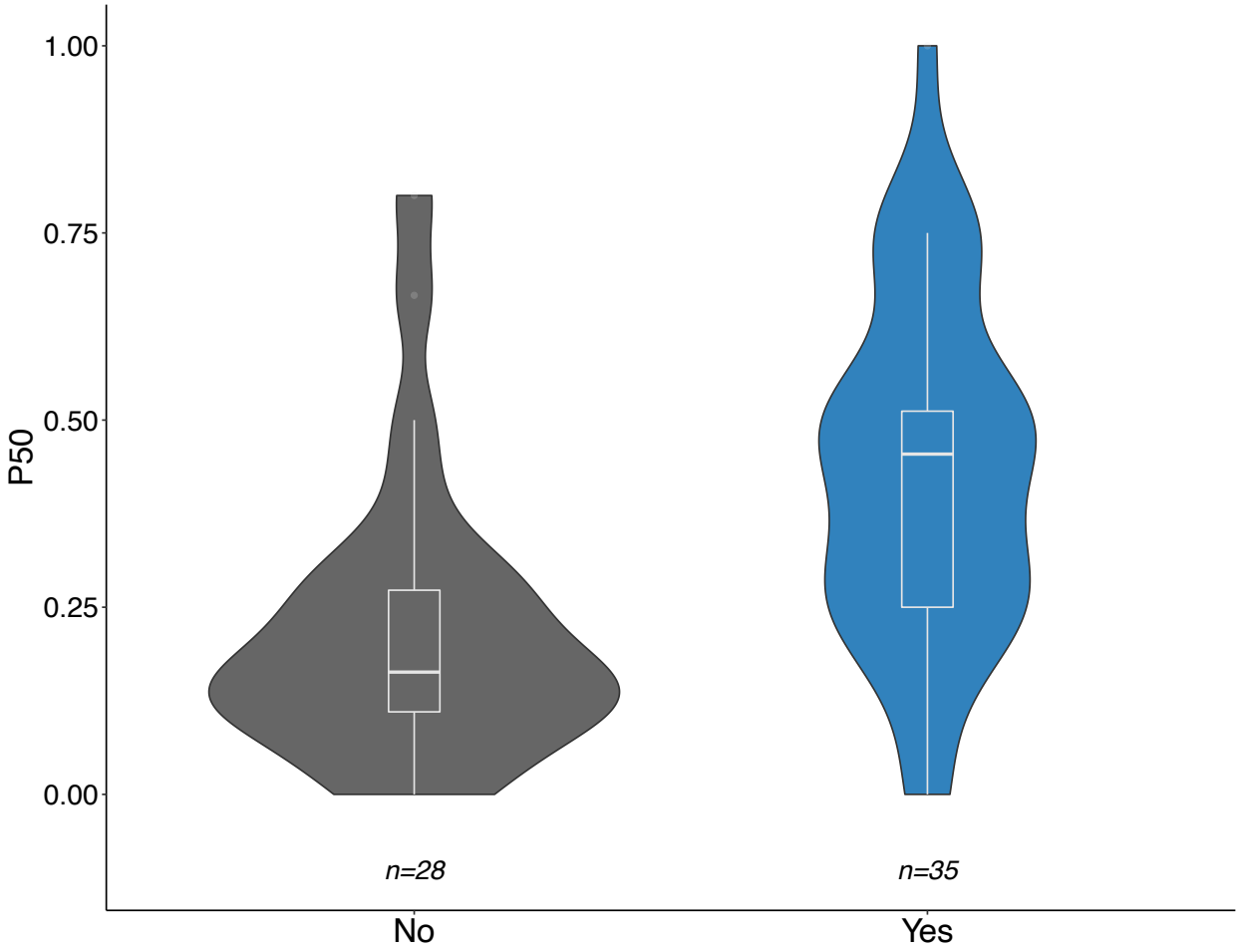


Fig. S8. Distribution of P50 from surveys in cultured aquatic animals in which history of antimicrobial use prior to sampling was recorded. Analysis is limited to surveys that explicitly identified that antimicrobials were ("yes") or were not ("no") applied to the sampled animals prior to sampling (n = 63 individual point prevalence surveys). The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. One-way ANOVA p=2E-04.

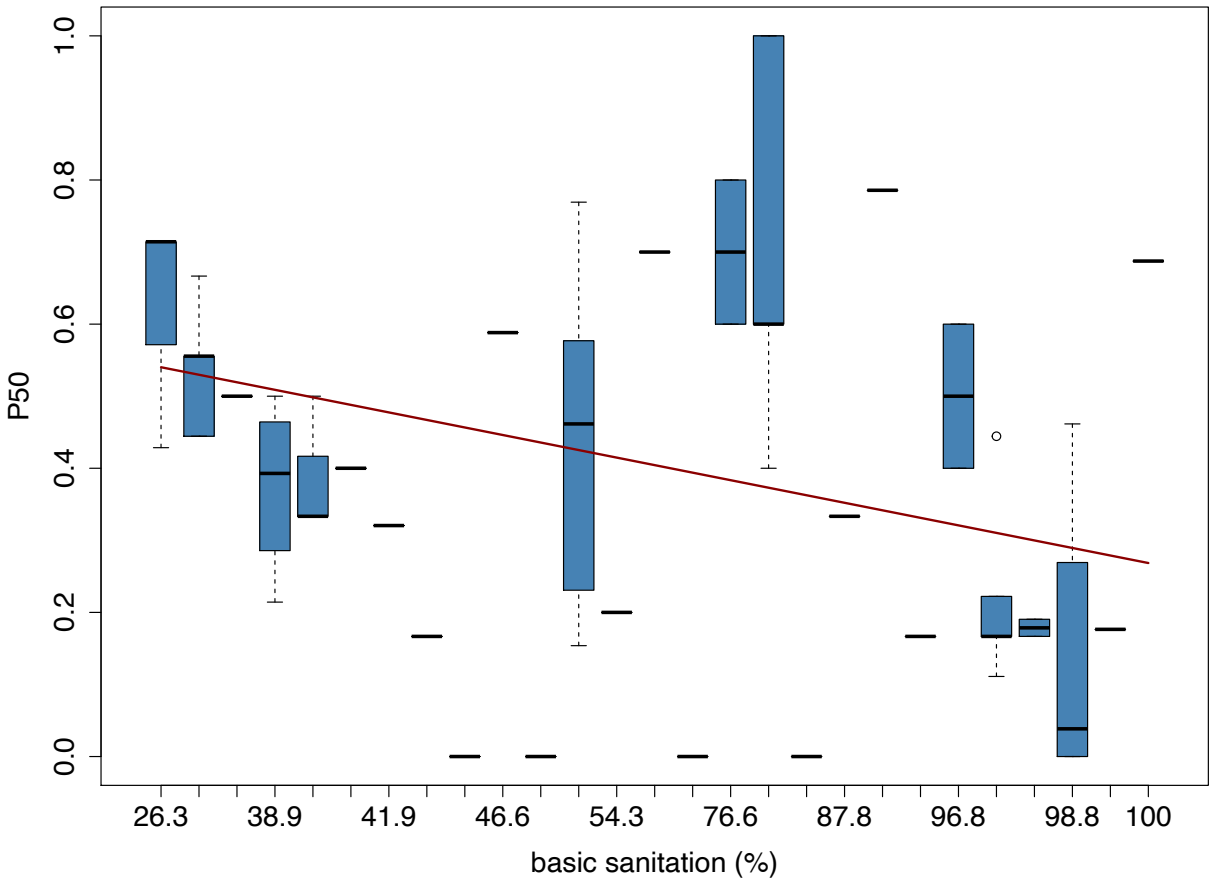


Fig. S9. P50 correlation with World Bank basic sanitation index[18]. The correlation is shown for surveys from wild caught aquatic animals in both freshwater and marine environments (n = 80). The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. Regression line is fitted using a generalized linear model with a solid line indicating statistical significance (p=0.024).

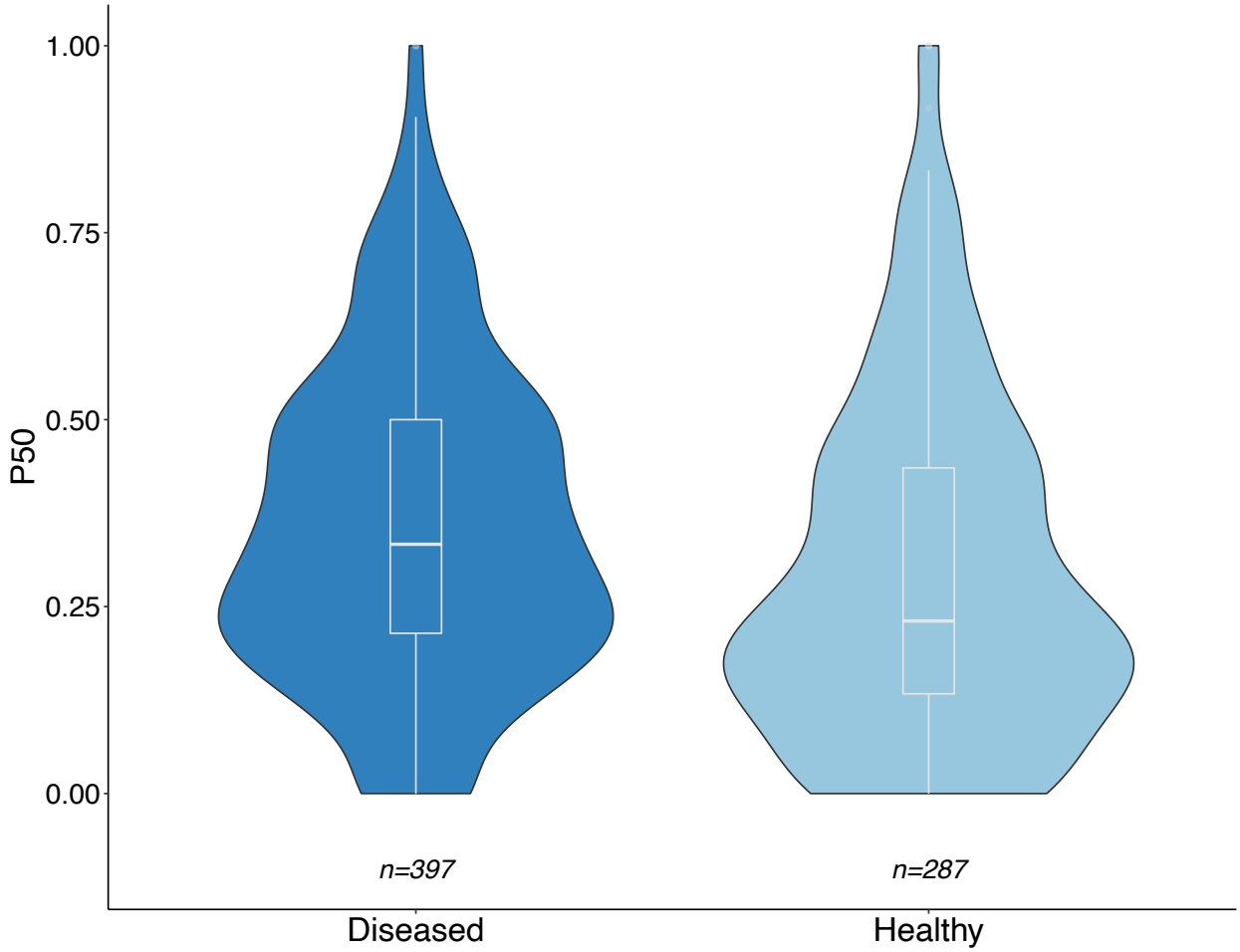


Fig. S10. Distribution of P50 from diseased and healthy animals. The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. n = 684 individual point prevalence surveys. One-way ANOVA p=6.05E-05.

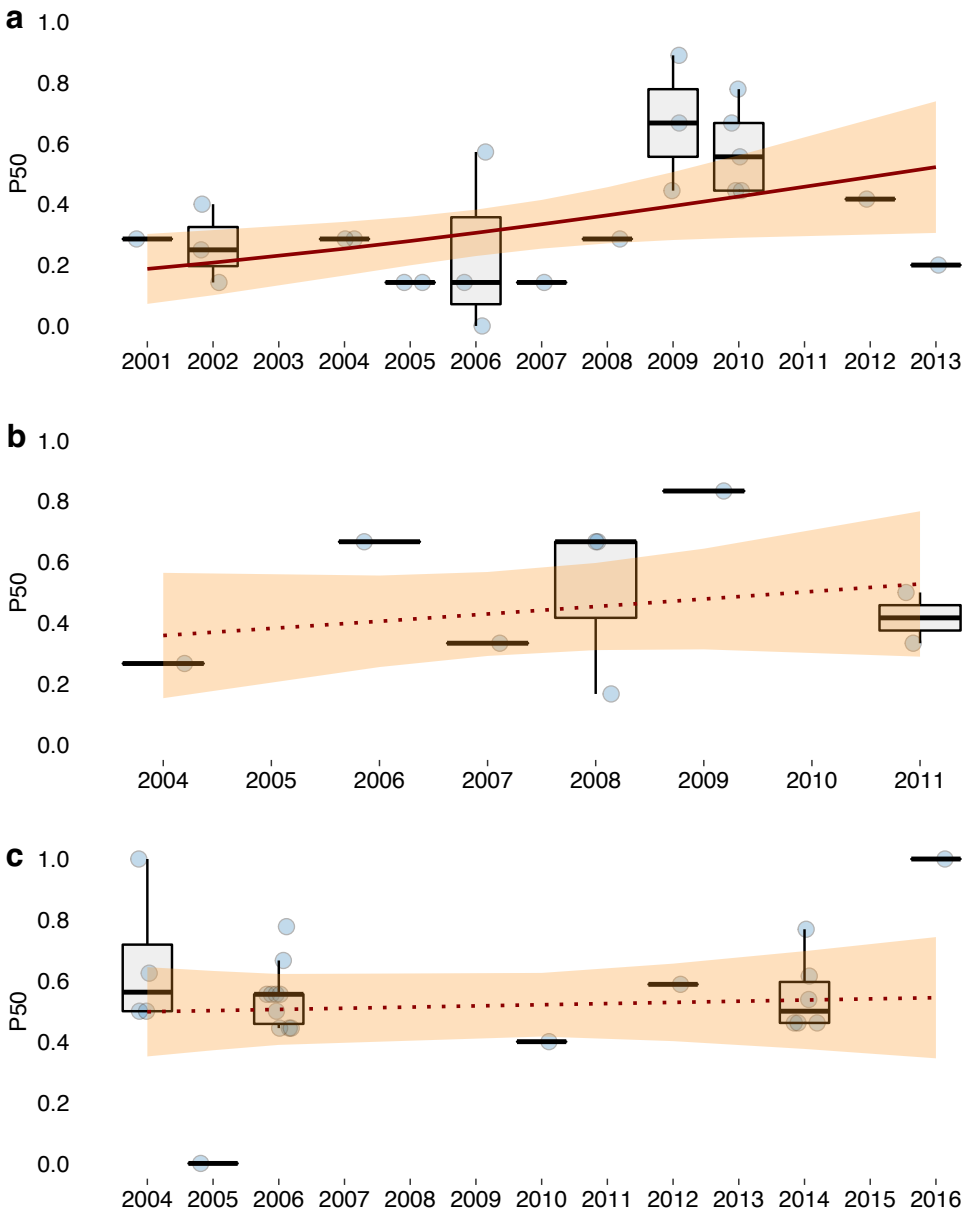


Fig. S11. Annual P50 trends for select pathogen-animal-country pairings. (a) Surveys of *Yersinia* spp. sampled from trout in Turkey (n = 23; p=.049); **(b)** Surveys of *Edwardsiella* spp. sampled from catfish in Vietnam (n = 9; p=.388); and **(c)** Surveys of Enterobacteriaceae sampled from fish in India (n = 24; p=.751). The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. Each survey is represented by a dot with horizontal jitter for visibility. Regression lines are fit using generalized linear model regressions, with a solid line in panel (a) indicating statistical significance (p=.049); 95% confidence intervals are shown in shaded areas.

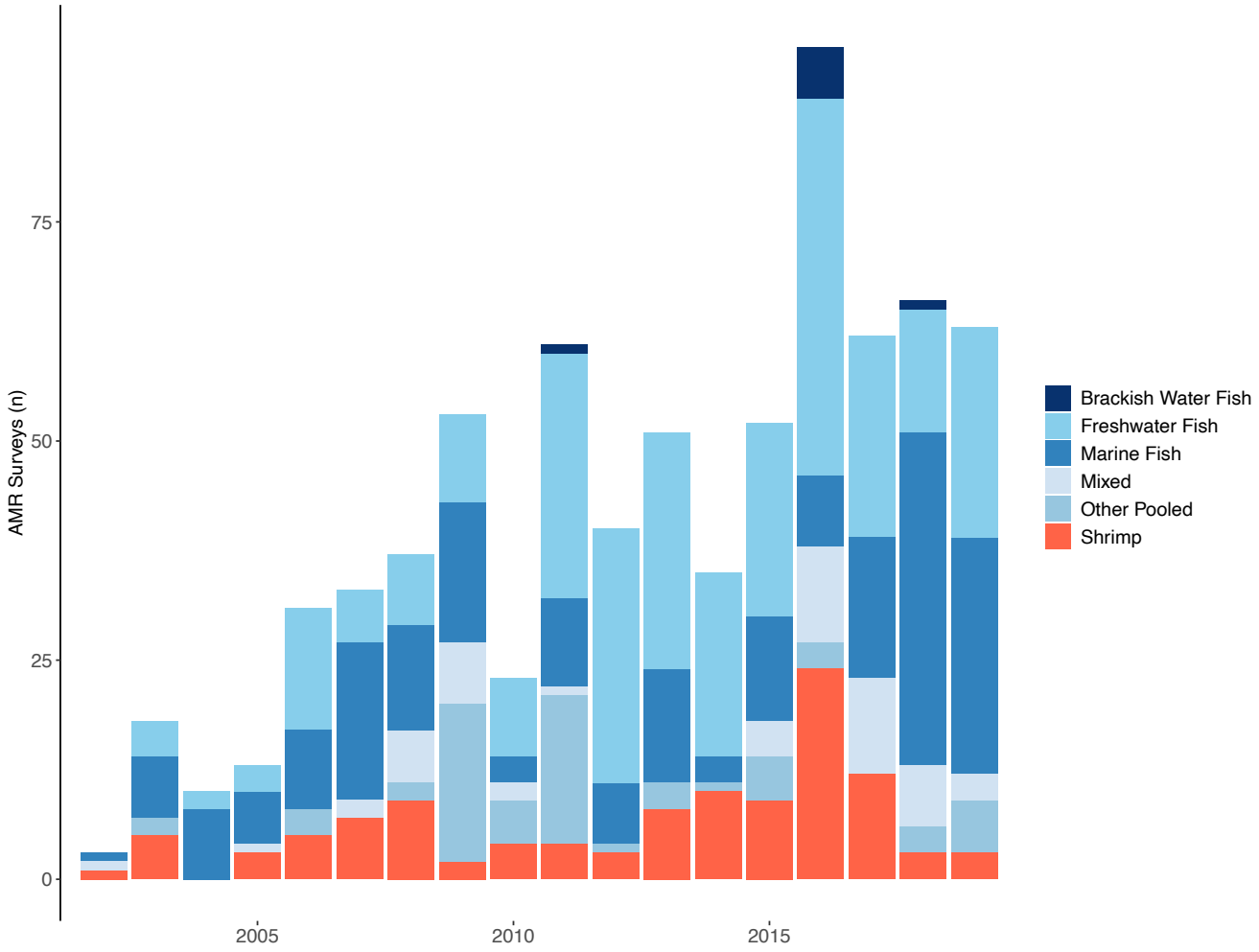


Fig. S12. **Species groups represented in resistance surveys from aquatic animals.** Year of publication is shown.

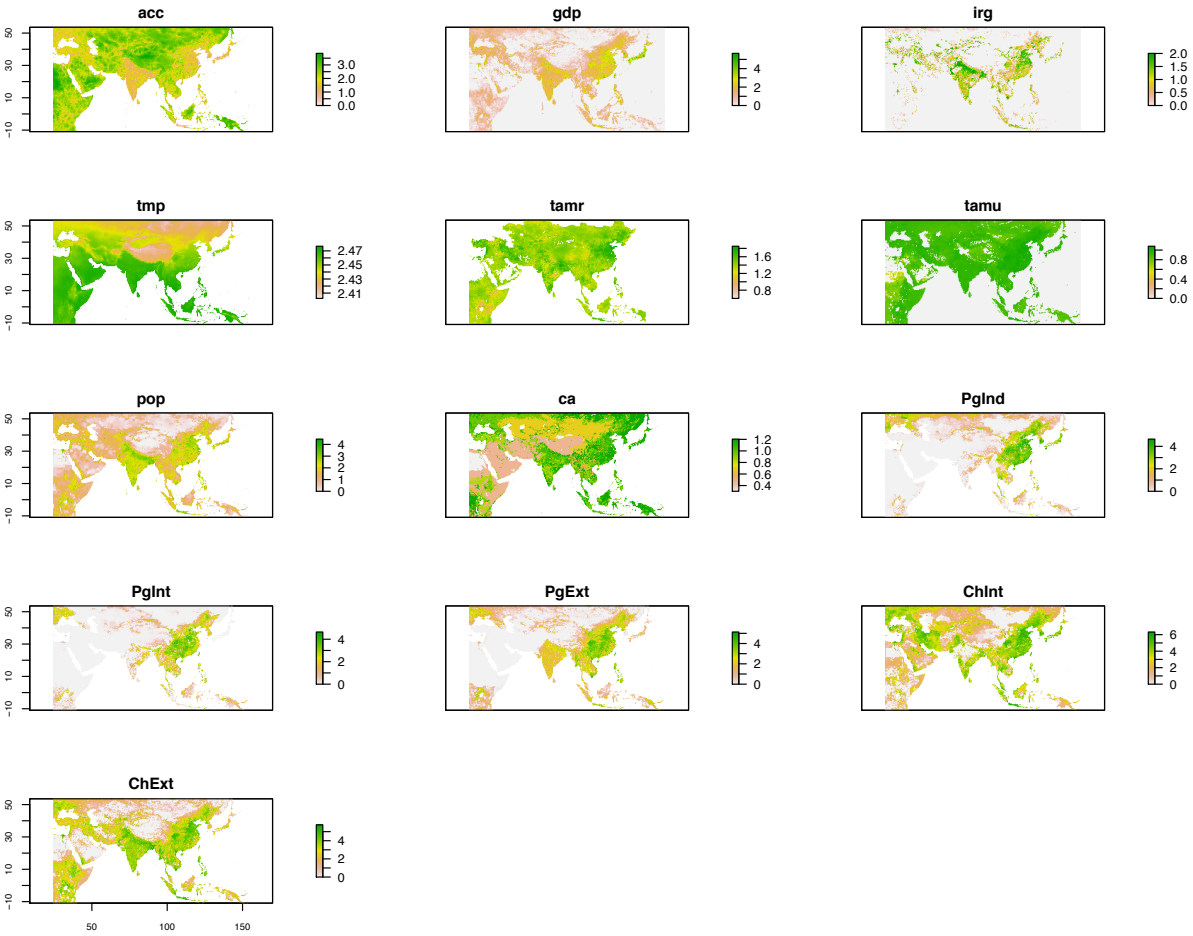


Fig. S13. Freshwater environmental and anthropogenic covariates used to train child models. All covariates are log10 scaled. Accessibility to cities (acc); gross domestic product (gdp); irrigated land percentage (irg); minimum monthly temperature (tmp); terrestrial livestock P50 (tamr); terrestrial livestock antimicrobial use (tamu); human population density (pop); and population densities of cattle (ca), pigs raised in intensively (PgInd), pigs raised semi-intensively (PgInt), pigs raised extensively (PgExt), chickens raised intensively (ChInt), and chickens raised extensively (ChExt). (See Table S2 for details)

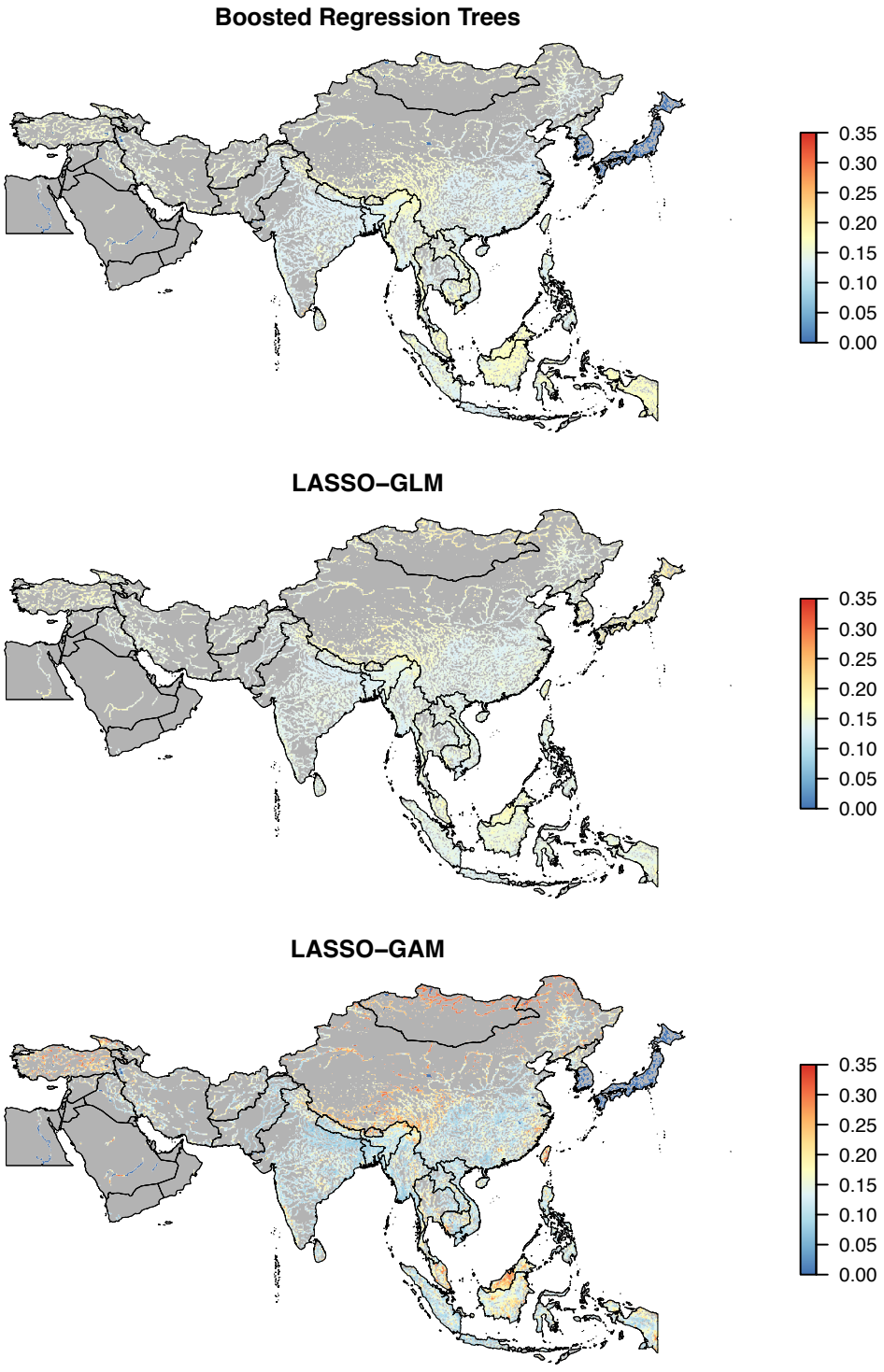


Fig. S14. Freshwater predicted P50 maps for each child model trained to environmental and anthropogenic covariates. Child models are boosted regression trees (BRT); least absolute shrinkage and selection operator applied to logistic regression (LASSO-GLM); and overlapped grouped LASSO penalties for General Additive Models selection (LASSO-GAM).

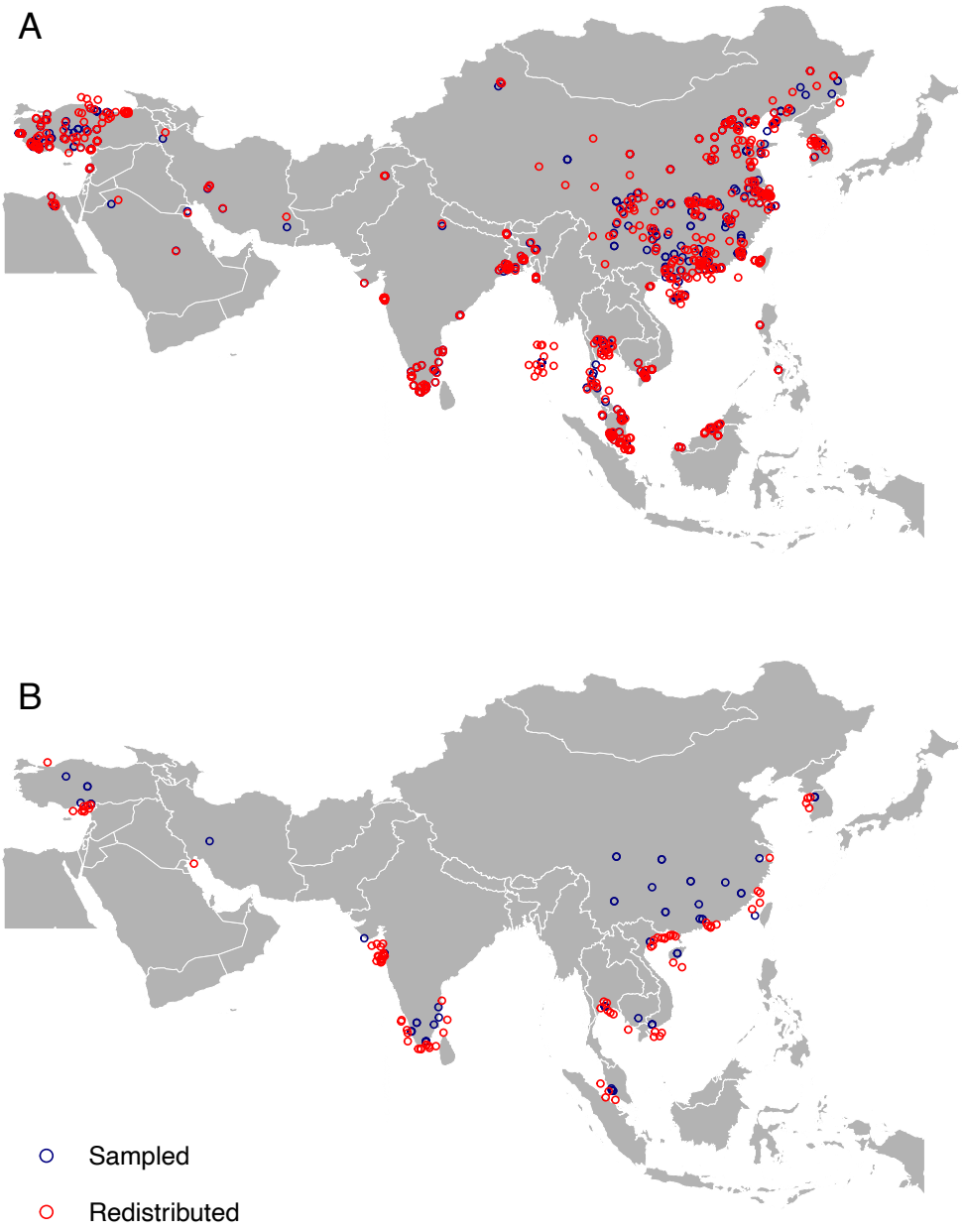


Fig. S15. Redistribution of surveys in (A) freshwater and (B) marine models. Freshwater surveys without precise sampling coordinates were redistributed at random within a geographic uncertainty range of the smallest available administrative unit. Surveys were redistributed on each of 10 bootstrap runs; a single bootstrap run is displayed. In some instances, given sampling locations (blue) are superimposed where multiple surveys were conducted at the same location. In the marine model (B), only post-harvest surveys were redistributed to open water (Supplementary Note 4, Marine Protocol) and are displayed.

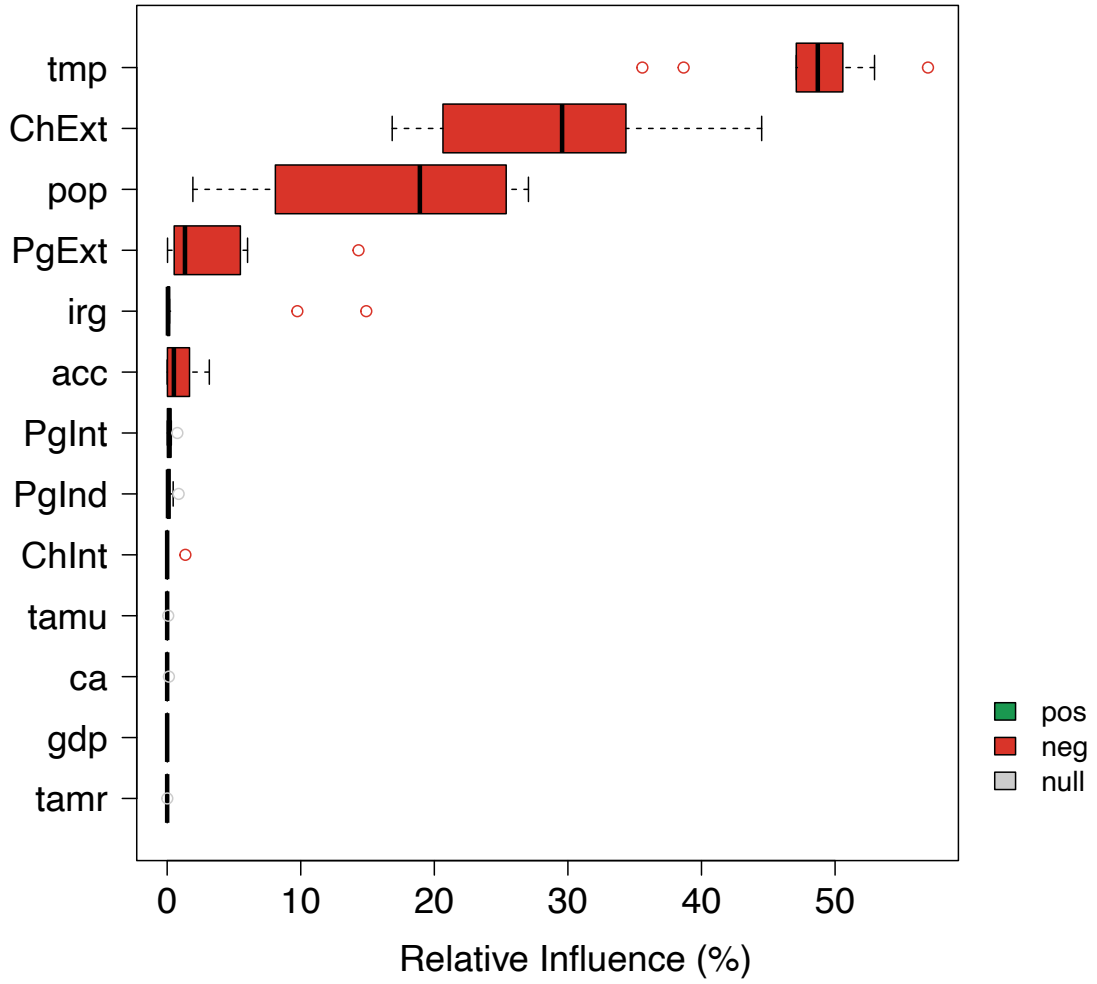


Fig. S16. Relative influence of covariates on resistance in freshwater boosted regression tree (BRT) model. Distributions of relative influence for each covariate in freshwater BRT models across 10 bootstrap runs (n = 10; for acronyms, see Supplementary Table S2). Coefficients after regularization from LASSO-GLM models were assigned as positive or negative for each covariate in each run. Positive, negative, and null associations with resistance for each covariate reflect the mean value of bootstrap runs where the coefficients were positive exceeding 30%, where the coefficients were negative, and where the coefficients were positive but in less than 30% of runs, respectively. The vertical box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. No positive associations (coefficients positive exceeding 30%) were identified.

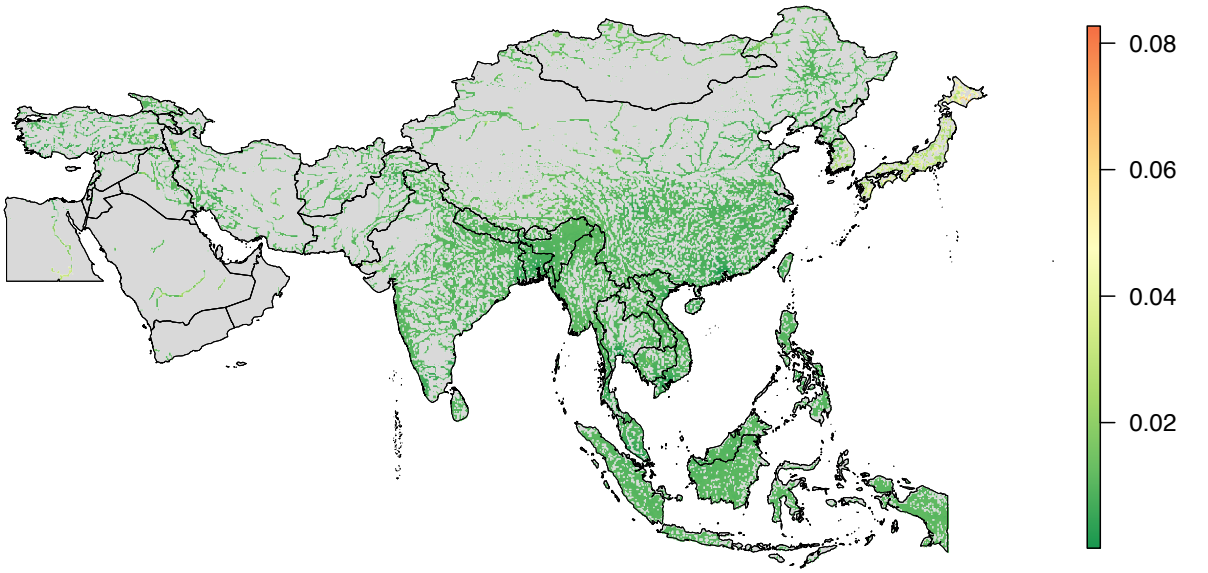


Fig. S17. Standardized kriging variance (spatial interpolation uncertainty) on P50 predictions in freshwater environments. Variance is standardized such that it equals zero at the location of observations.

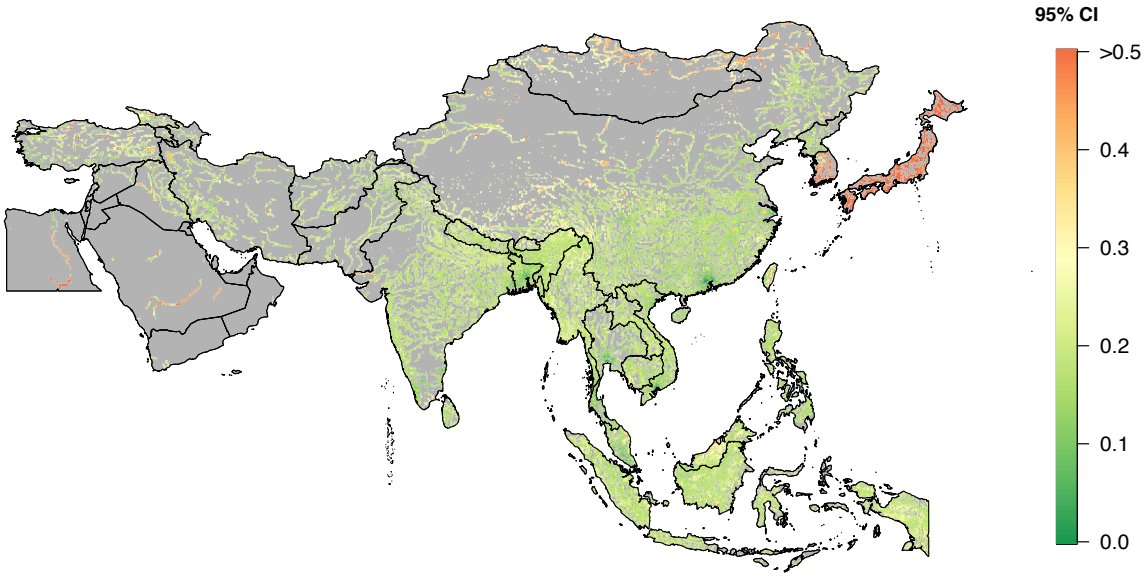


Fig. S18. 95% confidence interval on P50 predictions in freshwater environments.

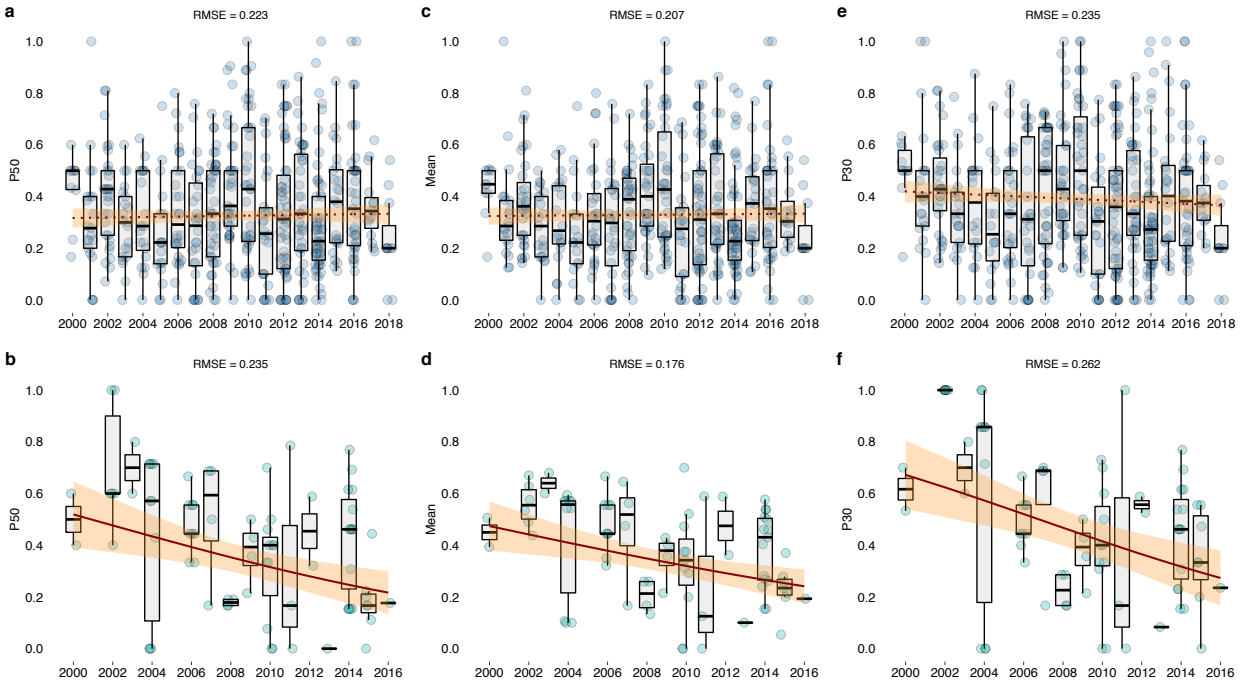


Fig. S19. Annual trends in multi-drug resistance. Surveys from cultured aquatic animals (top row, *n* = 558); surveys from wild caught aquatic animals (bottom row, *n* = 81). **(a,b)** the proportion of drugs with resistance greater than 50% (P50) in each survey; **(c,d)** mean resistance calculated as resistant isolates divided by the number of isolates * number of antibiotics tested in each survey; and **(e,f)** the proportion of drugs with resistance greater than 30% (P30) in each survey. The horizontal box lines represent the first quartile, the median, and the third quartile. Whiskers denote the range of points within the first quartile − 1.5× the interquartile range and the third quartile + 1.5× the interquartile range. Each survey is represented by a dot with horizontal jitter for visibility. Regression lines are fit using generalized linear model regressions, with a solid line indicating statistical significance (**(b)** p=0.003; **(d)** p=0.002; **(f)** p=0.001); 95% confidence intervals are shown in shaded areas. Goodness of fit is assessed using root mean square error (RMSE).

Table S1. Antimicrobial classes and corresponding compounds recommended by The Clinical & Laboratory Standards Institute (CLSI) for susceptibility testing of *Aeromonas* (M45)[19], *Streptococcus* (M100)[20], and *Vibrio* (M45), and by the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR) for *E. coli*.[21] *Compounds not represented in the dataset.

Antimicrobial Class	*Aeromonas* spp.	*E. coli*	*Streptococcus* spp. (β-Hemolytic /Viridans)	*Vibrio* spp.
Aminoglycosides	Amikacin Gentamicin	Gentamicin		Amikacin Gentamicin
Amphenicols	Chloramphenicol	Chloramphenicol	Chloramphenicol	Chloramphenicol
Carbapenems	Ertapenem Imipenem Meropenem	Imipenem Meropenem	Ertapenem Meropenem	Imipenem Meropenem
Cephalosporins	Cefepime Cefotaxime Cefoxitin Ceftazidime Ceftriaxone Cefuroxime	Cefepime Cefotaxime Cefoxitin Ceftazidime Ceftriaxone	Cefepime Cefotaxime Ceftaroline* Ceftriaxone	Cefazolin Cefepime Cefotaxime Cefoxitin Ceftazidime Cefuroxime
Glycopeptides			Vancomycin	
Glycylcyclines		Tigecycline*		
Lincosamides			Clindamycin	
Lipopeptides			Daptomycin*	
Macrolides		Azithromycin	Azithromycin Clarithromycin Erythromycin	Azithromycin
Monobactams	Aztreonam			
Nitrofurans		Nitrofurantoin		
Oxazolidinones			Linezolid	
Penicillins	Piperacillin-Tazobactam	Amoxicillin Ampicillin Temocillin*	Ampicillin Penicillin	Amoxicillin-Clavulanic Acid Ampicillin Ampicillin-Sulbactam Piperacillin Piperacillin-Tazobactam
Polymyxins		Colistin		
Quinolones	Ciprofloxacin Levofloxacin	Ciprofloxacin Nalidixic acid Pefloxacin	Levofloxacin Ofloxacin	Ciprofloxacin Levofloxacin Ofloxacin
Sulfonamides and dihydrofolate reductase inhibitors	Sulfamethoxazole-Trimethoprim	Sulfamethoxazole Sulfamethoxazole-Trimethoprim Sulfamonomethoxine Sulfisoxazole Sulfonamides Trimethoprim		Sulfadiazine Sulfamethoxazole Sulfamethoxazole-Trimethoprim Sulfamonomethoxine Sulfisoxazole Sulfonamides Trimethoprim

				Trimethoprim-Sulfadiazine
Tetracyclines	Tetracycline	Tetracycline	Tetracycline	Doxycycline Tetracycline

Table S2. Freshwater environmental and anthropogenic covariates used to train child models.

Name	Acronym	Year	Resolution	Source
Accessibility to Cities	acc	2015	30 arc second	Weiss D et al. Accessibility to Cities https://malariaatlas.org/research-project/accessibility_to_cities/
Gross Domestic Product	gdp	2010	30 arc second	World Bank https://datacatalog.worldbank.org/dataset/gross-domestic-product-2010
Irrigated Land Percentage	irg	2005	0.083333 decimal degrees	Siebert S et al. Global Map of Irrigation Areas http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version
Minimum Monthly Temperature	tmp	2020	2.5 minutes	WorldClim https://www.worldclim.org/data/worldclim21.html
Terrestrial Livestock P50	amr	2018	0.083333 decimal degrees	Van Boeckel T et al. https://resistancebank.org
Terrestrial Livestock AMU	amu	2017	0.083333 decimal degrees	Van Boeckel T et al.
Human Population Density	pop	2020	2.5 minutes	Center for International Earth Science Information Network - CIESIN - Columbia University. 2018. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11
Terrestrial Livestock Production	Ca PgInd PgInt PgExt ChInt ChExt	2010	0.083333 decimal degrees	Robinson TP et al. Global distribution of ruminant livestock production systems V5 (5 minutes of arc) Gilbert M et al. Global distribution of chickens and pigs raised in extensive, semi-intensive and intensive systems in 2010 (5 minutes of arc) http://www.fao.org/livestock-systems/production-systems/en/

Table S3. Root mean square error (RMSE) weightings in the marine ensemble model. Weights are calculated as the inverse of the RMSE of each constituent model divided by the sum of RMSE for all models [$\text{Weights} = 1/(RMSE_i/\sum_{i=1}^{3} RMSE_i)$], and expressed as their relative proportion.

	RMSE	Model weights
Inverse distance weighted	0.2359776	31.7
Natural neighbor	0.2255336	33.2
Ordinary kriging	0.2131811	35.1

Table S4. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 main checklist.

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Introduction, Line 105
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Line 94-108
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Line 105-106
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Line 392-396; SI Line 33-45
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Line 389-391; SI Line 50-55
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	SI Line 65-161
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	SI Line 45-48
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Line 400-411; SI Line 167-180; legend

Topic	No.	Item	Location where item is reported
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Line 400-411; SI Line 167-180; legend
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Line 407-411; SI Line 238-245; legend
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	SI Line 45-48
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Line 428-430; 452-453
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	Line 391-394
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Line 413-425
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Line 436-443; 452-453
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Line 436-440; 445-453; SI Line 259-321; 458
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Line 436-441

Topic	No.	Item	Location where item is reported
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Line 430-435; 348-351; SI Line 272-283
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Line 407-411; SI Line 45-48
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Line 441-443;461; 522-524
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Line 112-113; SI Line 163-165; Fig. S1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Supplementary data 2; Fig. S1
Study characteristics	17	Cite each included study and present its characteristics.	Supplementary data 2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	At screening/eligibility; Fig. S1
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Supplementary data 1; Fig. 1; Fig. 2
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Line 112-119
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Line 121-157
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Line 439-441
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	SI Line 272-283; Fig. S19

Topic	No.	Item	Location where item is reported
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Line 112-119; 359-366
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Line 122-157; 168-169; Fig. 1; Fig. 2; Fig. S18
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Line 191-203
	23b	Discuss any limitations of the evidence included in the review.	Line 330-362
	23c	Discuss any limitations of the review processes used.	Line 363-366
	23d	Discuss implications of the results for practice, policy, and future research.	Line 372-384
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	SI Line 55
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	SI Line 179-180
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	n/a
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Line 729-731
Competing interests	26	Declare any competing interests of review authors.	Line 738-739
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Supplementary data 1; Line 564-570

Supplementary References

1. International Office of Epizootics. *Aquatic animal health code.* (2019).
2. Page, M. J. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Med.* **18**, e1003583 (2021).
3. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, eaaw1944 (2019).
4. Schar, Daniel *et al.* Aquatic animal antimicrobial resistance point prevalence survey database legend. (2021) doi:10.5281/ZENODO.4609884.
5. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing of Bacteria Isolated From Aquatic Animals.* vol. CLSI supplement VET04. (Clinical and Laboratory Standards Institute, 2020).
6. U.S. FDA. The National Antimicrobial Resistance Monitoring System: Strategic Plan 2021 - 2025. (2020).
7. Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *The Lancet* **390**, 2171–2182 (2017).
8. Bhatt, S. *et al.* Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* **14**, 20170520 (2017).
9. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many?: *How to use pseudo-absences in niche modelling? Methods Ecol. Evol.* **3**, 327–338 (2012).
10. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
11. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).
12. Chouldechova, A. & Hastie, T. Generalized Additive Model Selection. *ArXiv150603850 Stat* (2015).

13. Hijmans, R. J. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* **93**, 679–688 (2012).
14. Lehner, B. & Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **296**, 1–22 (2004).
15. Messager, M. L., Lehner, B., Grill, G., Nedeva, I. & Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **7**, 13603 (2016).
16. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems: GLOBAL RIVER HYDROGRAPHY AND NETWORK ROUTING. *Hydrol. Process.* **27**, 2171–2186 (2013).
17. Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. H. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6**, 187–202 (2009).
18. World Bank. *People using at least basic sanitation services (% of population)*. data.worldbank.org.
19. *Methods for antimicrobial dilution and disk susceptibility testing of infrequently isolated or fastidious bacteria*. (Clinical and Laboratory Standards Institute, 2016).
20. WEINSTEIN, M. P. *M100-performance standards for antimicrobial susceptibility testing, 28th edition.* (CLINICAL AND LABORATORY, 2018).
21. WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance & World Health Organization. *Critically important antimicrobials for human medicine: ranking of antimicrobial agents for risk management of antimicrobial resistance due to non-human use*. (2017).

Supplementary Materials for

Predictive mapping of antimicrobial resistance of *Escherichia coli*, *Salmonella* and *Campylobacter* in food-producing animals in Europe between 2000 and 2021.

Authors:
Ranya Mulchandani[1,*], Cheng Zhao[1], Katie Tiseo[1], João Pires[1], Thomas P. Van Boeckel[1,2,*].

Affiliations:
[1]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland.
[2]One Health Trust, Washington DC.

*Correspondence to: ranya.mulchandani@usys.ethz.ch; thomas.vanboeckel@env.ethz.ch

Methods

Literature Review

We conducted a systematic literature review on antimicrobial resistance (AMR) prevalence in livestock and livestock products in Europe (Figure S1). We used three databases: PubMed, ISI Web of Science, and Scopus. Our original search focused on four pathogens commonly found in animals and their products*: Escherichia coli*, *Staphylococcus aureus, Campylobacter* spp., and non-typhoidal *Salmonella* spp. The searches were conducted at different time periods between May 2019 and January 2022, and included studies published between 2000 and 2021.

The general format for our literature queries was:

(Resistance) AND (Bacterial Species) AND (Animal/Sample Type) AND (Country).

The key words used for the literature review on PubMed, ISI Web of Science, and Scopus were: ("antibiotic resistance" OR "antimicrobial resistance" OR resistance OR resistencia OR "resistencia aos antibioticos" OR resistencia OR "resistencia a antibioticos" OR susceptibility OR susceptibilidade OR suscetibilidade OR antibiogram OR “antibiotic susceptibility testing” OR antibacteriano OR antibiotic OR antimicrobial OR antibiotic OR antibacterial OR antimicrobiano OR antibiograma OR antibiotic) AND (Escherichia OR “E. coli” OR coliform OR salmonella OR “salmonella spp.” OR “S. aureus” OR staphylococcus OR “Staphylococcus spp.” OR “MRSA” OR “MSSA” OR campylobacter OR “campylobacter spp.” OR “C. jejuni” OR “C. coli”) AND (animal OR food OR “food producing” OR meat OR cow OR cattle OR beef OR bovine OR buffalo OR pig OR piggeries OR pork OR “chicken” OR “flock” OR “broiler” OR “layer” OR “egg” OR “poultry” OR “avian” OR milk OR dairy OR cheese) AND (France OR Spain OR Netherlands OR Denmark OR Sweden OR Italy OR Greece OR Germany OR French OR Spanish OR Dutch OR Danish OR Swedish OR Italian OR Greek OR German OR Norway OR Norwegian OR Finland OR Finnish OR Poland OR

Polish OR "United Kingdom" OR England OR English OR Romania OR Romanian OR Bulgaria OR Bulgarian OR Iceland OR Icelandic OR Hungary or Hungarian OR Portugal OR Portuguese OR Austria OR Austrian OR Czechia OR "Czech Republic" OR Czechian OR Ireland OR Irish OR Lithuania OR Lithuanian OR Latvia OR Latvian OR Croatia OR Croatian OR Slovakia OR Slovakian OR Estonia OR Estonian OR Switzerland OR Swiss OR Moldova OR Moldovan OR Belgium OR Belgian OR "North Macedonia" OR Macedonia OR Macedonian OR Slovenia OR Slovenian OR Cyprus OR Luxembourg OR Malta OR Maltese). In PubMed, this query was put directly into the search bar. On Scopus, this search was conducted using TS=(key words given above), where TS stands for our search topic. In the ISI Web of Science, the search was conducted using TITLE-ABS-KEY=(key words given above). Here, TITLE-ABS-KEY stands for title, abstract, and key words.

In PubMed, Scopus, and ISI Web of Science, an initial search on eight European countries (Italy, Germany, the Netherlands, Spain, France, Greece, Denmark, and Sweden) was conducted on January 7, 2020, for Point Prevalence Survey (PPS) published between 2000 and 2019. These searches yielded 14,445 results. Titles and abstracts were screened manually. After removing duplications, reviews, meta-analyses, book chapters, and papers irrelevant to our topic of interest, we had 1,265 potentially relevant manuscripts. At this point, papers were read and removed if geographic data was unavailable, no antimicrobial susceptibility testing was performed, the study focused on sick animals, the survey focused on animals at the country-wide level, results were pooled between different animal species or sample types, or resistance prevalences were pooled between different pathogen types. From these, 191 papers were extracted, yielding 4853 resistance estimates.

Next, in PubMed, Scopus, and ISI Web of Science, a search for the remaining European countries was conducted on April 23, 2020, for PPS published between 2000 and 2019. This search yielded 54,591 results. Titles and abstracts were screened in the same manner as the first

search, and the same non-relevant results were removed. After this step, we had 745 potentially relevant manuscripts. From these, 98 were extracted, yielding 1567 resistance prevalences.

In PubMed, Scopus, and ISI Web of Science, a search for all European countries was conducted on January 7, 2021, for PPS published in 2020. This search yielded 6,005 results. Titles and abstracts were reviewed in the same manner as the previous two searches, and the same non-relevant results were removed. After this step, we had 253 potentially relevant manuscripts. From these, 34 were extracted, yielding 783 resistance prevalences.

A final literature search (identical to that of January 7, 2021) was run on January 10, 2022, to identify all PPS published in 2021. This search yielded an additional 6,598 results. As outlined previously, all the same steps for title and abstract screening were followed, leaving 110 potentially relevant manuscripts. Of these, 22 were extracted, yielding 606 resistance prevalences. Overall, this gave 345 papers with 7,809 resistance estimates of any antibiotic-pathogen combination.

As there was no mandated or routine reporting of *Staphylococcus aureus* to EFSA (there was only limited voluntary reporting of MRSA from 5 countries in 2018 and 6 countries in 2019) AMR estimates for *S. aureus* were subsequently excluded. Additionally, only countries reporting to EFSA were retained. The final number of manuscripts was 209.

Geographic localization of point prevalence surveys

Only PPS that reported geographic information were included in the study. The extracted information was recorded in "name_of_location" and "level_of_uncertainty" variables.

- "Name_of_location" contains the name(s) of the most precise location information available in the article. Where more than one location was reported, both names were recorded.

- "Level_of_uncertainty" contains the administrative level at which the sampling was performed (see Legend on resistancebank.org for full details). These data were then used to determine "Ycoord" and "Xcoord" variables.
- These data were then used to determine "Ycoord" and "Xcoord" variables. There were two ways in which these were generated:

1. **Samples taken from across an area/province** – the centroid of the province was obtained.
2. **Several sampling points across an area/region** – the middle point of all the sampling points was taken. This can be identified using variable "name_of_location" where more than one name is recorded.

Example:

DOI: 10.1155/2009/456573
Extract 1 from paper: "C. jejuni isolates were selected from a prevalence study of thermophilic campylobacters in livestock carried out in the Basque Country (Northern Spain)"
Extract 2 from paper: "…isolates were selected on the basis of isolation source (host, farm, and flock). Hence, the 72 isolates analysed by broth microdilution included 19 isolates from 12 poultry farms (18 flocks), 25 from dairy sheep (21 farms), and 28 isolates from cattle (14 beef cattle and 11 dairy cattle farms)"
Interpretation: Tested a specific subset of isolates from across Basque.
Level of uncertainty: Province
Name of location: Basque County
X/Y Coordinates: taken from the centroid of the Basque province.

Harmonization of antimicrobial resistance rates

The two most frequently used systems for antimicrobial susceptibility testing (AST) are Clinical and Laboratory Standards Institute (CLSI) and European Committee for Antimicrobial Susceptibility Testing (EUCAST). Each system uses breakpoints to classify susceptible and resistant phenotypes; these values are updated annually. Therefore, adjustment for breakpoint variation over time is essential.

In this study, we found 96% of records reported the guidelines used, while 72% of these records also reported the year of the guidelines used. The majority of records reported CLSI (52%), followed by EUCAST (29%), despite all studies originating from the European region. 4.4% of records did not report a guideline, and these records were excluded from subsequent analysis. The remaining records reported a mix of guidelines used in mentioned surveillance systems (e.g., DANMAP, NARMS, BSAC etc). For records where the guideline was reported, but no specific year, a date four years prior was assumed as this was the median lag between publication date of the survey and year of the guidelines. These assumptions were applied to maximise the amount of data retained for subsequent analyses.

The same harmonization procedure was then applied to all records as outlined in *Van Boeckel & Pires*, Science, 2019. This harmonization procedure resulted in 9% of records (262 out of 2888) being revised.

To assess the impact of using CBPs rather than ECOFFs, we changed the breakpoints used to ECOFFs rather than CBP, which resulted in 11% of the calculated P50s changing. Of these 38 P50s, the average absolute change was 18.9%. For these P50s, the majority (n=35) became larger, while five became smaller. Therefore, ~90% of the calculated P50s would remain the same if the breakpoints were changed, and the absolute change would be relatively small.

Desk review of national reports

We conducted a desk review of European countries to identify national reports that contain information on AMR in food-producing animals (Table S7). The contents of the reports were compared with EFSA, to determine if there was any further relevant data contained within these reports. Due to the limited additional data, with low comparability, these data were not extracted for this study.

Geospatial modelling of P50

During the first step, the P50 values (proportions) were transformed into presence and absence of resistance using a random binarization procedure. Concretely, each P50 value was duplicated 5 times, and compared with a random number between 0 and 1. P50 values higher than the random number were classified as presence of resistance, otherwise the values were classified as absence of resistance.

Sensitivity analyses and covariate importance

Sensitivity analyses were conducted by (a) restricting PPS to the same period as EFSA (2009-2020), (b) restricting to the six/seven most common drug-bug combinations and (c) by calculating P50 by class of drug rather than individual compound. For analysis (b), for *E. coli* and *Salmonella* the seven drugs included were: TET, AMP, SXT, CHL, CIP, GEN, CTX. The six most common drugs for *Campylobacter* were AMP, STR, GEN, CIP, TET, ERY.

The importance of covariates was calculated by sequentially removing each covariate from the modelling procedure and comparing the changes in the mean AUC across 10 Monte Carlo simulations.

Results

Descriptive analysis

A total of 81,639 records were identified from the literature search (Figure S1). Following de-duplication, title, abstract and subsequent full paper screen, a total of 209 studies with geographic information had data extracted. From 209 PPS where geographic information was reported, 2,849 AMR estimates were extracted, providing 368 P50s.

From the EFSA reports, 2,996 P50s at country-level (33,802 AMR estimates) were calculated from data collected between 2009 and 2020. The numbers of countries reporting to EFSA each year ranged from 23 in 2009, 20 in 2011, to 31 countries reporting annually from 2015 onwards.

Figures

Figure S1: PRISMA Flow Diagram

Figure S2: Geospatial modelling framework

Figure S3: (a) Prediction uncertainty calculated from the variation of predicted P50 values across the ten bootstraps **(b)** Hotspot map for 31 countries (light blue indicates hotspot areas, the top 95% percentile) **(c)** Standard deviation in P50 estimates per country

Figure S4: Sensitivity analyses of geospatial modelling for (a) date restriction to 2009-2020 only (b) 6-7 most common drug-bug combinations and (c) P50 by class of drug rather than individual compound.

Tables

Table S1: Suggested antimicrobials, by bacteria, for inclusion for antimicrobial susceptibility testing (AST) for surveillance of AMR in foodborne bacteria[1].

Antimicrobial classes	***Salmonella, E. coli***	***Campylobacter***
Aminoglycosides	Gentamicin	Gentamicin Streptomycin
Amphenicols	Chloramphenicol	
Carbapenems	Imipenem Meropenem	
Cephalosporins II	Cefoxitin	
Cephalosporins III	Cefatoxime (or Ceftriaxone) Ceftazidime	
Cephalosporins IV	Cefepime	
Glycopeptides		
Glycylcyclines	*Tigecycline*	
Lincosamides		*Clindamycin*
Lipopeptides		
Macrolides	Azithromycin	Erythromycin*
Nitrofurans	*Nitrofurantoin*	
Oxaxolidinones		
Penicillins	Ampicillin Amoxicillin *Temocillin*	Ampicillin
Polymyxins	Colistin	
Quinolones	Ciprofloxacin Nalidixic acid *Pefloxacin*^	Ciprofloxacin *Nalidixic acid*
Rifamycins		
Streptogramins		
Sulfonamides	Sulfisoxazole[#]	
Tetracyclines	Tetracycline	Tetracycline[~]
Trimethoprim	Trimethoprim	

Antimicrobials italicized are second priority

* Resistance towards erythromycin reflects azithromycin resistance

^ To screen for ciprofloxacin resistance in *Salmonella* spp. when disk diffusion is used.

[#] Trimethoprim-sulfamethoxazole can be used instead of using sulfisoxazole or trimethoprim alone

[~] Doxycycline may be used instead of tetracycline

[1] World Health Organization, "Integrated surveillance of antimicrobial resistance in foodborne bacteria: application of a one health approach: guidance from the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR)," World Health Organization, Geneva, 2017

Table S2: Environmental and anthropogenic covariates use to train child models

Name	Acronym	Year	Original Resolution	Source	Unit
Travel time to cities	acc	2015	30-arcsec resolution	Weiss et al 2018 (1) https://www.map.ox.ac.uk/accessibility_to_cities/.	minute
Antimicrobial use in animals	use	2013	0.083333 decimal degrees	Van Boeckel et al 2017 (2) http://science.sciencemag.org/content/357/6358/1350.full	Log10[(mg/pixel)+1]
Yearly average of minimum monthly temperature	tmp	1970-2000	2.5 minutes	Worldclim (3) http://worldclim.org/version2	°C * 10
Percentage irrigated areas	irg	2015	0.083333 decimal degrees	Global Map of Irrigation Areas (GMIA) (4) https://zenodo.org/record/6886564#.YuZ1HS8RpN0	%
Population density of cattle, chickens, pigs, and sheep (GLW version 4)	ca_v4 ch_v4 pg_v4 sh_v4	2015	0.083333 decimal degrees	Gridded Livestock of the World v4 (5) (https://www.nature.com/articles/sdata2018227)	Log10[(Heads/pixel) +1]
Percentage of tree coverage	veg	2013	0.008333 decimal degrees	Hansen et al 2013 (6) https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.2.html	%

1. D. J. Weiss, *et al.*, A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
2. T. P. Van Boeckel, *et al.*, Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences* **112**, 5649–5654 (2015).
3. S. E. Fick, R. J. Hijmans, WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology* **37**, 4302–4315 (2017).
4. S. Siebert, *et al.*, Development and validation of the global map of irrigation areas (2005).
5. Gilbert, M., Nicolas, G., Cinardi, G. et al. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Sci Data* 5, 180227 (2018). https://doi.org/10.1038/sdata.2018.227
6. M. C. Hansen, *et al.*, High-resolution global maps of 21st-century forest cover change. *science* **342**, 850–853 (2013).

Table S3: Absolute difference between resistance prevalence for antimicrobials by data source (point prevalence survey (PPS) vs European Food Safety Authority (EFSA)) between 2018 and 2020, and their WHO designation of antimicrobial importance*.

	E. coli	*Salmonella*	*Campylobacter*	WHO Grouping
AMP	19.11	8.27	-	Critically important
CAZ	10.41	-	-	Critically important
CHL	7.51	15.2	-	Highly important
CIP	-0.02	-17.4	1.71	Critically important
CST	10.66	3.41	-	Critically important
CTX	10.41	1.87	-	Critically important
ERY	-	-	12.1	Critically important
FOX	-7.96	-	-	Highly important
GEN	9.09	0.96	7.66	Critically important
IPM	5.53	-	-	Critically important
NAL	8.7	-2.35	0.78	Critically important
STR	-	-	4.99	Critically important
TET	16.07	3.81	1.87	Highly important

*A ratio <1 indicated a lower 3-year mean P50 using PPS data, and a ratio >2 meant a more than double 3-year mean P50 from PPS data compared to EFSA.

Table S4: Comparison between maps produced using all extracted data, maps produced using restricted number of drugs, maps produced when P50 is calculated by class of drug (rather than individual compound), and maps produced using only surveys published between 2009 and 2020. Mae: mean absolute error; Cor: Pearson correlation coefficient.

	E. coli	*Salmonella*	*Campylobacter*
(a) Restricted by year (2009-2020)			
No. of surveys	123	66	74
Mae	0.85%	0.75%	1.4%
Cor	0.994	0.986	0.995
(b) Restricted pathogen-antimicrobial combinations			
No. of surveys	153	97	111
Mae	1.5%	0.46%	0.63%
Cor	0.984	0.994	0.999
(c) P50 calculated at class level			
No. of surveys	156	99	113
Mae	1.0%	0.5%	0.49%
Cor	0.992	0.993	0.999

Table S5: Importance of covariates for mapping the distribution of AMR, indicating mean AUC of the full model, and the decrease in mean AUC after each covariate was removed from the modelling procedure.

	E. coli	*Salmonella*	*Campylobacter*
Full model	0.635	0.606	0.536
Travel time to cities	0.03	0.001	0.02
Antimicrobial use in animals	0.019	0.037	0.037
Yearly average of minimum monthly temperature	0.033	0.042	0.034
Percentage irrigated areas	0.016	0.034	0.027
Population density of cattle	0.024	0.036	0.023
Population density of chicken	0.025	0.041	0.023
Population density of pigs	0.017	0.03	0.032
Population density of sheep	0.03	0.029	0.032
Percentage of tree coverage	0.106	0.078	0.024

Table S6: Desk review of European countries reporting AMR in zoonotic and foodborne bacteria (2007-2020)

	Country	**EU Member State**	**Reporting to EFSA***	**PPS extracted^**	**National level integrated surveillance[2]** Comparison of report content & frequency of data reporting compared to EFSA reports
1	Albania		✓		
2	Austria	✓	✓	✓	
3	Belgium	✓	✓	✓	
4	Bulgaria	✓	✓	✓	
5	Cyprus	✓	✓		
6	Czech Republic	✓	✓	✓	
7	Denmark	✓	✓	✓	✓ DANMAP (Last accessed 21 Feb 2022). Established 1995. Pathogen & animal types: • *Salmonella Typhimurium* (pig) • *Salmonella Derby* (pig) • *Campylobacter jejuni* (chicken, cow) • *E. coli* (chicken, cattle) Reporting content: Same pathogens & animal types are reported to EFSA each year, with the same sample sizes. Additional data available: Last published report contains 2020 data.
8	Estonia	✓	✓	✓	

[2] R. Schrijver , M. Stijntjes and J. Rodríguez-Baño, "Review of antimicrobial resistance surveillance programmes in livestock and meat in EU with focus on humans," *Clinical Microbiology and Infection,* vol. 24, no. 6, pp. 577-590, 2018.

9	Finland	✓	✓	✓	✓ FINRES-VET (Last accessed 21 Feb 2022) Pathogen & animal types: • *Campylobacter jejuni* (chicken, cow) • *Salmonella* spp. (pooled animal types) Reporting content: Report less in report than to EFSA. Additional data available: Last published report contains 2020 data.
10	France	✓	✓	✓	✓ ONERBA (Last accessed 21 Feb 2022) Pathogen & animal types • *E. coli* (cattle, turkeys, pig) Reporting content: Only report mandatory data (e.g., in 2018 only reported on turkeys & chickens, and for the requested sample size). The ONERBA published report in 2018 contains larger samples sizes & contain additional data on pigs. Additional data available: Last published report only contains 2018 data, however historical reports contain additional animal types & larger sample sizes compared to EFSA.
11	Greece	✓	✓	✓	
12	Germany	✓	✓	✓	**?** GERMAP (no report publicly available since 2015)
13	Croatia	✓	✓		
14	Hungary	✓	✓	✓	
15	Iceland		✓		
16	Ireland	✓	✓	✓	
17	Italy	✓	✓	✓	
18	Lithuania	✓	✓		
19	Luxembourg	✓	✓		
20	Latvia	✓	✓		

21	Malta	✓	✓		
22	The Netherlands	✓	✓	✓	✓ MARAN (NETHMAP) (Last accessed 21 Feb 2022) Pathogen & animal types: • *Salmonella* spp. (pooled animal types) • *E. coli* (pigs, chicken, cow, turkey) Reporting content: Only trends reported in prose. Data not in an extractable format. Additional data available: N/A – no extractable data available.
23	Norway		✓		✓NORM-VET (Last accessed 21 Feb 2022) Pathogen & animal types • *Salmonella* spp., but animals are pooled • *Campylobacter jejuni* and *Campylobacter coli* (chicken, turkey, pigs) • *E. coli* (chicken, turkey, cattle, pigs, goats) Reporting content: Animal types differ year-on-year, in-line with EFSA requirements; report same sample sizes. E.g., in 2020, reported *E. coli* in chicken and turkeys while in 2019, reported *E.coli* in cattle and pig. Additional data available: Last published report contains 2020 data.
24	Poland	✓	✓	✓	
25	Portugal	✓	✓	✓	
-	Republic of North Macedonia		✓		*Excluded from geospatial analysis due to small numbers in EFSA data.*
26	Spain	✓	✓	✓	

27	Sweden	✓	✓	✓	✓ SVARM (Report - SWEDRES) (Last accessed 21 Feb 2022) Pathogen & animal types: • ESBL-producing *E.coli* (chicken) (no AST) • *Salmonella* spp. (pooled animals) • *Campylobacter jejuni* (chicken) and *coli* (pig) Reporting content: Report *Campylobacter* in-line with EFSA requirements. Additional data available: Last published report contains 2020 data.
28	Switzerland		✓	✓	ARC-Vet (Last accessed 21 April 2022) Pathogen & animal types • E. coli (pig, cattle) • Campylobacter coli (pig) Reporting content - Additional data available Last published report contains 2019 data.
29	Slovenia	✓	✓		
30	Slovakia	✓	✓	✓	
31	The United Kingdom	✓	✓	✓	✓ VARSS Pathogen & animal types: • *E. coli* (chicken, turkey, pigs) • *Salmonella* (chicken, turkey) • *Campylobacter jejuni* (chicken, turkey) Reporting content: Animal types & pathogen in-line with EFSA requirements, with same sample size. Additional data available: Last published report contains 2020 data.

*Last published report in April 2022 contains data from 2019/2020
^Where at least one PPS extracted per country, either published or data collected between 2000 and 2021

Supplementary Information for

Antimicrobial resistance in food animals: priority drugs maps to guide global surveillance.

Authors:
Cheng Zhao[1], Yu Wang[1], Ranya Mulchandani[1], Thomas P. Van Boeckel[1,2,*]

Affiliations:
[1]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland.
[2]One Health Trust, Washington DC.

*Correspondence to: thomas.vanboeckel@env.ethz.ch

Supplementary Methods

Literature review and data extraction

We searched for point-prevalence surveys (PPS) published between 2000 and 2019 reporting antimicrobial resistance in healthy food animals in low- and middle-income countries, focusing on *Escherichia coli* and nontyphoidal *Salmonella* spp.. The literature search was conducted in three rounds from four databases - PubMed, Scopus, ISI Web of Science, and China National Knowledge Infrastructure. The first round was conducted on 28.03.2019 from the first three aforementioned databases, and extracted data from all papers published between January 2000 and December 2018. The extracted data and details of literature review were published in Van Boeckel and Pires *et al.* 2019. The second round of literature search was conducted on 11.03.2020 from all four databases, and included surveys published between January 2000 and December 2019 exclusively for China. The extracted data and details of literature review were published in Zhao *et al.* 2020. The third round of literature search was conducted on 12.01.2022 from the first three aforementioned databases, and included all papers published between January 2019 and December 2019 in low- and middle-income countries apart from China. The search queries used for the third round of literature review was the same as in Van Boeckel and Pires *et al.* 2019.

All three rounds of literature review were conducted with the following procedure (Supplementary Table 1). First, we screened in total 44,325 titles and abstracts, and excluded 40,702 non-PPS publications. We read 3,623 manuscripts in full, and excluded strain surveys, surveys on diseased animals, surveys conducted on a mixture of animal species, surveys without subnational geographic information, and other non-PPS surveys. After the exclusion, there were 1,360 PPS suitable for AMR mapping purposes. We further excluded animal species with small sample sizes such as camel and buffalo, and excluded drug-pathogen combinations not considered in this analysis such as *Campylobacter* and erythromycin. After the exclusion, 1,088 PPS that reported resistance prevalence in *Escherichia coli* and nontyphoidal *Salmonella* spp to 7 antimicrobials (listed in Methods section) were retained for

the analyses. All data used in the current analyses are available in the supplementary file, and can also be downloaded at https://resistancebank.org.

Antimicrobial susceptibility testing in the PPS was conducted using either diffusion methods or dilution methods. The majority of PPS used diffusion methods, including disk diffusion (79%) and E-test (0.2%). The rest of PPS used dilution methods, including broth dilution (14%), agar dilution (5%), and automated devices such as VITEK2 (2%). Among the PPS, there was no systematic difference in the measurements between these two families of methods[1]. In each PPS, antimicrobial susceptibility testing results are compared with breakpoints to determine resistance, which are provided by laboratory guidelines and revised annually. Only 18% of records reported the breakpoints used. However, the majority (93%) of PPS mentioned the name of laboratory guidelines used, and 66% among these also mentioned the year of the guideline. The guidelines mentioned by the PPS included guidelines published by the Clinical & Laboratory Standards Institute (96%), the European Committee on Antimicrobial Susceptibility Testing (3%), and the French Society of Microbiology (1%). We adjusted for variations of breakpoints used between surveys, using a method developed by Van Boeckel and Pires *et al.* 2019 in section "Harmonization of Antimicrobial Resistance Rates" in the Supplementary Material of the reference publication[1]. The adjustment resulted in 635 (2%) resistance prevalence being revised.

Imputation of missing data on resistance prevalence for mapping priority antimicrobials

Missing resistance prevalence data in the point-prevalence surveys were imputed using Multivariate Imputation by Chained Equations (MICE)[2]. Using MICE, a set of plausible values for the missing resistance prevalence could be inferred from the distribution of reported resistance prevalence data, using specified imputation models. The prediction accuracy of three imputation models were compared: Bayesian linear regression (BLR), LASSO regression (LASSO-GLM), and feed-forward neural network (NN). For NN, we selected the optimal combination of hyperparameters, by comparing the root-mean-square error (RMSE) of the imputed values created using NN models with 500 different hyperparameters. These hyperparameters were drawn randomly from the following ranges: the number of nodes of the hidden layer between 1 to 272, dropout rate between 0.2 and 0.8, and learning rate between 0.00001 and 0.1. The lowest value of RMSE was generated with 145 nodes on the hidden layer, a dropout rate of 0.4, and a learning rate of 0.0001.

The comparison of imputation methods was conducted as following. First, we selected a subset of 272 surveys, which contained no missing values of resistance prevalence for the 7 antimicrobials listed in the Methods section. Second, we conducted 50 Monto Carlo simulations to estimate the accuracy of each imputation method. Concretely, for each simulation, we randomly removed 2 out of 7 reported antimicrobial resistance prevalence in each survey, and conducted 4-fold spatial cross validation to impute these deleted values back. These 4 spatial folds were determined based on the continents of the survey locations: America, Africa, western Asia, and eastern Asia. Finally, we compared the RMSE of the imputed missing values of each fold for all Monto Carlo simulations, by running MICE with different imputation methods. The prediction accuracy of LASSO-GLM (RMSE 26.6) outperformed BLR (RMSE 28.7) and NN (RMSE 27.0). Additionally, adding an ad-hoc step of predictive mean matching, and including additional covariates in the imputation process did not improve the prediction accuracy.

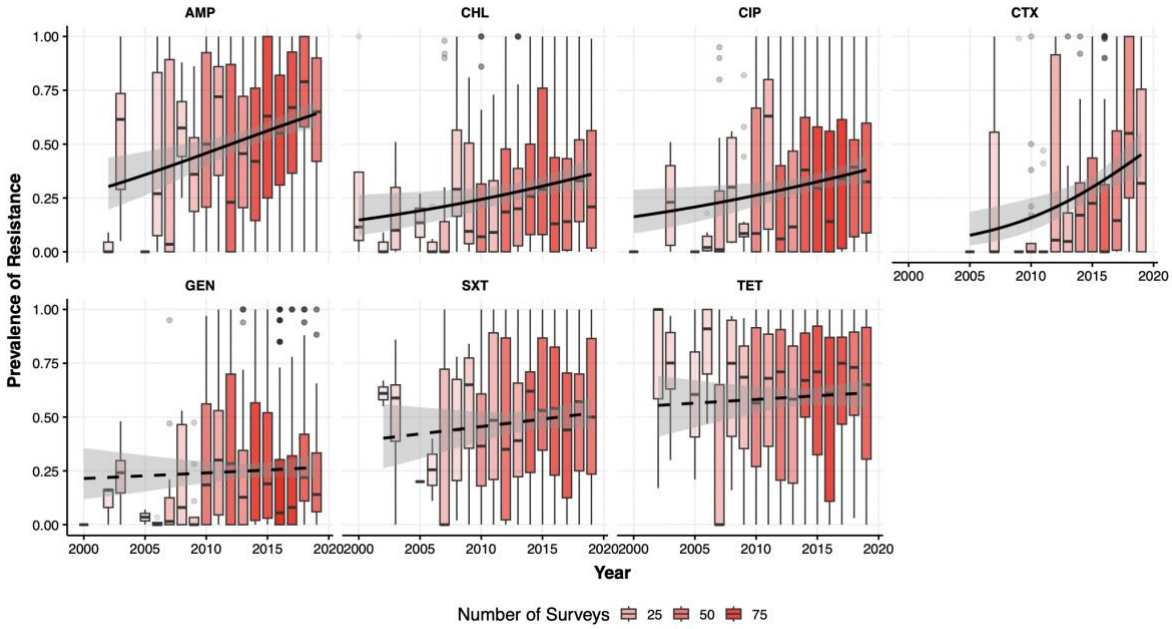
We conducted imputation on a subset of 806 PPS that reported at least 4 out of the 7 drugs. These PPS contained 1,411 resistance profiles – some PPS reported resistance profiles for multiple animal species or for multiple sample types. Using MICE combined with LASSO-GLM, we imputed 2,117 (21%) missing values out of 9,877 resistance prevalence. The number of imputed resistance prevalence was 720 for cefotaxime (51%), 375 (27%) for sulfamethoxazole-trimethoprim, 306 (22%) for chloramphenicol, 202 (14%) for ampicillin, 196 (14%) for tetracycline, 195 (14%) for ciprofloxacin, and 123 (9%) for gentamicin. We conducted 10 multiple imputations, each with 25 iterations.

Mapping resistance prevalence for each antimicrobial

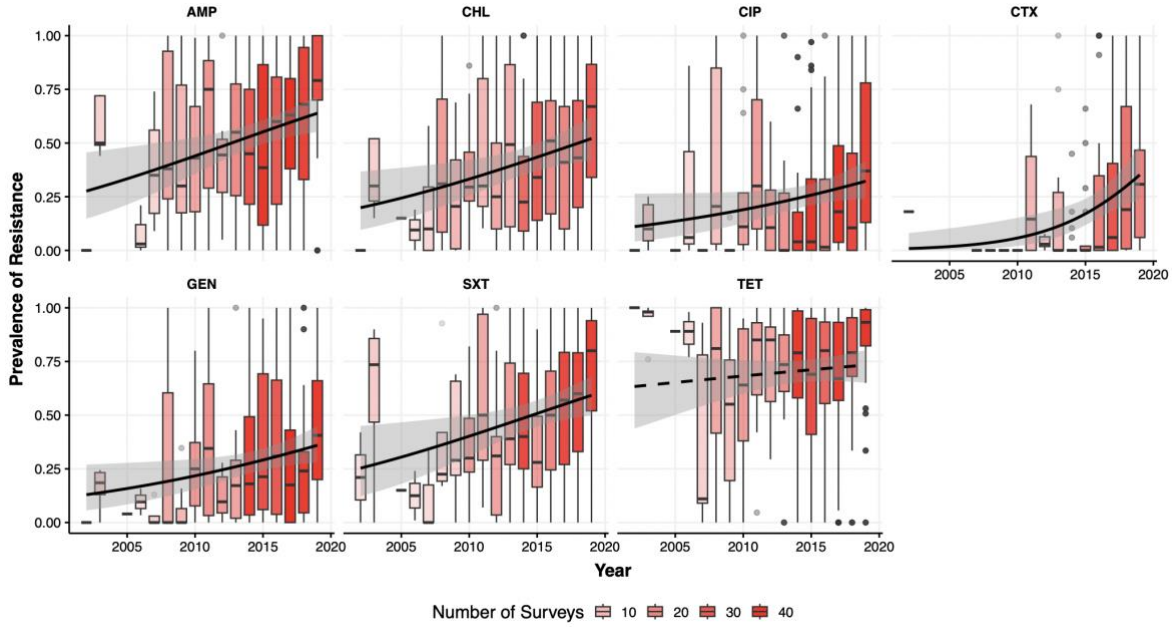
We mapped the prevalence of resistance for each antimicrobial using Gaussian process stacked generalization[3]. The mapping procedure included two steps. In the first step, we trained three 'child models' to predict resistance prevalence based a set of environmental and anthropogenic covariates (Supplementary Table 3). For each antimicrobial, we also included its estimated amount of use divided by the estimated biomass of food animals in 2020[4] as a covariate in the corresponding child models. The child models included boosted regression trees[5] (BRT), least absolute shrinkage and selection operator applied to linear regression[6] (LASSO-GLM), and feed-forward neural network implemented in Keras[7] (FFNN). The models were trained using four-fold spatial-cross validation (Supplementary Figure 15). For the BRT model, we applied a tree complexity of 3 with 50 initial trees, a learning rate of 0.0005, and a step size of 50. For the NN model, we applied one hidden layer with 31 nodes, a dropout rate of 0.49 and a learning rate of 0.01, using adaptive moment estimation optimizer, and the rectified linear activation function for each layer.

In the second step, the child model predictions were stacked using Gaussian process regression, fitted using the integrated nested Laplace approximations (INLA)[8]. This second step allowed to simultaneously capture the influence of environmental and anthropogenic covariates, as well as the residual spatial correlation. INLA is a deterministic method for Bayesian inference in latent Gaussian modelling, and is comparatively faster than other inference methods such as Markov chain Monte Carlo. The INLA formula included the child model predictions of resistance prevalence as fixed effects, and the spatial autocorrelation as a random effect. The coefficients of the fixed effects were constrained between 0 and 1, such that the coefficients approximately sum to one[3]. The residual spatial correlation was modelled as a Gaussian Markov random field (GMRF) with a Matern covariance function. The prior of the range for the covariance function was set at 4.06 decimal degrees, or roughly 487 km at equator, based on previous work on spatial correlation of AMR[9]. We constructed the mesh – on which the GMRF representation was built – using a cutoff of 0.005 decimal degrees, a maximum edge of 1 and 4 decimal degrees, an offset of 0.25 and 1.5 decimal degrees, for the inner domain and outer extension respectively.

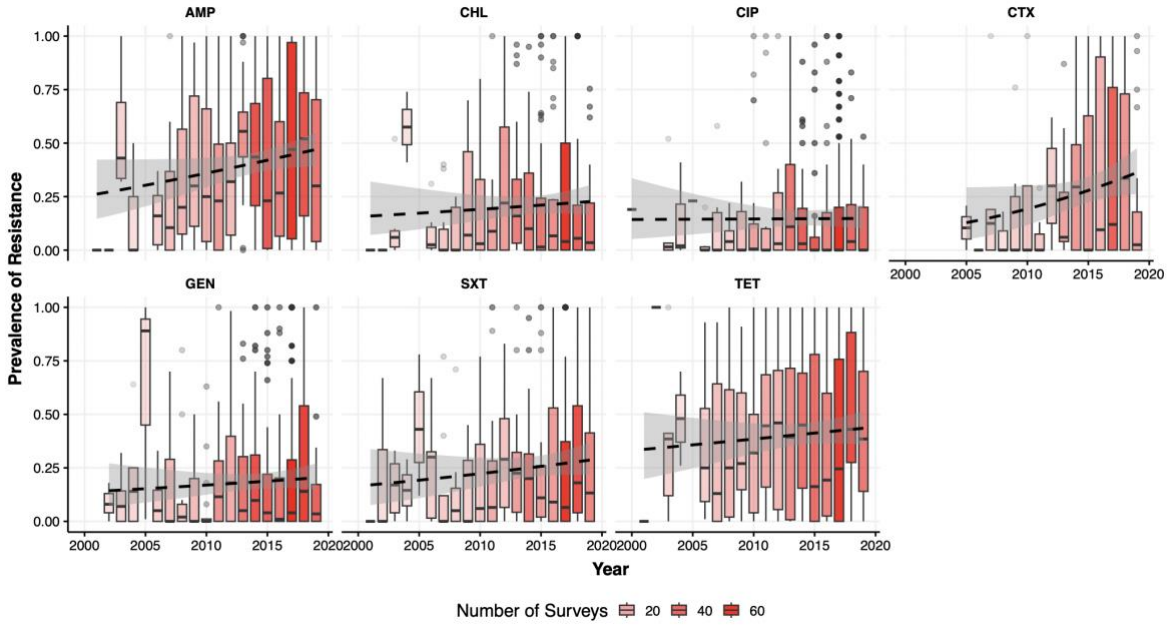
Supplementary Figures



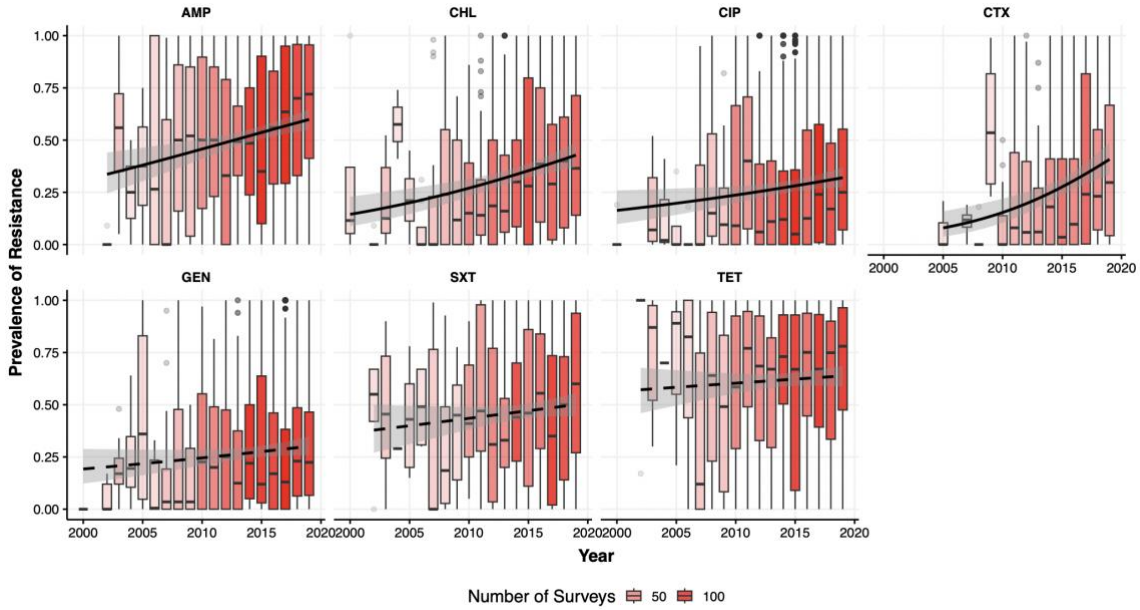
Supplementary Figure 1. **Chicken**: temporal trends of the prevalence of resistance, for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were significant (p value < 0.05) for AMP, CHL, CIP, and CTX.



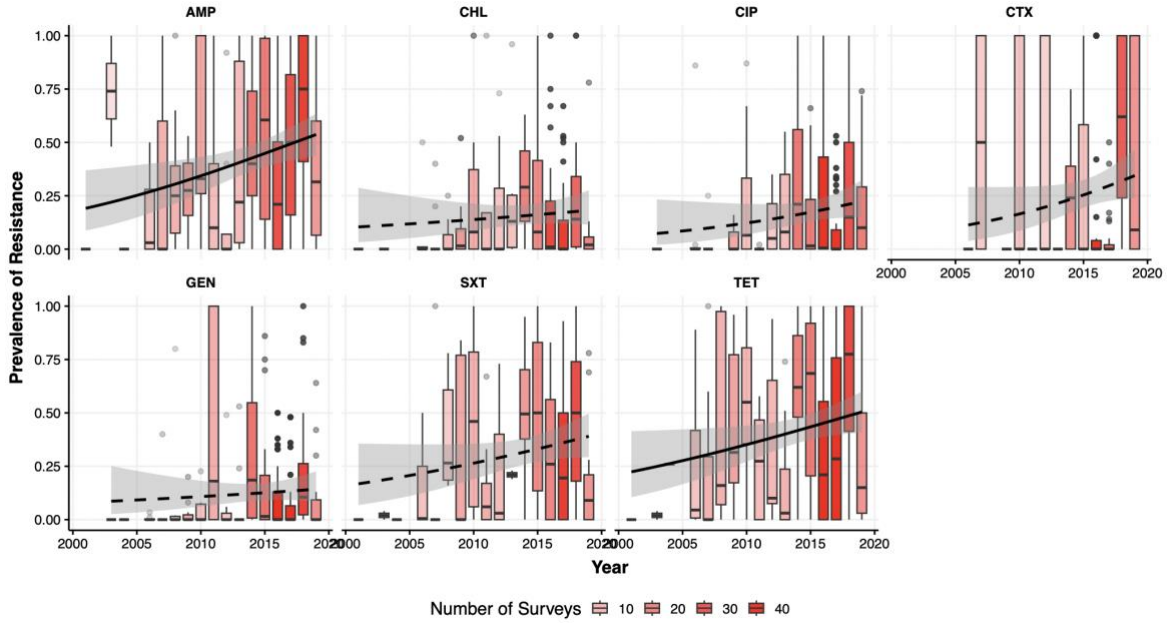
Supplementary Figure 2. **Pigs**: temporal trends of the prevalence of resistance, for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were significant (p value < 0.05) for all antimicrobials except TET.



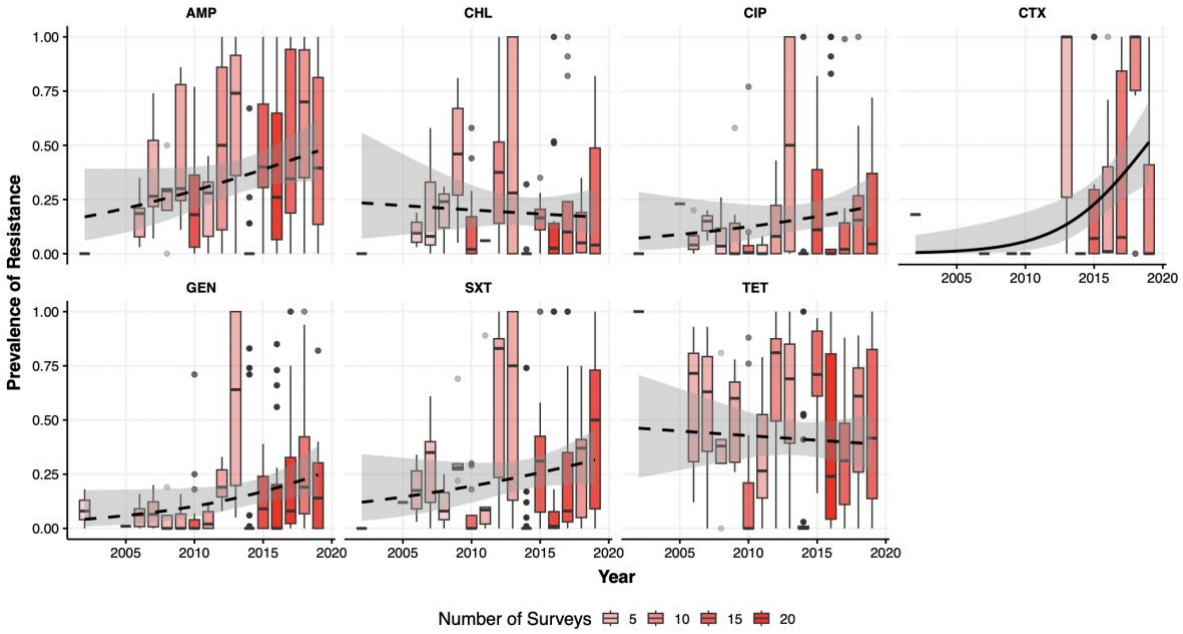
Supplementary Figure 3. **Cattle**: temporal trends of the prevalence of resistance, for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were not significant ($p > 0.05$) for all antimicrobials.



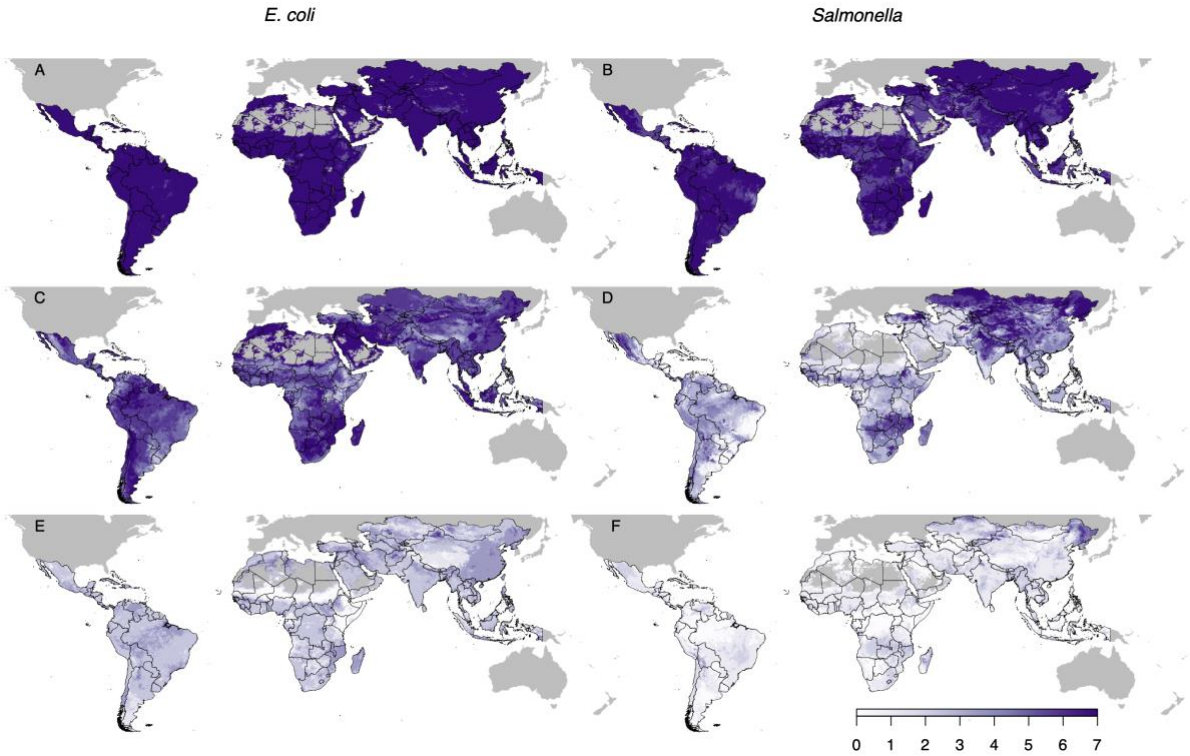
Supplementary Figure 4. **Asia**: temporal trends of the prevalence of resistance for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were significant ($p < 0.05$) for AMP, CHL, CIP, and CTX.



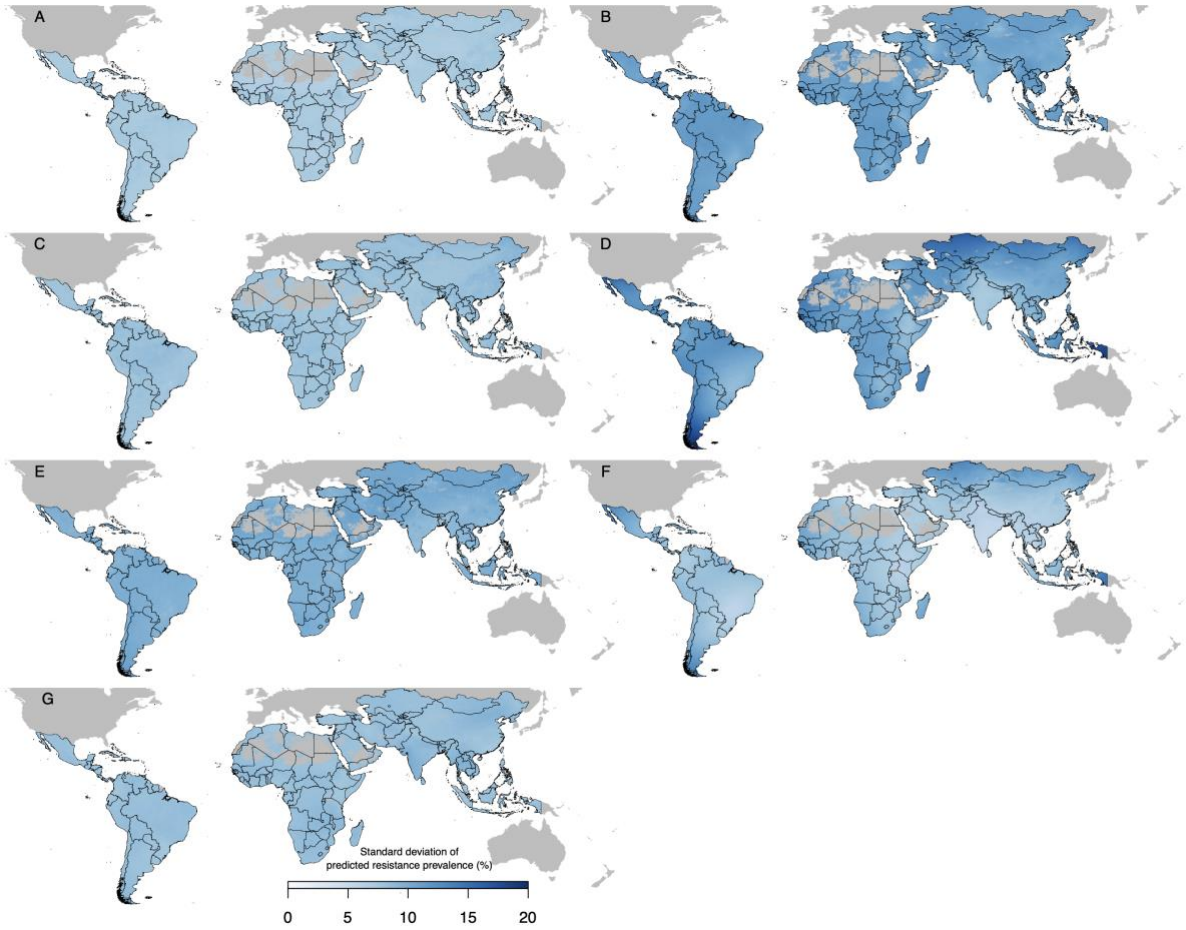
Supplementary Figure 5. **Africa**: temporal trends of the prevalence of resistance for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were significant ($p < 0.05$) for TET and AMP.



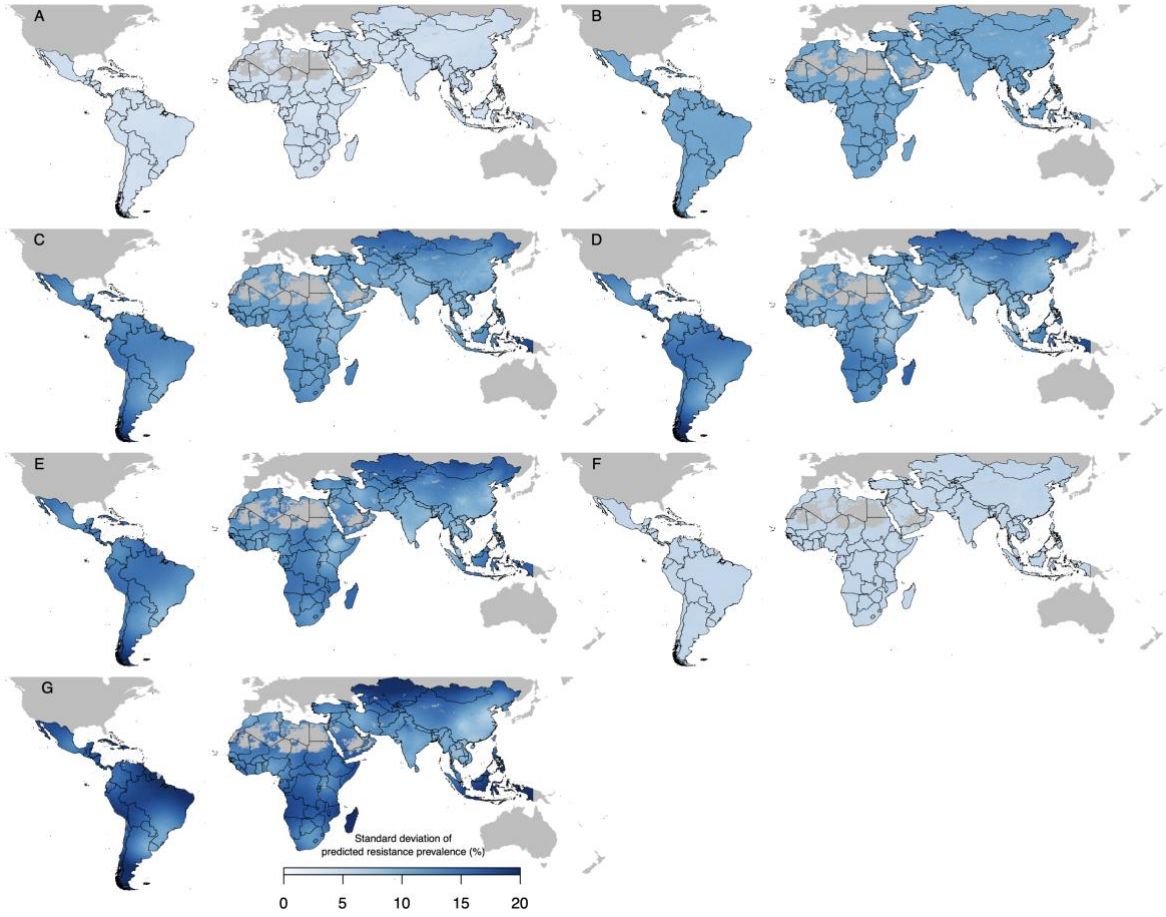
Supplementary Figure 6. **America**: temporal trends of the prevalence of resistance for ampicillin (AMP), chloramphenicol (CHL), ciprofloxacin (CIP), cefotaxime (CTX), gentamicin (GEN), sulfamethoxazole-trimethoprim (SXT), and tetracycline (TET). Solid lines represent significant temporal trends ($p < 0.05$), and dashed lines represent nonsignificant trends. Transparency levels of the red colors were proportional to the number of surveys published each year. Temporal trends were significant ($p < 0.05$) for CTX.



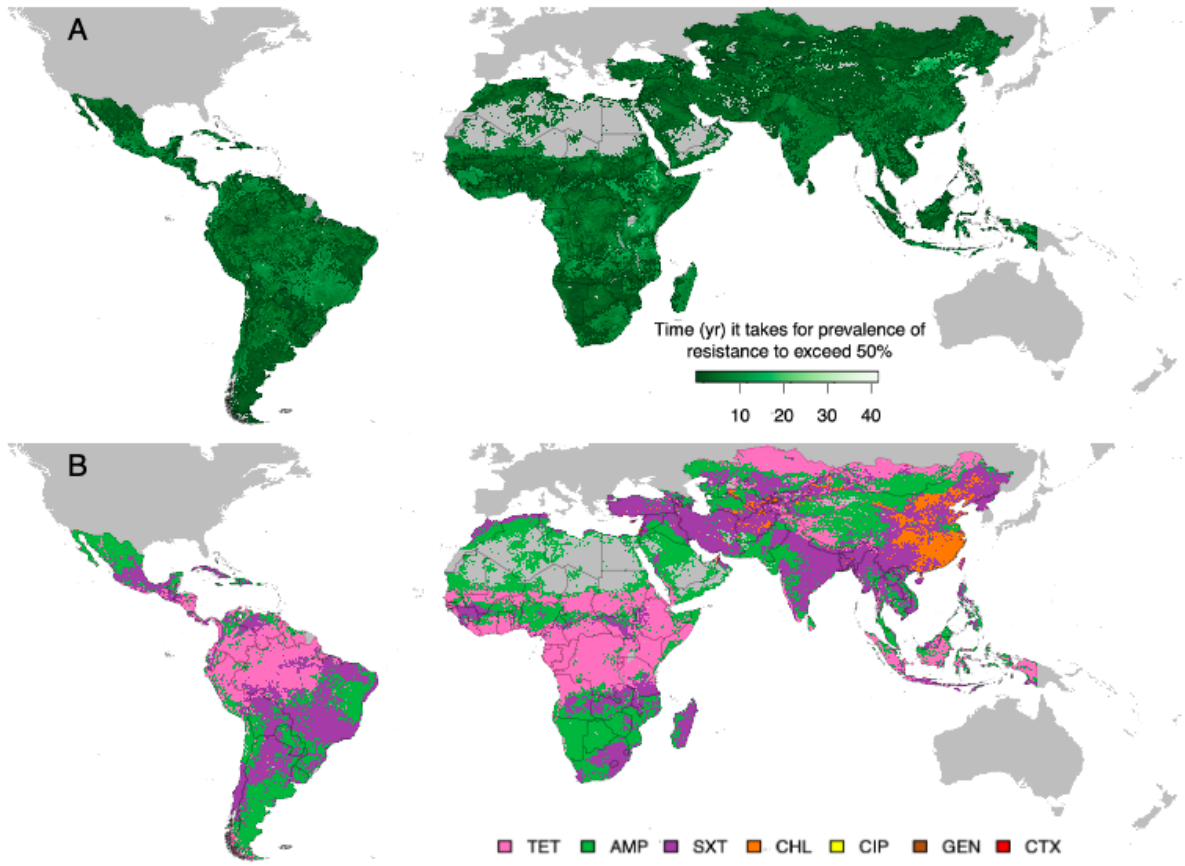
Supplementary Figure 7. The number of antimicrobials (out of 7) with resistance higher than 10% (N10: A, B), 25% (N25: C, D) and 50% (N50: E, F) in *E. coli* and *Salmonella*.



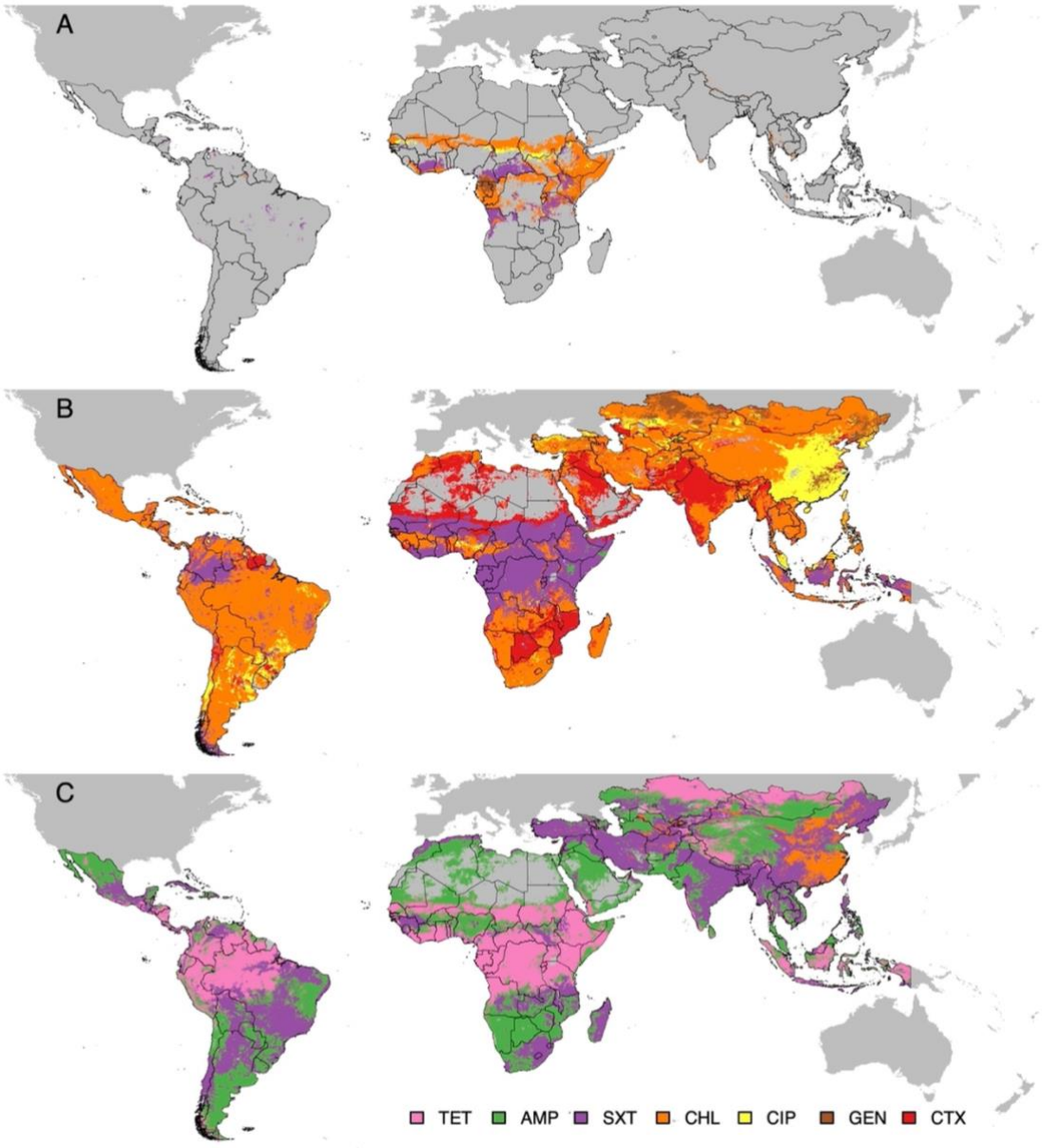
Supplementary Figure 8. Uncertainty of the predictions of resistance prevalence in *E. coli*, for tetracycline (TET), ampicillin (AMP), sulfamethoxazole-trimethoprim (SXT), chloramphenicol (CHL), ciprofloxacin (CIP), gentamicin (GEN), and cefotaxime (CTX). Shades of blue indicates the standard deviation on the predictions of resistance prevalence for each antimicrobial.



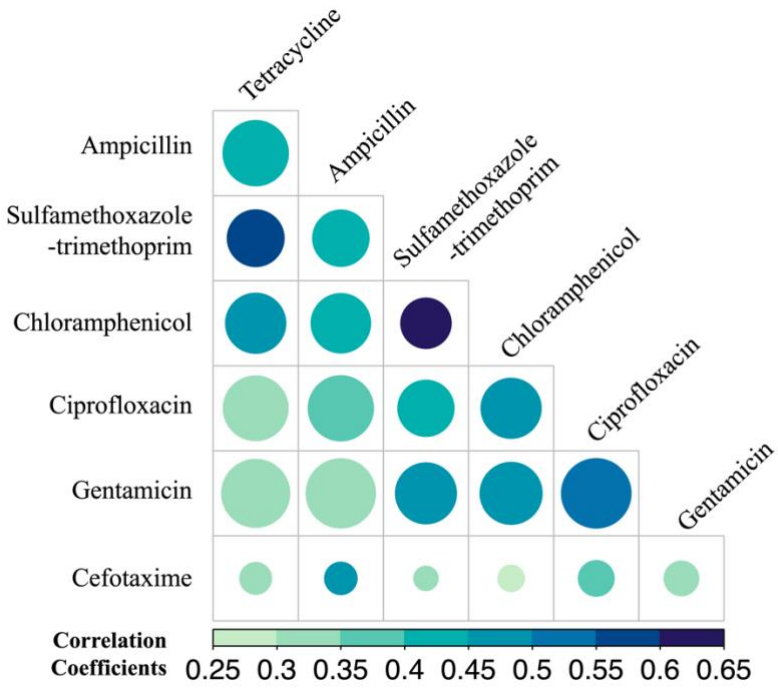
Supplementary Figure 9. Uncertainty of the predictions of resistance prevalence in *Salmonella*, for tetracycline (TET), ampicillin (AMP), sulfamethoxazole-trimethoprim (SXT), chloramphenicol (CHL), ciprofloxacin (CIP), gentamicin (GEN), and cefotaxime (CTX). Shades of blue indicates the standard deviation on the predictions of resistance prevalence for each antimicrobial.



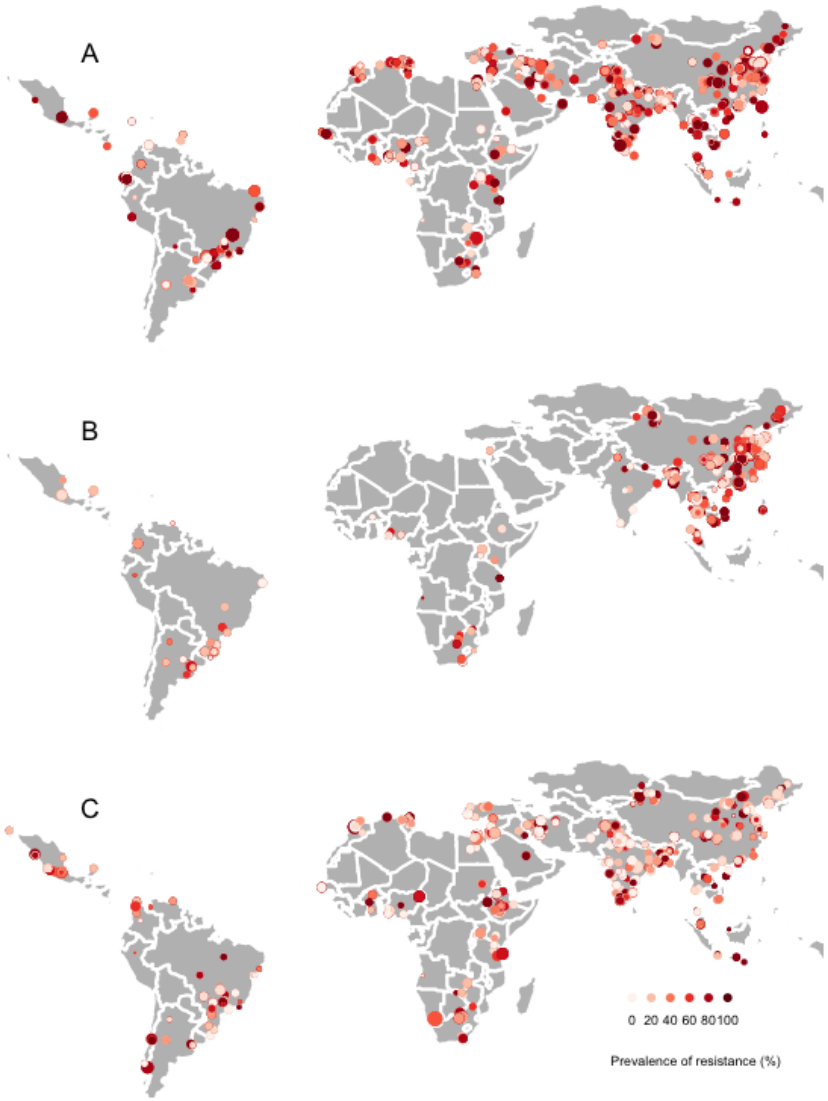
Supplementary Figure 10. Estimated time (years) that it takes for the prevalence of resistance to exceed 50% (A), for the predicted antimicrobial with the highest probability of its resistance prevalence exceeding 50% in the future (B).



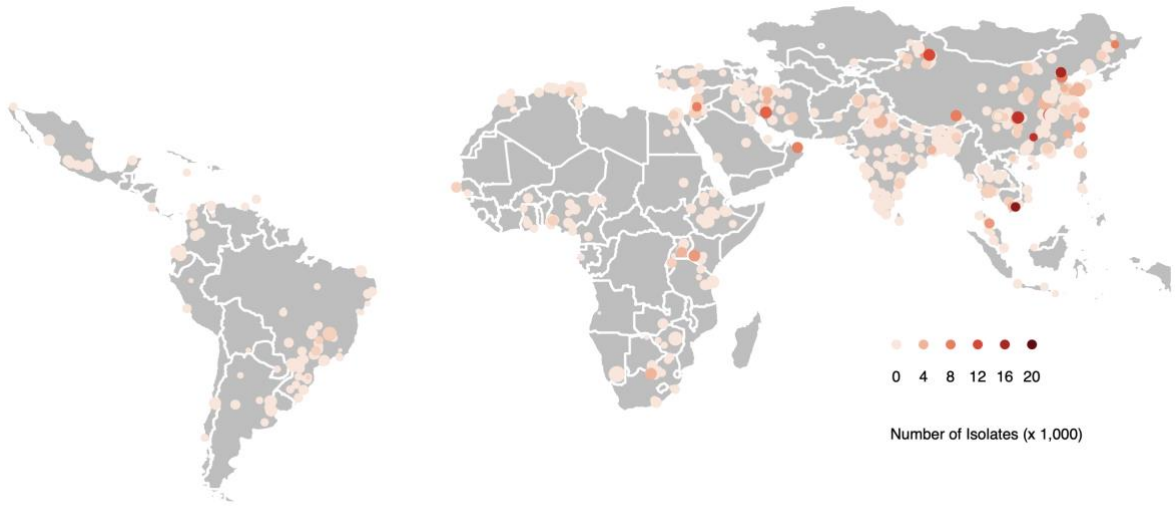
Supplementary Figure 11. Geographic distribution of antimicrobials with the highest probability of their resistance prevalence exceeding critical levels (A: 20%; B: 35%; C: 50%) in the future. TET: tetracycline; AMP: ampicillin; SXT: sulfamethoxazole-trimethoprim; CHL: chloramphenicol; CIP: ciprofloxacin; GEN: gentamicin; CTX: cefotaxime.



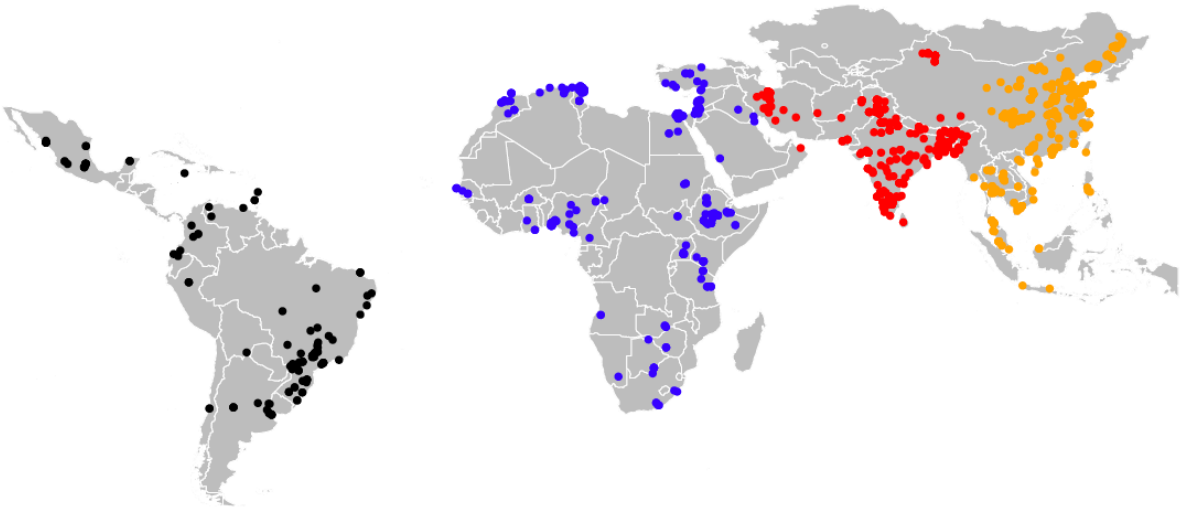
Supplementary Figure 12. Correlation coefficients between prevalence of antimicrobial resistance for 7 antimicrobial classes across 1,015 point prevalence surveys from food animals. Circle sizes are proportional to sample sizes.



Supplementary Figure 13. Geographic locations of point-prevalence surveys reporting resistance prevalence of *E. coli* and *Salmonella* isolated from poultry (A), pigs (B), and cattle (C). Sizes of the circle were in proportion to the log10 transformed sample sizes of each survey. Colors of the circles represented the prevalence of resistance reported in each survey.



Supplementary Figure 14. Geographic locations of point-prevalence surveys reporting resistance prevalence of *E. coli* and *Salmonella*. Sizes of the circle were in proportion to the log10 transformed sample sizes of each survey. Colors of the circles represented the number of bacterial isolates used to test the prevalence of resistance in each survey. The average number of isolates in each survey was 71 in Africa, 98 in America, and 94 in Asia.



Supplementary Figure 15. Distribution of spatial folds used for the four-fold spatial cross-validation procedure of the child models.

Supplementary Table 1. Extraction of point prevalence surveys (PPS) of antimicrobial resistance in *E. coli* and *Salmonella*, and exclusion criteria.

Exclusion Criteria	Literature Review Round #1	Literature Review Round #2	Literature Review Round #3
	n_{hits} = 32,030	n_{hits} = 8,481	n_{hits} = 3,814
Reviews, meta-analysis, and other non-PPS studies	- 30,038	- 7,401	- 3,263
	$n_{screened}$ = 1,992	$n_{screened}$ = 1,080	$n_{screened}$ = 551
Strain Surveys	NA	-115	-33
Diseased Animals	NA	-164	-21
Mixed Samples	NA	-97	-92
No Geographic Data	NA	-13	-54
Others	NA	-370	-238
	n_{PPS} = 926	n_{PPS} = 321	n_{PPS} = 113
Small Species Sample Size	-25	-6	-4
Drug-Pathogen Combinations Not Considered	-156	-45	-36
	n_{PPS_used} = 745	n_{PPS_used} = 270	n_{PPS_used} = 73

Supplementary Table 2. Coefficients associated with logistic regressions on temporal trends of resistance prevalence for Africa, Asia and America. Significant ($p < 0.05$) coefficient values are shown in bold.

	All		Africa (n = 1,673)		Asia (n = 6,148)		America (1,023)	
	estimate	standard error	estimate	standard error	estimate	standard error	estimate	standard error
TET	0.027	0.014	**0.07**	**0.033**	0.016	0.018	-0.017	0.043
AMP	**0.074**	**0.015**	**0.088**	**0.034**	**0.063**	**0.018**	0.087	0.047
SXT	**0.045**	**0.016**	0.064	0.037	0.029	0.019	0.071	0.053
CHL	**0.058**	**0.017**	0.036	0.046	**0.078**	**0.02**	-0.024	0.059
CIP	**0.051**	**0.018**	0.081	0.054	**0.046**	**0.02**	0.076	0.063
GEN	**0.037**	**0.016**	0.035	0.053	0.031	0.018	0.118	0.061
CTX	**0.155**	**0.03**	0.109	0.059	**0.148**	**0.037**	**0.319**	**0.107**

Supplementary Table 3. Environmental and anthropogenic covariates.

Name	Acronym	Year	Original Resolution	Source	Unit
Travel time to cities	acc	2015	30-arcsec resolution	Weiss et al 2018 [10] https://www.map.ox.ac.uk/accessibility_to_cities/.	minute
Antimicrobial use in animals 2013	use_2013	2013	0.083333 decimal degrees	Van Boeckel et al 2017 [11] http://science.sciencemag.org/content/357/6358/1350.full	Log10[(mg/pixel)+1]
Antimicrobial use in animals 2020	use_2020	2020	0.083333 decimal degrees	Mulchandani et al 2023[4] https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0001305	Log10[(mg/pixel)+1]
Yearly average of minimum monthly temperature	tmp	1970-2000	2.5 minutes	Worldclim [12] http://worldclim.org/version2	°C * 10
Percentage Irrigated areas	irg	2005	0.083333 decimal degrees	Global Map of Irrigation Areas (GMIA) [13] http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/	%
Population density of cattle, chickens, pigs, and sheep	ca_v4 ch_v4 pg_v4 sh_v4	2015	0.083333 decimal degrees	Gridded Livestock of the World v4 https://www.nature.com/articles/sdata2018227	Log10[(Heads/pixel) +1]
Percentage of tree coverage	veg	2013	0.083333 decimal degrees	Hansen et al 2013 [14] https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.2.html	%
Average pesticide application rate	pest	2015	0.083333 decimal degrees	PEST-CHEMGRIDS [15] https://sedac.ciesin.columbia.edu/data/set/ferman-v1-pest-chemgrids	kg/ha per year
Atmospheric ammonia	amm	2008-2016	0.01 decimal degrees	Van Damme et al 2018 [16] https://www.nature.com/articles/s41586-018-0747-1	10^{16} molecules cm^{-2}
Gross Domestic Product (GDP) in Purchasing Power Parity	gdp	2005	1 decimal degrees	G-Econ [17] https://sedac.ciesin.columbia.edu/data/set/spatialecon-gecon-v4	Billion US dollars
Fourier coefficients of Precipitation	wd1920	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA
Fourier coefficients of Middle Infra-red	wd1903	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA
Fourier coefficients of Normalised Difference Vegetation Index	wd1914	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA
Fourier coefficients of Enhanced Vegetation	wd1915	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA

Index					
Fourier coefficients of Day Land Surface Temperature	wd1907	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA
Fourier coefficients of Night Land Surface Temperature	wd1908	2001-2019	0.0083333 decimal degrees	Scharlemann et al 2008 [18] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0001408	NA

Supplementary Table 4. The average estimated time for resistance prevalence to exceed 50% across all pixels on the map, for each antimicrobial class and weighted by the distribution of animals' biomass.

Antimicrobial Class	Time for resistance prevalence to exceed 50% (years)
Tetracyclines	5.1
Penicillins	1.7
Sulfonamides	7.1
Amphenicols	10.8
Quinolones	12.4
Aminoglycosides	NA
Cephalosporins	4.1

Supplementary Table 5. Coefficients of LASSO regressions predicting the possibility that resistance prevalence of an antimicrobial will exceed 50% in the future, given the preceding resistance profile. The antimicrobials included cefotaxime (CTX), sulfamethoxazole-trimethoprim (SXT), chloramphenicol (CHL), tetracycline (TET), ampicillin (AMP), ciprofloxacin (CIP), and gentamicin (GEN). Proportion_PPS: the proportion of point prevalence surveys reporting an increased resistance prevalence to over 50% for an antimicrobial, out of all alternative antimicrobials; Proportion_AMU: the proportion of usage (kg) of an antimicrobial out of all alternative antimicrobials; N50: the number of antimicrobials with resistance above 50% in the preceding resistance profile. The abbreviations of the other covariates are explained in Supplementary Table 3. Covariates for which coefficients were 0 for all antimicrobials were removed from the table.

	TET	AMP	SXT	CHL	CIP	GEN	CTX
(Intercept)	-7.63	-2.921	-3.624	-5.311	-4.069	-4.183	-3.628
proportion_PPS	19.36	5.497	2.151	1.734	-0.96	-0.915	-1.215
proportion_AMU	0	0	0.346	0	0	0	0
N50	0.013	0.364	0.936	1.103	1	0.934	0.989
use_2020	0	0	0	0.079	0	0	-0.045
use_2013	0	0	0.01	0.004	-0.007	0	-0.215
ch_v4	0	0	0	0	0	0	-0.076
sh_v4	0	0	-0.022	0	0	0.039	0
Pest	0	0	0.002	0	0	-0.013	-0.001
amm	0	0	0	-0.004	0	0	0.056
gdp	0	0	0	0.058	0	0	-0.134
wd1920a0	0	0	0	0	-0.121	0	0
wd1920a3	-0.096	0	0	0	0	0	0
wd1920d2	0	0	0	-0.014	0.117	0	0
wd1920d3	0	0.11	-0.079	-0.057	0	0	0
wd1920dd	0	0	0	0	0.035	0	0
wd1920mn	0	-0.066	0	0.003	0	0	0
wd1920vr	0	0	0	0.016	0	0	0
wg1903a1	0	0	0	-0.095	0	0	0
wg1903a2	0	0	0	0	0.027	0	0
wg1903d3	0	0	0	0	0.084	0	0
wg1903dd	0	0	0	0	-0.176	0	0
wg1903mn	0	0.17	-0.073	0	0	0	0
wg1903vr	0	0	0	0	0	0	0.054
wg1907a0	0	0.134	0	0	0	0	0
wg1907a2	0	0	0	-0.026	0	0.002	0
wg1907a3	0	0	0	0	0	0.091	0
wg1907d2	0	0	-0.072	0	0	0	0
wg1907d3	0	-0.002	0	0.008	0	0	0
wg1907mn	0	0	-0.083	0	0	0	0
wg1908a2	0	0	0	0	0	0.176	0
wg1908a3	0	0	0.08	0.099	0	0	-0.307
wg1908d2	0	0	0	-0.058	0	0	0
wg1914da	0	0	0	0	0	0.102	0.004
wg1914dd	0	0	0	-1.562	0	0	2.818
wg1914mx	0	0	0	-0.163	0	0	0
wg1914vr	0	0	0	0	0.119	0	0

wg1915a2	0.03	0	0	-0.059	0	0	0
wg1915d1	-0.051	0	0	0	0	0	0
wg1915d3	0.086	0	0	0	0	0	0

Supplementary Table 6. The top 20 antimicrobial compounds with reported resistance prevalence and the corresponding antimicrobial classes in point prevalence surveys.

Antimicrobial Compound	Number of Reported Resistance Prevalence	Antimicrobial Class
Gentamicin	1,452	Aminoglycosides
Ampicillin	1,347	Penicillins
Ciprofloxacin	1,345	Quinolones
Tetracycline	1,271	Tetracyclines
Chloramphenicol	1,133	Amphenicols
Sulfamethoxazole-Trimethoprim	1,068	Sulfonamides
Streptomycin	937	Aminoglycosides
Nalidixic acid	852	Quinolones
Cefotaxime	726	Cephalosporins
Amikacin	688	Aminoglycosides
Kanamycin	664	Aminoglycosides
Amoxicillin-Clavulanic acid	614	Penicillins
Ceftriaxone	521	Cephalosporins
Enrofloxacin	473	Quinolones
Ceftazidime	456	Cephalosporins
Amoxicillin	454	Penicillins
Norfloxacin	449	Quinolones
Cefalotin	351	Cephalosporins
Colistin	330	Polymixins

Supplementary Table 7. Estimated parameters of the fitted INLA models predicting the geographic distribution of resistance prevalence of each antimicrobial. The table showed the mean value and standard deviation for the range of the spatial random effect, and the coefficients of three child models. The antimicrobials included cefotaxime (CTX), sulfamethoxazole-trimethoprim (SXT), chloramphenicol (CHL), tetracycline (TET), ampicillin (AMP), ciprofloxacin (CIP), and gentamicin (GEN).

	CTX	SXT	CHL	TET	AMP	CIP	GEN
range.mean	3.39	5.65	3.82	3.25	2.4	4.52	3.35
range.sd	1.58	1.31	1.47	1.57	1.52	3.92	1.5
beta_BRT.mean	0.47	0.36	0.38	0.25	0.29	0.34	0.45
beta_BRT.sd	0.06	0.06	0.05	0.05	0.04	0.04	0.06
beta_LASSO.mean	0.4	0.41	0.37	0.39	0.42	0.32	0.36
beta_LASSO.sd	0.06	0.05	0.05	0.04	0.05	0.04	0.05
beta_NNR.mean	0.41	0.31	0.36	0.37	0.35	0.43	0.39
beta_NNR.sd	0.06	0.04	0.05	0.05	0.05	0.05	0.05

Supplementary Table 8. Common resistance profiles in point prevalence surveys, with 2, 3, and 4 antimicrobials with resistance higher than 50% (N50). The total number of surveys for each N50 category, and the number of surveys reporting each resistance profile were shown in the brackets.

Number of antimicrobials with resistance higher than 50%	Resistance profiles
N50 = 2 (n = 201)	TET-AMP (n=97)
	TET-SXT (n=45)
N50 = 3 (n = 161)	TET-AMP-SXT (n=60)
	TET-SXT-CHL (n=25)
	TET-AMP-CIP (n=17)
	TET-AMP-CTX (n=14)
N50 = 4 (n = 138)	TET-AMP-SXT-CHL (n=66)
	TET-AMP-SXT-CIP (n=12)
	TET-SXT-CHL-GEN (n=11)
	TET-AMP-SXT-CTX (n=8)

Supplementary Table 9. Number of point-prevalence surveys conducted on chicken, pigs, and cattle in each year.

Year	Chicken	Pigs	Cattle
2000	1	0	1
2001	0	0	1
2002	2	2	2
2003	10	4	6
2004	0	0	5
2005	2	1	2
2006	6	3	6
2007	12	4	11
2008	11	12	14
2009	12	8	11
2010	31	14	18
2011	17	16	22
2012	50	12	19
2013	33	16	34
2014	63	38	46
2015	56	32	42
2016	85	31	33
2017	70	39	60
2018	50	31	41
2019	59	40	35

Supplementary Table 10. Number of point-prevalence surveys conducted on chicken, pigs, and cattle in each country.

Country ISO3	Chicken	Pigs	Cattle
AGO	1	1	1
BFA	2	1	2
BWA	1	1	3
CMR	1	0	0
DZA	8	0	2
EGY	17	0	12
ETH	11	2	26
GAB	1	0	0
GHA	3	0	2
GMB	1	0	0
KEN	5	2	2
MAR	6	0	4
NAM	0	0	1
NGA	16	4	10
SDN	0	0	2
SEN	1	0	1
TCD	1	0	0
TUN	13	0	8
TZA	3	1	7
UGA	4	2	3
ZAF	5	6	10
ZMB	1	0	2
ZWE	2	0	1
RWA	1	0	0
IND	84	20	93
BGD	36	0	11
NPL	8	1	1
BTN	1	1	0
PAK	10	0	4
IRN	33	0	21
IRQ	6	0	2
ISR	2	0	2
LBN	3	1	4
QAT	1	0	0
OMN	1	0	0
JOR	0	0	3
SAU	2	0	1
TUR	11	0	8
ARG	5	8	5
BOL	1	0	0
BRA	28	13	27
COL	1	1	4
ECU	6	0	0
PER	2	1	1
VEN	1	1	2

CHL	0	0	3
CRI	1	0	0
MEX	6	5	15
GRD	2	0	0
LCA	0	0	0
JAM	1	0	0
CHN	169	186	75
IDN	3	0	5
KHM	2	3	0
LAO	0	2	2
MYS	6	2	4
MMR	1	0	0
PHL	1	3	0
SGP	1	0	0
THA	14	26	10
VNM	14	9	7
LKA	1	0	0
KWT	1	0	0
PRY	1	0	0

Supplementary Table 11. The percentage of 10 x 10 km pixels in each country where the uncertainty of the predicted priority antimicrobial is higher than 0.4.

Country ISO3	Percentage of pixels with uncertainty > 0.4	Country ISO3	Percentage of pixels with uncertainty > 0.4	Country ISO3	Percentage of pixels with uncertainty > 0.4
AFG	0.210069	GNB	0.164634	PAK	0.082604
AGO	0.030731	GNQ	0	PAN	0.109756
ARE	0.152968	GTM	0.103553	PER	0.141867
ARG	0.032319	GUY	0.120746	PHL	0.159052
ARM	0.148741	HND	0.034509	PRI	0.064103
AZE	0.059736	HTI	0.086806	PRK	0.766391
BDI	0.068729	IDN	0.142597	PRY	0.031902
BEN	0.081277	IND	0.071204	PSX	0.383721
BFA	0.059798	IRN	0.131975	QAT	0.163934
BGD	0.023966	IRQ	0.08467	RWA	0.0625
BLZ	0.073643	ISR	0.343333	SAH	1
BOL	0.068859	JAM	0.089431	SAU	0.039423
BRA	0.092262	JOR	0.032558	SDN	0.213807
BRN	0.16129	KAB	0.055046	SDS	0.347961
BTN	0.039623	KAS	1	SEN	0.033037
BWA	0.002074	KAZ	0.064847	SGP	1
CAF	0.064724	KEN	0.018052	SLE	0.285542
CHL	0.120278	KGZ	0.159517	SLV	0.057613
CHN	0.119058	KHM	0.091589	SLO	0.15408
CIV	0.049419	KWT	0.081448	SOM	0.05876
CMR	0.047654	LAO	0.059844	SUR	0.110259
CNM	0.333333	LBN	0.086331	SWZ	0.40708
COD	0.079909	LBR	0.016187	SYR	0.034994
COG	0.14932	LBY	0.019029	TCD	0.019425
COL	0.16678	LKA	0.087649	TGO	0.161919
CRI	0.1198	LSO	0.779412	THA	0.089627
CUB	0.130506	MAR	0.069437	TJK	0.112322
CYN	0.566667	MDG	0.095638	TKM	0.026463
CYP	0.352941	MEX	0.067058	TLS	0.108974
DJI	0.4375	MLI	0.066136	TTO	0.212766
DOM	0.062609	MMR	0.102656	TUN	0.031632
DZA	0.027011	MNG	0.094879	TUR	0.084819
ECU	0.1261	MOZ	0.055876	TWN	0.129841
EGY	0.030746	MRT	0.032622	TZA	0.093735
ERI	0.011065	MWI	0.097397	UGA	0.090836
ESB	0	MYS	0.107326	URY	0.093725
ETH	0.034781	NAM	0.016376	UZB	0.098051
GAB	0.026928	NER	0.025251	VEN	0.086985

GEO	0.136406	NGA	0.100411	VNM	0.120401
GHA	0.050751	NIC	0.063336	YEM	0.020265
GIN	0.140271	NPL	0.059487	ZAF	0.039901
GMB	0.210084	OMN	0.119071	ZMB	0.088797
				ZWE	0.017105

Reference

1. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, 1266-+ (2019).
2. Van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**, 1–67 (2011).
3. Bhatt, S. *et al.* Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* **14**, 20170520 (2017).
4. Mulchandani, R., Wang, Y., Gilbert, M. & Van Boeckel, T. P. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. *PLOS Glob. Public Health* **3**, e0001305 (2023).
5. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
6. Tibshirani, R. Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* **73**, 273–282 (2011).
7. Chollet, F. Keras. (2015).
8. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71**, 319–392 (2009).
9. Zhao, C. *et al.* Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China. *Nat. Food* **2**, 596–602 (2021).
10. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
11. Van Boeckel, T. P. *et al.* Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci.* **112**, 5649–5654 (2015).
12. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
13. Siebert, S. *et al.* Development and validation of the global map of irrigation areas. (2005).
14. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *science* **342**, 850–853 (2013).
15. Maggi, F., Tang, F. H. M., la Cecilia, D. & McBratney, A. PEST-CHEMGRIDS, Global Gridded Maps of the Top 20 Crop-specific Pesticide Application Rates from 2015 to 2025. *Sci. Data* **6**, 1–20 (2019).
16. Van Damme, M. *et al.* Industrial and agricultural ammonia point sources exposed. *Nature* **564**, 99–103 (2018).
17. Nordhaus, W. D. Geography and Macroeconomics: New Data and New Findings. *Proc. Natl. Acad. Sci. U. S. Am. PNAS* **103**, (2005).
18. Scharlemann, J. P. *et al.* Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PloS One* **3**, e1408 (2008).

Supplementary Information for

A hybrid Neural Network-SEIR model for forecasting intensive care occupancy in Switzerland during COVID-19 epidemics.

Riccardo Delli Compagni[1], Zhao Cheng[1], Stefania Russo[2],Thomas P. Van Boeckel[1,3*]

[1] Health Geography and Policy Group, ETH Zürich, Switzerland

[2] Ecovision Lab, Photogrammetry and Remote Sensing, ETH Zürich, Switzerland

[3] Center for Diseases Dynamics Economics and Policy, Washington DC, United States

*Correspondence to: thomas.vanboeckel@env.ethz.ch

Supporting Information

Number of pages: 3; Number of Tables: 0; Number of Figures: 2;

3. Results

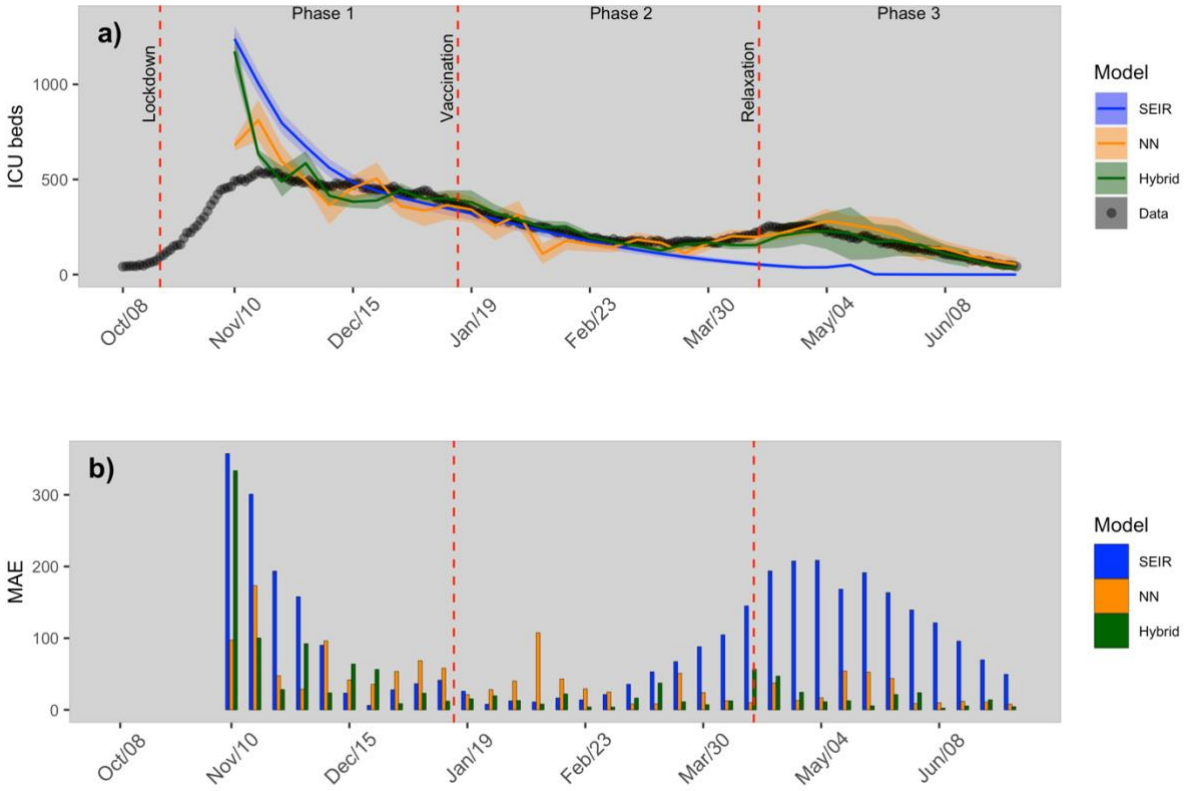


Fig S1. a) Predictions 7-days ahead of intensive occupancy at the national-level for the three models (shaded areas represent 95% confident intervals); b) corresponding Mean Absolute Error (MAE) calculated on test data.

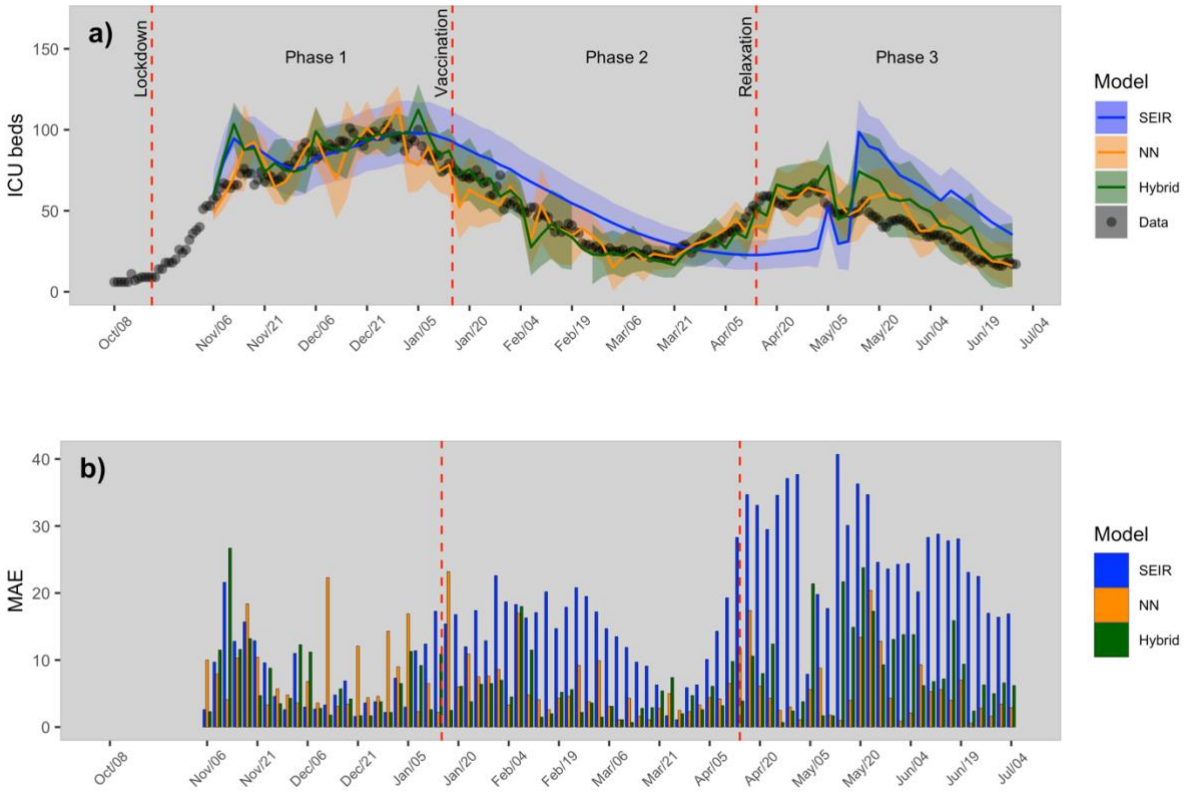


Fig S2. a) Predictions 3-days ahead of intensive occupancy at cantonal level (canton of Zurich) for the three models (shaded areas represent 95% confident intervals); b) corresponding Mean Absolute Error (MAE) calculated on test data.