DISS. ETH NO. 29931

Open-access geodata as a resource against infectious diseases

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by

NICOLA CRISCUOLO

MS in Environmental Sciences, University of Salerno, Italy

born on 09.06.1990

accepted on the recommendation of:
Dr. Thomas P. Van Boeckel
Prof. Nicolas Ray
Dr. Ioannis Magouras
Prof. Dr. Alex Hall

2024

Acknowledgements

When I met Prof. Thomas Van Boeckel during my job interview in 2019, I vividly recall him asking me about my familiarity with Geographic Information Systems. I confidently replied, “Yes, I can use QGIS, and for more complex tasks I often use the ‘unofficial’ version of ArcGIS”. There was a brief pause during which I questioned my response, but Thom replied “Yeah, I was also using a similar version when I was a student”. Besides the interest I had in working with him because of his research topic, I recall that moment as the second time when I told myself “Yeah, we might do cool things together”.

Now that I have reached the end of the Ph.D., my first acknowledgment must go to Thom. After several interviews to enter the field of public health, Thom was the one who decided to give me a chance. For this, I will always be grateful to him. Most importantly, I thank him for teaching me that there’s (almost) always a way to solve apparently impossible problems or, in the worst-case scenario, to attack them from a different angle. Finally, I thank him for all the efforts he put with me to teach me that “science needs good writing”, and how important it is in our field to capture the interest of people with synthetic, well-structured, and powerful messages. I know, Thom, I am still not very good at that, but I'll do my best to improve.

Secondly, I must acknowledge Cheng Zhao (now Dr. Zhao). Cheng is not only one of the smartest and strongest persons I know but also the kindest colleague and friend someone could ever desire to share an office with. Thanks for your constant help, support, and suggestions, for the food, and for the stressful moments we shared almost always together. For all your smiles, thank you.

I must thank also to Dr. João Pires. He was the first person I met in Zurich and if he had not told me that the door of our building was open, I would have been late my very first day of work. Thank you João, and also for every advice beyond the ones about the Ph.D. journey. Thanks to Prof. Laura Huber, the first Professor I have ever seen going inside a public fountain in front of a police station to wash her feet. Also, a big thank you to Alex and Ranya for bringing all their energy and their laughs to the group, and for being always available when in need of brainstorming. Guys, me, you, and Thom will be tied forever after the Canadian near-death experience. Alex, one day the “Morgan-Criscuolo” distance will be a big thing, you have my word. And thanks to Katie, for our morning coffee and for telling me about the existence of “The Last Podcast on the Left”.

A big thank you also to all the colleagues and friends I met in Zurich, and in particular to Sergei for immediately involving me in his life. Thanks to Sara, for having shared a big part of this journey with me. Thank you also to Ambra, Cristina, Gilbert, Soushieta, Giulia, Sumithra, Fokko, Berit, Paco, Judith, Claudia, Alejandro, Helene, and Sonja. Thanks to Lucca, Marta, Nico, and Chloe for the breathtaking table football matches.

Now, time for the long-distance ones. A huge thank you to Luigi, Sara, Mirko, Elena, Sara, and Antonio. You all serve as a constant reminder of the kind of people I want to surround myself with, and how important it is to never ignore my roots and always keep them strong.

Last but not least, there are the persons I am truly most grateful to: my mother Norma, my father Nino, and my sister Francesca. They have been pillars of support in every life decision, providing constant assistance during challenging moments, and expressing their unconditional love through countless simple gestures that often went unnoticed. It is thanks to you that I had the privilege to study and work in the sciences, and everything that this journey has brought me is owed to your enduring patience, encouragement, and love.

Table of Contents

Abstract

In the era of big data, open-access information online is ever increasing. Harnessing its potential can find useful application in many scientific sectors, including public health. One pressing concern within this realm is the well-being of food animals, whose importance extends beyond economics to encompass human health. Infectious diseases, antimicrobial resistance (AMR), and a limited access to healthcare professionals currently pose a threat to food animals and the 1.3 billion people worldwide who rely on them for their subsistence. This doctoral thesis explores the potential of using open-access data, in conjunction with geospatial models and user-friendly platforms, to provide new resources that can aid decision-makers in improving the health of animals raised for food.

Chapter 2 introduces the open-access platform *resistancebank.org*, an online repository that centralizes 2,045 point prevalence surveys (PPS) reporting AMR prevalence estimates in food animals. Launched in 2019 as a *shiny* application (R programming language), the platform consists of 42,891 resistance prevalence estimates of foodborne pathogens sampled from terrestrial and aquatic species in low- and middle-income countries. Besides individual PPS, the platform provides access to AMR maps at the 10x10 km^2 resolution, country-level reports of AMR, and allow users to upload the results of their PPS through an online form or an Excel template.

Chapter 3 illustrates how open-access addresses of veterinarians were used to identify areas where veterinarians are farther than 1 hour of travel time from food animals (i.e., "veterinary coldspots"). First, the assembling of a global address book of 303,745 veterinarians' addresses using web-scraping technique is presented. Log-Gaussian Poisson Regression models and spatial covariates were then used to predict the global distribution of veterinarians at the 10x10 km^2 resolution. The resulting map showed that 43% of veterinarians are in high-income countries, which, however, account for just 21.2% of the global production of food animals. As a consequence, the highest percentages of all food animals in coldspots were identified in Asia (44.1%), Latin America (27.7%), and Africa (18.7%).

In Chapter 4, AMR and veterinary capacity were used as case studies to present analyses about the optimal allocation of health facilities in underserved areas. Two geographic approaches were used to identify the locations of an additional 5% of the national number of i) health facilities (hospitals and clinical laboratories) to equip with laboratories for antimicrobial susceptibility tests (ASTs) in five African countries and ii) veterinarians predicted in Chapter 3, with the aim to reduce the number of food animals in coldspots. The allocation of laboratories testing for AMR resulted in a coverage of ~21 million people that can reach a laboratory within 1 hour. The allocation of veterinarians in countries with coldspots ensured veterinary care for 32.6% of the global number of cattle, chickens, and pigs living in coldspots.

In Chapter 5, the focus shifts on the global health crisis sparked by COVID-19. As for many students worldwide, the pandemic affected the trajectory of my Ph.D. However, thanks to the

skillset acquired while working on the project presented in Chapter 2, I was actively involved in research efforts aimed in monitoring the spread of the disease and the availability of healthcare resources in Switzerland. I set-up the open-access platform *icumonitoring.ch* used to summarize bi-weekly forecasts of COVID-19 cases, hospitalizations, intensive care units' occupancy, and ventilators' availability at the regional-, cantonal-, and hospital-level in Switzerland. Chapter 5 also includes my contribution in providing the Swiss Federal Office of Public Health with daily geocodings of COVID-19 cases. The output of this process was used to produce daily hotspots' maps of COVID-19 cases and to highlight how hospitalizations, deaths, and the prevalence of positive COVID-19 cases were higher among the population living in the poorest neighbors of Switzerland.

Zusammenfassung

Im Zeitalter von Big Data nimmt der Umfang an open-access Informationen im Internet stetig zu. Die Nutzung ihres Potenzials kann in vielen wissenschaftlichen Bereichen Anwendung finden, einschließlich des öffentlichen Gesundheitswesens. Ein drängendes Anliegen in diesem Bereich ist das Wohlergehen von Nutztieren, dessen Bedeutung über die Wirtschaft hinaus bis hin zur menschlichen Gesundheit reicht. Infektionskrankheiten, Antibiotikaresistenz (AMR) und der begrenzte Zugang zu Gesundheitsfachkräften stellen derzeit eine Bedrohung für Nutztiere und die 1,3 Milliarden Menschen weltweit dar, die auf sie für ihren Lebensunterhalt angewiesen sind. Diese Doktorarbeit untersucht das Potenzial der Nutzung von open-access Daten in Verbindung mit geografischen Modellen und benutzerfreundlichen Plattformen, um neue Ressourcen bereitzustellen, die Entscheidungsträgern bei der Verbesserung der Gesundheit von Nutztieren helfen können.

Kapitel 2 stellt die Open-Access-Plattform resistancebank.org vor, ein Online-Repository, das 2.045 Punktbefragungen (PPS) zur Prävalenz von AMR in Nutztieren zentralisiert. Die Plattform wurde 2019 als Shiny-Anwendung (Programmiersprache R) gestartet und enthält 42.891 Prävalenzschätzungen von Resistenzen gegenüber durch Lebensmittel übertragenen Krankheitserregern, die bei terrestrischen und aquatischen Arten in Ländern mit niedrigem und mittlerem Einkommen entnommen wurden. Neben individuellen PPS bietet die Plattform Zugriff auf AMR-Karten mit einer Auflösung von 10x10 km², länderspezifische Berichte zu AMR und ermöglicht Benutzern das Hochladen der Ergebnisse ihrer PPS über ein Online-Formular oder eine Excel-Vorlage.

Kapitel 3 zeigt, wie Adressen von Tierärzten verwendet wurden, um Gebiete zu identifizieren, in denen Tierärzte mehr als 1 Stunde Reisezeit von Nutztieren entfernt sind (sogenannte "tierärztliche Versorgungslücken"). Zunächst wird die Erstellung eines globalen Adressverzeichnisses von 303.745 Tierarztadressen mithilfe von Web-Scraping-Techniken vorgestellt. Log-Gaußsche Poisson-Regressionsmodelle und räumliche Kovariablen wurden dann verwendet, um die globale Verteilung von Tierärzten mit einer Auflösung von 10x10 km² vorherzusagen. Die resultierende Karte zeigte, dass 43% der Tierärzte in wohlhabenden Ländern tätig sind, die jedoch nur 21,2% der weltweiten Produktion von Nutztieren ausmachen. Als Folge wurden die höchsten Prozentsätze aller Nutztiere in Versorgungslücken in Asien (44,1%), Lateinamerika (27,7%) und Afrika (18,7%) identifiziert.

In Kapitel 4 wurden AMR und tierärztliche Kapazität als Fallstudien verwendet, um Analysen zur optimalen Verteilung von Gesundheitseinrichtungen in unterversorgten Gebieten vorzustellen. Zwei geographische Ansätze wurden verwendet, um in fünf afrikanischen Ländern die Standorte für zusätzliche 5 % der nationalen Anzahl an i) Gesundheitseinrichtungen (Krankenhäuser und klinische Laboratorien), ausgestattet mit Laboratorien für antimikrobielle Empfindlichkeitstests (ASTs) und ii) Tierärzten, vorausberechnet in Kapitel 3, zu ermitteln, mit dem Ziel, die Zahl der Nutztiere in Versorgungslücken zu verringern. Die Zuweisung von Laboren zur AMR-Testung führte zu

einer Abdeckung von rund 21 Millionen Menschen, die innerhalb von 1 Stunde ein Labor erreichen können. Die Zuweisung von Tierärzten in Ländern mit Versorgungslücken gewährleistete tierärztliche Betreuung für 32,6% der weltweiten Anzahl von Rindern, Hühnern und Schweinen in Versorgungslücken.

In Kapitel 5 liegt der Fokus auf der globalen Gesundheitskrise, die durch COVID-19 ausgelöst wurde. Wie für viele Studierende weltweit beeinflusste die Pandemie den Verlauf meiner Promotion. Dank der erworbenen Fähigkeiten während der Arbeit an dem Projekt in Kapitel 2 war ich aktiv an Forschungsanstrengungen beteiligt, die darauf abzielten, die Ausbreitung der Krankheit und die Verfügbarkeit von Gesundheitsressourcen in der Schweiz zu überwachen. Ich habe die Open-Access-Plattform icumonitoring.ch eingerichtet, die zweiwöchentliche Vorhersagen zu COVID-19-Fällen, Krankenhauseinweisungen, Auslastung der Intensivstationen und Verfügbarkeit von Beatmungsgeräten auf regionaler, kantonaler und Krankenhausebene in der Schweiz zusammenfasst. Kapitel 5 enthält auch meinen Beitrag zur Bereitstellung von täglichen Geocodierungen von COVID-19-Fällen an das Schweizer Bundesamt für Gesundheit. Die Ergebnisse dieses Prozesses wurden verwendet, um tägliche Karten der COVID-19-Hotspots zu erstellen und darauf hinzuweisen, wie Krankenhausaufenthalte, Todesfälle und die Prävalenz positiver COVID-19-Fälle unter der Bevölkerung in den ärmsten Vierteln der Schweiz höher waren.

Chapter 1

Introduction

1.1 Open-access data and infectious diseases

Infectious diseases, with their capacity to spread across the globe within days (*1–3*), are a serious threat to public health. In recent years, the challenge posed by infectious diseases has been further exacerbated by the intersecting forces of globalization (*4*), climate change (*5*), and the intensification of agricultural production (*6*). In 2019, among the ~56.5 million deaths that occurred worldwide, ~7.9 million (13.9%) were associated with infectious diseases (*7*). In many low- and middle-income countries (LMICs), where 83% of the world's population lives (*8*), infectious diseases remain among the leading causes of death (*9*), causing ~7.3 million fatalities in 2019 (*7*).

Data that feed into surveillance systems is the cornerstone of interventions aimed at reducing the burden of infectious diseases. In particular, the systematic collection and analysis of epidemiological, clinical, and environmental data can help i) developing prevention measures such as vaccines and antiviral drugs, ii) detecting clusters of infections and planning rapid interventions (*10–12*), and iii) establishing historical precedents to improve preparedness for future outbreaks (*13*).

In HICs, these approaches led to establishing surveillance systems for diseases of global importance like influenza (*14*) and antibiotic-resistant infections (*15–17*). However, in LMICs, the development of surveillance systems is heavily undermined by scarce economic resources. An example of the disparity between HICs and LMICs concerns health expenditures per capita. In 2019 the average health expenditure per capita in HICs was 3,269 US$ compared to the 281 US $ per capita of LMICs (*18*). Similarly, important disparities also exist in the availability of health workforces, with 36.6 doctors per 10,000 people in HICs against only 1.5 per 10,000 in low-income countries (*19*). These trends are also reflected, for some LMICs, in the founding date of public health agencies tasked with preventing and managing infectious diseases. For example, the Robert Koch Institute in Germany was founded in 1891, the Dutch National Institute for Public Health and Environment in 1910, and the Instituto Superiore di Sanità in Italy in 1934. In contrast, the Nigeria Centre for Disease Control (NCDC) was established as recently as 2011. Similarly, the United Kingdom National Health Service initiated universal healthcare in 1948, while South Africa adopted a similar stance only in 1994 (*20*). A parallel scenario unfolds in the foundation of continent-level public health agencies. For example, the European Centers for Disease Control and Prevention (CDC) emerged in 2005, while the United States National Institute of Health and the CDC were formally founded in 1887 and 1946, respectively. In contrast, the Africa CDC emerged only in 2016.
One of the consequences of the lack of resources for healthcare in LMICs translates into comparatively poorer medical care than in HICs, but also, and perhaps more importantly: a limited understanding of how diseases emerge and spread in these regions. For example, in 2013, an 86-day delay in the detection of the primary Ebola virus case in West Africa led to an outbreak that was over 20 times the size of all previous outbreaks, lasting nearly four times longer than any previous outbreak (*21*). In the absence of well-established surveillance

infrastructure using proxies and alternative indicators could help better track diseases in LMICs.

In this context, the same modern era of globalization that generates challenges to contain infectious diseases might also come with opportunities by creating solutions often hidden in plain sight: the existence of large open-access databases. Important categories of open data are to be distinguished. First, some open-access data is produced for public health purposes but is scattered across the scientific literature and/or different online platforms used for disease surveillance. Examples of these types of open-access data are point prevalence surveys (PPS) conducted in human and animal samples to collect prevalence estimates of antimicrobial resistance (*22*, *23*). One of the main challenges to working with this category of data is the need to harmonize them since they might have been produced using methodologies different across scientific studies.
Second, there are open-access data that are created with purposes different than informing public health efforts (i.e.: sales, trade, etc.) but that in fact hold considerable potential to inform such efforts. Today, a considerable volume of this category of data is uploaded daily on the internet. People all over the world contribute to expanding digital maps by uploading the locations of services they use or places they visit on platforms like OpenStreetMap (OSM) and Google Maps. Some of these data sources have already been used in public health applications: online national phonebooks have been used to inventory health facilities (*24*), while postal codes available from medical associations have been used to identify the location of physicians and investigate their geographic distribution (*25*). The main challenge of working with this category of open-access data is represented by their acquisition: thousands of this data are listed online, and manually extracting them can be time-consuming and increase the errors that can naturally occur during sampling processes.
Furthermore, other open-access data created with purposes different than informing public health efforts are maps of global environmental and anthropogenic variables. These data can supplement open- and restricted-access databases, in combination with mathematical models, to make inferences on the geographic distribution of infectious diseases and healthcare capacities. These maps are typically disseminated by international organizations such as NASA, the United States Geological Survey, the European Space Agency, and other academics.

In LMICs, these categories of open-access data represent an unprecedented opportunity to gain insights into infectious disease trends. Firstly, in the absence of surveillance systems that collect disease data, using open-access data can be the only available option to investigate disease trends. Secondly, findings produced through open-access data allow researchers, healthcare professionals, and the public to easily inspect, reproduce, and strengthen study results. Finally, they can potentially supplement existing public health efforts based on traditional data such as medical and disease records, pharmaceutical clinical trials, and public health surveys (*26*). For instance, they can i) increase statistical power to reduce disease model uncertainty, ii) provide greater spatial or temporal granularity to model disease trends, and iii) serve as early disease detection systems. One such example is the screening of social media posts mentioning symptoms of infectious diseases.

While open-access data holds considerable potential to improve infectious disease surveillance in resource-limited settings, transforming this data into actionable information for decision-makers comes with challenges. Unlike randomized control trials, census, laboratory experiments, and systematic surveillance campaigns, open-access data results from a data generation process that is heterogeneous in nature. Therefore, harnessing this potential will depend on our capacity to amalgamate considerable volumes of information to smooth out heterogeneities that result from the data generation process, and satisfactorily harmonize it from a statistical perspective.

Addressing these challenges is the aim of this doctoral thesis.

1.2 Studying infectious diseases using open-access resources

"Open-access data" are broadly defined as freely accessible and unrestricted data, text, software, and multimedia (*27*). This definition was coined in the early 2000s (*28–30*) when international organizations recognized the importance of open-access data in every public sector. Thanks to these initiatives, the public health sector can today rely on subject-specific open-access repositories such as PubMed Central (*31*), the International Society for Infectious Diseases' ProMED (*32*), and the WHO Global Health Observatory (*33*). According to the 2019 State of Open Data report, more than 70% of researchers use open-access data to inform their future research (*34*).

In human health, one of the most significant efforts associated with using open-access data to study infectious diseases is the Malaria Atlas Project (*35*) a centralized geodatabase of statistics about the prevalence, burden, and severity of malaria at the global-level. Assembling such a database required an extensive literature review that today accounts for >24,000 surveys of prevalence rates of malaria performed since 1985 (*36*). Other initiatives focused on the human immunodeficiency virus (HIV) to share prevalence estimates at the county- and regional-level (*37*) and sequences of the virus' genome (*38*). Furthermore, in China, databases of national web searches were used to identify the geographic distribution of HIV diagnoses within the country (*39*). Another example concerned a dengue outbreak in Brazil, where a sentiment analysis based on Twitter users' posts revealed a good geographic relationship with the locations of new infections (*40*).

Although open-access data are valuable for tracking infectious diseases in humans, they can also largely contribute to the study of infectious diseases in food animals. One of the reasons for focusing on food animals is to prevent the emergence of potential zoonoses, which can pose a threat to human health as well (*41*). The link between human and animal health is one of the pillars of the One Health approach, an interdisciplinary strategy that recognizes the interconnectedness between humans, animals, and the environment. Within this paradigm, considering animals in disease surveillance and prevention efforts becomes essential, as the health of people is closely tied to the health of animals and our shared environment (*42*).

Furthermore, keeping animals healthy is crucially important for the economy sector represented by food animals and for the 600 million small-scale farmers who rely on animals for subsistence (*43*).
Examples of the use of open data in animal health include country-to-country trade records for poultry and wild birds from the Food and Agricultural Organization (FAO) that were used to track the spread of the avian influenza virus in Asia, Africa, and Europe (*44*). In another study, clinical data from 1,000 farms that participated in a national public health initiative were used to train machine learning algorithms to automatically diagnose mastitis in dairy herds (*45*). In addition, in 2017, the agricultural gross domestic product (*46*) estimates were used in combination with reports of monthly outbreaks of foot and mouth disease (FMD) to estimate the global economic risk caused by FMD in 216 countries (*47*). Finally, in the United Kingdom, open-access data of cattle movements were used for assessing the risks of disease transmission of bovine tuberculosis and outperformed environmental and anthropogenic to that effect (*48*).

In parallel, an increasing number of open-access platforms are contributing to the dissemination of data to support investigations on infectious diseases in animals. Considering again herds' movements, in Switzerland, the animal movement database (Tierverkehrsdatenbank, TVD) uses ear tags to monitor and record the movements of farm animals, and facilitates real-time contact tracing during disease outbreaks (*49*). Another example is the Animal Diseases Information System (ADIS), a European project created in collaboration with the World Organisation for Animal Health (WOAH), designed to register and document the evolution of the situation of important infectious animal diseases (*50*). Furthermore, the Infectious Diseases of East African Livestock project centralizes clinical and epidemiological data on infectious diseases in calves (*51*). Finally, the Enhanced Infectious Diseases database (ENHanCEd), systematically collates information about pathogens in humans and animals into a single resource (*52*).

1.3 Maps for infectious diseases surveillance

Maps carry multiple functions in veterinary epidemiology. First, maps are a visual data representation, making them more accessible and comprehensible for decision-makers than numerical outputs. Second, maps can be used for monitoring the geographical progression of diseases and help prevent outbreaks in healthy animals. Third, statistical associations between maps of environmental variables (*53*) and maps of diseases can help identify environmental risk factors that may contribute to disease transmission. Fourth, maps can also be the outputs of geostatistical models used to predict incidence rates of diseases (e.g., malaria in LMICs (*35*)). This can be useful to leverage the uneven coverage of data in different parts of the world, particularly in low- and middle-income countries (LMICs). Finally, maps can support the process of resource allocation against diseases, for example, to improve the accessibility to health services.

One of the earliest use of maps to monitor epidemics can be traced back to the 17th century when maps were used to identify areas affected by the plague in the province of Bari (Italy).

Originally, these maps were purely descriptive and were used to highlight areas where new cases of diseases appeared. Perhaps the most notable map used for this purpose is the one created by the British physician John Snow to track cholera cases in London in 1854 (Fig. 1A). Using such a map, Snow was able to identify the source of the infection – a water pump in the neighborhood of Soho – and stop the epidemic. A century later, one of the first global-level maps showed the diffusion of the "Asian" influenza spread by the H2N2 virus that caused a global pandemic that killed between 1 and 4 million people globally (Fig. 1B). Over time, new mapping methods such as the choropleth and heat maps were used to describe the prevalence of diseases within different administrative country levels (Fig. 1C and 1D).

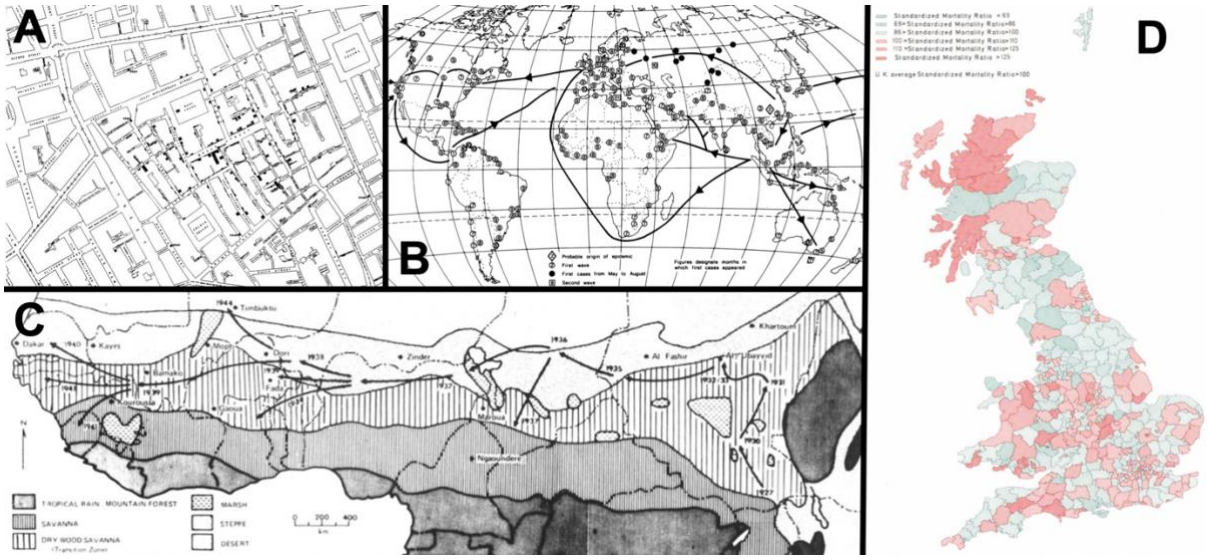

Fig. 1. Examples of historical maps used in public health. (A) Cholera cases (black boxes) tracked by John Snow in the Soho neighborhood of London in 1854 (*54*). (B) The global diffusion of "Asian" influenza between 1957 and 1958 (*55*). (C) Westward diffusion of cerebrospinal meningitis in sub-Saharan Africa prior to 1939 (*56*). (D) Standardized mortality ratio for breast cancer in Britain for women aged 15-64 years between 1980 and 1982 (*57*).

A significant milestone in the evolution of mapping methods was the introduction of the concept of spatial correlation, as articulated by Waldo R. Tobler in the form of the first law of geography (*58*). With increased computational capabilities, maps were further developed to incorporate this concept into statistical models applied to spatial elements. These advancements led to the development of the field of geostatistics, largely attributed to the contributions of Danie G. Krige and Georges Matheron (*59*). Their work formed the foundation for techniques that enable interpolating spatial data, capturing their correlation, and performing spatial predictions. These principles underlie different contemporary mapping methods. For example, the kriging technique is used to interpolate the values of a variable at observed locations and predict its value at unobserved locations, weighting predictions by the degree of correlation between observed values (*60*). Such a method was used for producing maps of species distribution (*61*) and maps of infectious diseases (*23*, *62–64*).

Another family of computational methods developed to study infectious diseases is network analysis, which is rooted in the graph theory developed in the 18th century (*65*). Network

analysis provides a means to examine interactions between spatial elements and was first introduced to enhance travel efficiency and reduce costs in urban planning. To date, this approach remains the cornerstone of the methods used to calculate travel times and generate maps of accessibility, amongst others for health facilities. Other methods in network analysis also include solvers for the "facility location problems" (*66*), which describe the challenges of maximizing access to a facility (i.e.: a hospital) for the largest population possible (*67*, *68*).

In this thesis, we explore the use of these methods in combination with open-access data for three case studies of global importance: antimicrobial resistance, COVID-19, and accessibility to veterinary healthcare.

1.4 Investigated case studies

Antimicrobial resistance

Antimicrobial resistance (AMR) is the ability of bacteria to evolve and withstand the effect of treatment with antibiotics (*69*). While the development of AMR is a naturally occurring phenomenon, excessive antibiotic use (AMU) accelerates this process (*70*). Infections with bacteria that are resistant to antibiotics (*71*) have serious consequences for human health. Currently, it is estimated that 1.3 million deaths occurring each year globally are attributable to antimicrobial-resistant infections (*72*).

However, to address the burden of AMR in humans and its relationship with AMU it is important to expand investigations beyond humans (*73*) and recognize the scale of AMU in animals. The antibiotics used in animals raised for food for the treatment and prevention of infectious diseases and, in 45 countries, for growth promotion (*74*), represent 73% of all antibiotics sold globally (*75*). The use of antibiotics in animals is driven by the increasing demand for animal protein (*76*), and contributes, along with other factors, to the increase of drug-resistant bacteria in food animals. These bacteria can then be transmitted to humans through the food chain or via environmental contamination. Therefore, regulating AMU in food animals has become a crucial component of the strategies for controlling AMR. While HICs can rely on surveillance systems to collect AMU and AMR data (*16*, *17*, *77*), in most of LMICs these systems are still in development. Reasons for the lack of systematic surveillance in these countries include the absence of laboratory facilities and a lack of qualified personnel for performing antimicrobial susceptibility testing (AST) (*78–80*).

In the absence of systematic surveillance, recent research efforts centralized open-access data scattered across the scientific literature to create resources intended as proxies of surveillance. For example, the global PPS project has established a global network of hospitals to assess and compare the quantity and quality of antibiotics prescribed and AMR prevalence in adults, children, and neonates worldwide (*22*). In a different effort, Hendriksen and colleagues attempted to monitor AMR worldwide using AMR gene abundances in urban sewage collected in 60 countries (*81*, *82*). In food animals, efforts focused on producing global predictions of

AMU using antibiotic sales (mostly from in HICs) (*83–85*), and collecting PPS of resistance prevalence estimates found in foodborne bacteria and summarized to predict current AMR trends in LMICs (*23*). Such results were integrated within geospatial models that returned global maps for the investigation of AMU and AMR trends at a finer geographical scale than the national one (i.e., at the 10x10 km^2 resolution) (*85*, *86*).

Further research initiatives based on online platforms hold the potential to improve proxies of surveillance based on PPS. Firstly, they can be used as repositories for centralizing existing PPS that were previously scattered across studies performed in different parts of the world. This can provide scientists with up-to-date databases that can support additional AMR investigations, such as studies about species-specific resistance and to identify new emerging AMR patterns in food animals. Secondly, online platforms could be used to summarize epidemiological information in outputs aggregated at different geographic levels, such as country-level reports. Third, they can become a focal point for the AMR community and engage researchers to contribute additional PPS to expand their existing database. Finally, they can overcome barriers imposed by publication and access fees of scientific journals by a free dissemination of their findings. Platforms conceived for these purposes are based on an architecture that requires communication with remote data storage services (Fig. 2).

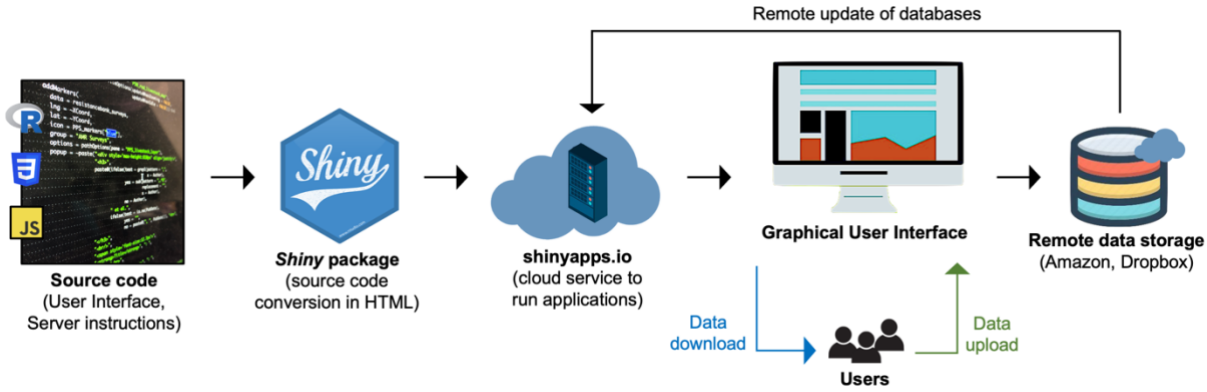

Fig. 2. Workflow of an open-access platform programmed in R. The workflow shows the operation of a platform conceived for data sharing. The source code of the platform is translated into HyperText Markup Language (HTML) to work on every browser and then uploaded on a cloud server that runs the platforms on the internet. Users can interact with the platform through a user-friendly Graphical User Interface, download outputs (e.g., datasets, documents, maps), and upload their data, which will be integrated into databases stored on remote data storage services (e.g., Amazon S3, Dropbox). These up-to-date databases will then be loaded each time new users access the platform.

These platforms can be coded in different programming languages (R, python, Java, etc.). Specifically, the R environment provides packages such as *shiny* (*87*) that translate the platform source code into HyperText Markup Language (HTML), used by web browsers. Furthermore, R also provides hosting services consisting of cloud servers to run platforms coded in R. This can considerably simplify the work of developers for configuring their local servers. In addition, platforms conceived for data sharing must rely on remote data storage services (e.g.,

Amazon, Dropbox, or Google) that will contain databases loaded upon users' access (Fig. 2, blue arrow). These services will be used to store data uploaded by the users (Fig. 2, green arrow) that will be subsequently combined with the existing databases.

Veterinarians

In the context of managing AMU and monitoring AMR in food animals, it is essential to recognize the role played by veterinarians. On the one hand, veterinarians could exacerbate AMR through unnecessary and non-targeted antibiotic prescriptions that could be partially driven by economic incentives, such as profits deriving from antibiotic sales (*88*). On the other hand, they play a key role in monitoring infectious diseases (*89*), promoting responsible farming practices, such as using narrow-spectrum antibiotics based on diagnostic tools (*90*), as well as performing regular check-ups on animals.

These aspects acquire greater importance when considering the livelihood of the 1.3 billion people (*91*) who rely on food animals for their subsistence. Of these, 600 million are small-scale farmers, and the majority of them currently live in extreme poverty (<2$ per day) (*92*). Therefore, keeping animals healthy is crucially important for their economic, sociocultural, environmental, and nutritional well-being. Lack of access to veterinarians, especially in these settings, can create multiple problems: for example, the absence of systematic guidance from veterinarians might result in a decrease in food animals' productivity due to the reliance of farmers on lay knowledge of animal health (*93*). This could lead to economic losses and food insecurity for the population (*94*), and potential misuse of antibiotics (*95*). Furthermore, lack of veterinary surveillance can increase the number of infections spreading undetected among animals (*96*), increasing the risk of epidemics that can impact food production and public health.

However, assessing the availability of veterinarians for food animals at a finer geographical scale than the national one is challenging. Currently, veterinary associations and organizations provide estimates of the number of veterinarians at the country-level (*97*), with only a few HICs reporting similar information at the provincial- or regional level (*98–102*). In addition, even fewer studies have investigated the national veterinarians' distribution (*103–105*) and their degree of accessibility (*99*). This situation can be attributed to multiple factors. Firstly, countries with limited resources may lack the means to systematically inventory fine-scale data of veterinarians. Secondly, the resolution of this data – when it exists – can vary significantly between regions, municipalities, and individual practices, and lacks standardization. Thirdly, many countries have stringent data privacy laws, such as the General Data Protection Regulation in Europe, which safeguards the personal information of healthcare professionals, including veterinarians. Similarly, professional associations and licensing bodies may have policies that limit the dissemination of information related to their members.

Without a fine-scale distribution of veterinarians at the global-level it is then challenging to quantify the differences in access to care at the sub-national level, especially in LMICs. Furthermore, outlining the global distribution of veterinarians could help support the

identification of “veterinary coldspots”. These represent areas where animals are farther than 1 hour of travel time from a veterinarian. This is a concept similar to the “golden hour” defined for humans, referring to the evidence that people who are cared within 1 hour after a traumatic event have a high chance of a positive health outcome (*106*). Mapping these coldspots globally could help in planning capacity-building efforts to expand the veterinary workforce.

In this context, open-access data could constitute an unprecedented opportunity to get a first outline of the global distribution of veterinarians. Such open-access data can be represented by the addresses of veterinarians scattered among diverse online platforms such as national phonebooks, tools to find veterinarians by postcode, registries of veterinary associations, and governmental reports. Their addresses can be used to obtain more granular information about veterinary capacities than the ones currently available at the national level.

However, assembling a global database of addresses to investigate the coldspots distribution comes with challenges. First, there is a need to identify exhaustively all potential platforms, and then techniques must be used to systematically collect such addresses to prevent errors occurring naturally during processes of manual data extraction and mitigate the long time required to extract thousands of addresses present within platforms of big countries. A solution to these challenges resides in web-scraping. Web scraping is a procedure using software that can be programmed in different languages (e.g., R, python, etc.) for automatic web data extraction (*107*). The history of web-scraping dates back to the 1990s, right after the World Wide Web was born (*108*). Since web pages are coded through the HTML language, web-scrapers were developed to read this HTML code (comparable to the text on book pages), identify specific strings of code (the words in the text), and assemble them into a database (Fig. 3). In addition, advanced software for web-scraping called web-crawlers (*109*), were developed to automatize the web pages navigation, which requires, for example, to press specific buttons and display additional text to collect.

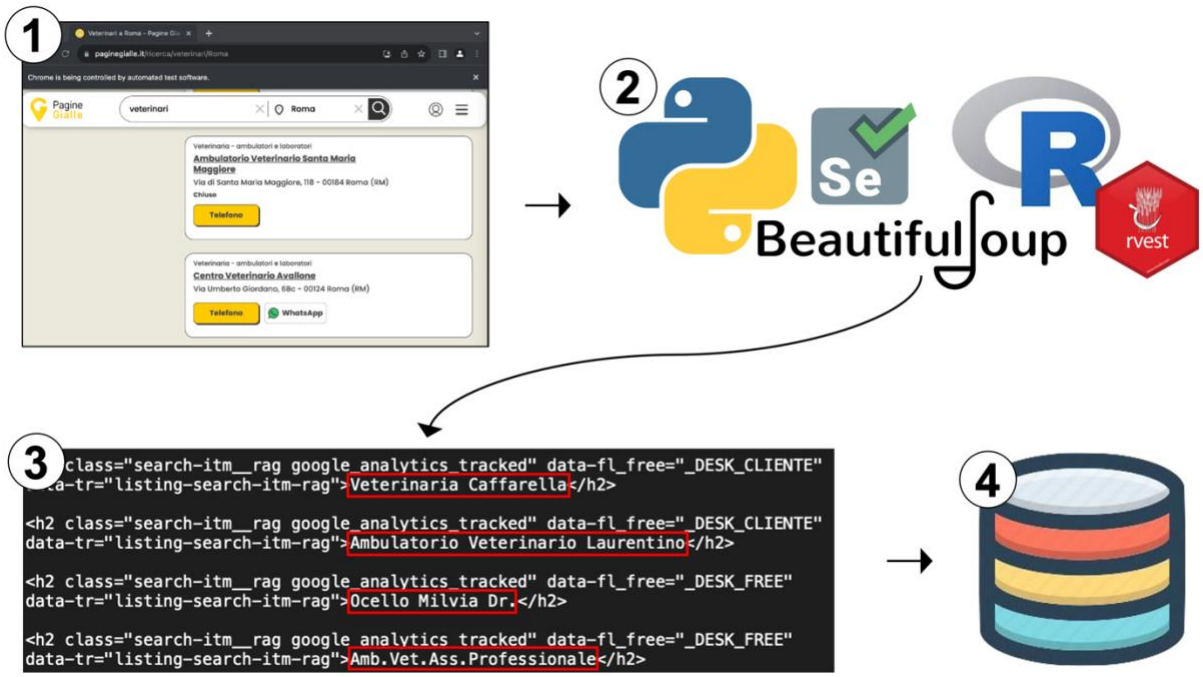

Fig. 3. Web-scraping workflow. 1) Launch of the web page containing the data to collect. 2) Web-scraping software defined through programming languages like R and python access the web page and read its HyperText Markup Language (HTML) source code. 3) Identification of the target text (red boxes) present within structures of the HTML code that repeats inside the web page (e.g., white blocks shown in the figure of Step 1). 4) This information is extracted and then assembled inside a database.

A database of veterinarians' addresses assembled through web-scraping represents the starting point for getting granular information about the workforce of veterinarians worldwide. In addition, these data could help supplement ongoing projects that aim to strengthen national veterinary capacities such as the Performance of Veterinary Services (PVS) pathway defined by the World Organisation for Animal Health (WOAH) (*110*). The PVS pathway offers countries the methods for a self-evaluation of their veterinary workforce. Therefore, providing them with a database of veterinarians' addresses could ease this process.
However, additional methods are required to identify what is the degree of accessibility of veterinarians to food animals living inside coldspots. Drawing from the definition of coldspots based on travel times, it is possible to use maps reporting the land coverage to calculate travel times to reach health services, such as veterinarians. These maps are called land cover maps (Fig. 4) and are openly disseminated at different resolutions (e.g., 100 m^2) by services like Copernicus Global Land Service (*111*) or NASA (*112*) at the global-level. These maps are created from satellite photographs obtained through sensors that can capture different land coverages (e.g., forests, water bodies, urban areas).

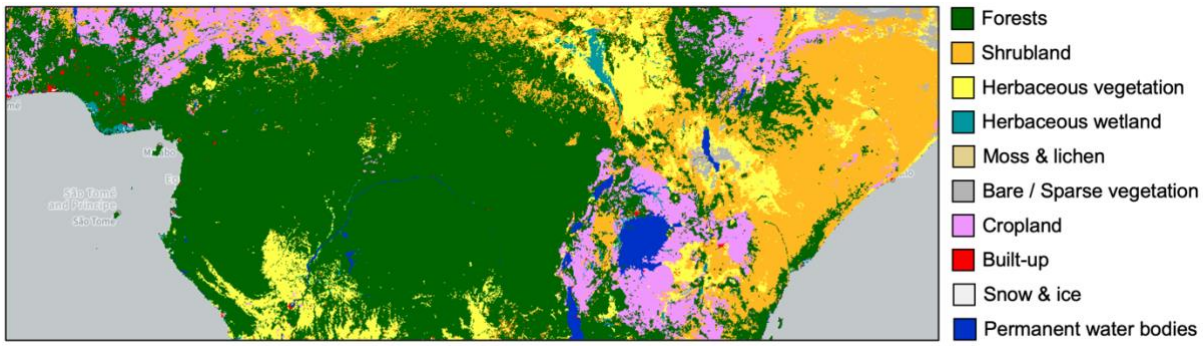

Fig. 4. Land cover map. Classes of physical land coverage in the North of Sub-Saharan Africa in 2019 at 100 m^2 resolution, disseminated by the Copernicus Global Land Service (*111*).

Next, land cover maps can be combined with maps of road networks available from OSM that are associated with a traveling speed. In a process known as "rasterization", these vectors can be converted into maps that match the resolution of land cover maps and are overlayed to them. Then, additional speed limits can be associated with the land cover classes where traveling by walking is possible (generally, the speed considered in this case is 5 km per hour (*113*)). Knowing the speed to travel across each class, these maps can be converted into friction surfaces, which are maps that contain information about the time required to cross a unit of length (e.g., meters) in each of their pixels. Notable examples of global friction surfaces available at 1 km^2 resolution are the ones produced by Weiss and colleagues to calculate travel time maps to cities with >50,000 inhabitants (*114*) or to health facilities (*115*). Other friction surfaces at a higher resolution (e.g., 100 m^2) have been so far assembled at the country-level and could provide more granular information about the population lacking accessibility to health facilities (*116*).

The identification of veterinary coldspots using these travel time maps can help in quantifying the number of food animals lacking prompt access to care. Furthermore, these maps can also support efforts to increase the number of animals with access to care by identifying ways to improve the network of veterinarians. One of the purposes of this network analysis, known as the facility location problem (*66*), is to maximize the population brought within a certain travel time threshold from facilities by allocating supplementary ones (*117*). This approach is defined as the maximal covering location problem (MCLP) (*118*). MCLP-based approaches used with maps of travel times and population density allow to identify locations of additional facilities that maximize the population coverage. Therefore, they can help support an efficient increment of veterinarians to reduce coldspots.

SARS-CoV-2

Open-access data can be essential to supplement or make up for the lack of official data during health emergencies. As of October 2023, the COVID-19 caused nearly 7 million deaths worldwide (*119*). In Switzerland, the first COVID-19 wave started in February 2020, leading the Swiss government to implement measures such as social distancing, and banning large events to limit the spread of the virus. Compared to countries like Italy, France, and Spain,

where a national lockdown was imposed (*120*), these measures were less severe, and they were eased during the summer of 2020. However, in October 2020, the second, and deadliest wave of COVID-19 led to stricter sanitary measures in Switzerland (*121*).

Along with these measures, open-access data and platforms available at the national- and global-level played a crucial role in monitoring the spread of the virus and planning national resource allocation (*122–124*). In Switzerland, COVID-19-related data such as the number of daily deaths, daily reported cases, and daily hospitalizations were initially daily centralized, and grouped by canton, by the open-access platform OpenZH (*125*). In addition, confidential time series of hospitalizations in intensive care units (ICU), at the hospital-level, were reported twice a day by the Information and Operation System (IES) managed by the Coordinated Sanitary Service of Switzerland.

In this context, ICU occupancy data could be used as a proxy to outline the pandemic trajectory at its initial stage (when COVID-19 tests were not yet available) and disseminate such information through user-friendly platforms to competent authorities and the public. Therefore, with the support of the Swiss Armed Forces, our research group supplemented restricted access data (IES) with open-access data (OpenZH) to provide forecasts of ICU occupancy, in addition to deaths, hospitalizations, and availability of ventilators at the hospital-, cantonal-, and regional-level of Switzerland. Although hospital-level data remained confidential, their forecasts were aggregated at the cantonal- and regional-level. In addition, data from IES and OpenZH were both included in the models forecasting COVID-19 cases to increase their predictive power.

Providing these forecasts to decision-makers became the priority during the pandemic, and open-access platforms (Fig. 2) can be useful for this purpose. Different platforms were developed in 2020 to support global (*119*, *123*) and local (*126*) efforts aiming at maximizing the spread of information useful for limiting the burden of COVID-19. In this context, a platform releasing data about ICU occupancy could not only be used as a proxy for tracking the evolution of the pandemic but also for actively supporting competent authorities in managing the flow of patients being hospitalized. However, it is essential that this work is not hindered by the spread of misinformation upon potential leaks of sensible data, such as the ICU capacities of each Swiss hospital. Therefore, this type of open-access platform should be used to display sensible data with the public only when released in an aggregated format (e.g., at the cantonal- or regional-level), and use modules to encrypt such data (e.g., hospital-level ICU occupancy) to display them through an authentication system, based on username and password, only with competent authorities.

Besides platforms, the spread of COVID-19 was also investigated using maps in combination with open-access data (*127*, *128*). In Switzerland, our group established a collaboration with the Federal Office of Public Health (FOPH) to provide daily maps of COVID-19 hotspots across the country. For this reason, in June 2020, I went with my Ph.D. supervisor to Bern, where we had a meeting to establish a collaboration with members of the FOPH who had access to the databases of every person in Switzerland who tested for COVID-19 antigens. The

databases also contained their residential address. Therefore, we offered to develop a pipeline for the routine geocoding of these addresses using the dedicated Google Application Programming Interface (API). Due to the sensitive content of these databases, using an "anonymization by dilution" that mixes people's addresses with the addresses of public facilities like restaurants and shops can prevent potential hacks in the API framework to rebuild the original database of COVID-19 positive cases.

Once geocoded, these addresses can be used to produce descriptive maps showing the locations of new daily infections and how their hotspots change or increase over time (*129*). This can support the investigation of the spatial trends of a pandemic and assist timely interventions to allocate additional healthcare resources where hotspots are emerging. Furthermore, using geographic coordinates of COVID-19 cases in combination with socioeconomic factors has the potential to unravel important socioeconomic disparities among the neighborhoods of the country. An example is the availability of healthcare resources, such as COVID-19 tests and ICU admissions, at different socioeconomic strata of the population.

1.5 Thesis structure

The research questions of this thesis develop around the potential of using open-access data and platforms as resources to improve the surveillance of food animals and human health. Based on such premises, this thesis is organized around four projects:

1. **Open-access repository for AMR data sampled in food animals**. This project describes the first version of the open-access platform *resistancebank.org*. When launched in 2019, the platform centralized 1,285 PPS reporting AMR prevalence estimates obtained from four foodborne pathogens (*Escherichia coli*, non-typhoidal *Salmonella* spp., *Staphylococcus aureus*, and *Campylobacter* spp.) sampled in cattle, chickens, pigs, and sheep across LMICs. A subsequent integration of PPS in September 2021 led to an upgrade of the platform, which in its current version contains 2,045 PPS performed on pathogens sampled both in terrestrial and aquatic food animals. Currently, *resistancebank.org* summarizes the information of these PPS through different, open-access outputs: 10x10 km^2 maps of AMR, species-specific resistance prevalence estimates to each antibiotic tested per PPS, and country-level reports with socio-economic information and AMU estimates, along with AMR prevalence estimates for different drug-pathogen combinations.

2. **Identification of coldspots of veterinary capacity**. This project illustrates how web-scraping techniques were used to amalgamate 303,745 addresses of veterinarians available in the public domain. These addresses, converted in geographic locations, were used with geospatial models and spatial covariates (i.e., maps of human population density, gross domestic product, travel time to cities, and food animals' density) to predict the global distribution of veterinarians at the 10x10 km^2 resolution. The project then describes how such a map was used to identify areas lacking accessibility to

veterinarians, comparable to the "medical deserts" for the human population. These areas, called "veterinary coldspots", were defined as areas where food animals are farther than 1 hour of travel time from the nearest veterinarian, and were observed predominantly in LMICs.

3. **Optimal allocation of public health facilities**. Building upon the topics of projects 1 and 2, this project illustrates how open-access data can be used to investigate the spatial allocation of public health facilities. The aim was to quantify the physical accessibility to facilities within 1 hour of travel time when targeted approaches were used to increase the national number of facilities by 5%. Using maps reporting traveling costs within each area of a country, two case studies related to AMR and veterinary capacities are presented in this project:

    - The first identifies locations of public health facilities (hospitals and diagnostic laboratories) to equip for AST in Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi to maximize the population living within 1 hour of motorized travel time from these facilities. Increasing their availability can increase the number of AST performed and better frame how AMR trends in humans and animals evolve.

    - The second explores the allocation of veterinarians to reduce the number of food animals in coldspots. However, this case study focuses more on approximating the allocation approach that maximizes the food animals removed from coldspots to reduce the computational time of the allocations when working with global-level maps.

4. **Open-access resources against COVID-19**. This project was conceived during the waves of the COVID-19 pandemic that hit Switzerland (2020-2021). It describes the development of open-access resources used to monitor and mitigate the effects of COVID-19. Specifically, data related to the forecasts of ICU occupancy were summarized in the online platform *icumonitoring.ch*, aiding decision-makers in managing the large flow of hospitalizations caused by the pandemic.
    This chapter also describes my involvement in another "operational support" project related to COVID-19 in Switzerland. We used the geographic coordinates of positive COVID-19 cases to investigate the relationship between socioeconomic indicators of the neighborhoods where they lived with the number of COVID-19 tests performed among these neighborhoods, as well as the number of positive cases, hospitalizations, and deaths.

References

1. G. Magiorkinis, K. Angelis, I. Mamais, A. Katzourakis, A. Hatzakis, J. Albert, G. Lawyer, O. Hamouda, D. Struck, J. Vercauteren, et al. The global spread of HIV-1 subtype B epidemic. *Infection, Genetics and Evolution* **46**, 169–179 (2016). doi:10.1016/j.meegid.2016.05.041
2. G. Tsoucalas, A. Kousoulis, M. Sgantzos. The 1918 Spanish Flu Pandemic, the Origins of the H1N1-virus Strain, a Glance in History. *European Journal of Clinical and Biomedical Sciences* **2**, 23–28 (1918). doi:10.11648/j.ejcbs.20160204.11
3. S. Platto, Y. Wang, J. Zhou, E. Carafoli. History of the COVID-19 pandemic: Origin, explosion, worldwide spreading. *Biochemical and Biophysical Research Communications* **538**, 14–23 (2021). doi:10.1016/j.bbrc.2020.10.087
4. B. D. Gushulak, D. W. MacPherson. Globalization of infectious diseases: The impact of migration. *Clinical Infectious Diseases* **38**, 1742–1748 (2004). doi:10.1086/421268
5. C. Mora, T. McKenzie, I. M. Gaw, J. M. Dean, H. von Hammerstein, T. A. Knudson, R. O. Setter, C. Z. Smith, K. M. Webster, J. A. Patz, et al. Over half of known human pathogenic diseases can be aggravated by climate change. *Nature Climate Change* **12**, (2022). doi:10.1038/s41558-022-01426-1
6. B. G. Henning. Standing in Livestock's 'Long Shadow': The Ethics of Eating Meat on a Small Planet. *Ethics and the Environment* **16**, 63 (2011). doi:10.2979/ethicsenviro.16.2.63
7. Institute for Health Metrics and Evaluation (IHME) - Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results (2022). Available at: https://vizhub.healthdata.org/gbd-results/. Accessed: 22/12/2023.
8. The World Bank. Population, total - Low & middle income (2022). Available at: https://data.worldbank.org/indicator/SP.POP.TOTL?end=2022&locations=XO&start=2012. Accessed: 11/10/2023.
9. The World Health Organization. The top 10 causes of death (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death#:~:text=People living in a low,remain in the top 10. Accessed: 13/11/2023.
10. S. Mazeri, J. L. Burdon Bailey, D. Mayer, P. Chikungwa, J. Chulu, P. O. Grossman, F. Lohr, A. D. Gibson, I. G. Handel, B. M. Barend, et al. Using data-driven approaches to improve delivery of animal health care interventions for public health. *Proceedings of the National Academy of Sciences of the United States of America* **118**, (2021). doi:10.1073/pnas.2003722118
11. B. I. Ajuwon, K. Roper, A. Richardson, B. A. Lidbury. One Health Approach: A Data-Driven Priority for Mitigating Outbreaks of Emerging and Re-Emerging Zoonotic Infectious Diseases. *Tropical Medicine and Infectious Disease* **7**, 0–5 (2022). doi:10.3390/tropicalmed7010004
12. N. Papachristou, G. Kotronoulas, N. Dikaios, S. J. Allison, H. Eleftherochorinou, T. Rai, H. Kunz, P. Barnaghi, C. Miaskowski, P. D. Bamidis. Digital Transformation of Cancer Care in the Era of Big Data, Artificial Intelligence and Data-Driven Interventions: Navigating the Field. *Seminars in Oncology Nursing* **39**, 151433 (2023). doi:10.1016/j.soncn.2023.151433
13. C. Potter, T. Harwood, J. Knight, I. Tomlinson. Learning from history, predicting the future: The UK Dutch elm disease outbreak in relation to contemporary tree disease threats. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 1966–1974 (2011). doi:10.1098/rstb.2010.0395
14. J. R. Ortiz, V. Sotomayor, O. C. Uez, O. Oliva, D. Bettels, M. McCarron, J. S. Bresee,

A. W. Mounts. Strategy to enhance influenza surveillance worldwide. *Emerging Infectious Diseases* **15**, 1271–1278 (2009). doi:10.3201/eid1508.081422
15. EFSA. Antimicrobial resistance (2023). Available at: https://www.efsa.europa.eu/en/topics/topic/antimicrobial-resistance. Accessed: 06/10/2023.
16. CDC. National Antimicrobial Resistance Monitoring System for Enteric Bacteria (NARMS) (2023). Available at: https://www.cdc.gov/narms/index.html. Accessed: 06/10/2023.
17. Government of Canada. Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) (2023).
18. World Health Organization. Global Health Expenditure Database - Current Health Expenditure (CHE) per Capita in US$ (2019). Available at: https://apps.who.int/nha/database/ViewData/Indicators/en. Accessed: 11/11/2023.
19. World Health Organization. The Global Health Observatory - Global Health Workforce Statistics Database (2019). Available at: https://www.who.int/data/gho/data/themes/topics/health-workforce. Accessed: 11/11/2023.
20. Wikipedia. Healthcare in South Africa. Available at: https://en.wikipedia.org/wiki/Healthcare_in_South_Africa. Accessed: 15/11/2023.
21. M. J. Matson, D. S. Chertow, V. J. Munster. Delayed recognition of Ebola virus disease is associated with longer and larger outbreaks. *Emerging Microbes and Infections* **9**, 291–301 (2020). doi:10.1080/22221751.2020.1722036
22. Global PPS. Antimicrobial consumption and resistance (2023). Available at: https://www.global-pps.com/project/. Accessed: 10/11/2023.
23. T. P. Van Boeckel, J. Pires, R. Silvester, C. Zhao, J. Song, N. G. Criscuolo, M. Gilbert, S. Bonhoeffer, R. Laxminarayan. Global trends in antimicrobial resistance in animals in low- And middle-income countries. *Science* **365**, (2019). doi:10.1126/science.aaw1944
24. C. Wolff, S. Roesel, G. Lipskaya, M. Landaverde, A. Humayun, N. Withana, N. Ramamurty, O. Tomori, S. O. Okiror, M. Salla, et al. Progress toward laboratory containment of poliovirus after polio eradication. *Journal of Infectious Diseases* **210**, S454–S458 (2014). doi:10.1093/infdis/jit821
25. M. B. Rosenthal, A. Zaslavsky, J. P. Newhouse. The geographic distribution of physicians revisited. *Health Services Research* **40**, 1931–1952 (2005). doi:10.1111/j.1475-6773.2005.00440.x
26. J. S. Ross, R. Lehman, C. P. Gross. The Importance of Clinical Trial Data Sharing: Toward More Open Science. *Circ Cardiovasc Qual Outcomes* **5**, 238–240 (2012). doi:10.1161/CIRCOUTCOMES. 112.965798
27. UNESCO. What is Open Access? (2021). Available at: https://en.unesco.org/open-access/what-open-access#:~:text=Open access (OA) means free,video%2C and multi-media. Accessed: 01/10/2023.
28. Budapest Open Access Initiative. Declaration about Open Access Data (2002). Available at: https://www.budapestopenaccessinitiative.org/read/. Accessed: 14/10/2023.
29. Wikipedia. Bethesda Statement on Open Access Publishing (2023). Available at: https://en.wikipedia.org/wiki/Bethesda_Statement_on_Open_Access_Publishing#:~:text=The group made a definition,for any responsible purpose%2C subject. Accessed: 14/10/2023.
30. Wikipedia. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2022). Available at:

https://en.wikipedia.org/wiki/Berlin_Declaration_on_Open_Access_to_Knowledge_in_the_Sciences_and_Humanities. Accessed: 14/10/2023.
31. National Library of Medicine. PubMed Overview (2023). Available at: https://pubmed.ncbi.nlm.nih.gov/about/. Accessed: 14/10/2023.
32. International Society for Infectious Diseases. ProMED (2023). Available at: https://promedmail.org. Accessed: 17/10/2023.
33. WHO. The Global Health Observatory (2023). Available at: https://www.who.int/data/gho. Accessed: 14/10/2023.
34. P. H. Ayris, I. Mark Hrynaszkiewicz, B. Fane, G. Baynes, E. Farrell. The State of Open Data 2019. *Springer Nature* 1–21 (2019). doi:10.6084/m9.figshare.9980783
35. The Malaria Atlas Project. About MAP (2023). Available at: https://malariaatlas.org/about-map/. Accessed: 29/10/2023.
36. C. L. Moyes, W. H. Temperley, A. J. Henry, C. R. Burgert, S. I. Hay. Providing open access data online to advance malaria research and control. *Malaria Journal* **12**, 1 (2013). doi:10.1186/1475-2875-12-161
37. UNAIDS. The AIDS data repository (2023). Available at: https://adr.unaids.org/en/. Accessed: 29/10/2023.
38. Los Alamos National Laboraotories. HIV databases (2023). Available at: https://www.hiv.lanl.gov/content/index. Accessed: 29/10/2023.
39. Q. Zhang, Y. Chai, X. Li, S. D. Young, J. Zhou. Using internet search data to predict new HIV diagnoses in China: A modelling study. *BMJ Open* **8**, (2018). doi:10.1136/bmjopen-2017-018335
40. J. Gomide, V. Almeida, A. Veloso, F. Bnevenuto, W. Meira, F. Ferraz. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. *WebSci '11: Proceedings of the 3rd International Web Science Conference* 1–8 (2011). doi:10.1145/2527031.2527049
41. K. E. Jones, N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, P. Daszak. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008). doi:10.1038/nature06536
42. Centers for Disease Contol and Prevention. One Health Basics (2023). Available at: https://www.cdc.gov/onehealth/basics/index.html#:~:text=What is One Health%3F,animals and our shared environment. Accessed: 22/12/2023.
43. WOAH. GBADs - The Global Burden of Animal Diseases (2023). Available at: https://gbads.woah.org. Accessed: 12/10/2023.
44. A. M. Kilpatrick, A. A. Chmura, D. W. Gibbons, R. C. Fleischer, P. P. Marra, P. Daszak. Predicting the global spread of H5N1 avian influenza. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19368–19373 (2006). doi:10.1073/pnas.0609227103
45. R. M. Hyde, P. M. Down, A. J. Bradley, J. E. Breen, C. Hudson, K. A. Leach, M. J. Green. Automated prediction of mastitis infection patterns in dairy herds using machine learning. *Scientific Reports* **10**, 1–8 (2020). doi:10.1038/s41598-020-61126-8
46. The World Bank. Agriculture, forestry, and fishing, value added (% of GDP) (2017). Available at: https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS. Accessed: 07/11/2023.
47. D. W. Shanafelt, C. A. Perrings. Foot and mouth disease: The risks of the international trade in live animals. *OIE Revue Scientifique et Technique* **36**, 839–865 (2017). doi:10.20506/rst36.3.2719
48. M. Gilbert, A. Mitchell, D. Bourn, J. Mawdsley, R. Clifton-Hadley, W. Wint. Cattle movements and bovine tuberculosis in Great Britain. *Nature* **435**, 491–496 (2005). doi:10.1038/nature03548

49. Swiss Federal Office of Public Health. Tierverkehrsdatenbank (2023). Available at: https://www.blw.admin.ch/blw/de/home/nachhaltige-produktion/tierische-produktion/tvd.html. Accessed: 15/11/2023.
50. The European Union. ADIS - Animal Disease Information System (2023). Available at: https://food.ec.europa.eu/animals/animal-diseases/animal-disease-information-system-adis_en. Accessed: 09/11/2023.
51. R. Callaby, C. Pendarovski, A. Jennings, S. T. Mwangi, I. Van Wyk, M. Mbole-Kariuki, H. Kiara, P. G. Toye, S. Kemp, O. Hanotte, et al. IDEAL, the Infectious Diseases of East African Livestock project open access database and biobank. *Scientific Data* **7**, 1–11 (2020). doi:10.1038/s41597-020-0559-7
52. University of Liverpool. ENHanCEd Infectious Diseases (EID2) database (2023). Available at: https://eid2.liverpool.ac.uk. Accessed: 29/10/2023.
53. NASA. Earth Data - Open Access for Open Science (2023). Available at: https://www.earthdata.nasa.gov. Accessed: 29/10/2023.
54. J. Snow. On the mode of communication of cholera. (1855).
55. G. C. Schild, G. M. Howe. Influenza. *A world geography of human diseases* (1977).
56. H. J. Jusatz, E. Rodenwaldt. Cerebrospinal meningitis. *World atlas of epidemic diseases* **111**, 39–44 (1961).
57. H. C. Hopps, R. J. Cuffey, R. J. Morenoff, J. Richmond, W. L. Sidley. Computerised mapping of disease and environmental data. *U.S. Department of Defence* (1968).
58. E. Geography. A Computer Movie Simulating Urban Growth in the Detroit Region Author (s): W . R . Tobler Source : Economic Geography , Jun ., 1970 , Vol . 46 , Supplement : Proceedings . International Geographical Union . Commission on Quantitative Methods (Jun ., 19. **46**, 234–240 (1970).
59. M. D. Ecker. Geostatistics: past, present, and future. *Environmetrics* (2011).
60. M. A. Oliver, R. Webster. Kriging: A method of interpolation for geographical information systems. *International Journal of Geographical Information Systems* **4**, 313–332 (1990). doi:10.1080/02693799008941549
61. K. Amiri, N. Shabanipour, S. Eagderi. Using kriging and co-kriging to predict distributional areas of Kilka species (Clupeonella spp.) in the southern Caspian Sea. *International Journal of Aquatic Biology* **5**, 108–113 (2017).
62. P. Goovaerts. Geostatistical analysis of disease data: Estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *International Journal of Health Geographics* **4**, (2005). doi:10.1186/1476-072X-4-31
63. R. Lecoustre. Analysis and Mapping of the Spatial Spread of African Cassava Mosaic Virus Using Geostatistics and the Kriging Technique. *Phytopathology* **79**, 913 (1989). doi:10.1094/phyto-79-913
64. M. Ali, P. Goovaerts, N. Nazia, M. Z. Haq, M. Yunus, M. Emch. Application of Poisson kriging to the mapping of cholera and dysentery incidence in an endemic area of Bangladesh. *International Journal of Health Geographics* **5**, 1–11 (2006). doi:10.1186/1476-072X-5-45
65. R. J. Wilson. History of Graph Theory. *Theory CRC Press* (2013).
66. H. A. Eiselt, V. Marianov. Foundations of Location Analysis. *Springer* (2011).
67. E. B. Laber, N. J. Meyer, B. J. Reich, K. Pacifici, J. A. Collazo, J. M. Drake. Optimal treatment allocations in space and time for on-line control of an emerging infectious disease. *Journal of the Royal Statistical Society. Series C: Applied Statistics* **67**, 743–789 (2018). doi:10.1111/rssc.12266
68. Y. Devi, S. Patra, S. P. Singh. A location-allocation model for influenza pandemic outbreaks: A case study in India. *Operations Management Research* **15**, 487–502 (2022). doi:10.1007/s12063-021-00216-w

69. M. A. Abushaheen, Muzaheed, A. J. Fatani, M. Alosaimi, W. Mansy, M. George, S. Acharya, S. Rathod, D. D. Divakar, C. Jhugroo, et al. Antimicrobial resistance, mechanisms and its clinical significance. *Disease-a-Month* **66**, (2020). doi:10.1016/j.disamonth.2020.100971
70. V. M. D'Costa, C. E. King, L. Kalan, M. Morar, W. W. L. Sung, C. Schwarz, D. Froese, G. Zazula, F. Calmels, R. Debruyne, et al. Antibiotic resistance is ancient. *Nature* **477**, 457–461 (2011). doi:10.1038/nature10388
71. C. G. Whitney, M. Farley M., J. Hadler, L. H. Harrison, C. Lexau, A. Reingold, L. Lefkowitz, P. R. Cieslak, M. Cetron, E. Zell, et al. Increasing prevalence of multidrug-resistant streptococcus pneumoniae in the United States. *The New England Journal of Medicine* **343**, 1917–1924 (2000).
72. C. J. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022). doi:10.1016/S0140-6736(21)02724-0
73. S. I. Hay, P. C. Rao, C. Dolecek, N. P. J. Day, A. Stergachis, A. D. Lopez, C. J. L. Murray. Measuring and mapping the global burden of antimicrobial resistance. *BMC Medicine* **16**, 1–3 (2018). doi:10.1186/s12916-018-1073-z
74. World Organisation for Animal Health. OIE Annual report on antimicrobial agents intended for use in animals. **3**, 1–129 (2018).
75. Organisation for Economic Co-operation and Development. Global antimicrobial use in the livestock sector. *Working Party on Agricultural Policies and Markets* 1–43 (2015).
76. M. Henchion, M. Hayes, A. M. Mullen, M. Fenelon, B. Tiwari. Future protein supply and demand: Strategies and factors influencing a sustainable equilibrium. *Foods* **6**, 1–21 (2017). doi:10.3390/foods6070053
77. E. Food, S. Authority. The European union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2017. *EFSA Journal* **17**, (2019). doi:10.2903/j.efsa.2019.5598
78. Food and Agriculture Organization. Taking a multisectoral one health approach: a tripartite guide to addressing zoonotic diseases in countries (2019). Available at: https://www.fao.org/3/ca2942en/CA2942EN.pdf. Accessed: 31/10/2023.
79. E. K. Rousham, L. Unicomb, M. A. Islam. Human, animal and environmental contributors to antibiotic resistance in low-resource settings: Integrating behavioural, epidemiological and one health approaches. *Proceedings of the Royal Society B: Biological Sciences* **285**, (2018). doi:10.1098/rspb.2018.0332
80. M. Nadimpalli, E. Delarocque-Astagneau, D. C. Love, L. B. Price, B. T. Huynh, J. M. Collard, K. S. Lay, L. Borand, A. Ndir, T. R. Walsh, et al. Combating Global Antibiotic Resistance: Emerging One Health Concerns in Lower-and Middle-Income Countries. *Clinical Infectious Diseases* **66**, 963–969 (2018). doi:10.1093/cid/cix879
81. R. S. Hendriksen, P. Munk, P. Njage, B. van Bunnik, L. McNally, O. Lukjancenko, T. Röder, D. Nieuwenhuijse, S. K. Pedersen, J. Kjeldgaard, et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications* **10**, (2019). doi:10.1038/s41467-019-08853-3
82. R. S. Hendriksen, V. Bortolaia, H. Tate, G. H. Tyson, F. M. Aarestrup, P. F. McDermott. Using Genomics to Track Global Antimicrobial Resistance. *Frontiers in Public Health* **7**, (2019). doi:10.3389/fpubh.2019.00242
83. T. P. Van Boeckel, C. Brower, M. Gilbert, B. T. Grenfell, S. A. Levin, T. P. Robinson, A. Teillant, R. Laxminarayan. Global trends in antimicrobial use in food animals. *PNAS* **112**, 5649–5654 (2015). doi:10.1073/pnas.1503141112

84. K. Tiseo, L. Huber, M. Gilbert, T. P. Robinson, T. P. Van Boeckel. Global trends in antimicrobial use in food animals from 2017 to 2030. *Antibiotics* **9**, 1–14 (2020). doi:10.3390/antibiotics9120918
85. R. Mulchandani, Y. Wang, M. Gilbert, T. P. Van Boeckel. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. *PLOS Global Public Health* **3**, 1–11 (2023). doi:10.1371/journal.pgph.0001305
86. T. P. Van Boeckel, J. Pires, R. Silvester, C. Zhao, J. Song, N. G. Criscuolo, M. Gilbert, S. Bonhoeffer, R. Laxminarayan. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, eaaw1944 (2019). doi:10.1126/science.aaw1944
87. W. Chang, J. Cheng, J. Allaire, Y. Xie, J. McPherson. shiny: Web Application Framework for R. R package version 1.3.2. https://CRAN.R-project.org/package=shiny. (2019).
88. K. Grave, H. C. Wegener. Comment on: Veterinarians' profit on drug dispensing. *Preventive Veterinary Medicine* **77**, 306–308 (2006). doi:10.1016/j.prevetmed.2006.01.010
89. L. H. Kahn. Confronting Zoonoses, Linking Human and Veterinary Medicine. *Emerging Infectious Diseases* **12**, 556–561 (2006).
90. F. R. Ungemach, D. Müller-Bahrdt, G. Abraham. Guidelines for prudent use of antimicrobials and their implications on antibiotic usage in veterinary medicine. *International Journal of Medical Microbiology* **296**, 33–38 (2006). doi:10.1016/j.ijmm.2006.01.059
91. M. Herrero, D. Grace, J. Njuki, N. Johnson, D. Enahoro, S. Silvestri, M. C. Rufino. The roles of livestock in developing countries. *Animal* **7**, 3–18 (2013). doi:10.1017/S1751731112001954
92. The World Bank. A Year in the Lives of Smallholder Farmers (2016). Available at: https://www.worldbank.org/en/news/feature/2016/02/25/a-year-in-the-lives-of-smallholder-farming-families#:~:text=There are an estimated 500 million smallholder households globally%2C amounting,less than %242 a day. Accessed: 01/11/2023.
93. M. Balehegn, A. Duncan, A. Tolera, A. A. Ayantunde, S. Issa, M. Karimou, N. Zampaligré, K. André, I. Gnanda, P. Varijakshapanicker, et al. Improving adoption of technologies and interventions for increasing supply of quality livestock feed in low- and middle-income countries. *Global Food Security* **26**, 100372 (2020). doi:10.1016/j.gfs.2020.100372
94. A. S. Barratt, K. M. Rich, J. I. Eze, T. Porphyre, G. J. Gunn, A. W. Stott. Framework for estimating indirect costs in animal health using time series analysis. *Frontiers in Veterinary Science* **6**, (2019). doi:10.3389/fvets.2019.00190
95. A. Radyowijati, H. Haak. Improving antibiotic use in low-income countries: An overview of evidence on determinants. *Social Science and Medicine* **57**, 733–744 (2003). doi:10.1016/S0277-9536(02)00422-7
96. R. T. de Melo, D. A. Rossi, G. P. Monteiro, H. Fernandez. Veterinarians and One Health in the Fight Against Zoonoses Such as COVID-19. *Frontiers in Veterinary Science* **7**, 1–5 (2020). doi:10.3389/fvets.2020.576262
97. Federation of Veterinarians of Europe. Survey of the veterinary profession in Europe (2019). Available at: https://fve.org/cms/wp-content/uploads/FVE_Survey_2018_WEB.pdf. Accessed: 01/04/2020.
98. S. M. Neal, M. J. Greenberg. Putting Access to Veterinary Care on the Map: A Veterinary Care Accessibility Index. *Frontiers in Veterinary Science* **9**, 1–8 (2022). doi:10.3389/fvets.2022.857644
99. M. Berrada, Y. Ndiaye, D. Raboisson, G. Lhermie. Spatial evaluation of animal health

care accessibility and veterinary shortage in France. *Scientific Reports* **12**, 1–10 (2022). doi:10.1038/s41598-022-15600-0
100. Federazione Nazionale Ordini Veterinari Italiani. Iscritti Ordine Veterinario (2023). Available at: https://www.fnovi.it/iscritti-ordine. Accessed: 15/11/2021.
101. Istituto Nacional de Estadìstica. Profesionales sanitarios colegiados (2016). Available at: https://www.ine.es/jaxi/Datos.htm?path=/t15/p416/a2016/&file=s04002.px#!tabs-mapa. Accessed: 15/11/2021.
102. Swiss Federal Office of Public Health. Register of Medical Professions (2022). Available at: https://www.healthreg-public.admin.ch/medreg/search. Accessed: 15/12/2021.
103. J. M. Richards. An Ecological Analysis of the Geographic Distribution of Veterinarians in the United States. *Journal of Vocational Behavior* **11**, 216–231 (1977). doi:10.1016/0001-8791(77)90008-2
104. S. Truchet, N. Mauhe, M. Herve. Veterinarian shortage areas: what determines the location of new graduates? *Review of Agricultural, Food and Environmental Studies* **98**, 255–282 (2017). doi:10.1007/s41130-018-0066-9
105. M. R. Olfert, M. Jelinski, D. Zikos, J. Campbell. Human capital drift up the urban hierarchy : veterinarians in Western Canada. 551–570 (2012). doi:10.1007/s00168-011-0448-2
106. American College of Surgeon. ATLS - Advanced Trauma Life Support Program for Doctors. *American College of Surgeons* (2008).
107. V. Singrodia, A. Mitra, S. Paul. A Review on Web Scrapping and its Applications. *2019 International Conference on Computer Communication and Informatics, ICCCI 2019* 1–6 (2019). doi:10.1109/ICCCI.2019.8821809
108. Wikipedia. Web scraping (2023). Available at: https://en.wikipedia.org/wiki/Web_scraping#:~:text=The history of web scraping,the size of the web. Accessed: 10/11/2023.
109. M. A. Khder. Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing and its Applications* **13**, 144–168 (2021). doi:10.15849/ijasca.211128.11
110. World Organisation for Animal Health. Performance of Veterinary Service (PVS) Pathway (2022). Available at: https://www.woah.org/en/what-we-offer/improving-veterinary-services/pvs-pathway/#:~:text=The PVS Pathway empowers national,inefficiencies and opportunities for innovation. Accessed: 20/11/2022.
111. Buchhorn, M. Copernicus Global Land Service: Land Cover 100m, epoch "2019", Globe (2020). Available at: https://lcviewer.vito.be/download. Accessed: 08/11/2023.
112. NASA. Earth Observatory (2023). Available at: https://earthobservatory.nasa.gov. Accessed: 10/11/2023.
113. D. J. Weiss, A. Nelson, K. Gligorić, S. Bavadekar, E. Gabrilovich, J. Rozier, H. S. Gibson, T. Shekel, C. Kamath, A. Lieber, et al. Global maps of travel time to healthcare facilities. *Nature Medicine* **26**, (2020). doi:10.1038/s41591-020-1059-1
114. D. J. Weiss, A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, E. Poyart, S. Belchior, N. Fullman, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018). doi:10.1038/nature25181
115. E. N. Hulland, K. E. Wiens, S. Shirude, J. D. Morgan, A. Bertozzi-Villa, T. H. Farag, N. Fullman, M. U. G. Kraemer, M. K. Miller-Petrie, V. Gupta, et al. Travel time to health facilities in areas of outbreak potential: Maps for guiding local preparedness and response. *BMC Medicine* **17**, 1–16 (2019). doi:10.1186/s12916-019-1459-6
116. F. Hierink, O. Oladeji, A. Robins, M. F. Muñiz, Y. Ayalew, N. Ray. A geospatial

analysis of accessibility and availability to implement the primary healthcare roadmap in Ethiopia. *Communications Medicine* **3**, 1–10 (2023). doi:10.1038/s43856-023-00372-z
117. ArcMap. Location-allocation analysis (2021). Available at: https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm. Accessed: 31/10/2023.
118. R. Church, C. Revelle. The maximal covering location problem. *Papers of the regional science association* **32**, 101–118 (1972). doi:10.1007/BF01942293
119. WHO. WHO Coronavirus (COVID-19) Dashboard (2023). Available at: https://covid19.who.int. Accessed: 05/10/2023.
120. NBC News. 'It's silent outside': Italy, France and Spain adapt to life under coronavirus lockdown (2020). Available at: https://www.nbcnews.com/news/world/it-s-silent-outside-italy-france-spain-adapt-life-under-n1160171. Accessed: 27/10/2023.
121. Wikipedia. COVID-19 pandemic in Switzerland. Available at: https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Switzerland. Accessed: 27/10/2023.
122. E. Dong, H. Du, L. Gardner. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533–534 (2020). doi:10.1016/S1473-3099(20)30120-1
123. Google. COVID-19 Open Data Repository (2022). Available at: https://health.google.com/covid-19/open-data/. Accessed: 05/10/2023.
124. UNESCO. Open access to facilitate research and information on COVID-19 (2021). Available at: https://en.unesco.org/covid19/communicationinformationresponse/opensolutions. Accessed: 05/10/2023.
125. Kanton Zürich. SARS-CoV-2 open government data reported by the Swiss Cantons and the Principality of Liechtenstein (20239. Available at: https://github.com/openZH/covid_19. Accessed: 31/10/2023.
126. Daniel Probst. COVID-19 Info Switzerland (2023). Available at: https://corona-data.ch. Accessed: 14/11/2023.
127. L. Rosenkrantz, N. Schuurman, N. Bell, O. Amram. The need for GIScience in mapping COVID-19. *Health and Place* **67**, 102389 (2021). doi:10.1016/j.healthplace.2020.102389
128. I. Franch-Pardo, B. M. Napoletano, F. Rosete-Verges, L. Billa. Spatial analysis and GIS in the study of COVID-19. A review. *Science of the Total Environment* **739**, 140033 (2020). doi:10.1016/j.scitotenv.2020.140033
129. T. P. Van Boeckel, N. G. Criscuolo. COVID-19 in Switzerland: a geographic approach to targeted testing. 1–3 (2020).

Chapter 2

resistancebank.org, an open-access repository for surveys on antimicrobial resistance in animals

Authors:
Nicola G. Criscuolo[1,*], João Pires[1], Cheng Zhao[1], Thomas P. Van Boeckel[1,2]

Affiliations:
[1]Institute for Environmental Decisions, ETH Zürich, Zürich, Switzerland.
[2]Center for Disease Dynamics, Economics and Policy, New Delhi, India.

*Correspondence to: nicola.criscuolo@usys.eth.z.ch

Published in:
Scientific Data 8:189 (2021). DOI: 10.1038/s41597-021-00978-9

Abstract

Antimicrobial resistance (AMR) is a growing threat to the health of humans and animals that requires global actions. In high-income countries, surveillance systems helped inform policies to curb AMR in animals. In low- and middle-income countries (LMICs), demand for meat is rising, and developing policies against AMR is urgent. However, surveillance of AMR is at best nascent, and the current evidence base to inform policymakers is geographically heterogeneous. We present *resistancebank.org*, an online platform that centralizes information on AMR in animals from 1,285 surveys from LMICs. Surveys were conducted between 2000 and 2019 and include 22,403 resistance prevalence estimates for pathogens isolated from chickens, cattle, sheep, and pigs. The platform is built as a *shiny* application that provides access to individual surveys, country-level reports, and maps of AMR at 10x10 kilometers resolution. The platform is accessed via any internet browser and enables users to upload surveys to strengthen a global database. *resistancebank.org* aims to be a focal point for sharing AMR data in LMICs and to help international funders prioritize their actions.

Introduction

Antimicrobials are essential drugs that have helped considerably reducing infectious diseases mortality. However, in recent years, their overuse in human medicine and animal production[1–3] has caused a rise in antimicrobial resistance[4–6] (AMR). Globally, 73% of all antimicrobials are used in animals to prevent and treat infections[7], but also to improve weight gain and productivity on farms[8]. The rise of antimicrobial use and resistance in animals is a growing concern for the future of animal health, and for the livelihood of billions of people who rely on animals for subsistence[9,10]. In addition, in recent years a growing body of evidence suggested that antimicrobial-resistant bacteria can be transferred between animals to humans[11–15], and cause drug-resistant infections in humans[16]. As for other infectious diseases of global importance[17–19], the rise of AMR in animals is a health challenge that requires close monitoring to coordinate international actions.

In high-income countries, trends in AMR in animals are monitored via systematic surveillance [20] by organizations such as the European Food Safety Authority (EFSA) in Europe, the National Antimicrobial Resistance Monitoring System for Enteric Bacteria (NARMS) in the United States, or the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) in Canada[21]. However, in low- and middle-income countries (LMICs), where demand for meat (and antimicrobials) is rising rapidly[2,10], systematic surveillance systems remain largely absent[22]. International actions to set-up or scale-up surveillance systems in LMICs have been initiated[23,24]. However, these may take years to be fully operational to inform policymakers. In the short term, efforts to target investments in LMICs against AMR could be informed by point-prevalence surveys[25] (PPS). Hundreds of PPS on foodborne pathogens are conducted each year across LMICs. In the absence of systematic surveillance systems, these could be used to document trends in AMR in food animals. In 2019, PPS that were initially scattered across the veterinary scientific literature were systematically reviewed to produce a

first global map of AMR in food animals at a sub-national level[25]. In addition, the findings of systematic reviews of PPS could be used to identify hotspots of resistance in animals in LMICs where stewardship efforts should be focused, or to identify areas poorly surveyed, where recruiting local epidemiologists may help improve the assessments of the AMR situation.
However, systematic surveys are time-consuming, need to be repeated frequently, and may require access to publications outside of the public domain. In addition, valuable information on AMR may be missed in systematic reviews due to i) linguistic barriers, ii) expensive publishing fees of international journals for researchers in LMICs, and iii) data availability restrictions from industry-program sponsor or governmental monitoring programs. An open-access platform for reporting results of PPS on AMR in real-time could help overcome these limitations and empower local communities of researchers. A platform designed with an intuitive interface may also encourage data sharing in the AMR community. This would not only improve the circulation of knowledge between researchers but also strengthen estimates of the AMR burden and provide up-to-date information to policymakers who allocate resources for intervention.
Online platforms have enabled data sharing in multiple scientific fields[26–29]. In epidemiology, they are used, amongst others, to integrate translational medicine data[30], exchange datasets of high-risk tumors[31], or disseminate estimates of the burden of malaria[32–34]. In AMR epidemiology, online platforms have been introduced to report drug-resistant infections in humans[32]. Thus far a comparable tool is missing for reporting AMR levels in animals. The development of such a tool to encourage data aggregation and visualization has been recognized as a priority by international donors[35,36] and organizations[37,38].
Here, we introduce *resistancebank.org*, an online platform for surveys and maps of AMR in animals. First, we present a database of PPS reporting AMR prevalence estimates globally. Second, we introduce local indicators of AMR burden available for download: maps and country-level reports. Third, we provide a step-by-step guide of the User Interface (UI), as seen by the visitor of the platform to upload their data in *resistancebank.org*.

Methods

Database

We conducted a systematic literature search in January 2019 and extracted information on resistance prevalence estimates from 1,285 PPS. The search for PPS on AMR in food animals from LMICs was conducted in three bibliographic databases: PubMed, Scopus, and Web of Science. We targeted four indicator bacteria recommended by the World Health Organization's Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR). Titles and abstracts were deduplicated and screened for PPS. Books, meta-analysis, reviews, and PPS reporting on sick animals were excluded, according to the AGISAR guidelines[39]. In addition, we included data available in paper journals, Ph.D. and MSc thesis, and conference proceedings after field visits to five veterinary schools in India. All relevant data to AMR surveillance were screened across all manuscripts, including sampling size, animal hosts, bacterial species, sampling latitude and longitude. The resistance prevalence estimates from antimicrobial susceptibility testing (AST) were aggregated by individual location/host/bacteria

combinations. For each study, all tested antimicrobials and the number of isolates included in each assay were recorded. For the geospatial analysis, only antimicrobials recommended by the AGISAR were used. A description of the database variables reported in Table 1 can be found on *resistancebank.org*. For a detailed explanation of the methods used for the literature search see the Supplementary Materials of Van Boeckel & Pires, 2019[25].

Table 1. **Metadata of the variables available in *resistancebank.org***.

DOI	Digital object identifier or PubMed ID
Author	Name of the first author
ISO3	Country identifier of the International Organization of Standardization.
YCoord, XCoord	Latitude and longitude of the PPS in decimal degrees
StartDate, EndDate	Starting and ending date of field samplings (day/month/year)
Species	Animal host species
SampleOrigin	Type of biological sample used to isolate bacteria from
Method	Experimental Methodology for the AST
Pathogens	Bacterial species
Strain	Bacterial subtype
Nsamples	Number of samples collected
Prev	Pathogen prevalence
NIsolates	Number of isolates used for AST
Class	Classification of antimicrobials based on their chemical structure
Compound	Antibiotic molecule
ATC.Code	Anatomical Therapeutic Chemical identifier of the compound
Rescom	Percentage of isolates resistant for a given drug-pathogen combination
Concg	Concentration/amount of antimicrobial used for AST
Guidelines	Guideline used for antimicrobial AST (document and year)
Breakpoint	Resistance breakpoint used for interpreting AST results
Remarks	Comments relative to the publication.

The database contains 22,403 resistance prevalence estimates (*n*) extracted from PPS conducted on foodborne pathogens in LMICs between 2000 to 2019 (Fig. 1). The pathogens isolated for AST include *Escherichia coli* (*n* = 9,206, 41.09%), non-typhoidal *Salmonella* spp. (*n* = 7,080, 31.60%), *Staphylococcus aureus* (*n* = 4,828, 21.55%), and *Campylobacter* spp. (*n* = 1,290, 5.76%). The PPS were conducted in 72 out of the 135 countries classified as LMICs by The World Bank[40]. Overall, 61.4% of all PPS listed on *resistancebank.org* were conducted in Asia, 24.1% in Africa, and 14.5% in Central and South America (Fig. 1); 49.41% of PPS were collected in India, China, Brazil, and Iran. Across all studies, bacteria were isolated from poultry (38.35%), cattle (37.56%), pigs (15.28%), and sheep (8.81%). In addition, in 55.66% of the PPS, bacteria were isolated from food products, in 43.82% from living or slaughtered animals, and in the remaining 0.52% from drag swabs (e.g., fecal samples, eggshells). The platform includes resistance prevalence estimates for 143 antimicrobials grouped in 37 different families based on their chemical structure[41]. Amongst all the resistance prevalence

estimates available in *resistancebank.org*, 13,163 (59%) are for drug-pathogen combinations recommended for AST by the AGISAR[39].

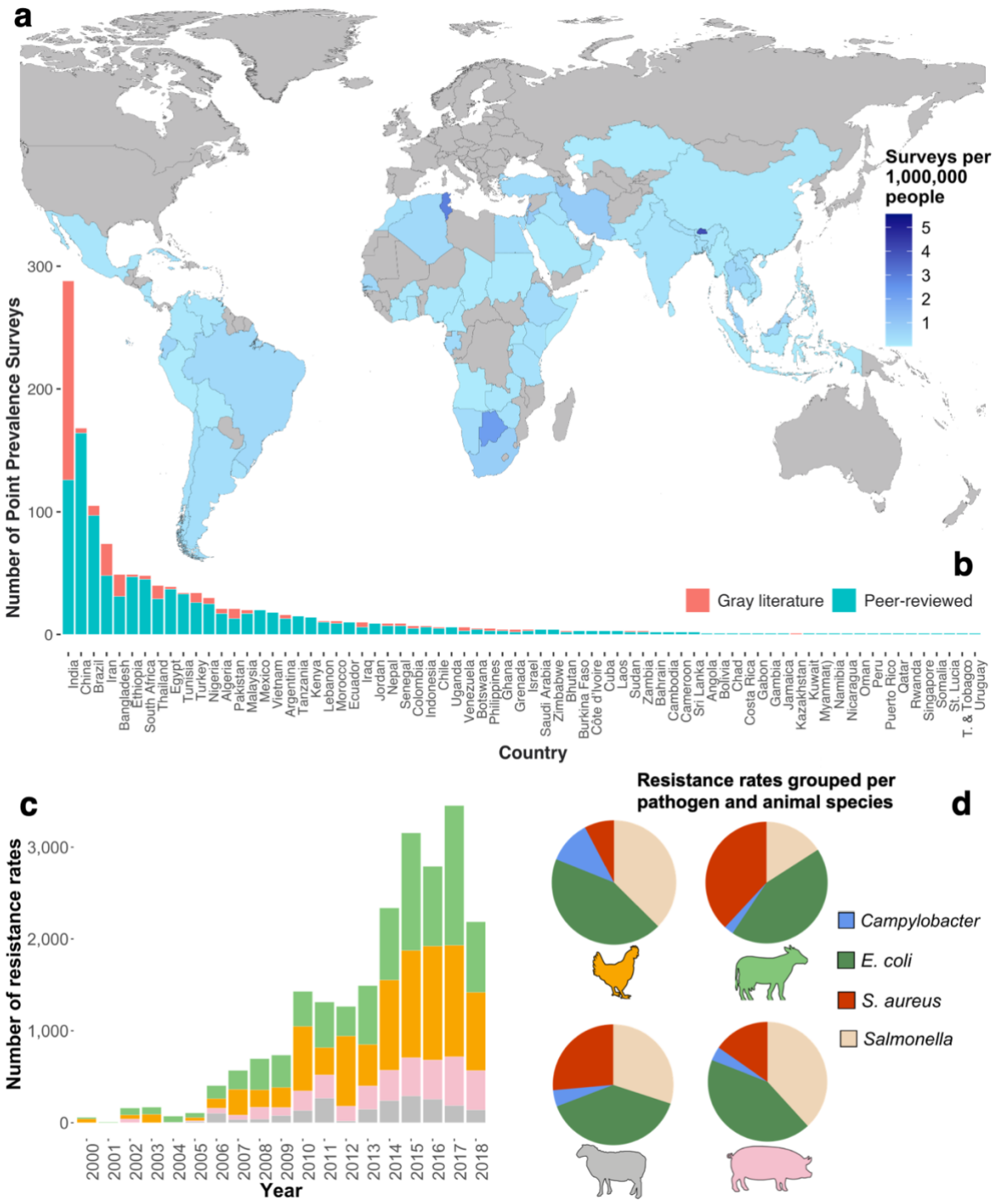

Fig. 1. Metadata of *resistancebank.org*. a) Number of surveys per 1 million people in LMICs. b) Point-prevalence surveys grouped by country and literature source (peer-reviewed vs grey literature). c) Number of resistance prevalence estimates grouped by year and animal species. d) Resistance prevalence estimates grouped by pathogen and animal species.

Software implementation

We coded *resistancebank.org* in the open-access R language[42], in combination with JavaScript and CSS code for the UI (full code[43] available at https://github.com/hegep-eth/resistancebank.org). We used the functionalities of the *shiny* package[44] to translate the R code used for the software architecture into HTML language to produce an online platform accessible from all major web browsers: Safari, Google Chrome, Microsoft Edge, Internet Explorer, and Opera. Once completed, we deployed the application on the shinyapps.io servers (https://www.shinyapps.io), a cloud service that, with our configuration, can guarantee simultaneous access to the platform to 2,500 users.

We used the *leaflet* R package to display spatial data[34,45,46] and add geographic layers in the UI. For displaying the AMR maps, we used a geographic information system software (QGIS 2.18[47]) to produce raster tiles (light square images in .png format) at ten different zoom levels. We stored 2.8 million tiles on a GitHub Pages website linked, as an online resource, to the *leaflet* object used to define the maps. Depending on the zoom and map position, the platform loads just the necessary maps tiles to ensure smooth navigation across zoom levels.

For remote data collection and storage, we used the R packages *rdrop2* and *aws3* to interface the platform with cloud storage services, respectively Dropbox and Amazon Web Services. With every new submission, *reistancebank.org* uploads the corresponding .csv file in an online folder, emptied every time a human operator approves the new submissions. The central database (and its main sub-datasets), the plots present in each pop-up window of the geographic markers, and all of the files downloadable from *resistancebank.org* are stored remotely too. We used functions implemented in packages of the ROpenSci project (e.g., *europepmc*) to gather the bibliographic information of the collected PPS.

Finally, we used a reactive R Markdown document to generate AMR country reports downloadable from *resistancebank.org* in .pdf format. We connected the Markdown file (.Rmd format) to the source code of *resistancebank.org* through an input parameter, i.e., a country name present in the application that depends on a specific object (datasets, functions, etc.). This country name can be specified by the users directly in the UI (e.g., selection of an LMIC through a drop-down menu). Based on that choice, the R Markdown document uses just a specific set of data (or functions) associated with one country to produce its corresponding country report.

Results

User interface

The UI of *resistancebank.org* is organized around an interactive map (Fig. 2). The red map shows spatial variations in a summary metric of resistance: P50, the proportion of antimicrobials tested for which bacteria have developed a resistance higher than 50%[25]. This index has been predicted by geospatial models at 10x10 km resolution for every LMIC using PPS[25]. Updates of the P50 map will be conducted on an annual basis.

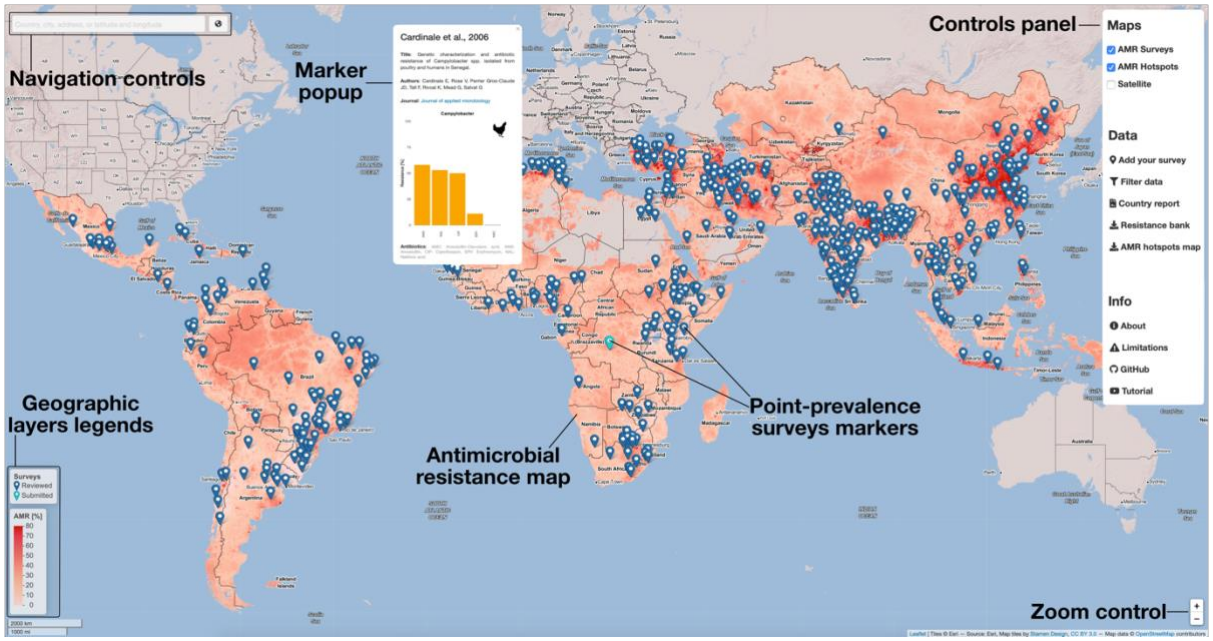

Fig. 2. User Interface. Antimicrobial resistance map (red) showing the proportion of antimicrobials with resistance higher than 50% (P50). The point-prevalence surveys are represented as blue geographic markers linked to a pop-up window displaying resistance prevalence estimates and bibliographic information. The Controls panel enables activating individual geographic layers, downloading the database of point-prevalence surveys, uploading new point-prevalence surveys, and displaying country reports.

The panel in the top-right corner is divided into three sections: the "Maps" section controls the activation of every geographic layer available on *resistancebank.org* (PPS, AMR map, satellite base map). In the "Data" section, users can download the database of PPS (.csv format) described in Table 1 via the "Resistance bank" button, as well as the P50 raster (.tif format) via the "AMR hotspots map" button. The "Country report" button displays a panel containing a summary of the AMR reports available for each country (Fig. 3). From this panel, it is possible to download the country report for each country where at least one PPS has been reported. Along with the report, users can also download a country-based subset of the central database (.csv format) containing all the data used to generate this output.

Through the "Add your survey" menu, users can choose amongst two modalities for uploading their PPS data in *resistancebank.org*: either by filling an online form, or an Excel template (see next section). The last button in the "Data" section, "Filter data" allows users to filter the database for countries, animal species and animal sample origin, pathogens and if their combination with the antimicrobials aligns with the AGISAR guidelines. In addition, it is also possible to filter the database for an individual antimicrobial class defined by the World Health Organization based on its importance for human medicine[41] (e.g. 3rd generation cephalosporine). Users can then download the filtered results or display them on the map. In the latter case, when users filter for the antimicrobial class, the geographic marker intensity colour of each survey will vary based on the average resistance prevalence estimate for the selected antimicrobial class.

Finally, the "Info" panel links to the software's GitHub repository (https://github.com/hegep-eth/resistancebank.org) and a YouTube video (https://youtu.be/TpMQ_3JLJ2I) illustrating how to use the different sections of the platform.

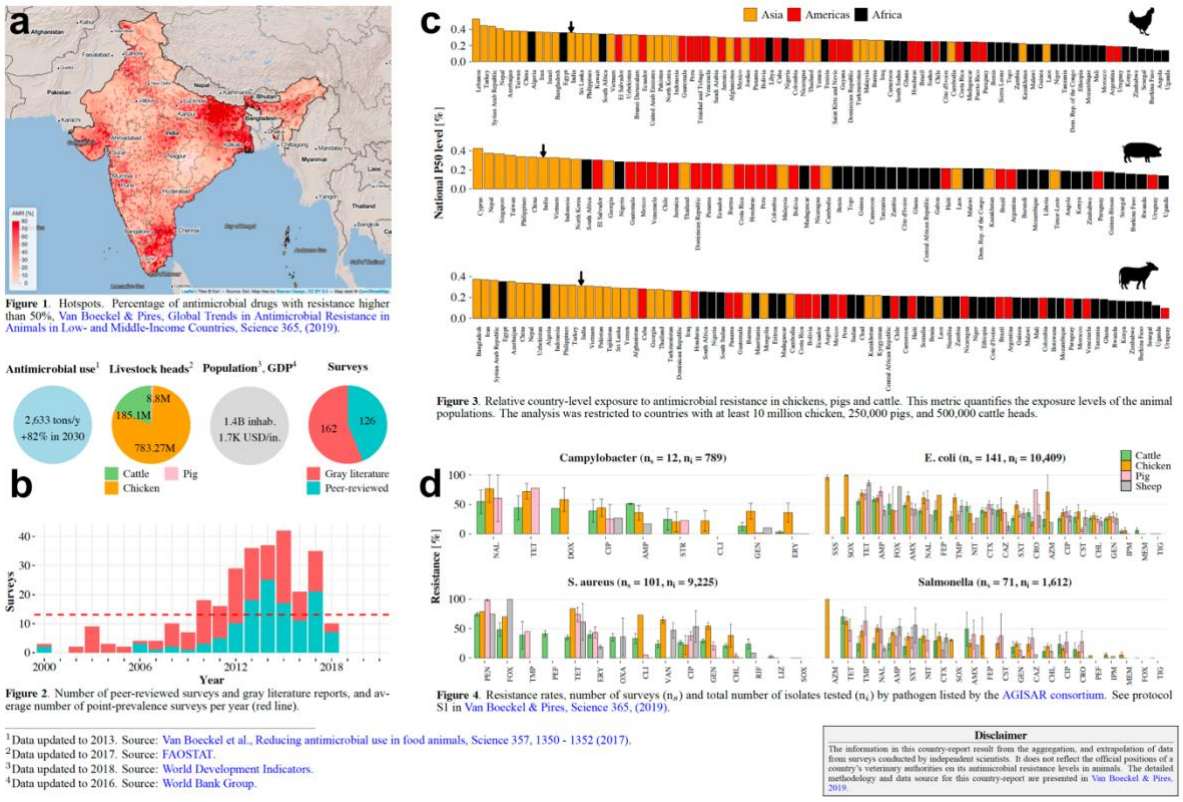

Figure 1. Hotspots. Percentage of antimicrobial drugs with resistance higher than 50%, Van Boeckel & Pires, Global Trends in Antimicrobial Resistance in Animals in Low- and Middle-Income Countries, Science 365, (2019).

Figure 2. Number of peer-reviewed surveys and gray literature reports, and average number of point-prevalence surveys per year (red line).

Figure 3. Relative country-level exposure to antimicrobial resistance in chickens, pigs and cattle. This metric quantifies the exposure levels of the animal populations. The analysis was restricted to countries with at least 10 million chicken, 250,000 pigs, and 500,000 cattle heads.

Figure 4. Resistance rates, number of surveys (n_s) and total number of isolates tested (n_l) by pathogen listed by the AGISAR consortium. See protocol S1 in Van Boeckel & Pires, Science 365, (2019).

[1]Data updated to 2013. Source: Van Boeckel et al., Reducing antimicrobial use in food animals, Science 357, 1350 - 1352 (2017).
[2]Data updated to 2017. Source: FAOSTAT.
[3]Data updated to 2018. Source: World Development Indicators.
[4]Data updated to 2016. Source: World Bank Group.

Disclaimer
The information in this country-report result from the aggregation, and extrapolation of data from surveys conducted by independent scientists. It does not reflect the official positions of a country's veterinary authorities on its antimicrobial resistance levels in animals. The detailed methodology and data source for this country-report are presented in Van Boeckel & Pires, 2019.

Fig. 3. Country report downloadable from *resistancebank.org*. a) Country-level antimicrobial resistance map. b) Socio-demographic indicators (antimicrobial use and its projected increase in 2030, livestock heads, population, and gross domestic product per inhabitant) and the number of point-prevalence surveys grouped by year and type of paper (peer-review or grey literature). c) Country-level exposure to antimicrobials in chickens, pigs, and cattle (Supplementary Materials of Van Boeckel & Pires, 2019[25]). d) Resistance prevalence estimates grouped per drug-pathogen combinations listed by the Advisory Group on Integrated Surveillance of Antimicrobial Resistance.

Uploading new data

One of the objectives of *resistancebank.org* is to provide up-to-date data visualizations for policymakers continuously collecting evidence from potential users who conducted a PPS in LMICs. To this end, users can input their AST data *on resistancebank.org* using an online form or a pre-filled Excel template. The Excel template can be downloaded and subsequently uploaded from the platform via the "Upload template" button under the "Add your survey" section. Before integrating a new survey, the platform executes automatic verifications for possible misspelled words and typos. If necessary, typos are corrected after comparison with a set of correct words provided for every template field. If *resistancebank.org* recognizes errors that can't be corrected automatically (e.g., numerical values outside an appropriate range, such as AMR prevalence estimates higher than 100%), the platform will invite users to revise their

inputs. If users wish to use an online form instead of an Excel template, a step-by-step user guide is provided in the next paragraph ("Use Scenario").
Following the submission of the template or the form, the application automatically performs bibliographic research in the NCBI and PubMed databases to control if the users have provided a valid Digital Object Identifier (DOI) of their study. The application will also automatically extract information on the author(s) name, study title, publication year, journal name, and a link to the journal website associated with the survey submitted. Following submission, a new temporary geographic marker is added on the map, while awaiting further verification by a human operator who is notified of a new submission by email. This marker is light blue, to differentiate it from dark blue markers corresponding to confirmed studies (see Fig. 2 or the YouTube video). The human operator verifies the new data through an internal auxiliary software developed to support *resistancebank.org*. These verifications include a critical interpretation of the resistance values reported and breakpoints values used for each drug-pathogen combination. The human operator may contact the authors of the study to request corrections/clarifications, and then give its final approval to a submission and merge it with the other surveys in the database. After the upload of a new survey, a near-real-time update of the platform is triggered such as to update all outputs (database available for download and country-level reports). This final step enables *resistancebank.org* to present only the most recent aggregated AMR results based on the PPS available in the scientific literature.

Use scenario
We describe examples of the possible use of the platform by a user who wants to upload his/her AMR survey conducted, hypothetically, on a farm near New Delhi, India. The subsequent steps (visually represented in Fig. 4 and the YouTube video) aim to give an overview of the functionalities of *resistancebank.org* and the procedure to submit a new survey.

1. A user launches the online platform by connecting to https://resistancebank.org.
2. The user starts exploring the AMR map and the PPS geographic marker (in the Controls panel section "Maps" these two geospatial layers are both active when the application starts). The user can zoom in on the desired location or type the location name in the navigation bar in the top left part of the UI to explore P50 levels near New Delhi. If the user has recorded the precise coordinates of the study, the input text box can also accept latitude and longitude (separated by a space) in decimal degrees.
3. The map view is now centered in New Delhi. The user can start exploring the P50 levels around the city and the PPS information aggregated at the animal species level present in the geographic marker. This window also contains a URL to connect the user to the journal webpage of the study to retrieve additional information besides the ones presented in *resistanbank.org*.
4. Detailed information about the national AMR situation in India is available in the panel accessible through the "Country report" button. Once "India" is selected from the drop-down menu, the country report is ready for download, together with the database of the PPS data collected just in India.
5. The user decides to upload a PPS in *resistancebank.org*. He/she clicks on the "Add your survey" button, a new panel will open, and then the user can decide to upload data

through an online form or an Excel template. The hyperlink ".xlsx template" triggers the download of a .zip folder that contains the template and a guide on how to fill it. For this example, we will use the online form.

6. In the upper section of the form, the user can input his contact information. The second section of the form concerns bibliographic information; if a DOI is available, fields as title, journal, and publication year will be automatically filled. In this section, the user must specify the study region, the location (e.g., city, address, or latitude and longitude), and the sampling scheme adopted for the study (if it is a routine, a longitudinal study, a one-time research survey, or a study mandated by public authorities). In the AMR section, every row corresponds to an AST. The mandatory fields are animal species, the animal sample origin (e.g., different swabs, stools, meat, eggshells, gut, dairy products, etc.), the pathogen, the AST method, the antibiotic tested, and the relative value of resistance. Other non-mandatory fields are the number of isolates (at least 10 to be valid for a survey or template submission), the prevalence, the strain used in the quality control of the AST performance, the report of the use of antimicrobials in the farms where animal samples were collected, and the breakpoints. If the researcher tested more than one antibiotic, he/she can add a row that will keep the information from the previous row, except for the new compound name and its resistance value.
7. Once uploaded, the form panel closes automatically, notifying the user of the submission. The light blue geographic marker will appear on the location specified in the form. The geographic marker will remain in a temporary status until a human operator accepts the new submission and merges it with the existing database stored in *resistancebank.org*.

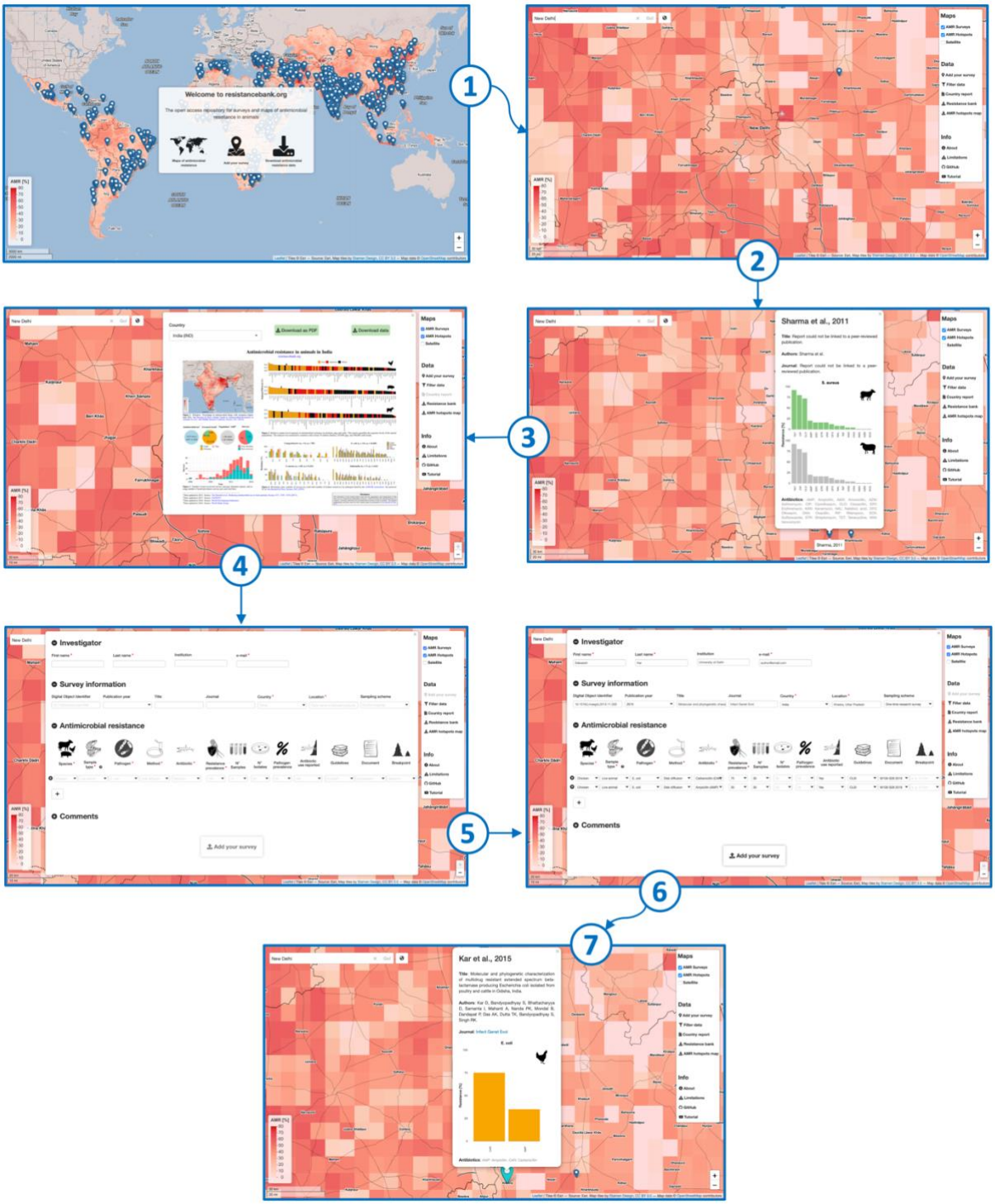

Fig. 4. Workflow of the submission of a point-prevalence survey on *resistancebank.org*. Panel 1 shows the home screen of *resistancebank.org*. The steps from 2 to 4 show how the users can explore the visual outputs of the platform, i.e., the antimicrobial resistance map at (10x10 kilometers), the resistance prevalence estimates present in the pop-up window of a point-prevalence survey represented as a geographic marker, and the country report. Panels 5 and 6 show the steps to input new data. Panel 7 shows temporary geographic markers, bar charts, and bibliographic information when the submission of a survey is complete.

Discussion

Thus far, a large body of evidence on AMR trends in animals in LMICs was scattered across the veterinary literature. *resistancebank.org* is a starting point to integrate this information. The platform is a surrogate but not a substitute for state-of-the-art systematic surveillance systems[48,49]. The goals of the platform are to summarize current knowledge on AMR in animals and to provide a tool for strengthening its evidence-based surveillance with additional PPS in the future. Furthermore, it overcomes barriers associated with traditional scientific publication (publication fees and access fees), thereby improving the visibility of researchers from LMICs, empowering local communities of scientists, and encouraging national networks coordinators to release their findings onto the website.
Locally, *resistancebank.org* could be used to encourage epidemiological investigations by field officers from LMICs in areas of particular interest. Globally, *resistancebank.org* offers the opportunity to support the actions of international funders such as the Bill & Melinda Gates Foundation, the Fleming Fund, the Food and Agriculture Organization, and the World Organization for Animal Health. Specifically, areas identified as hotspots of resistance (P50 > 0.4) could be used to investigate the effects of stewardship campaigns, and alternatives to antimicrobials, such as vaccines and probiotics[50,51].

Before *resistancebank.org*, different studies have centralized AMR data to describe their large-scale trends both in foodborne and human pathogens[5,25,52,53]. However, the use of an online platform has multiple advantages over individual studies. First, given its open-access nature, downloading and uploading information can be done free by anyone. Second, the diversity of outputs: we provide maps, summary reports for policymakers but also detailed data about resistance prevalence estimates in individual surveys, with the possibility to filter them at a national- and microbiological-level to better target interventions. Third, the information on *resistancebank.org* is continuously updated in near-real-time. Fourth, the platform provides a much-needed -and thus far missing- focal point of data for a community of researchers studying the epidemiology of AMR in animals. For humans, online platforms that display AMR trends do exist: the Surveillance Atlas of Infectious Diseases[54] developed by the European Centre for Disease Prevention and Control, and *resistancemap*[55], developed by the Center for Disease Dynamics, Economics and Policy, display, respectively, European and global yearly AMR trends in common human pathogens. However, unlike *resistancebank.org*, these platforms lack the high spatial resolution of the data, since they aggregate trends at the country-level. While such trends are informative, the granular information underlying them is unfortunately not available in open access. Furthermore, these platforms do not include a possibility for uploading new surveys or dataset by external users. Similar platforms focus on the genetic determinants of AMR[56,57], and how these affect the spread of pathogens. These include, amongst others, Microreact[58] and Nextstrain[59], and are complementary to the phenotypic information provided on *resistancebank.org*. For animals, WHONET[60] (https://whonet.org) stores users' AST results obtained from individual bacterial isolates. However, this platform is not currently available for every operating system, and -to the best of our knowledge- does not include geographic information on AMR trends in a centralized context. In contrast,

resistancebank.org can be used with any internet browser. Our platform has been developed to complement current tools available for AMR surveillance in animal production which requires international attention given its potential implications on human health, animal health, and the long-term sustainability of the livestock sector.

Limitations

The data presented in *resistancebank.org* come with limitations. The first set of limitations concerns the quality of the event-based surveillance data and comparability across surveys. In the human population, most of studies focus on diagnostic samples taken mainly from sick patients. In contrast, in animals, surveillance relies on different data collection contexts: sampling of living versus dead animals, sampling of animal food products, outbreak investigation, sample collection required by food regulatory authorities, etc. These different sampling contexts, which are inherent to event-based surveillance, represent a challenge to the harmonization and the interpretation of resistance prevalence estimates reported on this platform. In particular, the surveys listed on *resistancebank.org* may differ in terms of i) sampling strategy (random or convenient), ii) animal breeds and farming systems, iii) the number of isolates tested per survey, testing, and iv) the degree of aggregation used for reporting antimicrobial prevalence estimates in each survey (population versus isolate-level information). For these reasons, in *resistancebank.org*, we allow users to specify additional surveys information such as the sampling scheme, the guidelines, and breakpoints used for AST, the quality control strains used, etc., to include as much information about these factors that may affect the interpretation of the resistance prevalence estimates reported.

The second set of limitations concerns the attempt to summarize trends in resistance across drug-pathogen combinations using P50: the proportion of drugs tested in a survey with resistance prevalence estimates higher than 50%. From a practical perspective, P50 expresses the probability of providing treatments that work out of a portfolio of treatment options, when antimicrobial therapy is indicated for a medical condition. Multiple summary metrics have been proposed[61–64] and debated[65] to aggregate resistance prevalence estimates to multiple drug-pathogen combinations. As with every attempt to reduce this complexity, P50 comes with sources of uncertainty. First, the number of drugs tested in each survey can differ, and this can typically be influenced by the methods used for antimicrobial susceptibility testing in different laboratories (diffusion vs dilution methods), although a good agreement has been shown between the methods[66]. Second, in some surveys, screening for resistance of second-line drugs such as imipenem may be conducted on a subset of the isolates and introduce bias in P50 estimates. In this study, sub-sampling for second-line antimicrobials was limited to 34 out of 1,940 estimates of P50. Third, P50 reflects the number of compounds with resistance higher than 50% rather than the number of classes of compounds with resistance higher than 50%. Therefore, *resistancebank.org* also provides resistance prevalence estimates for classes of compounds considered medically important by the WHO[41]. The P50 is a summary metric intended to help resource allocation against AMR in countries where systematic surveillance is limited. However, because of the non-systematic nature of the data P50 summarizes,

comparisons of resistance prevalence estimates for individual drug classes should be preferred for informing public health strategies.

The third set of limitations concerns the intensity of our data collection efforts between countries. Our online literature search was supplemented by field officers who collected PPS on paper during visits to veterinary schools. However, these field visits could only be conducted in India, where our collaboration network is extensive. Collaboration with international organizations could help leverage a larger network of field officers to supplement the information currently in *resistanbank.org*. We conducted the literature searches in six languages (English, Mandarin Chinese, Spanish, French, Portuguese, and German). Although these languages are spoken by 46.6% of the world population[67], further inquiries in other languages could help supplement our database. Finally, the computational cost of re-running the geospatial model is currently preventing instantaneous updates of the AMR map on a global level and should be the focus of future research efforts to move from yearly updates of our maps to daily updates. For the reasons listed above, *resistancebank.org* is an imperfect surrogate to systematic surveillance systems. It is a platform reporting large-scale trends in AMR meant to help international funders to target their efforts in the short term and facilitate the development of a global systematic surveillance system in the long term.

Data availability

All data presented in this article are available at https://resistancebank.org.

Code availability

The scripts and the files of *resistancebank.org* are available in its GitHub repository (https://github.com/hegep-eth/resistancebank.org). We developed the platform through the R software (version 4.0.2) and the RStudio Integrated Development Environment (IDE; version 1.3.1073). In its current version *resistancebank.org* relies on the following dependencies: *aws.s3* (0.3.21), *dplyr* (1.0.2), *europepmc* (0.4), *ggimage* (0.2.8), *ggmap* (3.0.0), *ggplot2* (3.3.2), *gmailr* (1.0.0), *gsubfn* (0.7), *leaflet* (2.0.4.1), *lemon* (0.4.5), *raster* (3.3-13), *rcrossref* (1.0.0), *rdrop2* (0.8.2.1), *rgdal* (1.5-16), *rmarkdown* (1.15), *scales* (1.1.1), *shiny* (1.5.0), *shinyBS* (0.61), *shinyjs* (1.1), *shinyWidgets* (0.5.3), *stringr* (1.4.0), *svglite* (1.2.3.2), *sp* (1.4-5).

Acknowledgements

This work was supported by the Branco Weiss Fellowship, and the Swiss National Science Foundation (SNF).

Author contributions

Thomas Van Boeckel and Nicola Criscuolo conceived the work. Nicola Criscuolo cleaned the database, developed *resistancebank.org,* and wrote the first version of the manuscript. João Pires and Cheng Zhao conducted literature reviews. All authors critically revised and edited the manuscript.

Competing interest

The author(s) declare no competing interests.

References

1. Goossens, H. Antibiotic consumption and link to resistance. *Clin. Microbiol. Infect.* **15**, 12–15 (2009).
2. Van Boeckel, T. P. *et al.* Global trends in antimicrobial use in food animals. *PNAS* **112**, 5649–5654 (2015).
3. Klein, E. Y. *et al.* Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3463–E3470 (2018).
4. Chen, C. C. *et al.* Organ-level quorum sensing directs regeneration in hair stem cell populations. *Cell* **161**, 277–290 (2015).
5. Versporten, A. *et al.* Antimicrobial consumption and resistance in adult hospital inpatients in 53 countries: results of an internet-based global point prevalence survey. *Lancet Glob. Heal.* **6**, e619–e629 (2018).
6. Chantziaras, I., Boyen, F., Callens, B. & Dewulf, J. Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: A report on seven countries. *J. Antimicrob. Chemother.* **69**, 827–834 (2014).
7. Van Boeckel, T. P. *et al.* Reducing antimicrobial use in food animals. *Science (80-.).* **357**, 1350–1352 (2017).
8. Organisation for Economic Co-operation and Development. *Global antimicrobial use in the livestock sector*. *Working Party on Agricultural Policies and Markets* (2015).
9. The High Level Panel of Experts on Food Security and Nutrition (FAO). *Sustainable agricultural development for food security and nutrition*. (2016).
10. Thornton, P. K. Livestock production: recent trends, future prospects. *Philos. Trans. R. Soc. B.* **365**, 2853–2867 (2010).
11. Tang, K. L. *et al.* Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. *Lancet Planet. Heal.* **1**, e316–e327 (2017).
12. Leverstein-van Hall, M. A. *et al.* Dutch patients, retail chicken meat and poultry share the same ESBL genes, plasmids and strains. *Clin. Microbiol. Infect.* **17**, 873–880 (2011).
13. Smith, T. C. *et al.* Methicillin-Resistant Staphylococcus aureus in Pigs and Farm Workers on Conventional and Antibiotic-Free Swine Farms in the USA. *PLoS One* **8**, 1–5 (2013).
14. Woolhouse, M., Ward, M., Van Bunnik, B. & Farrar, J. Antimicrobial resistance in humans, livestock and the wider environment. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).
15. Ward, M. J. *et al.* Time-scaled evolutionary analysis of the transmission and antibiotic resistance dynamics of Staphylococcus aureus clonal complex 398. *Appl. Environ. Microbiol.* **80**, 7275–7282 (2014).
16. Liu, C. M. *et al.* Escherichia coli ST131-H22 as a foodborne uropathogen. *MBio* **9**, (2018).
17. Hay, S., Guerra, C. A., Tatem, A. J., Noor, A. M. & Snow, R. W. The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis* **4**, 327–336 (2004).
18. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504–507

(2013).
19. Grobusch, M. P. & Kapata, N. Global burden of tuberculosis: where we are and what to do. *Lancet Infect. Dis.* **18**, 1291–1293 (2018).
20. Perez, F. & Villegas, M. V. The role of surveillance systems in confronting the global crisis of antibioic-resistant bacteria. *PMC* **28**, 375–383 (2017).
21. Acar, J. F. & Moulin, G. Integrating animal health surveillance and food safety: The issue of antimicrobial resistance. *OIE Rev. Sci. Tech.* **32**, 383–392 (2013).
22. Laxminarayan, R., Sridhar, D., Blaser, M., Wang, M. & Woolhouse, M. Achieving global targets for antimicrobial resistance. *Science (80-.).* **353**, 874–875 (2016).
23. United States Agency for International Development (USAID). Combating antimicrobial resistance. https://www.usaid.gov/global-health/health-systems-innovation/health-systems/combating-antimicrobial-resistance. (2020).
24. The Fleming Fund. https://www.flemingfund.org. (2020).
25. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science (80-.).* **365**, (2019).
26. Howison, J., Deelman, E., McLennan, M. J., Da Silva, R. F. & Herbsleb, J. D. Understanding the scientific software ecosystem and its impact: Current and future measures. *Res. Eval.* **24**, 454–470 (2015).
27. Nguyen, P. *et al.* The CHRS Data Portal , an easily accessible public repository for PERSIANN global satellite precipitation data. *Nat. Publ. Gr.* 1–10 (2018). doi:10.1038/sdata.2018.296
28. Criscuolo, N. *et al.* High Biodiversity Arises from the Analyses of Morphometric, Biochemical and Genetic Data in Ancient Olive Trees of South of Italy. *Plants* **8**, (2019).
29. Criscuolo, N. G. & Angelini, C. StructuRly: A novel shiny app to produce comprehensive, detailed and interactive plots for population genetic analysis. *PLoS One* **15**, 1–12 (2020).
30. Emam, I. *et al.* PlatformTM, a standards-based data custodianship platform for translational medicine research. *Sci. data* **6**, 149 (2019).
31. Depuydt, P. *et al.* Meta-mining of copy number profiles of high-risk neuroblastoma tumors. *Sci. Data* **5**, 1–9 (2018).
32. Freifeld, C. C., Mandl, K. D., Reis, B. Y. & Brownstein, J. S. HealthMap: Global Infectious Disease Monitoring through Automated Classification and Visualization of Internet Media Reports. *J. Am. Med. Informatics Assoc.* **15**, 150–157 (2014).
33. Pfeffer, D. A. *et al.* MalariaAtlas: An R interface to global malariometric data hosted by the Malaria Atlas Project. *Malar. J.* **17**, 1–10 (2018).
34. Tomlinson, S., South, A. & Longbottom, J. Malaria Data by District: An open-source web application for increasing access to malaria information. *Wellcome Open Res.* **4**, 151 (2019).
35. Tacconelli, E. *et al.* Surveillance for control of antimicrobial resistance. *Lancet Infect. Dis.* **18**, e99–e106 (2018).
36. The Fleming Fund. Country Grants. https://www.flemingfund.org/grants-funding/country-grants/. (2020).
37. Food and Agriculture Organization of the United Narions. *Monitoring and surveillance of antimicrobial resistance in bacteria from healthy food animals intended for*

consumption. *Regional Antimicrobial Resistnce Monitoring and Surveillance Guidelines* (2019).
38. Laxminarayan, R. *et al.* The Lancet Infectious Diseases Commission on antimicrobial resistance: 6 years later. *Lancet Infect. Dis.* **3099**, 1–10 (2020).
39. WHO. *Integrated Surveillance of Antimicrobial Resistance in Foodborne Bacteria: Application of a One Health Approach*. *World Health Organization* (2017).
40. The World Bank. World Bank Country and Lending Groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519. (2019).
41. WHO. *Critically Important Antimicrobials for Human Medicine*. (2018).
42. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. (2019).
43. Criscuolo, N. G., Pires, J., Zhao, C. & Van Boeckel, T. P. Source code for: resistancebank.org - An open-access repository for surveys of antimicrobial resistance in animals. Zenodo https://doi.org/10.5281/zenodo.4604894. (2021).
44. Chang, W., Cheng, J., Allaire, J., Xie, Y. & McPherson, J. shiny: Web Application Framework for R. R package version 1.4.0. https://CRAN.R-project.org/package=shiny. (2019).
45. Calderwood, J. *et al.* Hotspot mapping in the Celtic Sea: An interactive tool using multinational data to optimise fishing practices. *Mar. Policy* **116**, 1–12 (2019).
46. Moraga, P. SpatialEpiApp: A Shiny web application for the analysis of spatial and spatio-temporal disease data. *Spat. Spatiotemporal. Epidemiol.* **23**, 47–57 (2017).
47. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project, version 3.14. http://qgis.osgeo.org. (2020).
48. Hammerum, A. M. *et al.* Danish integrated antimicrobial resistance monitoring and research program. *Emerg. Infect. Dis.* **13**, 1632–1639 (2007).
49. Veldman, K. *et al. Monitoring of antimicrobial resistance and antibiotic usage in animals in the Netherlands in 2018. Combined with NETHMAP-2019: Consumption of antimicrobial agents and antimicrobial resistance among medically important bacteria in the Netherlands.* (2019).
50. Vyas, U. & Ranganathan, N. Probiotics, prebiotics, and synbiotics: Gut and beyond. *Gastroenterol. Res. Pract.* **2012**, (2012).
51. Gaggìa, F., Mattarelli, P. & Biavati, B. Probiotics and prebiotics in animal feeding for safe food production. *Int. J. Food Microbiol.* **141**, S15–S28 (2010).
52. Okeke, I. N. *et al.* AMR Resistance in developing countries. Part 1: recent trends and current status. *Lancet Infect Dis* **5**, 481–493 (2005).
53. Koluman, A. & Dikici, A. Antimicrobial resistance of emerging foodborne pathogens: Status quo and global trends. *Crit. Rev. Microbiol.* **39**, 57–69 (2013).
54. European Centre for Disease Prevention and Control. Surveillance Atlas for Infectious Diseases. https://atlas.ecdc.europa.eu/public/index.aspx. (2020).
55. The Center for Disease Dynamics Economics & Policy. ResistanceMap. https://resistancemap.cddep.org. Date accessed: February 19, 2020. (2018).
56. Hendriksen, R. S. *et al.* Using Genomics to Track Global Antimicrobial Resistance. *Front. Public Heal.* **7**, (2019).
57. Köser, C. U., Ellington, M. J. & Peacock, S. J. Whole-genome sequencing to control

antimicrobial resistance. *Trends Genet.* **30**, 401–407 (2014).
58. Argimón, S. *et al.* Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb. genomics* **2**, 1–11 (2016).
59. Hadfield, J. *et al.* NextStrain: Real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
60. Stelling, J. M. & Brien, T. F. O. Surveillance of Antimicrobial Resistance: The WHONET Program. *Clin. Infect. Dis.* **24**, S157–S168 (1997).
61. Havelaar, A. H. *et al.* A summary index for antimicrobial resistance in food animals in the Netherlands. *BMC Vet. Res.* **13**, 305 (2017).
62. Hughes, J. S. *et al.* How to measure the impacts of antibiotic resistance and antibiotic development on empiric therapy: New composite indices. *BMJ Open* **6**, (2016).
63. Magiorakos, A. P. *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: An international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**, 268–281 (2012).
64. Laxminarayan, R. & Klugman, K. P. Communicating trends in resistance using a drug resistance index. *BMJ Open* **1**, (2011).
65. Vandenbroucke-Grauls, C. M. J. E. *et al.* The proposed Drug Resistance Index (DRI) is not a good measure of antibiotic effectiveness in relation to drug resistance. *BMJ Glob. Heal.* **4**, 1–3 (2019).
66. Bengtsson, S., Bjelkenbrant, C. & Kahlmeter, G. Validation of EUCAST zone diameter breakpoints against reference broth microdilution. *Clin. Microbiol. Infect.* **20**, (2013).
67. Ethnologue. What are the top 200 most spoken languages? https://www.ethnologue.com/guides/ethnologue200. (2020).

Chapter 3

Mapping global coldspots of veterinary capacity

Authors:
Nicola G. Criscuolo[1], Yu Wang[1], Thomas P. Van Boeckel[1,2,3,*]

Affiliations:
[1]Department of Environmental Systems Science, ETH Zürich; Zurich, Switzerland.
[2]One Health Trust; Washington D. C., 20015, United States of America.
[3]Université Libre de Bruxelles (ULB), Brussels, Belgium.

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

Submitted

Abstract

Veterinarians play a vital role in providing healthcare, detecting zoonotic outbreaks, and safeguarding the livelihood of those relying on animals for subsistence. However, veterinary capacities are unequal between countries, and their geography is seldom documented despite significant implications for healthcare access. We used web-scraping techniques to sample 303,745 addresses of veterinarians from 115 countries. Once geocoded, we used geospatial models implemented in the Bayesian framework provided by the Integrated Nested Laplace Approximation to map the global distribution of veterinarians at 10x10 km^2. High population density and gross domestic product, and low travel time to cities were associated with a high density of veterinarians. The distribution of veterinarians was used in combination with density maps of cattle, chickens, and pigs to identify areas where veterinarians are farther than 1 hour of travel time from the animals, which we defined ("veterinary coldspots"). Low- and middle-income countries accounted for 93.8% of all the coldspots identified at the global-level.

Introduction

Maps have been instrumental in prioritizing interventions against infectious diseases of global importance (*1*). In human medicine, fine-scale maps have helped quantify the burden of diseases such as malaria and dengue, and facilitated geographically targeted campaigns to target insecticide bed nets in specific regions (*2*). Concomitantly, maps of healthcare facilities (*3*, *4*) have been used to make international comparisons of access to primary care and guide the deployment of important medicines such as antiretroviral therapy against HIV (*5*). In animal health, similar efforts have been conducted to map diseases that threaten the livelihood of those who raise animals for subsistence such as avian influenza (*6*), rift valley fever (*7*), or antimicrobial resistance (*8–11*). However, unlike for humans, mapping access to veterinary care has thus far received little attention.

Veterinarians are the first line of defense against all diseases of animal origin (*12*). Yet efforts to map veterinarians have thus far been limited to country-level census or, in a few countries, at province/state-level (United States, France, Italy, Spain, Switzerland) (*13–17*). The most notable global effort to inventory veterinary capacities – currently led by the World Organisation for Animal Health – is focused on country-level performance assessments (*18*). Although crucially important, these efforts can overshadow important geographic disparities in access to veterinary care. This is particularly true for low- and middle-income countries (LMICs) currently transitioning from extensive farming to intensive farming (*19*), a phase associated with an increased risk of emergence of zoonotic pathogens (*20*).

Important insight could be gained from assessing the state of the veterinary workforce at a fine spatial scale. Firstly, by identifying areas akin to medical deserts in human medicine (*21*) where long travel times to/for veterinarians are a major obstacle to delivering care. While travel time is only one dimension of the multifaceted challenge of access to care (i.e.: cost, education,

regulatory hurdles) it was recently identified in a meta-analysis (*22*) as a leading factor for insufficient animal care in LMICs. Maps could guide capacity building in these regions using travel time as a criterion for resource allocation. Secondly, mapping areas where veterinary capacities are currently insufficient could indirectly help strengthen surveillance for potential pandemic pathogens (*23*) in regions associated with a high risk of disease emergence (*24*, *25*).

In the last decade, platforms have appeared in many countries that enable internet users to find veterinarians in their vicinity using their postcodes/addresses (*26–29*). Whilst not primarily intended for public health, these platforms are an unprecedented opportunity to investigate the fine-scale geographic distribution of veterinarians and its determinants. However, to convert this data into actionable insights for capacity-building decisions multiple challenges must first be addressed. First, platforms listing addresses of veterinarians must be exhaustively inventoried. Second, web-scraping tools are required to extract the hundreds of thousands of addresses of veterinarians listed on webpages, and these must also be curated and geocoded. Third, statistical models that can capture variations in veterinarians' density, and integrate the influence of factors affecting that density, must be developed and validated regionally to ensure robust interpolation between data-rich regions and data-poor regions. Fourth, model predictions for the number of veterinarians in each country must be consistent with estimates from national and international databases. Finally, models should account for variations in the presence of veterinarians on online platforms between countries characterized by different levels of economic development and internet penetration.

In this study, we mapped the global distribution of veterinarians at 10x10 km^2 resolution using geospatial models in combination with a global address book of veterinarians assembled from open-access online platforms. Furthermore, we mapped regions where food animals live more than 1 hour away from veterinarians by motorized transport and classified these areas as "veterinary coldspots".

Building a global address book of veterinarians

We identified 87 online platforms listing the addresses of veterinarians across 115 countries (fig. S1), including OpenStreetMap and Google Maps, 16 national platforms listing veterinarians by postcode, 21 websites of veterinary councils, and 48 national phonebooks (fig. S2). We used web-scraping methods (Supplementary Materials) to extract addresses of 303,745 veterinary practices (Fig. 1 and fig. S3). Records were curated to i) remove practices different from the veterinary ones, ii) prevent duplicates across platforms, and iii) remove permanently closed practices. A majority of addresses were retrieved from the Americas (39.7%), Europe (38%), and Asia (18.2%), while Africa and Oceania accounted for 2.1%, and 2% of records, respectively. Veterinarians' specialization (companion *vs* food animals) was reported for 9.1% of veterinarians sampled. Amongst those, 85% cared for companion animals, and 15% cared for food animals.

Global distribution of veterinarians

We used a Log-Gaussian Poisson Regression model (*30*) to map veterinarians globally at 10x10 kilometers (fig. S4) via statistical associations between veterinarians' locations, and covariates such as human population density, gross domestic product, travel time to cities, and environmental variables (*31*, *32*) (table S1). Models were cross-validated in nine subregions (fig. S5) High-income countries (HICs) and low- and middle-income countries (LMICs) accounted, respectively, for 43% and 57% of all veterinarians predicted. Population density, gross domestic product, and low travel time to cities were positively associated with a high density of veterinarians and had the highest influence in predicting their distribution (table S2).

Mapped predictions of veterinarians were re-aggregated in each country and compared with the numbers of veterinarians reported from veterinary associations, international organizations, and peer-reviewed publications (table S3). National predicted numbers of veterinarians were within ±50% of estimates from national reports in 75.7% of countries (Supplementary Materials, fig. S6). In addition, records of graduates from 680 veterinary schools in 120 countries were used as upper bound to assess the plausibility of the re-aggregated predictions in each country (Supplementary Materials). Our predictions were lower than the national numbers of veterinary graduates in all countries except for Finland, Latvia, Cuba, and Uganda (fig. S6).

We also mapped the proportion and number of veterinarians specialized in food animals (*vs* companion animals) using a beta regression model (*33*) (Supplementary Materials, fig. S7). Pixels with a majority of veterinarians specialized in food animals were in Africa (54.6%), Oceania (43.4%), and Latin America (29.8%). In contrast, this proportion decreased to 18.9% in Europe, 13.5% in Asia, and 6% in North America.

A map of veterinarians specialized in food animals was compared with the global distribution of the livestock units (LSUs (*34*), Supplementary Materials) of food animals raised extensively (*19*) (Fig 1B). The aim was to identify veterinarians' availability in small-scale farms of people who rely on livestock for subsistence (*35*, *36*). For extensive farms in LMICs, one veterinarian cared on average for 5.4 LSUs while this number fell to 0.2 in HICs. At the country-level, we identified pixels with more than 5 LSUs of extensively raised food animals per veterinarian (hereafter referred to as "high animals' density areas") in 81.5% of Latin American countries, 92% of African countries, and 63.3% of Asian countries.

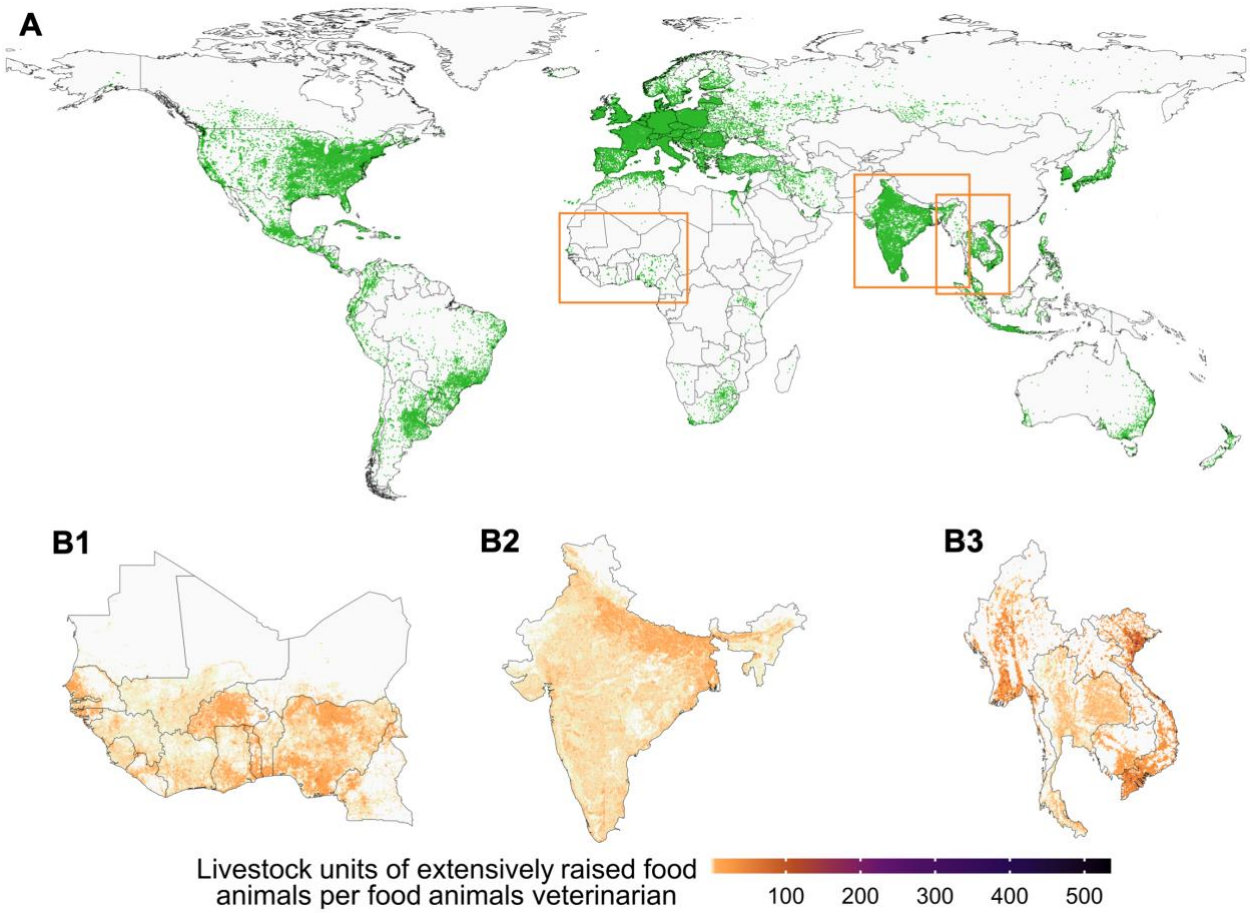

Fig. 1. Point pattern of the addresses of veterinarians and livestock units of extensively raised food animals per veterinarian. (A) Green dots represent the geographic locations of veterinarians, obtained by geocoding the addresses within online platforms. (B) Distribution of livestock units (*34*) of extensively raised food animals per veterinarian (specialized in food animals) at the 10x10 km^2 resolution in West Africa (B1), India (B2), and Mainland South-East Asia (B3).

Travel time to veterinary services

For each country, we calculated the travel time between the mapped predictions of veterinarians specialized in food animals and cattle, chickens, and pigs (*31*). We used a friction surface to estimate travel times based on land use characteristics (e.g., roads' speed limits, physical barriers, and elevation) (*37*). "Coldspots" of veterinary capacity were defined as 10x10 km^2 pixels where food animals were more than 1 hour away by motorized transport from the nearest veterinarian.

Global maps of coldspots (Fig. 2) showed that 189 million LSUs, lived more than 1 hour away from veterinarians. That is equivalent to 1.2 times the biomass of animals raised for food in the U.S. Asia had the highest percentage of animals living in coldspots (44.1%), followed by Latin America (27.7%), and Africa (18.7%). The highest percentages of all cattle in coldspots were in Brazil (22.1%), Sudan (8.1%), China (7.6%), Chad (7.5%), and Australia (4.5%). For

chickens, the highest percentages were in China (15%), Bolivia (8.5%), Russia (8.2%), Iran (7.5%), and Indonesia (6.9%), while for pigs were in China (49.9%), Myanmar (7.4%), Papua New Guinea (6%), Russia (5.5%), and Brazil (4.6%). Finally, at the species-level, LMICs accounted for 94% of cattle, 93.4% of chickens, and 99.4% of pigs in coldspots. Countries with the highest average travel time to reach an animal in coldspots were Somalia, China, Guyana, Sudan, Papua New Guinea, and Central African Republic. These patterns remained consistent when setting maximum travel time thresholds to define coldspots at 2 and 4 hours (fig. S8).

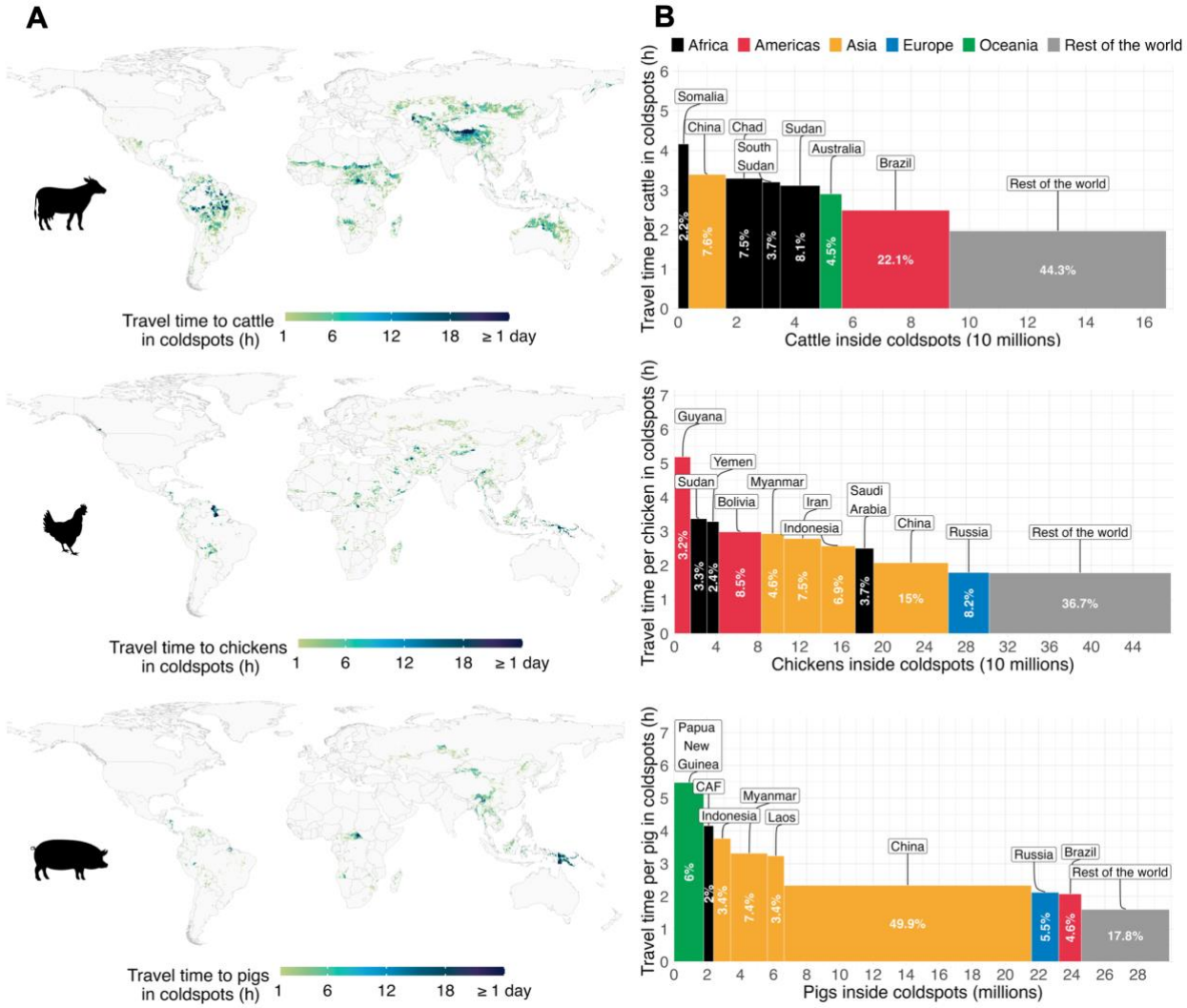

Fig. 2. Coldspots of veterinary capacity in cattle, chickens, and pigs. (A) Travel time for animals living at least 1 hour away by motorized transport from a veterinarian. (B) Average motorized travel time per animal living in coldspots per country. Numbers in the bars represent the percentages of animals per country living in coldspots, relative to the global number of animals in coldspots.

Except for North America and Europe (excluding Russia), regions with high percentages (>50%) of coldspots of veterinarians and low animal densities (<5 LSUs/10 km^2) were present in every continent (Fig. 3). Of even greater concerns were regions that combined the presence of coldspots with growing animals' densities (>5 LSUs/10 km^2). These were located

predominantly in Asia, and to a lesser extent in West Africa, and around the African Great Lakes. In Rwanda, Malawi, Bangladesh, Papua New Guinea, Vietnam, Philippines, and Haiti more than 75% of the aggregated LSUs of cattle, chickens, and pigs were present in coldspots and/or high animals' density areas.

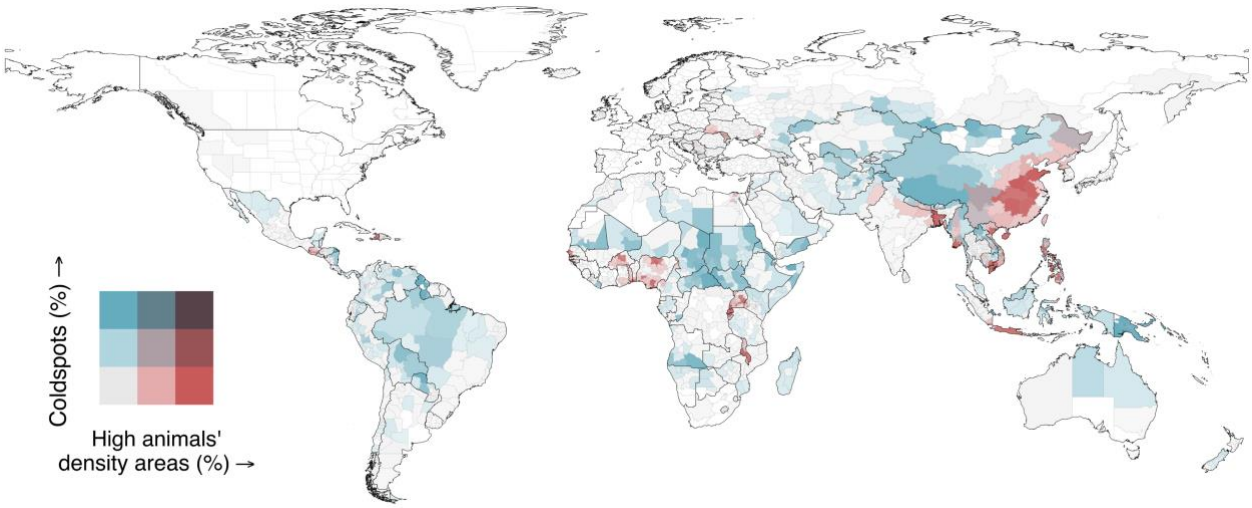

Fig. 3. Regional veterinary capacities available to food animals. Proportion of 10x10 km^2 pixels per administrative unit where food animals were more than 1 hour away (by motorized transport) from the nearest veterinarian ("coldspot"), combined with the proportion of pixels classified as high animals' density areas, i.e., pixels where the number of extensively raised food animals (converted in livestock units (*34*)) per veterinarian was higher than 5. In regions with a white background, the proportion of food animals in coldspots and/or high animals' density areas was lower than 1%.

Discussion

In this study, we assembled an address book of >300,000 veterinarians from 115 countries using data web-scraped from online veterinary registers. Although this data is irregular in nature, when used in combination with statistical models showed a good agreement with national estimates of the number of veterinarians. Furthermore, we showed that it can supplement international efforts that document veterinary capacities by providing insights into the geographic distribution of veterinarians at unprecedented resolution (10x10 km^2).

Using this resource, we could produce travel time maps to identify "veterinary coldspots", i.e., areas where veterinarians are >1 hour away from food animals. This could translate in the absence of regular veterinary controls in these areas that can affect the well-being of animals, but it could also escalate the potential for zoonotic diseases. For further investigating this dynamic, future research efforts could leverage our coldspots maps to examine their correlation with distributions of hotspots of emerging infectious diseases, as the ones provided by Jones and colleagues *(41)*. Since nearly 60% of such diseases have a zoonotic origin *(41)*, understanding the spatial overlap between veterinary coldspots and disease hotspots becomes crucial. However, a notable challenge in conducting such a comparative analysis lies in the fact

that, due to the lack of veterinarians, many of these zoonoses could spread undetected and lead to a biased correlation with the presence of coldspots.
Furthermore, it is essential to acknowledge that physical access is only one dimension of the multifaceted aspect of access to care. Our efforts to build a global dataset of veterinarians could be expanded, at each location, to incorporate information on other factors affecting access to care such as price, training level, number of personnel, disease burden (GBAD), etc. Assembling this information must necessarily take place through a multi-stakeholder approach: via international organization, but also potentially by involving individual practitioners in data gathering and validation. The tentative goal of this work is to provide the backbone to seed this process and illustrate the value of high-resolution data for capacity building in animal health.

In HIC, the predictions of veterinarians derived from web-scraped data mirrored past research from the United States, Canada, and France (*42–44*), showing that veterinarians predominantly cluster in affluent urban areas thereby reflecting the growing trend for specialization in pet care (80-95%) over food animal care (5-20%) (*45*). In contrast, in LMICs our predictions suggest that the proportion of food animal veterinarians is higher (25-50%) but their number is limited, and their geographic distribution is highly heterogenous leading to >93% of coldspots of access to care being in LMICs. Potential causes for limited access to care include insufficient number of veterinarians but also insufficient training and equipment for diagnostic laboratories (*46*), comparatively low salaries (*47*), and low standard of living in rural areas (*48*). The original data from LMICs supporting this assessment is currently limited: of the web-scraped veterinary register just 33.6% were from LMIC although these make up 79.8% of the global food production. In this context, our maps can provide a first (publicly available) baseline for increasing veterinary capacities but also a starting point for better documenting the veterinary workforce at a sub-national scale. For example, in China, the largest animal producer in the world, public data on the geographic distribution of veterinarians could not be identified for this study, and even national estimates of the number of veterinarians could only be found via press articles (*49*).

Limitations

As for any modeling study, our analysis and datasets come with limitations. First, our study focused on veterinarians – holders of a university degree in veterinary medicine – and did not include “paravets” who are semi-autonomous professionals predominantly present in LMICs whose qualifications vary considerably between countries. Unlike veterinarians, paravets were not systematically inventoried on online platforms and were therefore not included in these analyses. Second, the uncertainty associated with spatial interpolations of the veterinarians’ maps is reflected in confidence interval maps (fig. S5). These uncertainty levels reflect the spatial cross-validation procedure used to prevent regional overfitting. However, these do not reflect a comparison with independent field surveys since all data sources identified were included in the model training and cross-validation to produce the most accurate maps possible. Third, the distribution of coldspots in countries undergoing a rapid growth in intensive farming systems, such as those for cattle in Brazil *(50)*, could be overestimated. This could be attributed

to the fact that large intensive farms, in contrast to extensive farms, could rely on private veterinarians to manage the health of food animals (Supplementary Materials). Therefore, such veterinarians might have no incentive in being present in the online platforms we web-scraped, resulting in a biased distribution of coldspots. Fourth, estimates of the number of veterinary graduates were collected in each country for comparison with our maps. However, these did not account for veterinarians working in a country different from the country of graduation. Finally, for high-income countries such as Australia, our travel time analysis did not account for the fact that veterinarians may occasionally travel via airplane to reach remote farms (*51*).

Perspective

Web-scrapable data such as online registers of veterinarians' addresses vary considerably in completeness between countries. However, when supplemented with adequate predictive models, our analysis shows that it also carries considerable benefits to address questions at a spatial scale that are not covered by existing databases that document veterinary capacities.

This first attempt at mapping the global veterinary workforce can help outline two priorities for action. First, within-country disparities in access to care are considerable, and these should focus the attention of national and international funders. Second, maps of veterinarians can serve as advocacy material beyond the veterinary realm to strengthen surveillance against pathogens that can emerge in animals but affect humans. Considerable attention has been devoted to identifying regions at risk of new emerging pathogens but comparatively less emphasis has been placed on documenting the state of the veterinary workforce in these regions, and thus our ability to rapidly detect and control these pathogens. Our maps offer a starting to address these challenges.

References

1. B. J. Reich, M. Haran. Precision maps for public health. *Nature* **555**, 32–33 (2018). doi:10.1038/d41586-018-02096-w
2. A. Bertozzi-Villa, C. A. Bever, H. Koenker, D. J. Weiss, C. Vargas-Ruiz, A. K. Nandi, H. S. Gibson, J. Harris, K. E. Battle, S. F. Rumisha, et al. Maps and metrics of insecticide-treated net access, use, and nets-per-capita in Africa from 2000-2020. *Nature Communications* **12**, 1–12 (2021). doi:10.1038/s41467-021-23707-7
3. D. J. Weiss, A. Nelson, K. Gligorić, S. Bavadekar, E. Gabrilovich, J. Rozier, H. S. Gibson, T. Shekel, C. Kamath, A. Lieber, et al. Global maps of travel time to healthcare facilities. *Nature Medicine* **26**, (2020). doi:10.1038/s41591-020-1059-1
4. healthsites.io - Building an open data commons of health facility data with OpenStreetMap (2023). Available at: https://healthsites.io. Accessed: 18/05/2023.
5. C. G.S., T. F.C., B. T.W., N. M.L. Population uptake of antiretroviral treatment through primary care in rural South Africa. *BMC public health* **10**, 585 (2010).
6. M. Gilbert, N. Golding, H. Zhou, G. R. W. Wint, T. P. Robinson, A. J. Tatem, S. Lai, S. Zhou, H. Jiang, D. Guo, et al. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. *Nature Communications* **5**, 1–7 (2014). doi:10.1038/ncomms5116
7. A. C. A. Clements, D. U. Pfeiffer, V. Martin. Application of knowledge-driven spatial modelling approaches and uncertainty management to a study of Rift Valley fever in Africa. *International Journal of Health Geographics* **5**, 1–12 (2006). doi:10.1186/1476-072X-5-57
8. T. P. Van Boeckel, J. Pires, R. Silvester, C. Zhao, J. Song, N. G. Criscuolo, M. Gilbert, S. Bonhoeffer, R. Laxminarayan. Global trends in antimicrobial resistance in animals in low- And middle-income countries. *Science* **365**, (2019). doi:10.1126/science.aaw1944
9. D. Schar, C. Zhao, Y. Wang, D. G. J. Larsson, M. Gilbert, T. P. Van Boeckel. Twenty-year trends in antimicrobial resistance from aquaculture and fisheries in Asia. *Nature Communications* **12**, 6–15 (2021). doi:10.1038/s41467-021-25655-8
10. N. G. Criscuolo, J. Pires, C. Zhao, T. P. Van Boeckel. resistancebank.org, an Open-Access Repository for Surveys of Antimicrobial Resistance in Animals. *Scientific Data* **8**, 1–10 (2021). doi:10.1038/s41597-021-00978-9
11. C. Zhao, Y. Wang, K. Tiseo, J. Pires, N. G. Criscuolo, T. P. Van Boeckel. Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China. *Nature Food* **2**, 596–602 (2021). doi:10.1038/s43016-021-00320-x
12. V. Bellemain. The role of Veterinary Services in animal health and food safety surveillance, and coordination with other services. *Revue scientifique et technique (International Office of Epizootics)* **32**, 371–381 (2013).
13. S. M. Neal, M. J. Greenberg. Putting Access to Veterinary Care on the Map: A Veterinary Care Accessibility Index. *Frontiers in Veterinary Science* **9**, 1–8 (2022). doi:10.3389/fvets.2022.857644
14. M. Berrada, Y. Ndiaye, D. Raboisson, G. Lhermie. Spatial evaluation of animal health

care accessibility and veterinary shortage in France. *Scientific Reports* **12**, 1–10 (2022). doi:10.1038/s41598-022-15600-0
15. Federazione Nazionale Ordini Veterinari Italiani. Iscritti Ordine Veterinario (2023). Available at: https://www.fnovi.it/iscritti-ordine. Accessed: 15/11/2021.
16. Istituto Nacional de Estadìstica. Profesionales sanitarios colegiados (2016). Available at: https://www.ine.es/jaxi/Datos.htm?path=/t15/p416/a2016/&file=s04002.px#!tabs-mapa. Accessed: 15/11/2021.
17. Swiss Federal Office of Public Health. Register of Medical Professions (2022). Available at: https://www.healthreg-public.admin.ch/medreg/search. Accessed: 15/12/2021.
18. World Organisation for Animal Health. Performance of Veterinary Service (PVS) Pathway (2022). Available at: https://www.woah.org/en/what-we-offer/improving-veterinary-services/pvs-pathway/#:~:text=The PVS Pathway empowers national,inefficiencies and opportunities for innovation. Accessed: 20/11/2022.
19. M. Gilbert, G. Conchedda, T. P. Van Boeckel, G. Cinardi, C. Linard, G. Nicolas, W. Thanapongtharm, L. D'Aietti, W. Wint, S. H. Newman, et al. Income disparities and the global distribution of intensively farmed chicken and pigs. *PLoS ONE* **10**, 1–14 (2015). doi:10.1371/journal.pone.0133381
20. J. H. Spencer, M. L. Finucane, J. M. Fox, S. Saksena, N. Sultana. Emerging infectious disease, the household built environment characteristics, and urban planning: Evidence on avian influenza in Vietnam. *Landscape and Urban Planning* **193**, (2020). doi:10.1016/j.landurbplan.2019.103681
21. V. Lucas-Gabrielli, G. Chevillard. 'Medical deserts' and accessibility to care: what are we talking about? *Médecine/Sciences* **34**, 599–603 (2018). doi:10.1051/medsci/20183406022
22. E. LaVallee, M. K. Mueller, E. McCobb. A Systematic Review of the Literature Addressing Veterinary Care for Underserved Communities. *Journal of Applied Animal Welfare Science* **20**, 381–394 (2017). doi:10.1080/10888705.2017.1337515
23. Wikipedia. PREDICT (USAID) (2023). Available at: https://en.wikipedia.org/wiki/PREDICT_(USAID). Accessed: 15/05/2023.
24. J. P. Colella, B. R. Agwanda, F. A. A. Khan, B. John, A. C. Bonilla Carrión, N. U. de la Sancha, J. Dunnum, A. W. Ferguson, S. Greiman, E., P. K. Kiswele, et al. EcoHealth reframing of disease monitoring Build international biorepository capacity. *Science* **370**, 773–775 (2020).
25. C. J. Carlson, M. J. Farrell, Z. Grange, B. A. Han, N. Mollentze, A. L. Phelan, A. L. Rasmussen, G. F. Albery, B. Bett, D. M. Brett-Major, et al. The future of zoonotic risk prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, (2021). doi:10.1098/rstb.2020.0358
26. Royal College of Veterinary Surgeons. Find A Vet (2023). Available at: https://findavet.rcvs.org.uk/home/. Accessed: 10/02/2020.
27. Ordre National des Vétérinaires. Trouver un vétérinaire pour soigner mon animal (2023). Available at: https://extranet.veterinaire.fr/annuaires/soigner-mon-animal. Accessed: 10/03/2020.
28. Tierarzt-onlineverzeichnis. Tierarzt-Suche nach Bundesländern (2023). Available at: https://www.tierarzt-onlineverzeichnis.de/suche.

29. Dierenarts kliniek.nl. Alle dierenartsklinieken van Nederland (2023). Available at: https://dierenarts-kliniek.nl. Accessed: 12/04/2020.
30. F. E. Bachl, F. Lindgren, D. L. Borchers, J. B. Illian. inlabru: an R package for Bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution* **10**, 760–766 (2019). doi:10.1111/2041-210X.13168
31. M. Gilbert, G. Nicolas, G. Cinardi, T. P. Van Boeckel, S. O. Vanwambeke, G. R. W. Wint, T. P. Robinson. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data* **5**, 1–11 (2018). doi:10.1038/sdata.2018.227
32. G. Cinardi, D. Da Re, M. Gilber, T. P. Robinson, W. G. R. Wint. Gridded Livestock of the World - 2015 (GLW4). *Harvard Dataverse* (2022). doi:10.7910/DVN/LHBICE
33. F. Cribari-Neto, A. Zeileis. Beta regression in R. *Journal of Statistical Software* **34**, 1–24 (2010). doi:10.18637/jss.v034.i02
34. Eurostat. Livestock Unit (2022). Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Livestock_unit_(LSU). Accessed: 15/12/2022.
35. D. Temple, X. Manteca. Animal Welfare in Extensive Production Systems Is Still an Area of Concern. *Frontiers in Sustainable Food Systems* **4**, (2020). doi:10.3389/fsufs.2020.545902
36. J. Ilukor. Improving the delivery of veterinary services in Africa: insights from the empirical application of transaction costs theory in Uganda and Kenya. *Revue Scientifique et Technique de l'OIE* **36**, 279–289 (2017). doi:10.20506/rst.36.1.2628
37. The Malaria Atlas Project. Accessibility to Healthcare - Motorized friction surface (2020). Available at: https://malariaatlas.org/project-resources/accessibility-to-healthcare/. Accessed: 01/10/2022.
38. A. T. Murray. Maximal Coverage Location Problem: Impacts, Significance, and Evolution. *International Regional Science Review* **39**, 5–27 (2016). doi:10.1177/0160017615600222
39. H. A. Eiselt, V. Marianov. Foundations of Location Analysis. *Springer* (2011).
40. WOAH. GBADs - The Global Burden of Animal Diseases (2023). Available at: https://gbads.woah.org. Accessed: 12/10/2023.
41. K. E. Jones, N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, P. Daszak. Global trends in emerging infectious diseases. **451**, 990–994 (2008). doi:10.1038/nature06536
42. J. M. Richards. An Ecological Analysis of the Geographic Distribution of Veterinarians in the United States. *Journal of Vocational Behavior* **11**, 216–231 (1977). doi:10.1016/0001-8791(77)90008-2
43. M. R. Olfert, M. Jelinski, D. Zikos, J. Campbell. Human capital drift up the urban hierarchy: Veterinarians in Western Canada. *Annals of Regional Science* **49**, 551–570 (2012). doi:10.1007/s00168-011-0448-2
44. S. Truchet, N. Mauhe, M. Herve. Veterinarian shortage areas: what determines the location of new graduates? *Review of Agricultural, Food and Environmental Studies* **98**, 255–282 (2017). doi:10.1007/s41130-018-0066-9
45. Health for Animals. Global State of Pet Care: Stats, Facts and Trends - https://www.healthforanimals.org/wp-content/uploads/2022/07/Global-State-of-Pet-

Care.pdf. (2022).
46. J. N. Lasley, E. O. Appiah, K. Kojima, S. D. Blacksell. Global Veterinary Diagnostic Laboratory Equipment Management and Sustainability and Implications for Pandemic Preparedness Priorities1. *Emerging infectious diseases* **29**, 1–12 (2023). doi:10.3201/eid2904.220778
47. The World Bank - World Bank Country and Lending Groups (2022). Available at: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. Accessed: 15/03/2023.
48. Federation of Veterinarians of Europe. Shortage of veterinarians in rural and remote areas. *Summary report* 1–14 (2020).
49. The Guardian - China's new animal health rules alone won't stop zoonotic outbreaks, experts warn (2021). Available at: https://www.theguardian.com/environment/2021/jan/26/chinas-new-animal-health-rules-alone-wont-stop-zoonotic-outbreaks-experts-warn. Accessed: 11/09/2022.
50. P. Vale, H. Gibbs, R. Vale, M. Christie, E. Florence, J. Munger, D. Sabaini. The Expansion of Intensive Beef Farming to the Brazilian Amazon. *Global Environmental Change* **57**, 101922 (2019). doi:10.1016/j.gloenvcha.2019.05.006
51. A. Scott. The Flying Vet. *HarperCollinsPublishers Australia* (2023).

Supplementary materials

Collecting national estimates and addresses of veterinarians

Between March 2020 and June 2021, we conducted an online sampling campaign on a global scale to collect data about veterinarians. We searched for addresses of veterinarians to investigate their global distribution and for their national estimates as a reference for the comparison with our country-level predictions. Concretely, in each country, we searched for:

A. National estimates of the number of veterinarians.

B. Annual estimates of the number of graduates from veterinary schools.

C. Addresses or geographic coordinates of veterinary hospitals, clinics, and private practices.

D. Information about the animals treated in these veterinary practices, i.e., companion animals (cats, dogs, other pets) or food animals.

A. National estimates of veterinarians

In this phase, we consulted websites of international organizations such as the World Organization for Animal Health (WOAH), the Food and Agriculture Organization (FAO), the World Veterinary Association (WVA), and the World Small Animal Veterinary Association (WSAVA). In Europe, we retrieved data from the website of the European Board of Veterinary Specialization (EBVS) and the 2019 survey of the veterinary profession in Europe (*1*) compiled by the Federation of Veterinarians of Europe (FVE). For the rest of the world, we consulted websites of national associations and universities, governmental reports, and estimates by statistical agencies. We also searched national estimates on scientific reports, peer-reviewed publications, and online newspapers.
In addition, we sent 309 emails to veterinary associations and governmental agencies to request national estimates (response rate of 16.8%). From this search, we found 117 national-level estimates of the number of veterinarians from across 79 countries. We collected 79 estimates from international organizations and national associations, 13 from governmental agencies, 20 from online newspapers and blogs, and 8 from scientific literature (table S3).

B. Estimates of graduates from veterinary schools

We estimated the number of graduates who may still be practicing veterinary medicine in the year 2022 as the upper threshold for comparing the number of veterinarians we predicted at the national level. First, we defined a global list of veterinary schools available from the WOAH questionnaires on the veterinary educational establishments (*2*) and Wikipedia (*3*). In addition, we sent 659 emails to veterinary schools (response rate of 4.7%) and consulted the WOAH

questionnaires to collect school-specific data about the range of graduates per year and the average number of years required to obtain a degree in veterinary medicine. Also, we collected the foundation year of each school and the retirement age per country (*4*). For calculating the number of graduates, we made the following assumptions:

- 10% of students graduate each year from a veterinary school.
- Older schools have a bigger range of graduates per year than newer schools.
- The first valid year considered to estimate the number of graduates is given by the foundation date of the school plus the years required to obtain the degree.
- The last valid year considered for graduation is 2022.
- Veterinarians who are now retired based on a country's retirement age were excluded from the number of graduates estimated.

Then, for each school, we calculated the number of graduates, NG, as follows:

$$NG = \sum_{n=1}^{YA} 10\% \cdot G_{MAX} + (n-1) \cdot \left(\frac{G_{MAX} - G_{MIN}}{YA - 1}\right)$$

Where *YA* is the number of years, net of the retirement age, in which students graduated from the school, and G_{MAX} and G_{MIN} are, respectively, the maximum and minimum number of graduates per year. Finally, we summed NG for each national veterinary school, estimating the national-level number of graduates from 680 veterinary schools out of the 719 sampled across 120 countries.

C. - D. Addresses and specialization of veterinarians

For assembling the global database of veterinarians' addresses, we first identified the types of online sources listing their addresses or coordinates. We prioritized the data collection from:

i. Online platforms specifically designed to search veterinarians by postcode (e.g., https://findavet.rcvs.org.uk/home/).

ii. Websites of national phonebooks, like countries' Yellow Pages.

iii. Websites of national veterinary associations and governmental agencies as the ministries of agriculture.

iv. Open-access web maps like OpenStreetMap and Google Maps.

We retrieved websites of data sources listed in points (i) to (iii) by querying internet browsers both in English and the main country language. The key search string we used to identify platforms listing veterinarians' addresses combined the country name with terms such as "find a veterinarian", "veterinarians near me", "veterinarians by address", and "veterinarians by postcode". For finding national phonebooks, we combined the country name with the terms "national phonebook", "yellow pages", and "white pages", and used the list of World's Yellow Pages websites available from Wikipedia (*5*). Finally, for the websites of veterinary associations and governmental agencies, we combined the country name with the terms "veterinary association", "veterinary union", "veterinary council", "ministry of agriculture", and "national statistics institute". In addition, we used generic queries like "list of veterinarians" followed by the country name to find data potentially present in online sources not considered.

Next, we defined the key search terms to use on each online data source to display veterinarians' addresses. Specifically:

i. For online platforms listing veterinarians by postcode, we searched the complete list of veterinarians in the country, without specifying a city or a postcode in the website's search box.

ii. For national phonebooks, we obtained webpages containing lists of addresses through the search terms "veterinarian", "veterinarians", "vet", "veterinary clinic", "veterinary hospital", and "veterinary practice".

iii. For the websites of veterinary councils, we collected the online list or the downloadable PDF of the veterinarians registered.

iv. When sampling veterinarians' addresses through OpenStreetMap (*6*), we used the dedicated Application Programming Interface (API) available in the R package *osmdata* (*7*) to query the OpenStreetMap database one country at a time. As reported by the OpenStreetMap glossary, we built the API query by using the term "amenity" to subset the group of services listed on the database and the term "veterinary" as the key to refine the search. According to the glossary, a query built with these tags gives in return "places where a veterinary surgeon, also known as a veterinarian or vet, practices" (*8*).

    In contrast, the API to query the Google Maps database, available in the R package *googleway* (*9*), allowed for only three queries per zone, which are too few to list veterinarians in a country. For this reason, we defined multiple country-level queries for every city listed on the *opendatasoft* database (*10*) through strings containing the term "veterinarian" followed by the city and the country name. In addition, we restricted the search only to places offering veterinary care using the tag "veterinary_care" in the function to query Google's database, as reported in the API's user manual (*11*).

Each query of Google Maps always returned the names and addresses of veterinarians complete with geographic coordinates. In contrast, every OpenStreetMap query returned entries with

geographic coordinates, but only 37.7% of them were complete with names and addresses of veterinarians. Every other online source inspected provided only the addresses of veterinarians, and, where available, their name and specialization, as text strings. Therefore, we used different means to extract such information from PDFs and web pages:

- For PDFs, we used the R package *tabulizer* (*12*) to extract the file's tables containing the addresses of veterinarians.

- When addresses were only available on web pages, we sampled them through web-scraping (*13*). Specifically, we coded web-scraping software in the R and python programming languages through the packages *rvest* (*14*) and *BeautifulSoup* (*15*). In addition, we used the *selenium WebDriver* package (*16*) to automate web navigation through web-crawlers to visualize and sample data in multiple web pages of the same website. This practice was also used with websites embedding applications displaying data only upon users' interaction with specific buttons.

Finally, we used every address sampled as input for the Google's geocoding API (*17*) to retrieve its geographic coordinates if not already available (detailed point pattern examples in fig. S9).

Data curation of the addresses of veterinarians

Addresses of veterinarians from diverse sources led to collecting duplicated addresses or entries not specifically related to the veterinary medical profession. Therefore, assembling a database with unique geographic entries of veterinarians required screening for:

i. Duplicated addresses of veterinarians sampled from different online data sources.

ii. Entries related to the broader field of veterinary care where the presence of veterinarians is not required (e.g., veterinary pharmacies).

iii. Entries wrongly geocoded by Google's API because of missing information inside the text strings of the addresses.

For step (i), we first combined the database's strings reporting the name, the address, and the geographic coordinates of each entry. Then, we calculated the pairwise strings' similarity through the Levenshtein similarity (*18*). For this first screening, we set a threshold of 90% similarity to consider two strings as duplicates and remove one of them from the database. However, we also identified duplicates when two or more online data sources reported a veterinarian's name and address in different ways. For example, several websites reported strings with truncated addresses, buildings' numbers placed before the street name or vice versa, or missing addresses' names, as for data collected especially from OpenStreetMap. For this reason, we performed a second screening using geographic coordinates of sampled data

and classified an observation as duplicate if falling within an area of ~10 m radius built around another veterinary practice (fig. S10A).

In step (ii) we removed entries related to facilities different from veterinary practices. Hence, we first restricted the screening to entries where a facility's name didn't imply the presence of a veterinarian working in that facility (e.g., the store "Budget Pet Supplies" in England). Next, we used Google's language API (*19*) to translate more than 100 words (e.g., "pet shop", "pet supplies", "pet food", "kennel", "veterinary pharmacies", etc.) in every language available in the Google database. Then, we screened the remaining entries for patterns in the facilities' names matching the translated words and removed the matching entries from the database. Nevertheless, online data sources often list facilities that are different from veterinary practices but whose name doesn't contain information on the type of service they provide (e.g., the dogs' training center "Pets with Problems" in England). For this reason, we combined web-crawling and web-scraping software to perform automatic Google searches using names and addresses of the remaining veterinary practices to collect information on the service they provide, listed in the webpage of the Google results (fig. S10B). Then, we removed entries not corresponding to veterinarians. Since Google also lists the status of a facility, with this method we could identify veterinary practices permanently closed, excluding them from the database.

Step (iii) of data curation concerned entries of veterinarians' addresses wrongly geocoded by the Google API. Hence, for each country, we overlayed their shapefiles to spatial points of sampled addresses and retained in the database only the ones falling within the boundaries of the shapefiles.

Predicting the global distribution of veterinarians

We aggregated the geocoded addresses of veterinarians in pixels of 0.08333 decimal degrees, or approximately 10x10 km^2 at the equator. As a result, we obtained counts of veterinarians per pixel that we modeled through a geospatial analysis to predict the distribution of veterinarians at the global level.

Step 1. **Selection of spatial covariates**. We selected spatial covariates potentially related to the distribution of veterinarians. Previous country-level studies showed that veterinarians' distribution depends on the high population density and income levels of a region (*20*, *21*), and opportunities to share practices for early graduates (*22*). For this reason, we selected spatial covariates that could correctly represent this observed tendency of veterinarians to aggregate in urban areas. Specifically, we included in our stack maps of population density, Gross Domestic Product (GDP), and travel time to cities with more than 50,000 inhabitants (hereafter referred to as "major cities"). Besides socio-economic indicators, we also selected agricultural covariates potentially useful to represent the separation of veterinarians between rural and urban areas. Hence, we included the proportion of areas used for croplands and pastures, and the density of cattle, chickens, pigs, and sheep available from the 4th version of the Gridded Livestock of the World database (GLW4). All covariates, except the proportion of areas used for croplands and

pastures, were log10-transformed and all of them were resampled at the 10x10 km^2 resolution. Plots, measure units, and references of these covariates and support maps used in this study are available in fig. S11 and table S1.

Before incorporating covariates in the models, we quantified their correlation through a version of the Variance Inflation Factor (VIF) adapted for spatial objects (*23*). As for the VIF used in linear regression, we used a threshold of 10 to exclude correlated covariates (*24*) (VIF values in table S4). Furthermore, in a similar approach used to create gridded maps of the human population (*25–27*), we used the 10x10 km^2 resolution map of the world settlement footprint (WSF) (*28*) to outline areas where there are human settlements. We used such areas to identify the pixels for constraining both the modeling analysis and the predictions of veterinarians.

Step 2. **Geospatial model**. We modeled the counts of veterinarians as a Poisson variable using environmental and anthropogenic covariates as fixed effects. In addition, we used a Gaussian Random Field (GRF), discretized to finite elements called mesh through the Stochastic Partial Different Equation method (SPDE) (*29*, *30*), as a random effect to account for spatial autocorrelation (*31*, *32*). A statistical model with these characteristics is called Log-Gaussian Poisson Regression model (*33*) (LGPR). Given the global scale of the study, we fitted the LGPR models through the Integrated Nested Laplace Approximation algorithm (INLA) for computational efficiency (*33–36*).

Step 3. **Definition of areas for training models**. We defined the geographic areas where to train and validate the accuracy of LGPR models. Specifically, we divided the world into subregions according to the division proposed by the United Nations (*37*). However, we divided the "Europe and North America" subregion into three different subregions to have a comparable number of subregions based on high-income (HICs) and low- and middle-income countries (LMICs) as classified by the World Bank. As a result, from West to East, we defined 9 subregions: North America, Latin America, Europe, the Middle East, Sub-Saharan Africa, Russia, Central Asia, Eastern Asia, and Oceania (fig. S12).

Step 4. **Mesh definition**. In each subregion, we initialized the SPDE model by creating the mesh and assigning priors to the function used to capture the spatial autocorrelation of data. First, we used subregions' shapefiles to define the mesh reticulate where to apply the SPDE algorithm. As suggested by Lindgren et al., we defined a regular mesh inside the whole study area where we sampled data (*36*). Also, we defined a 2.5-decimal degree buffer zone outside the shapefile borders to capture spatial autocorrelation of data present along borders while avoiding issues related to the edge effect (*38*). Then, we defined the mesh resolution of the study area (inner mesh) by setting a maximum length of the triangle vertices of 0.15 decimal degrees. In contrast, for the mesh in the buffer zone (outer mesh), we allowed for a maximum length of 10 decimal degrees since this area doesn't influence predictions' accuracy (*39*). Finally, we set the minimum length of triangle vertices of the inner mesh to 0.01 decimal degrees to cover every part of the study area.

Step 5. **Covariance function**. We applied a Matérn covariance function to the mesh to capture the spatial autocorrelation of data by specifying its hyperparameters. Since we lacked information about spatial dependence among veterinarians, we assigned Penalized-Complexity priors (*40*) to the hyperparameters of the covariance function by setting a probability lower than 0.01 both for the range and its standard deviation of being higher, respectively, than 0.2 and 0.05 decimal degrees.

Step 6. **Importance of spatial covariates**. We used INLA to fit different LGPR models by including in the formula, besides the SPDE, one spatial covariate at the time. For each model, we calculated the Deviance Information Criterion (DIC) (*41*, *42*), whose lower values suggest better model performances (*43*). If a covariate didn't decrease the DIC, we excluded it from the LGPR formula. Then, of all the covariates selected, we retained in the final formula just the ones whose 95% credible intervals of their posterior means didn't cross zero (*44*). Once we identified the subregional LGPR model with the lowest DIC, we reran the models removing one covariate at a time to calculate the change in the DIC and identify the covariates with the highest effect on the outcome. Second, we compared the DIC of each best subregional LGPR model with and without the SPDE model to understand if adding a random effect to capture spatial autocorrelation to the models' formula decreased the DIC (*45*).

The models trained in the nine subregions with spatial covariates, in combination with the SPDE model, showed lower values of the DIC compared to models with covariates only, which suggests that accounting for spatial autocorrelation led to better model performances (fig. S13). According to their 95% credible intervals, travel time to major cities and GDP were significant in every subregional model, while population density in six of them (fig. S14 and table S5). In all models, travel time to major cities and GDP showed an inverse and direct relationship with the response, respectively. The proportion of areas used for croplands was significant in the models of seven subregions and had an inverse relationship with the response. Finally, the covariates of the density of food animals were selected in almost half of the subregional models but didn't show a clear relationship with the response.

For each subregion, the covariates resulting in the highest DIC decrease were population density, travel time to major cities, and GDP (table S2). When significant, population density always accounted for the highest share of the DIC decrease produced by a full model. In subregions where population density was discarded, travel time to major cities was the most influential covariate. As for the effects of spatial autocorrelation, the range of each SPDE model was always lower than 0.2 decimal degrees, except for the models trained in Sub-Saharan Africa and Central Asia, with ranges respectively of 0.285 and 0.254 decimal degrees.

Step 7. **Accuracy of predictions**. We assessed the accuracy of LGPR models by comparing values of predicted *vs* observed counts of veterinarians through the adjusted Coefficient of Determination based on the deviance residuals of the model (R^2_{DEV}) (*46*–

48). We validated the predictions' accuracy of every subregional LGPR model by computing the R^2_{DEV} between the predicted and the observed counts of veterinarians per pixel both in the subregions where we trained the models and in all external subregions.

The spatial cross-validation showed that each subregional model produced the most accurate predictions in the same subregion it was trained (R^2_{DEV} range: 0.68-0.96), fig. S5). In addition, models trained in North America, Europe, and Oceania showed moderate accuracy (*49*) when predicting veterinarians in each of the same high-income subregions (R^2_{DEV} range: 0.58-0.69). Other models showed moderate performances in adjacent subregions, such as the model of North America predicting in Latin America (R^2_{DEV}: 0.56) and the models of Central and Eastern Asia predicting in Russia (R^2_{DEV}: 0.70). The model trained in the Middle East produced the best performances in Eastern Asia (R^2_{DEV}: 0.56) and Oceania (R^2_{DEV}: 0.58), while the one trained in Sub-Saharan Africa was accurate only in the same subregion. Similarly, all subregional models performed poorly in Sub-Saharan Africa ($R^2_{DEV} < 0.33$).
For each subregion we quantified the uncertainty of predictions by mapping their standard deviation and 95% confidence intervals (*43*). The global map of the standard deviation of predictions (fig. S4B) showed values lower than 1 in 81.5% of the pixels. Pixels with the highest uncertainty of predictions were identified in Mexico, Saudi Arabia, and China.

Step 8. **Assembling the global map of veterinarians' distribution**. Once assessed the accuracy of the LGPR models, we used the predictions obtained by each subregional model that returned the highest R^2_{DEV} to assemble the global map of the distribution of veterinarians, regardless their specialization.

Step 9. **Validating predictions against national census of veterinarians**. We furtherly verified the goodness of our geospatial analysis by checking the agreement between the national estimates of veterinarians and their country-level numbers aggregated from the pixel-level predictions. Specifically, for each country, we summed the pixel-level predictions of the number of veterinarians and compared the result with veterinarians' national estimates sampled from veterinary associations, governmental agencies, peer-reviewed literature, and online newspapers, and with the estimated number of graduates from veterinary schools.

Investigating preferential sampling of addresses

In our study, we only used online data sources to collect the addresses of veterinarians. However, if veterinarians register online only in certain areas to compete for a pool of patients, our sampling campaign could be preferential.
Although preferential sampling can lead to biased predictions of the response (*50*), works have shown that the inclusion of spatial covariates and a random effect accounting for spatial autocorrelation can typically account for this bias (*51–53*). Hence, we selected five countries

(Austria, Belgium, Denmark, Netherlands, and Switzerland) as study areas to inspect for the presence of preferentially sampled addresses of veterinarians using two methods:

i. A joint model proposed by Pennino et al. based on a Log-Gaussian Cox Process (LGCP) to account for preferential sampling (*50*, *51*). Then, we used the R^2_{DEV} to compare the accuracy of the predictions of the LGCP with the accuracy obtained through the LGPR models fitted in the same study areas.

ii. A Monte Carlo test that targets the excess of clustered sampling locations in areas of high or low spatial autocorrelation of data by using the K-Nearest Neighbors algorithm (*54*, *55*).

The LGCP models accounting for preferential sampling showed a lower accuracy of predictions when compared to the LGPR models used to predict veterinarians' counts (fig. S15). This means that including the spatial covariates was sufficient to account for possible bias in our sampling.
This found confirmation also in the results obtained through the Monte Carlo test for preferential sampling (table S6). When considering a spatial autocorrelation field based only on the sampling locations, the test showed that, in each country, data were sampled in areas of high spatial autocorrelation (p-value < 0.05). However, when including a covariate such as population density to capture the structure of the sampling pattern, the correlation between the sampling locations and the resulting spatial field, calculated considering different neighbors for each sampled location (K), were less strong and their distribution was close to a randomly generated point pattern (range of the p-values for different values of K > 0.05).

Predicting the proportion of veterinarians specialized in food animals

We used the available information about the specialization of veterinarians to investigate the separation between veterinarians specialized in treating food animals and veterinarians specialized in treating companion animals, with the purposes of:

i. Modeling the distribution of the proportion of veterinarians specialized in treating food animals and predicting it in areas where we didn't find information about veterinarians' specialization.

ii. Understanding the effect of the covariates used to predict the global distribution of veterinarians on the proportion of food *vs* companion animals' veterinarians.

Step 1. **Data preparation**. We separated spatial points of veterinarians specialized in treating food animals from the rest of the veterinarians sampled. We found this information, only through governmental agencies and online platforms listing veterinarians by postcode in Mexico, Argentina, Chile, Great Britain, Belgium, Switzerland, Italy, South Africa, and Iran. Next, we aggregated these points in pixels with a resolution of 10x10 km^2 and calculated, for each pixel, the ratio between the count of

food animals' veterinarians and the other veterinarians sampled. Since 80.1% of the ratios computed was between 0 and 1, we used this share of values to predict the proportion of veterinarians specialized in food animals.

Step 2. **Generalized linear model specification**. We used the beta regression model (BRM) (*56*) to predict the proportion of food animals' veterinarians at the 10x10 km^2 resolution and to investigate its relationship with the spatial covariates selected in this study.

Step 3. **Model selection**. First, we used the best subset selection method (*57*) to compute BRMs with every possible combination of covariates. From the batch of BRMs, we extracted just the ones where every covariate was significantly related to the response according to the *z*-test on the regression coefficients (p-value < 0.05). Second, we identified the best BRM as the one with the lowest AIC value and analyzed plots of its residuals, which showed to be normally distributed and evenly spread around zero (fig. S16A and fig. S16B).

Step 4. **Effect of covariates**. The proportion of veterinarians specialized in food animals was related to the spatial covariates used to produce the global map of the distribution of veterinarians. However, their effect on the response was different from their effect on the distribution of all veterinarians. The proportion of food animals' veterinarians was highly related to the human population density and GDP (p-values respectively < 0.01 and < 0.001). However, their regression coefficients showed that these covariates had an inverse relationship with the proportion of veterinarians who specialized in food animals (table S7). In addition, also high values of the proportion of areas used for croplands were inversely related to the response, while high values of chickens and sheep density had a positive effect on it. Finally, the density of cattle was the only covariate that, although significant and typical of agricultural environment, was inversely related to the proportion of veterinarians who specialized in food animals.

Step 5. **Mapping the proportion of veterinarians specialized in food animals**. We used covariates values' and the BRM with the lowest AIC to predict the proportion of food animals' veterinarians in every other 10x10 km^2 pixel where we predicted the presence of veterinarians at the global level.

Step 6. **Comparing veterinarians' distribution with food animals' density and antimicrobial resistance**. The map of the proportion of food *vs* companion animals' veterinarians was multiplied for the map of the distribution of veterinarians to obtain the map of the distribution of food animals' veterinarians. This map was used to identify pixels where food animals outnumber food animals' veterinarians. Therefore, we compared the global distributions of food animals' veterinarians and food animals raised in extensive farm systems. Conversely to intensive farm systems, where food animals are housed in confined spaces that facilitate the regular monitoring of their health, extensive farm systems may involve less frequent contacts with veterinarians, who might rely more on

visual inspections during routine checks or when specific health concerns arise (*58*). For this task, we expressed the 10x10 km^2 resolution density maps of food animals raised extensively (*59*) in livestock units (LSUs), which is a mean to compare different food animals' productions according to the Eurostat glossary (*60*); for the conversion we used the coefficients of 0.014 for chickens, and 0.5 for pigs. Then, we summed the LSUs pixels' values for each species into a single layer, divided it for the global map of the distribution of veterinarians, and produced a 10x10 km^2 resolution global map of the geographic distribution of LSUs of extensively raised food animals per veterinarian.
In addition, we investigated if pixels with high LSUs of extensively raised food animals per veterinarian could be related to the distribution of antimicrobial resistance (AMR) in food animals. Therefore, we calculated the Pearson correlation between the distribution of LSUs of extensively raised food animals per food animals veterinarian and the distribution of AMR in cattle, chickens, and pigs available for LMICs (*61*). However, this analysis returned a Pearson correlation coefficient of 0.25.

Identification of coldspots of veterinary capacity in food animals

In the human population, the "golden hour" is defined as the hour immediately after a traumatic injury during which there is the highest probability that a prompt medical care can maximize chances of survival (*62*). Comparably, we defined a veterinarian coldspot as a pixel of 10x10 km^2 where cattle, chickens, and pigs, which together represent 84.2% of the biomass of animals farmed worldwide (*63*), are farther than 1 hour of motorized travel time (TT^M) from the nearest veterinarian.

For each country, we intersected the GLW4 maps at 1 km^2 resolution with the map of food animals' veterinarians disaggregated at a resolution of 1 km^2. Through this step, we identified pixels with cattle, chickens, and pigs but no veterinarians. We called these pixels "isolated farms" (IF) if they contained, according to the GLW4, at least 1 cattle head, 1 pig, and 10 chickens (since 1 chicken is present almost everywhere in the world). The IF map of cattle in Kenya is reported in fig. S17A; we will refer to this country throughout the manuscript to show small-scale examples of our results. Next, we defined the starting points of veterinarians that travel to IF as the centroids of the pixels with predicted veterinarians and we used a friction surface to calculate the cumulative cost, in hours, to travel from these centroids to every pixel of the map. For this study, we used the 2020 version of the global friction surface at 1 km^2 resolution produced by The Malaria Atlas Project (*64*, *65*) (table S1 and fig. S17B). This friction surface combines areas where it is only possible to walk with an average speed of 5 km/h with areas where the road networks allow for the use of motorized vehicles ("walking + motorized friction surface"). Then, through the R package *gdistance* (*66*, *67*), we combined veterinarians' coordinates, IF, and the friction surface to obtain a TT^M map of veterinarians traveling to every pixel of a country and extracted only IF where TT^M was higher than 1 hour (fig. S18A). We computed these maps separately for each country to exclude potential veterinarians traveling across national borders. Then, we aggregated these country-level maps at 10x10 km^2 resolution and assembled them to obtain global maps of coldspots of veterinary capacity for cattle, chickens, and pigs. The aggregation of coldspots maps from the 1 km^2

resolution (i.e., the resolution of the friction surface) to the 10x10 km^2 resolution caused a 1% area loss for cattle coldspots, 5% for chickens coldspots, and 7% for pigs coldspots.

In addition, we subset the GLW4 maps through the coldspots maps to quantify the number of food animals still living inside coldspots. Hereafter, we will refer to such animals as "isolated animals" (IA). Then, we weighted the TT^M of every coldspot based on its IA value. The metric, calculated for cattle, chickens, and pigs coldspots in each country, was defined as:

$$\overline{TT}_{IA} = \frac{\sum_i^{N_C}(TT_i^M \cdot IA_i)}{\sum_i^{N_C} IA_i}$$

Where *i* represents the *i*-pixel of the coldspots map and N_C the number of pixels with coldspots. We computed this metric with three maximum TT^M thresholds of 6, 4, and 2 hours for each coldspot. The reason was to understand if extreme TT^M values (e.g., $TT^M > 24$ h) could affect the sensitivity of $\overline{TT}_{IA}$.

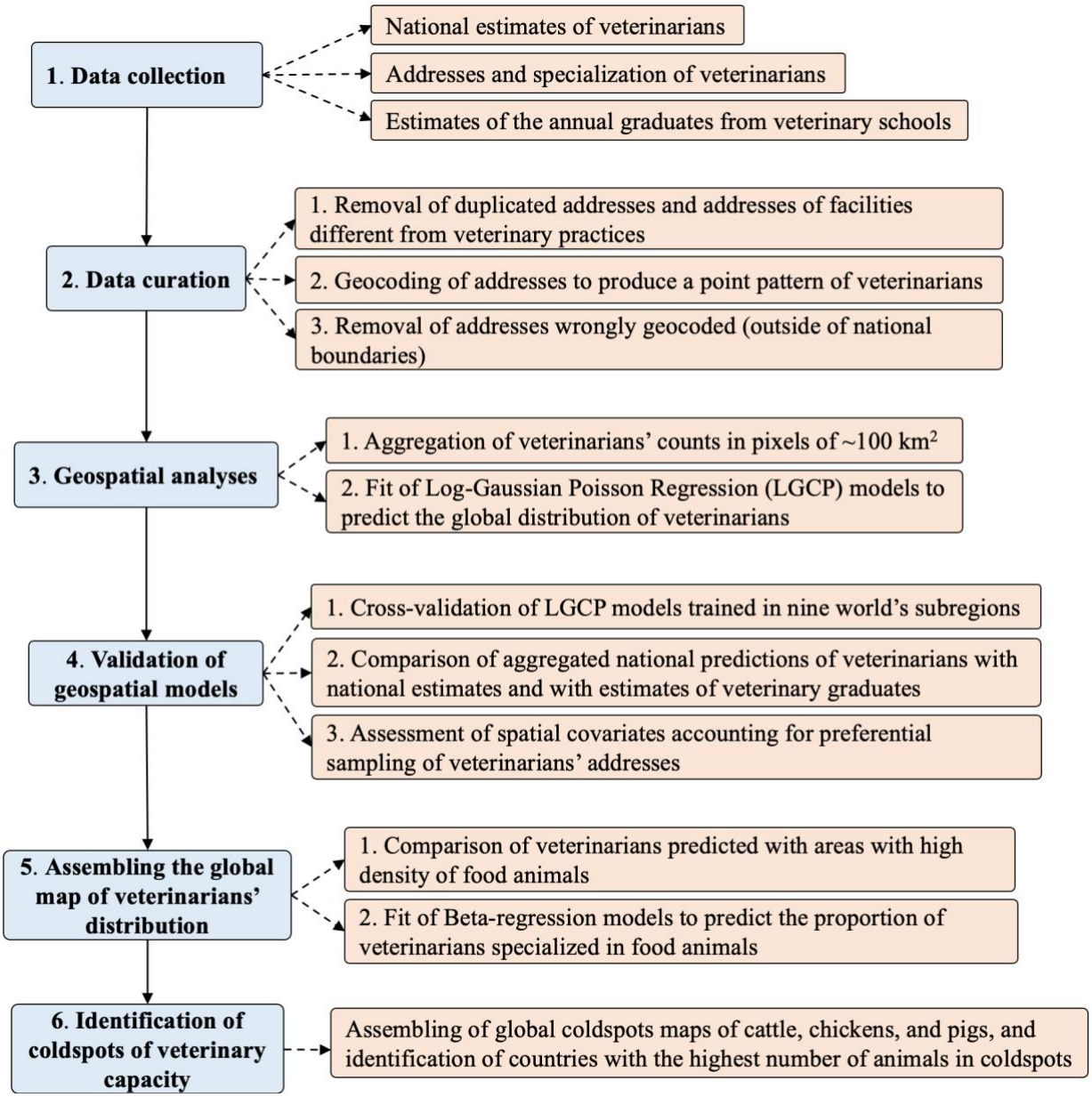

Fig. S1. Workflow used to map the coldspots of veterinary capacity.
Boxes blues report the main steps used to produce the outputs of this study. Boxes in oranges report the type of data collected and the analyses performed in each step. Coldspots were defined as 10x10 km^2 pixels where food animals were farther than 1 hour of motorized travel time from the nearest veterinarian.

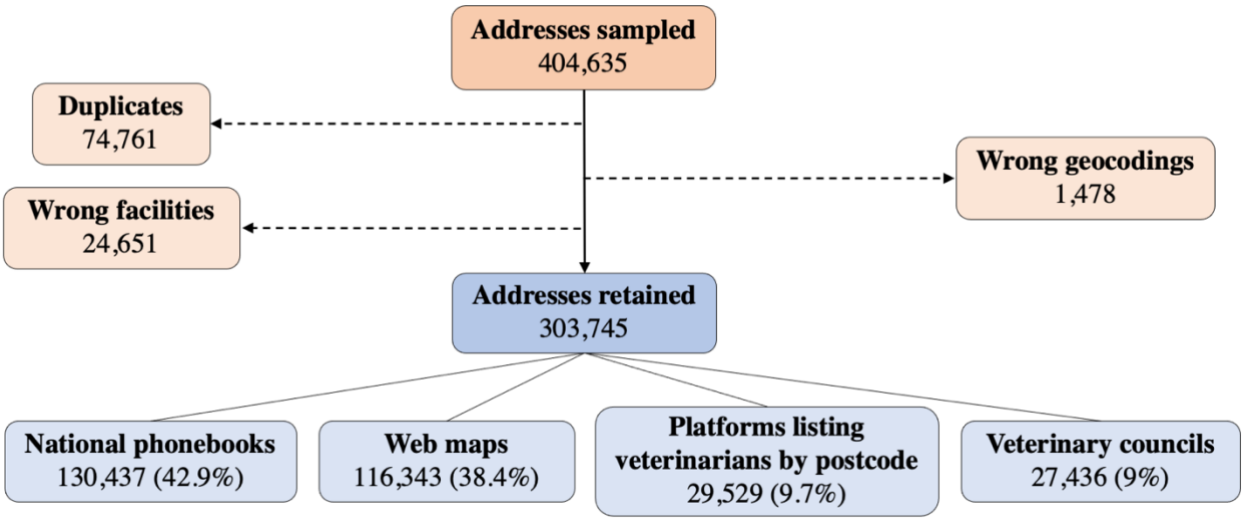

Fig. S2. Curation process for building an address book of veterinarians.
The orange boxes report the original number of addresses of veterinarians sampled from online sources and the number of duplicates, facilities different from veterinary practices, and veterinary practices wrongly geocoded identified in the data curation phase. The blue boxes report the number of addresses of veterinarians retained in the database and group them by the online sources where they were sampled.

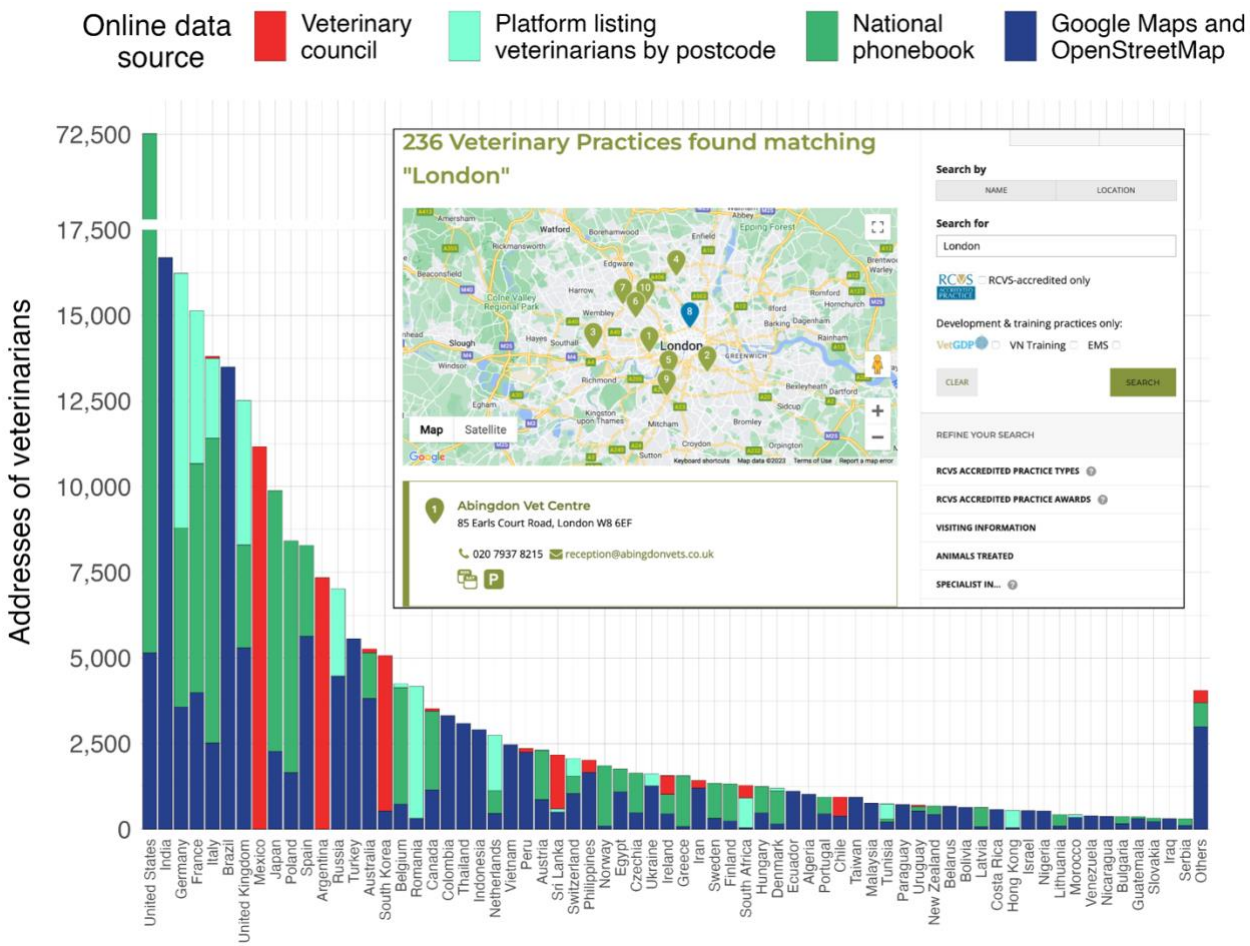

Fig. S3. Number of veterinarians sampled per country grouped by the online sources listing their addresses. We grouped countries with less than 300 sampled addresses of veterinarians in the category "Others". The image above the bars shows the webpage of the platforms of the Royal College of Veterinary Surgeons, in the United Kingdom (*68*), which is one of the online sources identified that allows users to search for the nearest veterinarians according to a postcode/address.

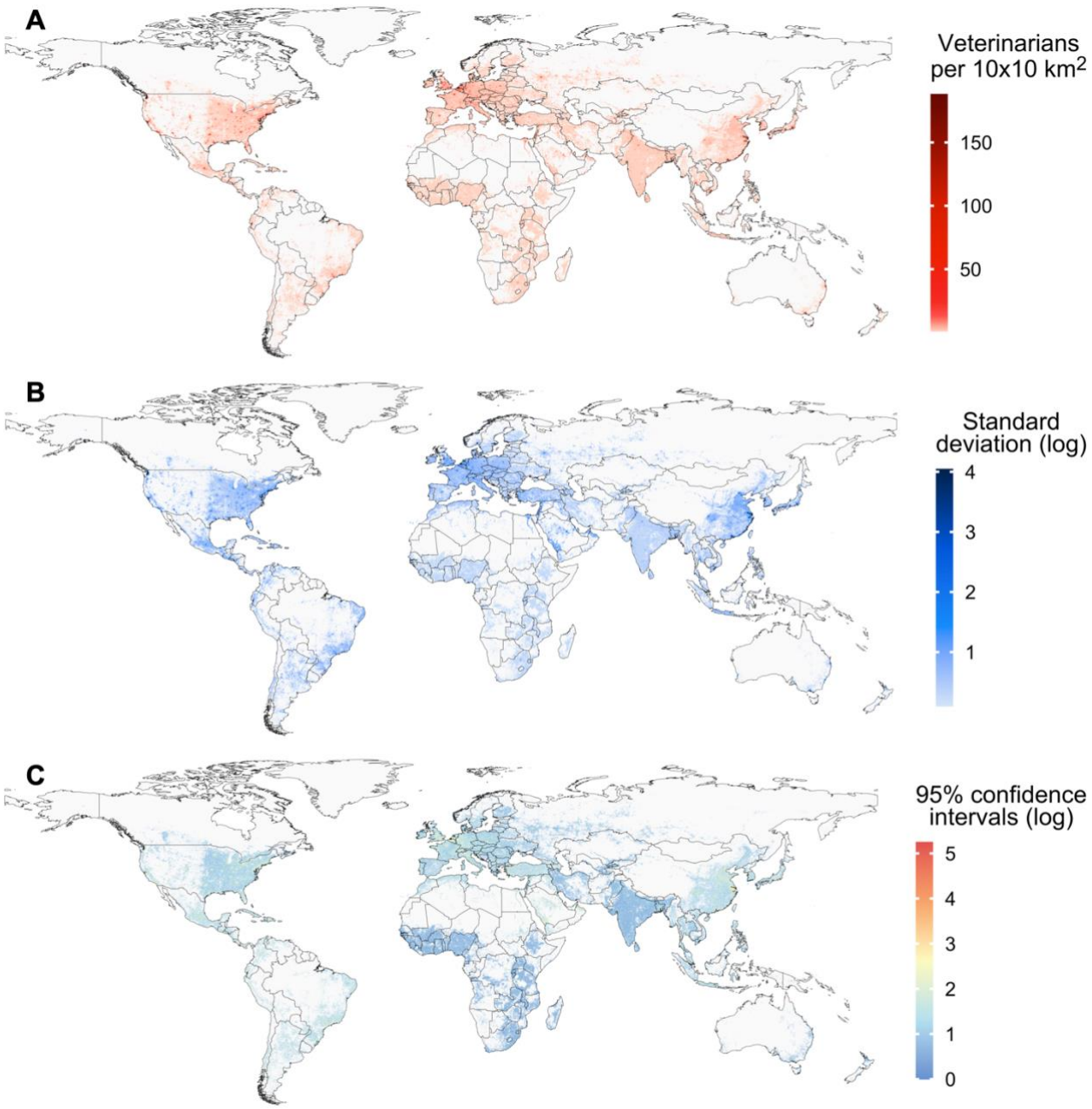

Fig. S4. Global maps of the distribution of veterinarians at the 10x10 km^2 resolution. (A) Global distribution of predicted veterinarians, regardless their specialization, (B) standard deviation of predictions, and (C) 95% confidence intervals of predictions.

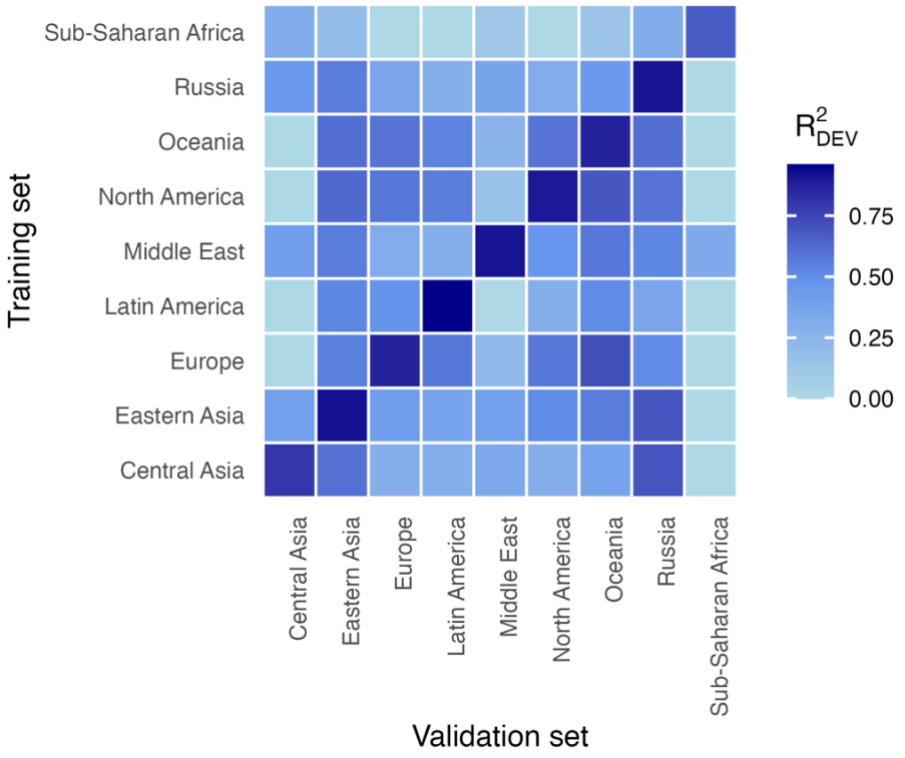

Fig. S5. Deviance R^2 to assess the predictions' accuracy of subregional models.
The matrix reports the values of the Deviance R^2 (R^2_{DEV}) (*46*) computed for the cross-validation of every subregional Log-Gaussian Poisson Regression model. Values on the diagonal represent the accuracy of predictions computed in the same subregion of the data used to train the models, while all other values represent the accuracy of predictions computed in external subregions.

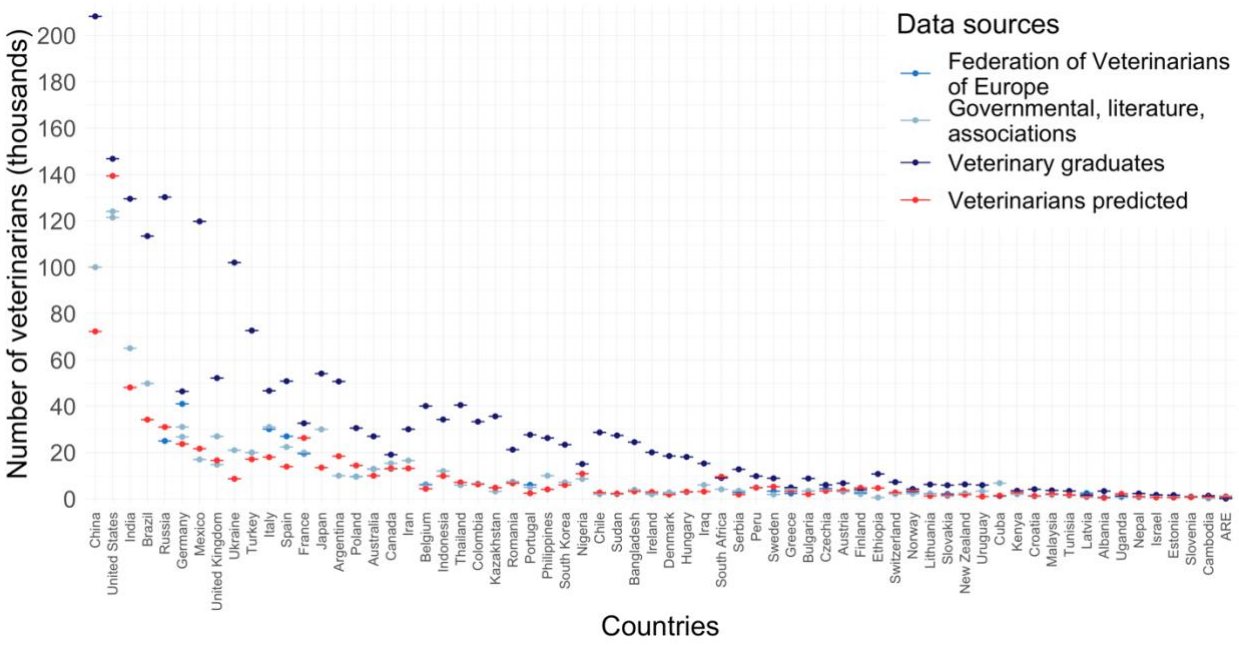

Fig. S6. Predictions and national estimates of veterinarians.
Predicted number of veterinarians aggregated at the country-level (red dots), national estimates of veterinarians from the Federation of Veterinarians of Europe (medium blue dots) and from governmental agencies, literature, and veterinary associations (light blue dots), and the estimated number of graduates from all veterinary schools of a country (dark blue dots). We reported the comparisons for countries with at least 1,000,000 inhabitants and where we retrieved both national estimates and the number of veterinary graduates.

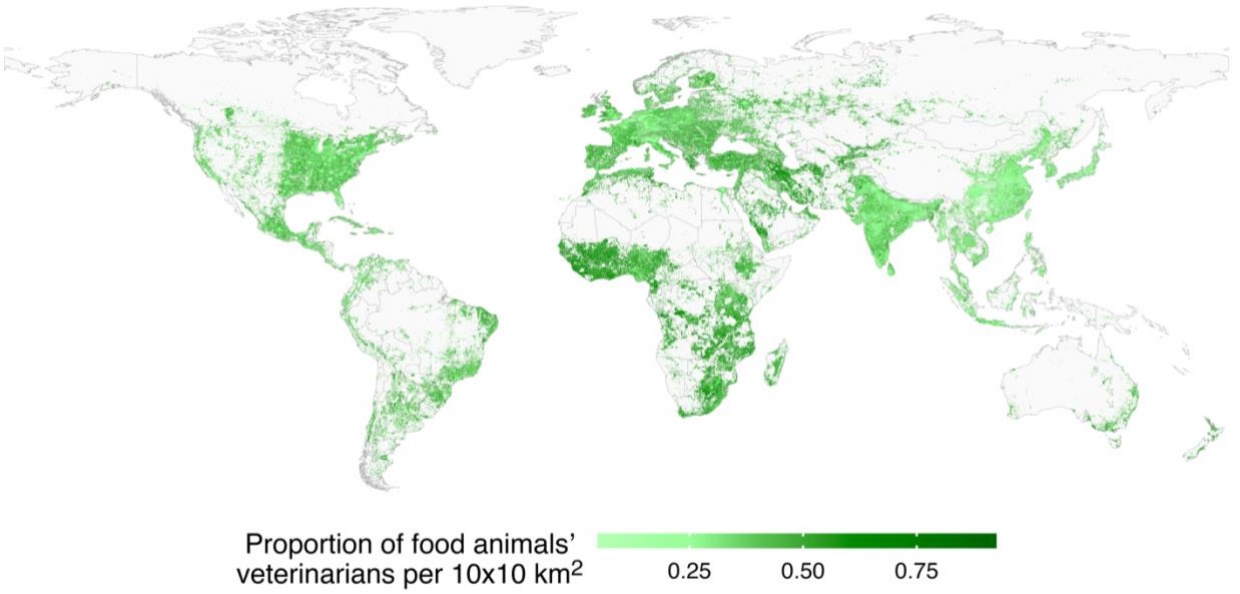

Fig. S7. Proportion of veterinarians specialized in food animals at 10x10 km^2 resolution. The proportion of veterinarians specialized in food animals was predicted through beta regression models in the same pixels where we predicted the global distribution of veterinarians (fig. S4).

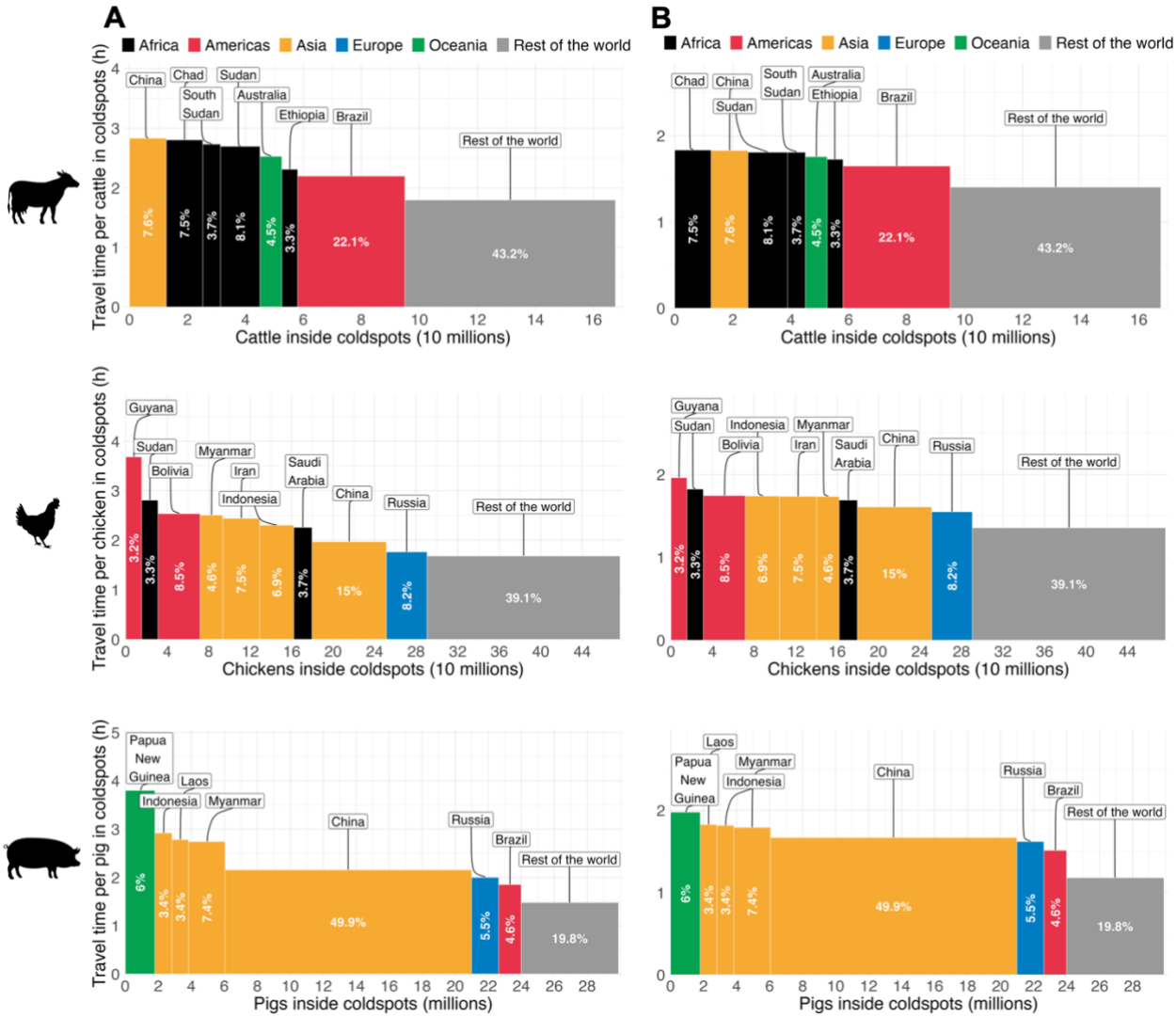

Fig. S8. Country-level average travel time per animal in coldspots based on two maximum travel time thresholds.

The barplots are equivalent to the ones shown in Fig. 2. However, in each computed coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of motorized travel time from the nearest veterinarian, we set a maximum threshold of motorized travel time of 4 hours (A) and 2 hours (B).

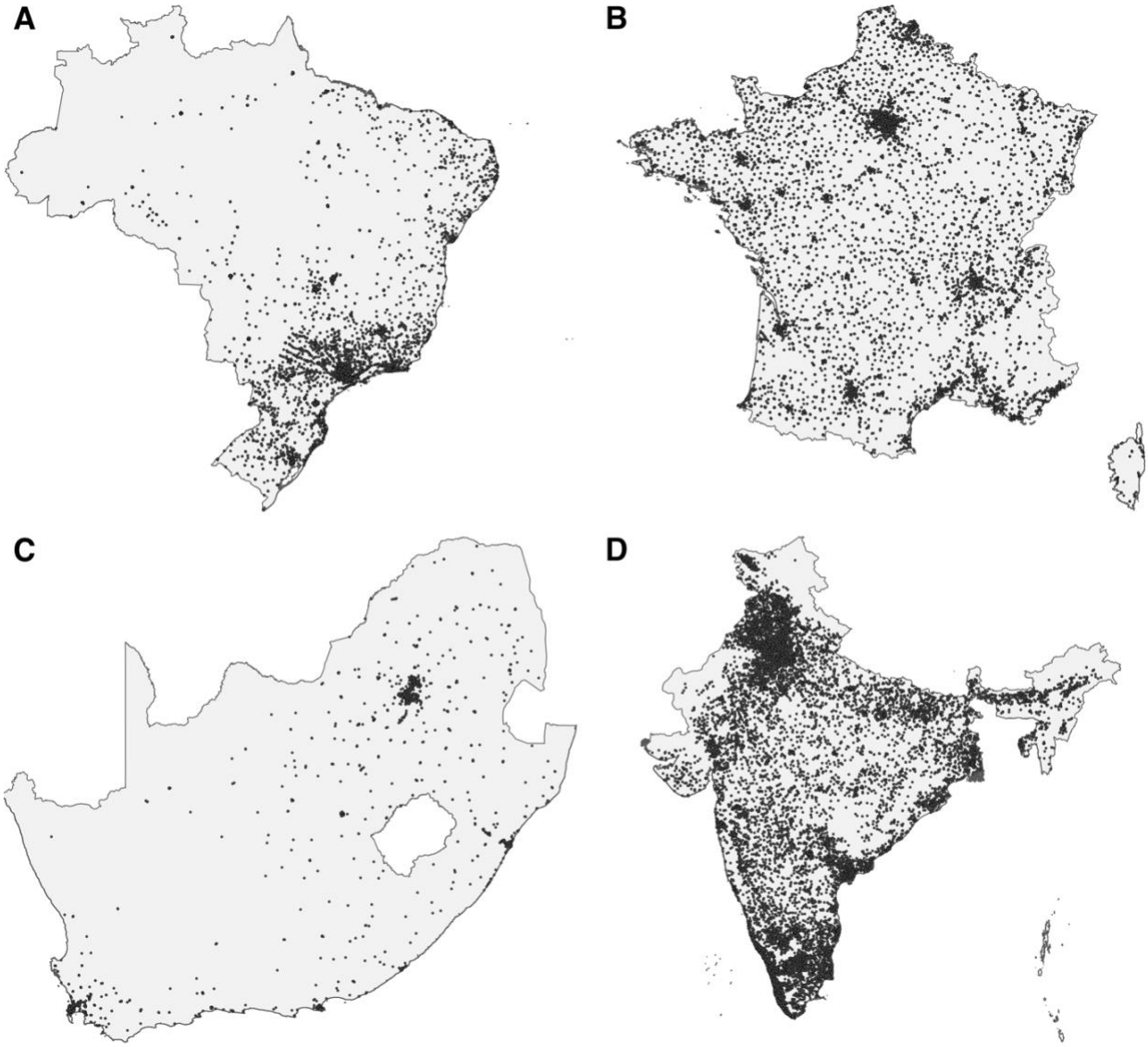

Fig. **S9**. **Geocoded addresses of veterinarians**.
The dots represent the geocoded addresses of veterinary practices (hospitals, clinics, private studies, etc.) in Brazil (A), France (B), South Africa (C), and India (D). We sampled these practices through online platforms listing veterinarians by postcode (*68*), online national phonebooks, websites of veterinary councils, and querying the databases of Google Maps and OpenStreetMap.

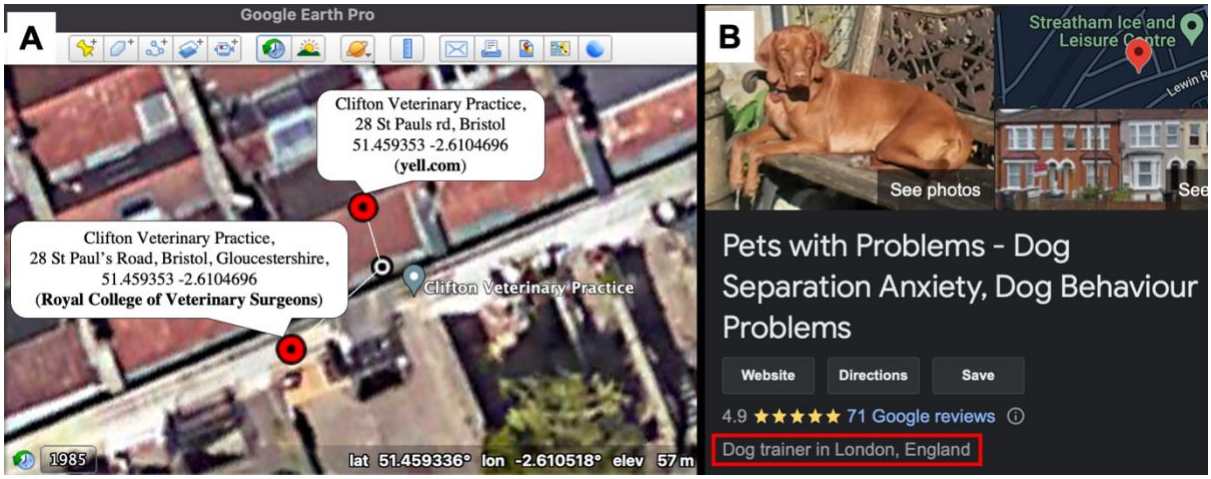

Fig. S10. Identification of duplicates and wrongly sampled addresses of veterinarians.
(A) Google Earth Pro's window shows the "Clifton Veterinary Practice", in England. We sampled information about this veterinary practice from two online sources: the Royal College of Veterinary Surgeons and the British national phonebook. The vignettes contain the name and address of each veterinary practice as reported online, and their geographic coordinates obtained from Google's geocoding API. Once assembled this information in two text strings, we calculated their affinity through the Levenshtein similarity (*18*). Although both strings point to the same veterinary practice, we obtained only 74% of similarity (lower than the 90% threshold we set to remove duplicates), because the sources report names of addresses in two different ways.
(B) Google's search page shows information about a dog training center in England which was sampled from the online national phonebook. If considering only its name as reported online (i.e., "Pets with Problems"), it is trivial to understand that this facility is different from a veterinary practice. Hence, we used web-crawling to perform Google searches using the name and address of sampled data to capture information about the type of service each facility provides (red box) and exclude the ones different from veterinary practices.

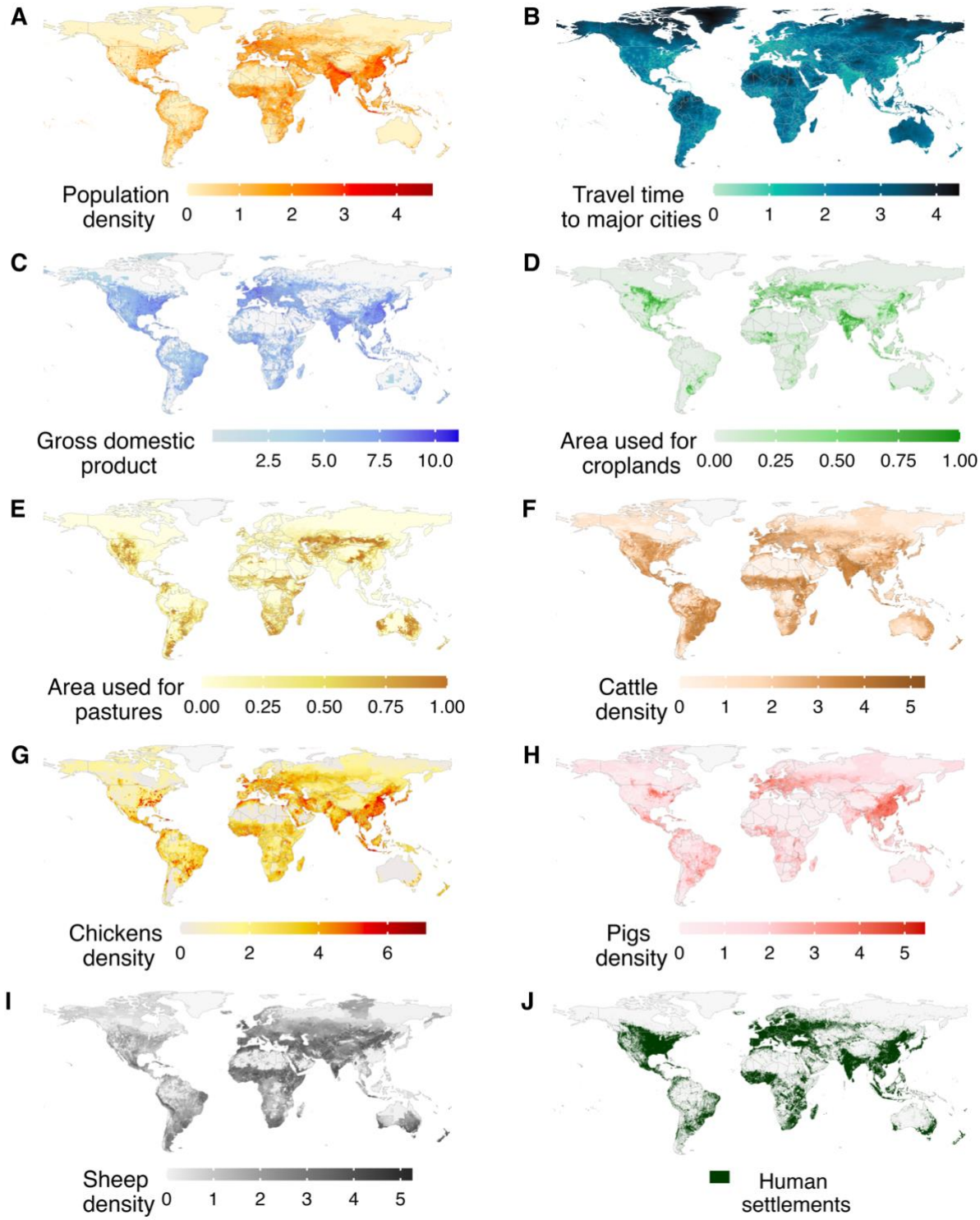

Fig. S11. Global maps of anthropogenic and agricultural covariates at 10x10 km^2 resolution.

Pixels' values of global maps of population density (A), travel time to major cities (B), gross domestic product (C), and density of food animals (F, G, H, I) were Log10-transformed. Pixels' values of the global maps of croplands (D) and pastures (E) represent a proportion. Pixels of the World Settlement Footprint (J) represent areas where there are human settlements.

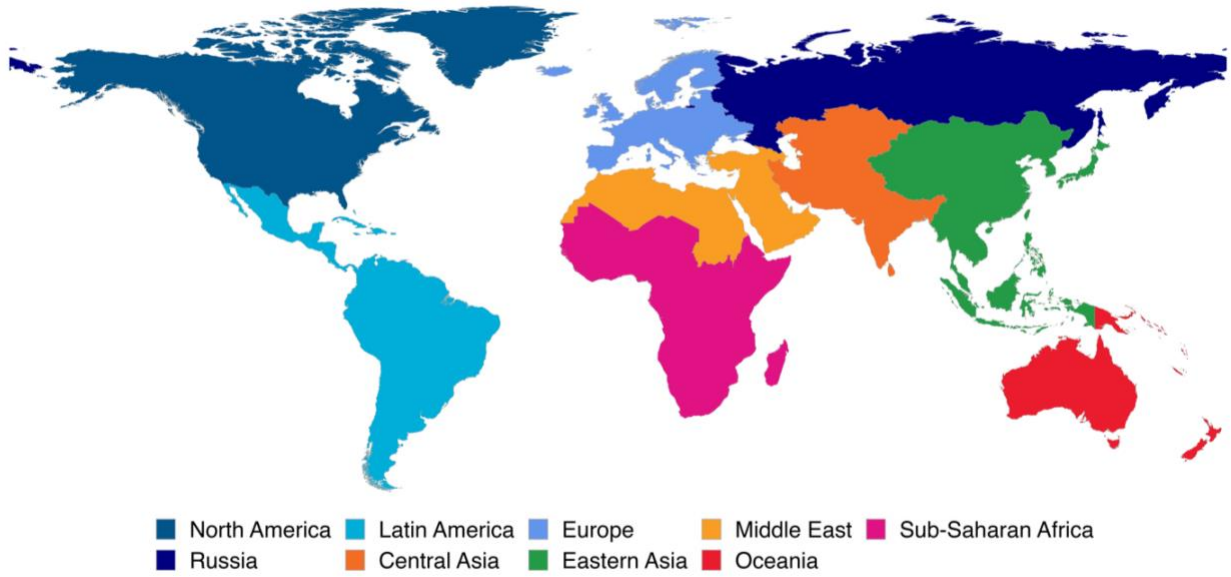

Fig. S12. World subregions used to train Log-Gaussian Poisson Regression models and to validate predictions of veterinarians' distribution.
The division into subregions, and the colors used to represent them, refers to the scheme that the United Nations used in their official reports (*37*). However, in our study, we further divided the subregion of "Europe and North America" into the subregions of North America, Europe, and Russia.

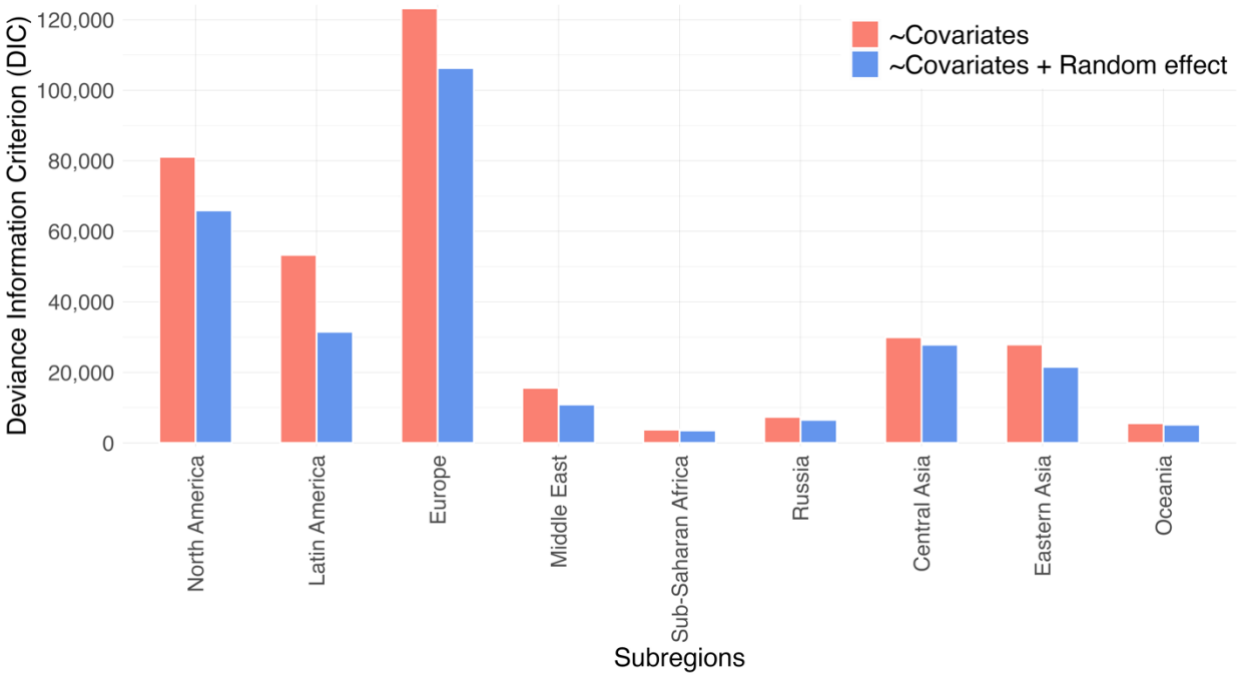

Fig. S13. Comparison between Deviance Information Criterion (DIC) values.
The DIC values derive from the subregional Log-Gaussian Poisson Regression models that returned the highest predictions' accuracy of the counts of veterinarians aggregated in 10x10 km^2 pixels (fig. S5). For each model, the DIC was computed with and without the random effect used to capture the spatial autocorrelation of the counts of veterinarians.

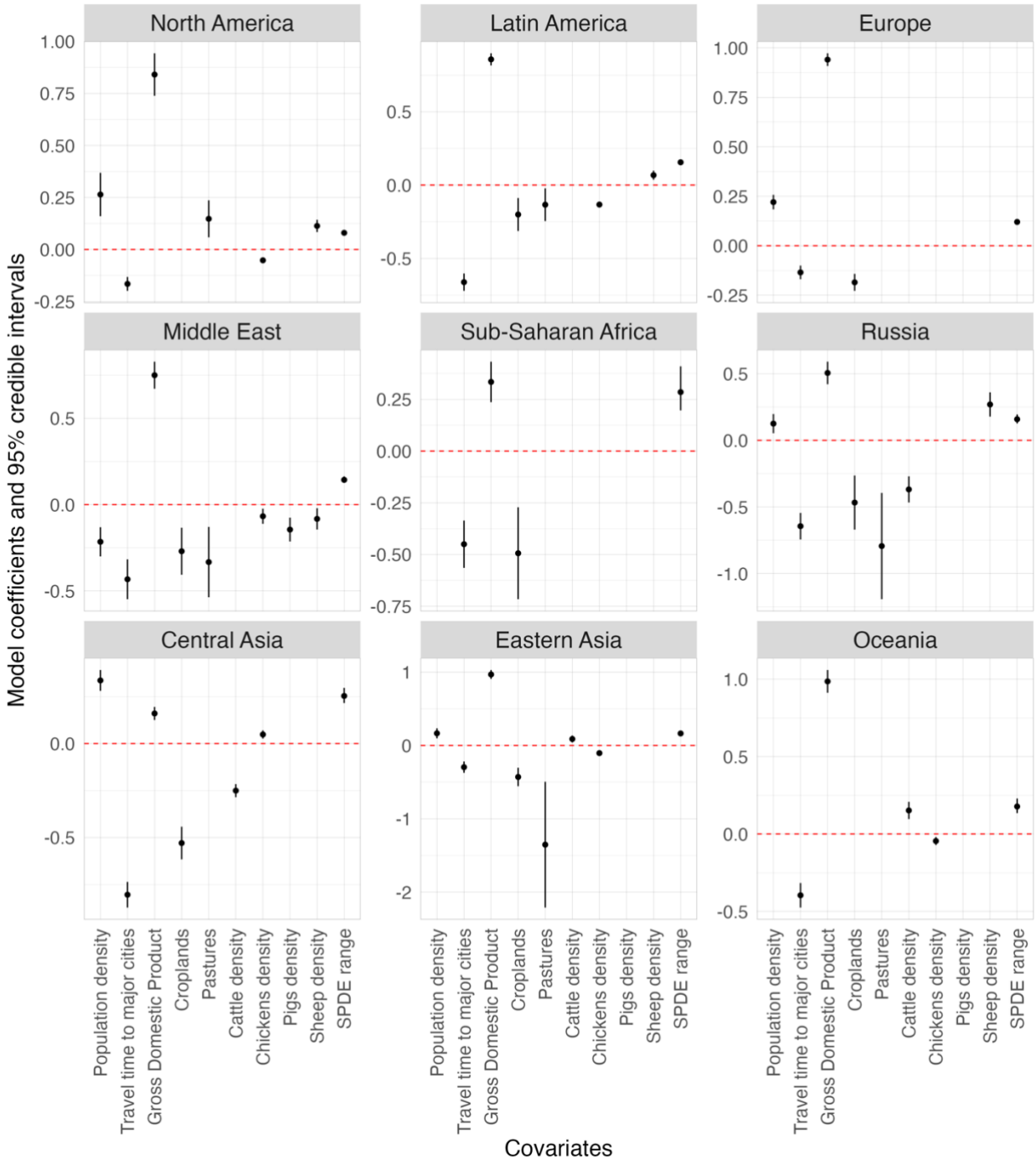

Fig. S14. Posterior means of covariates, range of random effect, and 95% credible intervals.

Posterior means and the 95% credible intervals of the spatial covariates, and the range of the random effect (for spatial autocorrelation), computed through each subregional Log-Gaussian Poisson Regression model that returned the highest predictions' accuracy of the counts of veterinarians aggregated in 10x10 km^2 pixels (fig. S5). The red dashed line at 0 shows that the credible interval of each covariate doesn't include zero and hence they are significant in predicting the distribution of veterinarians.

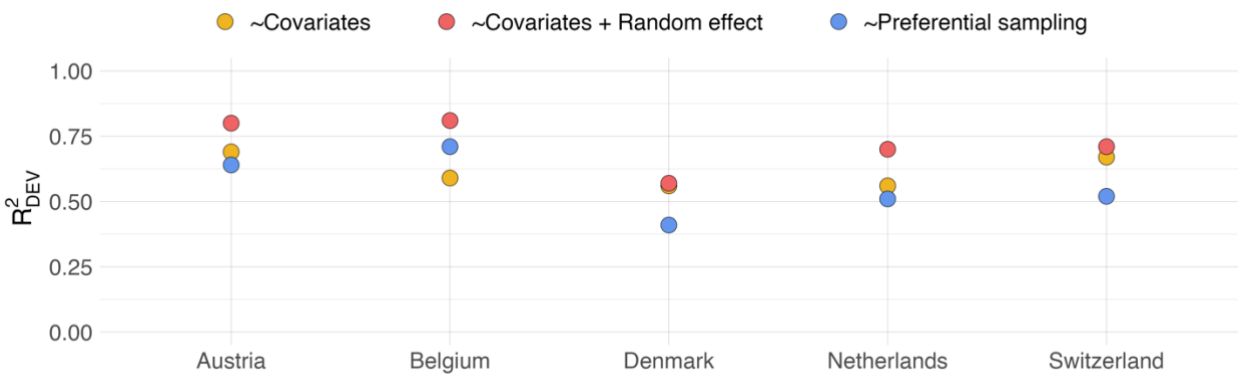

Fig. S15. Comparison of predictions' accuracy among geospatial models.
Values of the Deviance R^2 (R^2_{DEV}) (*46*) to assess the predictions' accuracy of counts of veterinarians aggregated in 10x10 km^2 pixels. The R^2_{DEV} was computed through a Log-Gaussian Poisson Regression model fitted only with spatial covariates (yellow), a model including covariates and the random effect to capture spatial autocorrelation (red), and a Log-Gaussian Cox Process model accounting for preferential sampling (blue) (*50*, *51*). The models were applied in five countries where we sampled addresses of veterinarians.

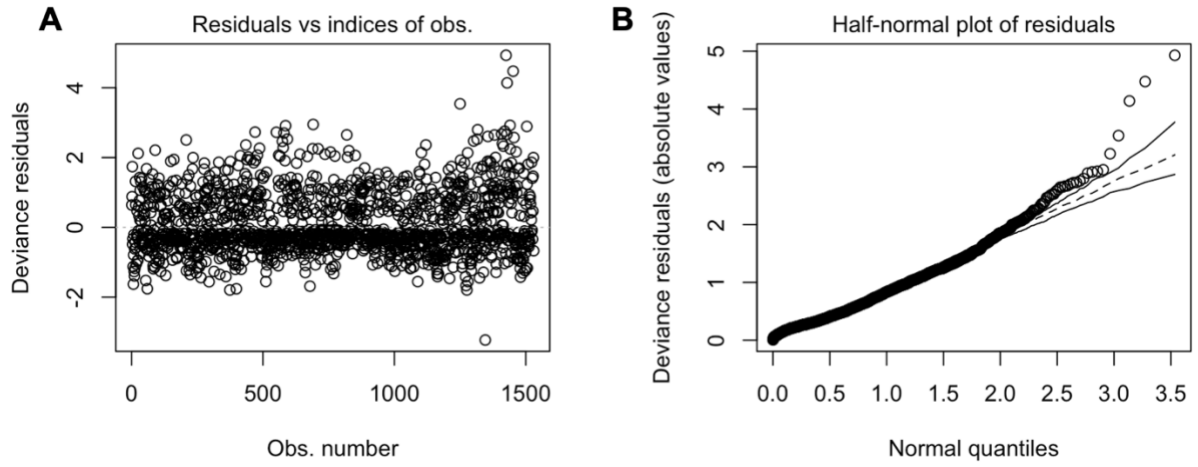

Fig. **S16**. **Diagnostics plots of the beta regression's residuals**.
From the beta regression model used to predict the proportion of veterinarians specialized in food animals we extracted the deviance residuals (*69*) of each fitted observation to inspect (A) their values and (B) compare their distribution with a theoretical normal distribution.

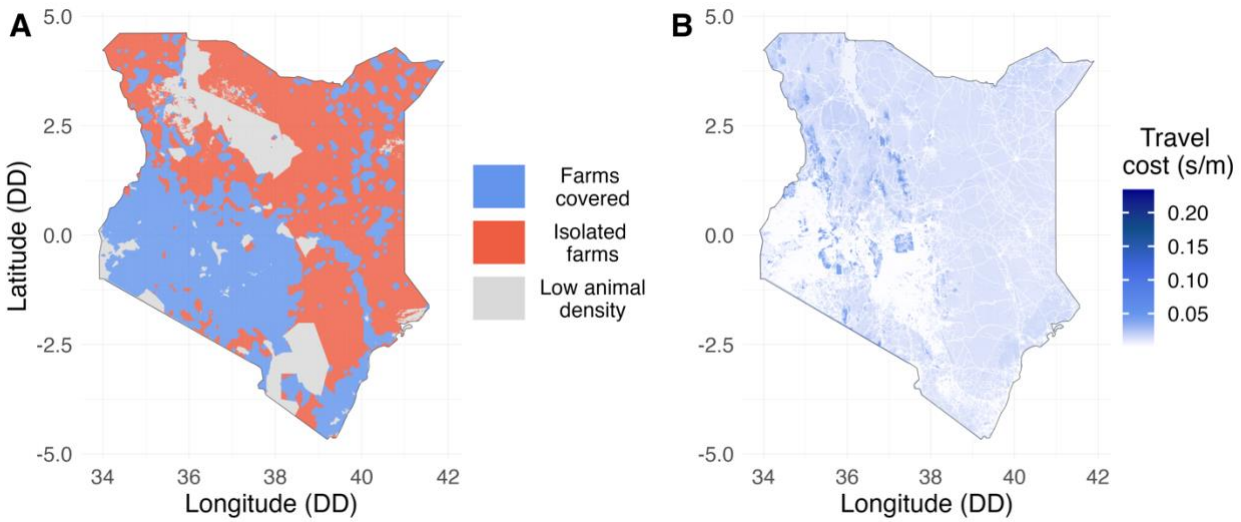

Fig. S17. Cattle's isolated farms and the friction surface of Kenya.
(A) Isolated farms represent pixels where we didn't predict veterinarians but where the density of cattle and pigs is higher than 1 per km^2 and the density of chickens is higher than 10 per km^2. Pixels in light blue represent areas where food animals are within 1 hour of motorized travel time from a veterinarian, while pixels in grey represent areas where the density of food animals is less than 1 per km^2. (B) The friction surface used in this study is a 1 km^2 resolution map that reports, for each pixel, the seconds to cross 1 meter through motorized vehicles (*65*).

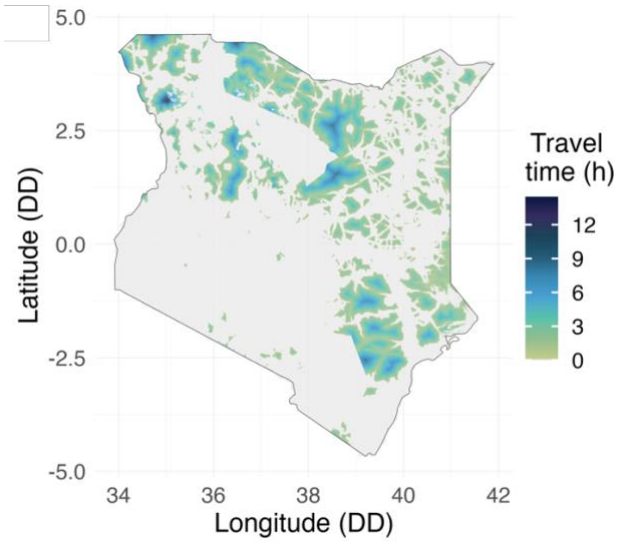

Fig. **S18**. **Coldspots map of cattle in Kenya**.

Coldspots of veterinary capacity were defined as 10x10 km^2 pixels where food animals were farther than 1 hour of motorized travel time from the nearest veterinarian. The map report coldspots of cattle identified in Kenya. Pixels where motorized travel time is less than 1 hour are represented in grey.

Table S1. **Anthropogenic and agricultural covariates used to predict the distribution of veterinarians and map veterinary coldspots**. Veterinary coldspots were defined as 10x10 km^2 pixels where food animals were farther than 1 hour of motorized travel time from the nearest veterinarian.

Name	Original resolution	Year	Unit	Reference
Population density	0.083333 decimal degrees	2015	Log10[(inh./10x10 km^2)+1]	Center for International Earth Science Information Network (CIESIN) (*70*)
Travel time to major cities	0.083333 decimal degrees	2015	Log10[(minutes)+1]	Weiss et al., 2018 (*71*)
Gross domestic product	0.083333 decimal degrees	2015	Log10[(US $)+1]	Kummu et al., 2018 (*72*)
Area used for croplands	0.083333 decimal degrees	2000	Proportion	Ramankutty et al., 2008 (*73*)
Area used for pastures	0.083333 decimal degrees	2000	Proportion	Ramankutty et al., 2008 (*73*)
Population density of cattle, chickens, pigs, and sheep	0.083333 decimal degrees	2015	Log10[(animals/10x10 km^2)+1]	Gilbert et al., 2018 (*74*, *75*)
World Settlement Footprint	0.083333 decimal degrees	2015	Proportion	Marconcini et al., 2020 (*28*)
Motorized friction surface	0.0083333 decimal degrees	2020	seconds/meter	The Malaria Atlas Project, 2020 (*65*)

Table S2. **DIC decrease after the addition of the significant spatial covariates**.
For each subregional model with the highest accuracy in predicting the distribution of veterinarians (fig. S5), the table reports the proportional decrease of the Deviance Information Criterion obtained by including each significant spatial covariate (i.e., whose posterior mean was included in their 95% credible intervals) to the models with only the intercept.

	North America	**Europe**	**Russia**	**Oceania**	**Latin America**	**Middle East**	**Sub-Saharan Africa**	**Central Asia**	**Eastern Asia**
Population density	88.87%	78.73%	94.01%			89.71%		86.49%	90.43%
Travel time to major cities	10.92%	19.35%	4.04%	93.84%	98.67%	8.06%	95.79%	10.11%	8.2%
Gross domestic product	0.02%	1.79%	1.12%	5.82%	1.12%	1.17%	3.01%	0.82%	1.02%
Area used for croplands		0.14%	0.35%		0.04%	0.12%	1.2%	1.57%	0.25%
Area used for pastures	0.04%		0.09%		0.01%	0.18%			0.05%
Cattle density			0.15%	0.21%				0.94%	0.01%
Chickens' density	0.01%			0.12%	0.16%	0.29%		0.08%	0.05%
Pigs' density						0.07%			
Sheep density	0.14%		0.23%		0.02%	0.4%			

Table S3. **Websites of veterinarians' addresses and national-level data**.
The table reports names, URLs, and references of online data sources. The column "Websites with addresses" reports the online data sources used for sampling addresses of veterinarians (excluding OpenStreetMap and Google Maps). The columns "Associations", "Governmental agencies", and "Literature, newspapers", report the sources used to identify national estimates of veterinarians.

ISO3	Websites with addresses	Associations	Governmental agencies	Literature, newspapers
AFG				AVMA
ALB	Bujqesia	FVE		
AND	Andorra Telecom			
ARE				Gulf News
ARG	SENASA		FAO	
AUT	Herold	FVE, EBVS		
AUS	Yellow Pages VET VOICE	AVA (*76*)		PetKeen
BEL	Trouver un vétérinaire	FVE, EBVS		
	Pages d'or			
BGD			Bangladesh Veterinary Board	
BGR	Golden Pages	FVE, EBVS		
BRA				Germiniani, 2004 (*77*)
CAN	canada411	CVMA		Whiting, 2021 (*78*)
	VCA animal hospitals			
CHE	Société des Vétérinaires Suisses Local.ch	FVE, EBVS		
CHL	Servicio Agrícola y Ganadero		USDA	
CHN				The Guardian
COL				El Tiempo
CUB	Páginas Amarillas	AVMA		
CYP	Yellow Pages			
CZE	Zlaté Stránky	FVE, EBVS		
DEU	GoYellow	FVE, EBVS		Vets Online
	TIERARTZ			
DNK	Netdyredoktor De Gule Sider	FVE, EBVS, DVA		
EGY	Yellow Pages			Arabmedicare
ESP	Páginas Amarillas	FVE, EBVS		
EST	Yellow Pages	FVE, EBVS		
ETH				Mayen, 2006 (*79*)
FIN	finder.fi	FVE, EBVS	University of Helsinki	
FRA	Ordre National des vétérinaires Pages Jaunes	FVE, EBVS		
GBR	Royal College of Veterinary Surgeons	FVE, EBVS		Statista
	Yell			
GRC	VRISKO	FVE, EBVS	University of Thessaly	
	Greek Yellow Pages			
GMB		WTG		
GTM	Paginas Amarillas			
GUY	Veterinary Board of Guyana			
HKG	VSBHK	HKVA		
HRV	Yellow Pages	FVE, EBVS		
HUN	Arany Oldalak	FVE, EBVS		
IDN			Universitas Gadjah Mada	
IRL	Golden Pages	FVE, EBVS		That's farming
	Veterinary Council			
IRN	Ministry of Agriculture			Hakimemehr
IRQ				Khamas, 2004 (*80*)
ISL	Já	FVE		
IND		IVA		
ISR		IVMA		
ITA	FNOVI	FVE, EBVS		
	Pagine Bianche			

	Pagine Gialle			
	Federazione Europea Ornitologi			
JAM	Jamaica Veterinary Board	JVMA		
JPN	Business directory			Pet Hospital
KAZ		KVA		
KEN				Otieno et al., 2016 (*81*)
KHM			FAO	
KOR	Korean Veterinary Medical Association	KVMA		
LKA	Sri Lanka Veterinary Association			
	Sri Lanka Veterinary Council			
LIE	Yellow Pages		Federal Veterinary Office	
	Welcome.li			
LSO				Boipabolo Temong
LTU	info.lt	FVE, EBVS		
LUX	yellow.lu	FVE, EBVS		
LVA	Viss	FVE, EBVS		
MAR	veterinaire.ma			
MDA	Yellow Pages			
MEX	DENUE			Statista
MKD	Food and Veterinary Agency	FVE		
MLT	Yellow			
MMR				Statista
MYS			Federal Legislation Portal	Statista
NAM	Veterinary Association of Namibia			
NPL				Wikipedia
NGA				Daily Trust
NLD	Dierenarts kliniek	FVE, EBVS		
	Gouden Gids			
NOR	Gules Ider	FVE, EBVS		
NZL	Yellow Pages		New Zealand Government	
PER	petID		ESAN	
PHL	Bureau of animal industry	PVMA		
POL	Yellow Pages	FVE, EBVS		
	Panorama Firms			
PRT	Página Amarelas	FVE, EBVS		
QAT	Yellow Pages Qatar			
ROU	Colegiul Medicilor Veterinari	FVE, EBVS		
RUS	Veterinarka	FVE		
RWA				allAfrica
SDN			Sudan Veterinary Council	
SGP	Singapore Veterinary Association			CN Asia
SOM				Catley et al., 1997 (*82*)
SRB	11811	FVE, EBVS		
SVK	Zlaté Stránky	FVE, EBVS		
SVN	Rumenstrani	FVE, EBVS		
SWE	Eniro	FVE, EBVS		
THA		WSAVA		
TLS		WOAH		
TUN	veterinaire.tn			Leaders
	Pages Jaunes			
TUR		WVA		
TWN				Wenxuan et al., 2021 (*83*)
UGA		WTG		
UKR	UA.REGION.INFO		Encyclopedia of Modern Ukraine	
URY	Páginas Amarellas			Laport et al., 2017 (*84*)
	NATIVACABAL			
USA	Yellow Pages	AVMA		VIN News
ZAF	South African Veterinary Association			Farmers Weekly
	South African Veterinary Council			

Table S4. **Variance Inflation Factor (VIF) of spatial covariates used to predict the distribution of veterinarians**.

	Population density	**GDP**	**Travel time to major cities**	**Area used for croplands**	**Area used for pastures**	**Cattle density**	**Chickens density**	**Pigs density**	**Sheep density**
VIF	3.22	3.08	2.9	1.46	1.36	2.11	2.48	1.72	1.69

Table S5. **Posterior means of the covariates' coefficients and range of the random effect**. For each subregion, the table reports the posterior means and standard deviation (in brackets) of each spatial covariate selected by each subregional Log-Gaussian Poisson Regression model. The last row reports the range and standard deviation of the random effect used to capture the spatial autocorrelation of the observed data.

	North America	**Europe**	**Russia**	**Oceania**	**Latin America**	**Middle East**	**Sub-Saharan Africa**	**Central Asia**	**Eastern Asia**
Intercept	-1.351	-1.971	2.588	-1.919	0.206	1.244	3.364	5.246	-2.177
	(0.346)	(0.125)	(0.424)	(0.374)	(0.211)	(0.394)	(0.47)	(0.206)	(0.31)
Population	0.264	0.22	0.125			-0.215		0.336	0.165
density	(0.053)	(0.019)	(0.037)			(0.043)		(0.028)	(0.035)
Travel time	-0.165	-0.135	-0.645	-0.396	-0.662	-0.433	-0.45	-0.805	-0.296
To major cities	(0.017)	(0.018)	(0.051)	(0.04)	(0.03)	(0.059)	(0.058)	(0.035)	(0.039)
Gross domestic	0.841	0.94	0.506	0.986	0.857	0.75	0.334	0.16	0.968
product	(0.052)	(0.016)	(0.043)	(0.037)	(0.021)	(0.04)	(0.05)	(0.018)	(0.031)
Area used for		-0.185	-0.467		-0.201	-0.27	-0.494	-0.53	-0.43
croplands		(0.022)	(0.104)		(0.057)	(0.07)	(0.113)	(0.044)	(0.064)
Area used for	0.147		-0.794		-0.134	-0.333			-1.353
pastures	(0.045)		(0.204)		(0.057)	(0.104)			(0.438)
Cattle density			-0.368	0.152				-0.251	0.088
			(0.05)	(0.028)				(0.018)	(0.024)
Chickens' density	-0.052			-0.045	-0.133	-0.133		0.048	-0.105
	(0.006)			(0.013)	(0.008)	(0.008)		(0.011)	(0.015)
Pigs' density						-0.145			
						(0.035)			
Sheep density	0.113		0.269		0.068	0.068			
	(0.015)		(0.046)		(0.016)	(0.016)			
SPDE model	0.08	0.12	0.158	0.178	0.156	0.143	0.285	0.254	0.164
range	(0.004)	(0.005)	(0.017)	(0.024)	(0.009)	(0.009)	(0.054)	(0.02)	(0.011)

Table S6. **Results of the** "**Fast Monte Carlo test for preferential sampling**".
The table reports the ranges of the *p*-values of each preferential sampling test (*85*) performed on five countries to inspect the relationship between the sampling pattern and the spatial autocorrelation of data.

	Austria Switzerland	Belgium	Denmark	Netherlands	
Intercept	<0.05	<0.05	<0.05	<0.05	<0.05
Population density	0.1 – 0.7	0.2 – 0.85	0.5 – 0.95	0.15 – 0.4	0.9 – 1

Table S7. **Coefficients, standard error, and *z*-statistics significance of each spatial covariate used in the beta regression model to predict the proportion of veterinarians specialized in food animals**.

Covariate	**Estimate**	**Std. Error**	***p*-value (*z*-statistics)**
Population density	-0.14873	0.04672	0.001455
Gross domestic product	-0.46963	0.05700	$<2.0^{-16}$
Area used for croplands	-0.27617	0.08885	0.001883
Cattle density	-0.13423	0.04140	0.001187
Chickens' density	0.08807	0.01752	5.0^{-7}
Sheep density	0.11030	0.02878	0.000127

References

1. Federation of Veterinarians of Europe. Survey of the veterinary profession in Europe (2019). Available at: https://fve.org/cms/wp-content/uploads/FVE_Survey_2018_WEB.pdf. Accessed: 01/04/2020.
2. World Organisation for Animal Health. Veterinary Educational Establishments in OiE Member Countries (2013). Available at: https://www.woah.org/fileadmin/Home/eng/Support_to_OIE_Members/vee/en_vee_list.php. Accessed: 01/05/2022.
3. Wikipedia. List of schools of veterinary medicine (2023). Available at: https://en.wikipedia.org/wiki/List_of_schools_of_veterinary_medicine. Accessed: 01/05/2022.
4. Wikipedia. Retirement age (2023). Available at: https://en.wikipedia.org/wiki/Retirement_age. Accessed: 01/05/2023.
5. Wikipedia. List of yellow pages (2023). Available at: https://en.wikipedia.org/wiki/List_of_yellow_pages. Accessed: 15/09/2020.
6. OpenStreetMap Foundation. OpenStreetMap (2023). Available at: https://www.openstreetmap.org. Accessed: 01/12/2021.
7. M. Padgham, R. Lovelace, M. Salmon, B. Rudis. Osmdata. *The Journal of Open Source Software* **2**, 305 (2017). doi:10.21105/joss.00305
8. OpenStreetMap Foundation. OpenStreetMap glossary - Tag:amenity=veterinary (2023). Available at: https://wiki.openstreetmap.org/wiki/Tag:amenity%3Dveterinary. Accessed: 01/12/2021.
9. D. Cooley. googleway: Accesses Google Maps APIs to Retrieve Data and Plot Maps (2022). Available at: https://cran.r-project.org/package=googleway. Accessed: 15/01/2021.
10. opnedatasoft. Geonames - All Cities with a population higher than 1000 (2023). Available at: https://public.opendatasoft.com/explore/dataset/geonames-all-cities-with-a-population-1000/table/?disjunctive.cou_name_en&sort=name. Accessed: 01/03/2021.
11. Google Maps Platform. Place Types (2023). Available at: https://developers.google.com/maps/documentation/places/web-service/supported_types. Accessed: 01/03/2021.
12. T. L. Jeeper. tabulizer: Bindings for Tabula PDF Table Extractor Library. (2022).
13. M. A. Khder. Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing and its Applications* **13**, 144–168 (2021). doi:10.15849/ijasca.211128.11
14. H. Wickham. rvest: Easily Harvest (Scrape) Web Pages. (2022).
15. Richardson, L. Beautiful Soup Documentation (2007). Available at: https://beautiful-soup-4.readthedocs.io/en/latest/. Accessed: 15/01/2021.
16. Software Freedom Conservacy. Selenium WebDriver (2023). Available at: https://www.selenium.dev/documentation/. Accessed: 15/05/2021.
17. D. Kahle, H. Wickham. ggmap: Spatial visualization with ggplot2. *The R Journal* **5**, 144–161 (2013). doi:10.32614/rj-2013-014

18. V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR* pp 845–848. (1965). doi:10.1016/S0074-7742(08)60036-7
19. M. Edmondson. googleLanguageR: Call Google's 'Natural Language' API, 'Cloud Translation' API, 'Cloud Speech' API and 'Cloud Text-to-Speech' API. (2020).
20. J. M. Richards. An Ecological Analysis of the Geographic Distribution of Veterinarians in the United States. *Journal of Vocational Behavior* **11**, 216–231 (1977). doi:10.1016/0001-8791(77)90008-2
21. M. R. Olfert, M. Jelinski, D. Zikos, J. Campbell. Human capital drift up the urban hierarchy: Veterinarians in Western Canada. *Annals of Regional Science* **49**, 551–570 (2012). doi:10.1007/s00168-011-0448-2
22. S. Truchet, N. Mauhe, M. Herve. Veterinarian shortage areas: what determines the location of new graduates? *Review of Agricultural, Food and Environmental Studies* **98**, 255–282 (2017). doi:10.1007/s41130-018-0066-9
23. B. Naimi, N. A. S. Hamm, T. A. Groen, A. K. Skidmore, A. G. Toxopeus. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**, 191–203 (2014). doi:10.1111/j.1600-0587.2013.00205.x
24. A. Field, J. Miles, Z. Field. Discovering statistics using R. *SAGE Publications Ltd* (2012).
25. WorldPop methods - Mapping Settlements. Available at: https://www.worldpop.org/methods/. Accessed: 01/12/2021. (2023).
26. F. J. Reed, A. E. Gaughan, F. R. Stevens, G. Yetman, A. Sorichetta, A. J. Tatem. Gridded population maps informed by different built settlement products. *Data* **3**, (2018). doi:10.3390/data3030033
27. F. R. Stevens, A. E. Gaughan, J. J. Nieves, A. King, A. Sorichetta, C. Linard, A. J. Tatem. Comparisons of two global built area land cover datasets in methods to disaggregate human population in eleven countries from the global South. *International Journal of Digital Earth* **13**, 78–100 (2020). doi:10.1080/17538947.2019.1633424
28. M. Marconcini, A. Metz-Marconcini, S. Üreyen, D. Palacios-Lopez, W. Hanke, F. Bachofer, J. Zeidler, T. Esch, N. Gorelick, A. Kakarla, et al. Outlining where humans live, the World Settlement Footprint 2015. *Scientific Data* **7**, 1–14 (2020). doi:10.1038/s41597-020-00580-5
29. F. Lindgren, H. Rue. An explicit link between Gaussian fields and Gaussian Markov random fields : the stochastic. 423–498 (2011).
30. D. L. Miller, R. Glennie, A. E. Seaton. Understanding the Stochastic Partial Differential Equation Approach to Smoothing. *Journal of Agricultural, Biological, and Environmental Statistics* **25**, 1–16 (2020). doi:10.1007/s13253-019-00377-z
31. H. Rue, L. Held. Gaussian Markov Random Fields - Theory and Applications. *CRC Press* (2005).
32. D. Simpson, J. B. Illian, F. Lindgren, S. H. Sørbye, H. Rue. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Biometrika* **103**, 49–70 (2015). doi:10.1093/biomet/asv064
33. F. E. Bachl, F. Lindgren, D. L. Borchers, J. B. Illian. inlabru: an R package for Bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution* **10**,

760–766 (2019). doi:10.1111/2041-210X.13168
34. H. Rue, S. Martino, N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **71**, 319–392 (2009). doi:10.1111/j.1467-9868.2008.00700.x
35. R. S. Bivand, V. Gómez-Rubio, H. Rue. Spatial data analysis with R-INLA with some extensions. *Journal of Statistical Software* **63**, 1–31 (2015). doi:10.18637/jss.v063.i20
36. F. Lindgren, H. Rue. Bayesian Spatial Modelling with R-INLA. *Journal Of Statistical Software* **63**, 1–25 (2015). doi:10.18637/jss.v063.i19
37. World regions in the SDG framework of the United Nations. Available at: https://unstats.un.org/sdgs/indicators/regional-groups/. Accessed: 08/12/2022. (2021).
38. F. Gao, W. Kihal, N. Meur, M. Souris, S. Deguen. Does the edge effect impact on the measure of spatial accessibility to healthcare providers? *International Journal of Health Geographics* **16**, 1–16 (2017). doi:10.1186/s12942-017-0119-3
39. A. J. Righetto, C. Faes, Y. Vandendijck, P. J. Ribeiro. On the choice of the mesh for the analysis of geostatistical data using R-INLA. *Communications in Statistics - Theory and Methods* **49**, 203–220 (2020). doi:10.1080/03610926.2018.1536209
40. G. A. Fuglstad, D. Simpson, F. Lindgren, H. Rue. Constructing Priors that Penalize the Complexity of Gaussian Random Fields. *Journal of the American Statistical Association* **114**, 445–452 (2019). doi:10.1080/01621459.2017.1415907
41. A. Berg, R. Meyer, J. Yu. Deviance Information Criterion for Comparing Stochastic Volatility Models. *Journal of Business and Economic Statistics* **22**, 107–120 (2004). doi:10.1198/073500103288619430
42. A. Van Der Linde. DIC in variable selection. *Statistica Neerlandica* **59**, 45–56 (2005). doi:10.1111/j.1467-9574.2005.00278.x
43. N. Lezama-Ochoa, M. G. Pennino, M. A. Hall, J. Lopez, H. Murua. Using a Bayesian modelling approach (INLA-SPDE) to predict the occurrence of the Spinetail Devil Ray (Mobular mobular). *Scientific Reports* **10**, 1–11 (2020). doi:10.1038/s41598-020-73879-3
44. P. Moraga. Geospatial Health Data Modeling and Visualization with R-INLA and Shiny. *Chapman & Hall/CRC Biostatistics Series* (2019). doi:10.1201/9780429341823
45. V. Gómez-Rubio. Bayesian Inference with INLA. *Chapman & Hall/CRC Press* (2020).
46. A. C. Cameron, F. A. G. Windmeijer. R-squared measures for count data regression models with applications to health-care utilization. *Journal of Business and Economic Statistics* **14**, 209–220 (1996). doi:10.1080/07350015.1996.10524648
47. M. Mittlböck. Calculating adjusted R2 measures for Poisson regression models. *Computer Methods and Programs in Biomedicine* **68**, 205–214 (2002). doi:10.1016/S0169-2607(01)00173-0
48. M. Mohebbi, R. Wolfe, A. Forbes. Disease mapping and regression with count data in the presence of overdispersion and spatial autocorrelation: A Bayesian model averaging approach. *International Journal of Environmental Research and Public Health* **11**, 883–902 (2014). doi:10.3390/ijerph110100883
49. D. S. Moore, W. Notz, M. A. Flinger. The basic practice of statistics. (2013).
50. M. G. Pennino, I. Paradinas, J. B. Illian, F. Muñoz, J. M. Bellido, A. López-Quílez, D.

Conesa. Accounting for preferential sampling in species distribution models. *Ecology and Evolution* **9**, 653–663 (2019). doi:10.1002/ece3.4789
51. P. B. Conn, J. T. Thorson, D. S. Johnson. Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. *Methods in Ecology and Evolution* **8**, 1535–1546 (2017). doi:10.1111/2041-210X.12803
52. A. Lee, A. Szpiro, S. Y. Kim, L. Sheppard. Impact of preferential sampling on exposure prediction and health effect inference in the context of air pollution epidemiology. *Environmetrics* **26**, 255–267 (2015). doi:10.1002/env.2334
53. A. E. Gelfand, S. K. Sahu, D. M. Holland. On the effect of preferential sampling in spatial prediction. *Environmetrics* **23**, 565–578 (2012). doi:10.1002/env.2169
54. J. Watson. A fast Monte Carlo test for preferential sampling. *arXiv* 1–34 (2020).
55. T. T. Zhao, Y. Feng, P. N. Doanh, S. Sayasone, V. Khieu, C. Nithikathkul, M. B. Qian, Y. T. Hao, Y. S. Lai. Model-based spatial-temporal mapping of opisthorchiasis in endemic countries of Southeast Asia. *eLife* **10**, 1–21 (2021). doi:10.7554/eLife.59755
56. F. Cribari-Neto, A. Zeileis. Beta Regression in *R*. *Journal of Statistical Software* **34**, (2010). doi:10.18637/jss.v034.i02
57. R. R. Hocking, R. N. Leslie. Selection of the Best Subset in Regression Analysis. *Technometrics* **9**, 531–540 (1967). doi:10.1080/00401706.1967.10490502
58. D. Temple, X. Manteca. Animal Welfare in Extensive Production Systems Is Still an Area of Concern. *Frontiers in Sustainable Food Systems* **4**, (2020). doi:10.3389/fsufs.2020.545902
59. M. Gilbert, G. Conchedda, T. P. Van Boeckel, G. Cinardi, C. Linard, G. Nicolas, W. Thanapongtharm, L. D'Aietti, W. Wint, S. H. Newman, et al. Income disparities and the global distribution of intensively farmed chicken and pigs. *PLoS ONE* **10**, 1–14 (2015). doi:10.1371/journal.pone.0133381
60. Eurostat. Livestock Unit (2022). Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Livestock_unit_(LSU). Accessed: 15/12/2022.
61. T. P. Van Boeckel, J. Pires, R. Silvester, C. Zhao, J. Song, N. G. Criscuolo, M. Gilbert, S. Bonhoeffer, R. Laxminarayan. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**, eaaw1944 (2019). doi:10.1126/science.aaw1944
62. American College of Surgeon. ATLS - Advanced Trauma Life Support Program for Doctors. *American College of Surgeons* (2008).
63. FAOSTAT. Crops and Livestock Products (2020). Available at: https://www.fao.org/faostat/en/#data/QCL. Accessed: 10/05/2023.
64. D. J. Weiss, A. Nelson, C. A. Vargas-Ruiz, K. Gligorić, S. Bavadekar, E. Gabrilovich, A. Bertozzi-Villa, J. Rozier, H. S. Gibson, T. Shekel, et al. Global maps of travel time to healthcare facilities. *Nature Medicine* **26**, 1835–1838 (2020). doi:10.1038/s41591-020-1059-1
65. The Malaria Atlas Project. Accessibility to Healthcare - Motorized friction surface (2020). Available at: https://malariaatlas.org/project-resources/accessibility-to-healthcare/. Accessed: 01/10/2022.
66. J. van Etten. R Package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software* **76**, (2017). doi:10.18637/jss.v076.i13

67. E. W. Dijkstra. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik* **1**, 269–271 (1959). doi:10.1007/BF01386390
68. Royal College of Veterinary Surgeons. Find A Vet (2023). Available at: https://findavet.rcvs.org.uk/home/. Accessed: 10/02/2020.
69. A. C. Davison, A. Gigli. Deviance residuals and normal scores plots. *Biometrika* **76**, 211–221 (1989). doi:10.1093/biomet/76.2.211
70. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. Available at: https://doi.org/10.7927/H49C6VHW. Accessed: 01/03/2022. (2015).
71. D. J. Weiss, A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, E. Poyart, S. Belchior, N. Fullman, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018). doi:10.1038/nature25181
72. M. Kummu, M. Taka, J. H. A. Guillaume. Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015. *Scientific Data* **5**, 1–15 (2018). doi:10.1038/sdata.2018.4
73. N. Ramankutty, A. T. Evan, C. Monfreda, J. A. Foley. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles* **22**, 1–19 (2008). doi:10.1029/2007GB002952
74. M. Gilbert, G. Nicolas, G. Cinardi, T. P. Van Boeckel, S. O. Vanwambeke, G. R. W. Wint, T. P. Robinson. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data* **5**, 1–11 (2018). doi:10.1038/sdata.2018.227
75. G. Cinardi, D. Da Re, M. Gilber, T. P. Robinson, W. G. R. Wint. Gridded Livestock of the World - 2015 (GLW4). *Harvard Dataverse* (2022). doi:10.7910/DVN/LHBICE
76. Australian Veterinary Association. Australian veterinary workforce survey 2018. (2019).
77. C. B. G. De Lourdes. Veterinary education in Brazil: Past history, current issues. *Journal of Veterinary Medical Education* **31**, 28–31 (2004). doi:10.3138/jvme.31.1.28
78. T. L. Whiting. Veterinary Practice — The Canadian multinational veterinary workforce. *Canadian Veterinary Journal* **62**, 1195–1201 (2021).
79. F. Mayen. A status report of veterinary education in Ethiopia: Perceived needs, past history, recent changes, and current and future concerns. *Journal of Veterinary Medical Education* **33**, 244–247 (2006). doi:10.3138/jvme.33.2.244
80. W. A. Khamas, A. Y. M. Nour. Veterinary medical education in Iraq. *Journal of Veterinary Medical Education* **31**, 301–309 (2004). doi:10.3138/jvme.31.4.301
81. M. Bwana Otieno, J. Orungo Onono. Current Trends in the Sectors of Interest for Veterinary Medicine Students and Job Placements of Veterinarians in Kenya Drugs. *Journal of Dairy, Veterinary & Animal Research* **3**, 73–78 (2016). doi:10.15406/jdvar.2016.03.00075
82. A. Catley, M. Said, M. Farah, I. Handule. Veterinary Services in the Somali National Regional State, Ethiopia: A Situation Analysis. (1997).
83. Wenxuan, C. Veterinarian So Easy? Discussion on the Current Situation of the Lack of Large Animal Veterinary Talents (2021). Available at: http://www.cqvip.com/qk/94496x/201710/673305493.html. Accessed: 10/08/2021.
84. Malvina M., Prieto L., Pablo E., S. Cánen. Animal Welfare in the Uruguayan Veterinary

Profession Field. *Journal of Agricultural Science and Technology A* **7**, 357–362 (2017). doi:10.17265/2161-6256/2017.05.008

85. J. Watson. A fast Monte Carlo test for preferential sampling. 1–34 (2020).

Chapter 4

Quantifying travel time to healthcare: a case study with clinical laboratories and veterinarians

Abstract

In humans and animals, access to healthcare is essential for treating, monitoring, and preventing infectious diseases, including antimicrobial resistance. In low- and middle-income countries, the limited availability of laboratories with antimicrobial susceptibility testing capacities (AST) precludes healthcare professionals from collecting bacterial samples and understanding trends in antimicrobial resistance. Similarly, lack of access to veterinary services limits the availability of treatment for food animals and systematic monitoring of their health. This, in turn, affects disease detection and threatens the livelihood of those relying on animals for subsistence. This study focuses on improving the accessibility to health services under a hypothetical 5% scale-up of capacities. The first case study explores increasing the number of AST laboratories in Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi. Our findings indicated that a 5% scale-up in the number of health facilities to be equipped for AST, if geographically targeted, could increase the population living within 1 hour of travel time from AST laboratories by 28%. Our second case study focuses on veterinary services globally. Here we showed that a geographically targeted 5% increase in the workforce of veterinarians would reduce the number of food animals living more than 1 hour away from a veterinarian worldwide by 32.6%.

Introduction

Timely access to healthcare is essential to facilitate the early detection of disease outbreaks and help contain their spread (*1–4*). However, in low- and middle-income countries (LMICs), limited road networks and infrastructures (bridges, tunnels, etc.) (*5*, *6*) as well as the inadequate distribution of health facilities (*7*), can seriously challenge access to healthcare services. Recently, a review by Kelly and colleagues showed that high travel time to healthcare was associated with adverse health outcomes in 77% of cases (*8*), including increased mortality in women giving birth (*9*) and in infants in need of primary care (*10*).

In LMICs, poor accessibility to health services also affects surveillance of diseases of global importance. A key example is antimicrobial resistance (AMR) (*11*). Sub-Saharan Africa has the highest mortality rate from AMR infections, with 24 deaths per 100,000 attributable to AMR (*12*), more than those attributable to the human immunodeficiency virus (HIV) (*13*). One of the key actions to decrease this burden is to track temporal AMR trends across the continent. However, this objective is challenged by the very limited number of health facilities equipped with laboratories that can perform antimicrobial susceptibility testing (AST) (*13*).

Parallels can also be drawn with food animals. An extended network of health services, in this case represented by veterinarians, is crucial to providing prompt access to care for animals, preserving their health, and ensuring food security. This is essential for the sustainability of the food animal sector which currently contributes to nearly 40% of the total agricultural output in

high-income countries (HICs) and about 20% in LMICs (*14*), as well as for supporting the livelihood of 1.3 billion people active in this sector worldwide (*15*).

However, until recently, little attention has been drawn to improving accessibility to AST laboratories and to improving access to veterinarians from the perspective of geographic distance. Improving access to these health services could be addressed through a scale-up in their national capacities. Nevertheless, the distribution of the human population, veterinarians, and animals is highly heterogeneous within countries. Therefore, geographically targeted approaches to scale-up health services may carry the greatest benefits in terms of improved access to care for the populations of humans and animals (*16*).

Maximizing the population using health services within a certain travel time threshold can be formalized mathematically as a maximal coverage location problem (MCLP) (*17*, *18*). These approaches have been applied in different public health contexts: for improving access to HIV-testing laboratories in South Africa (*19*), to maximize access to primary healthcare throughout Sub-Saharan Africa (*7*), or to enable access to antivenoms used against snakebites in Nepal (*20*). However, MCLP approaches are often associated with a prohibitive computational cost typical of the class of NP-hard problems (*21*). Therefore, previous analyses have focused on identifying approximations of MCLP to address this challenge while attempting to improve access to health services (*22*) or for integrating additional factors to physical distance such as the capacity for treatment of each facility (*23*).

In this study, we used two MCLP-based approaches to identify the optimal locations that would maximize the population living within 1 hour of travel time from health services under an increase in their capacities. Specifically, we increased by 5% the national health facilities to be equipped for AST and performed a 5% scale-up of national veterinary capacities. In the first case study, we focused on maximizing the population with access to AST laboratories in five African countries (Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi) (*13*) using AccessMod, a dedicated accessibility software developed by the GeoHealth group (University of Geneva) and the World Health Organization (*24*). For the second case study, we focused on improving access to veterinarians caring for food animals worldwide. Given the global scale of this second dataset (Chapter 3), we developed an approximation of the MCLP based on catchment radii to overcome the computational barriers. We used Kenya, Panama, Ecuador, Liberia, Eritrea, Honduras, Nicaragua, Costa Rica, and Cambodia as "test countries" to evaluate the validity of this approximation and show that it can produce meaningful results within a fraction of the computation time of the exact solution.

Methods

Data collection

Laboratories for antimicrobial susceptibility testing

We identified health facilities currently equipped with AST laboratories in Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi (*25*). Specifically, for each country, we obtained the 2022 reports of the "Mapping Antimicrobial Resistance and Antimicrobial Use Partnership" (MAAP) (*26–30*) published in July 2023 by the African Society for Laboratory Medicine (*31*). From these reports, we sampled the addresses of health facilities equipped for AST and retrieved their longitude and latitude (in decimal degrees) using Google's geocoding Application Programming Interface (API) available from the R package *ggmap* (*32*).

Next, for each country, we used the OpenStreetMap (OSM) database (*33*) to sample the coordinates of the health facilities that are currently not equipped with AST laboratories. These health facilities were necessary to select the potential candidates that, once equipped for AST, would maximize the human population with access to AST laboratories within 1 hour of travel time (hereafter referred to as "population covered"); we chose the threshold of 1 hour for consistency with the MAAP reports, that indicated the population currently covered by existing AST laboratories. Then, we used the R package *osmdata* (*34*) to query the OSM database and collect the locations of current health facilities able to perform AST: hospitals and clinical diagnostic laboratories (*35*). The country shapefiles downloaded from the Database of Global Administrative Areas (GADM) (*36*) were used to define the geographical extent of the areas where to query OSM. Then, for each area, we performed queries using the key string "healthcare" and the filtering string "laboratory" and "infectious diseases". Another query was performed using the key string "amenity" and the filtering string "hospital". These health facilities are hereafter referred to as "candidates".

Next, country-level maps at the ~100 m^2 resolution were collected to perform accessibility analysis in AccessMod. Firstly, Digital Elevation Model maps (DEM) (*37*) were used to initiate every analysis in AccessMod (mandatory requirement) (*38*). Secondly, maps of land cover categories (*39*) were used to produce friction surfaces, i.e., maps that express the cost to travel across each of its pixels (for example in seconds per meter traveled). Land cover maps were provided by the GeoHealth group (University of Geneva) (*40*), which merged them with the rasterized network of roads and rivers downloaded from OSM (*41*). Then, we collected tables of motorized travel speeds to associate with each category of the merged land cover maps. These tables were provided by Weiss et al. (*42*) and adapted for the African countries by Hierink et al. (*43*). In this way, we obtained friction surfaces that combine areas where it is only possible to walk with an average speed of 5 km/h (*42*) with areas where the road networks allow for the use of motorized vehicles ("walking + motorized" friction surface). This friction surface represents a scenario where people have to walk to the closest roads where it is possible to use motorized vehicles (e.g., public transport, cars), also assuming that such vehicles would be available without delay when reaching the roads (*20*). Finally, we quantified the increase in the population covered when equipping candidates for AST using national maps of the human population count (*44*).

Veterinarians

The data used to scale up national veterinary capacities were presented in Chapter 3. Specifically, we used the coldspots maps to select countries where supplementary veterinarians were needed. Coldspots were defined as 10x10 km^2 resolution maps reporting areas where cattle, chickens, and pigs are farther than 1 hour of travel time from the nearest veterinarians. In this case, we used the threshold of 1 hour of travel time for consistency with the concept of the "golden hour" used in human medicine, which represents the time when patients have the highest chance of survival if they receive medical care (*45*).
In addition, we used the World Settlement Footprint map (WSF) (*46*) to identify pixels where there are human settlements but currently no veterinarians. These 10x10 km^2 pixels represented the potential locations to select for scaling up the veterinarians in a country. Therefore, for consistency with the case study about AST laboratories, we will refer also to these pixels as "candidates".
Next, we computed travel time maps for each supplementary veterinarian allocated using the 1x1 km^2 resolution "walking + motorized" friction surface, available from the Malaria Atlas Project (*47*, *48*). From each of these travel time maps, we quantified the animal population covered by each supplementary veterinarian using density maps of cattle, chickens, and pigs available from the 4th version of the Gridded Livestock of the World (*49*, *50*).

Travel time analyses

Laboratories for antimicrobial susceptibility testing

AccessMod (version 5.8) (*51*, *52*) was used to identify the 5% of national candidates to be equipped for AST. Firstly, AccessMod was used to correct country-level maps of the human population. This step aimed at identifying if some map pixels corresponded to pixels of the merged land cover representing physical barriers where humans could not be found (e.g., rivers). Therefore, AccessMod was used to translate population counts found at physical barriers to their closest neighbors where it is more reasonable to find humans (e.g., pixels categorized as "built-up" (*53*)). Secondly, AccessMod was used to produce country-level friction surfaces by associating the merged land cover maps with the tables of travel time scenarios defined for each land cover typology. Once all the input files for AccessMod were ready, four different analyses were performed for each country:

1. Initial accessibility evaluation: we computed 100 m^2 resolution maps of travel time to reach the existing AST laboratories. Specifically, we calculated the cumulative time required to cross each friction surface pixel to reach existing AST laboratories. This calculation was done using the eight-directional least-cost path algorithm (*54*). Furthermore, no corrections were made for the slopes (isotropic movement), since this is usually applied to pedestrians and cyclists (*43*). These aspects were equally considered for the case study about veterinarians.

2. Zonal statistics: zonal descriptive statistics were produced to evaluate the regional coverage of the existing AST laboratory. The input files for AccessMod used in this

step were the travel time maps computed in Step 1 and the GADM shapefiles of the regional division for each country (*36*).

3. Geographic coverage: this step produced travel time maps (Step 1) for each candidate and returned their catchment area. The catchment area is defined as the area around a facility that reports population information (e.g., number of people) using its services (*55*, *56*). Along with the shapefiles of the catchment areas, AccessMod also returned the population covered by each candidate to potentially equip for AST.

4. Scaling up analysis: this analysis used the catchment areas computed in Step 3 to maximize the population covered using a recursive exhaustive approach (REA) that tests iteratively every candidate. The algorithm then chooses 5% of the supplementary candidates to be equipped for AST which cover the highest cumulative population.

After performing these analyses in each country we calculated i) the spatial distribution of the 5% supplementary candidates to be equipped for AST using kernel density (KDE) maps (*57*), ii) the population covered by each supplementary laboratory, iii) the number of supplementary candidates to be equipped for AST necessary to cover at least 95% of the population, iv) and identified regions with the highest number of supplementary laboratories allocated.

Veterinarians

In each country with coldspots of veterinary capacities (Chapter 3), we added 5% veterinarians (compared to the country's total) such as to minimize the number of animals living in coldspots (IA). As for the scaling up of AST laboratories, a recursive exhaustive approach (REA) that adds veterinarians in every possible candidate will bring the highest number of IA within 1 hour of travel time from a veterinarian (hereafter referred to as "animal population covered"). Since the candidates to test in this study were >250,000, we used the R package *gdistance* (*58*) to develop an approximation of the REA to decrease its computational time.

First, we selected nine "small" countries (Kenya, Panama, Ecuador, Liberia, Eritrea, Honduras, Nicaragua, Costa Rica, and Cambodia) to quantify the animal population covered using the REA. This was done by recording the IA decrease for each supplementary veterinarian of the additional 5% capacity added. Then, we developed a scaling up approach to allocate the same 5% of veterinarians in less time when compared to REA, while aiming for a comparable animal population covered. This approximation was called the "contiguity approach".

We defined the contiguity approach by sampling properties in the proximity area around each candidate. These properties can vary across countries since they depend on i) the density of IA, ii) the distribution of existing veterinarians, and iii) local values of the friction surface, due to the characteristics of the territory and traveling speed limits. These properties were used as a guide in the contiguity approach to choose the candidates where the supplementary 5% of veterinarians could cover a similar animal population covered by the REA. The steps to perform the contiguity approach were the following:

1. We randomly selected 100 candidates and allocated the 1st veterinarian in one of these pixels, re-computed the coldspots map, and compared it with the initial one to extract the catchment area, i.e., where the travel time decreased to <1 hour due to the presence of a new veterinarian (fig. S1A).

2. We calculated the average catchment area radius based on the distance of the allocated veterinarian, starting from each point of the catchment area perimeter (fig. S1A).

3. We repeated step (2) independently for all the randomly selected candidates to calculate an average radius of the catchment areas specific to each country.

4. For the 1st veterinarian to permanently allocate in the country, we defined a circle around every candidate available based on the average radius calculated in step (3).

5. In this circle, we sampled a) IA, calculated intersecting maps of food animals' density with the coldspot map b) the number of pixels classified as coldspots (N_C), c) the minimum distance of the candidate from other veterinarians already available in the country, and d) the average travel time. Property (a) prioritizes candidates where catchment area IA is the highest. Property (b) and (c) identify candidates that are far from other veterinarians and hence where a new allocation will not create clusters of veterinarians. Property (d) allows the identification of candidates where travel time in their catchment areas is the lowest; this permits a newly allocated veterinarian to travel farther and reach more coldspots.

6. For each candidate, we categorized the values of property (a) to (c) in five different groups by building an ordinal variable based on numerical ranges of equal width.

7. We assembled each variable in a database and sorted its rows starting with the groups of property (a), i.e., prioritizing allocation in candidates with a high IA (*7*). Then, the sorting of the rows corresponding to candidates surrounded by the highest IA continued based on the groups of the highest N_C values. The group of rows corresponding to candidates with high values of IA and N_C was then sorted based on their highest minimum distance from existing veterinarians (fig. S1B). Finally, among the sorted rows, which correspond to candidates with the highest IA, N_C, and farthest from existing veterinarians, we selected for the allocation the one whose catchment area presented the minimum average TT^M.

8. We allocated an additional veterinarian to this candidate, excluded it from the batch of candidates, and re-computed the coldspots maps. Then, we recorded the updated IA.

9. We repeated steps (4) to (8) for all the remaining veterinarians to allocate in the country by using each time the updated version of the coldspots map.

Furthermore, in the nine countries inspected, we compared the results of scaling up veterinarians using the REA and the contiguity approach (i.e., approaches that re-compute coldspots iteratively) with two supplementary approaches that allocate all veterinarians at once. We based the first approach on the administrative division of a country, prioritizing the allocation in regions with the highest number of IA ("administrative approach"). The second approach assigned veterinarians randomly within a country ("random approach"). For the random approach, we repeated the scaling up for 100 Monte Carlo simulations to avoid selecting the candidates minimizing IA by chance.

For each approach, including the 100 Monte Carlo allocations of the random approach, we calculated its performance in terms of isolated animals' coverage (IA_C) as follows:

$$IA_C\ (\%) = \frac{IA - IA_{Approach}}{IA - IA_{REA}} \cdot 100$$

Where IA_{REA} is the number of isolated animals living in coldspots after allocating veterinarians through the REA, while $IA_{Approach}$ is the number of isolated animals living in coldspots after allocating veterinarians through the contiguity, the administrative, or the random approach. Then, we inspected the results of each approach through 10x10 km^2 resolution KDE maps computed from the geographic coordinates of the supplementary veterinarians. Based on the results of these comparisons, we selected the fastest approach to achieving results like the REA and applied it in every country with coldspots of cattle, chickens, and pigs.

Results

AST Laboratories

Between 2016 and 2018, less than 24 AST laboratories were in operation across Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi (Table 1 and fig. S2). Malawi was the country with the least AST laboratories per million inhabitants (0.8 lab /1,000,000), while Gabon was the one with the best coverage (8.5 labs / 1,000,000). The proportion of the population living more than 1 hour away from an AST laboratory was 72.2% in Sierra Leone, 65.9% in Burkina Faso, 42.9% in Malawi, 34.5% in Senegal, and 33.9% in Gabon (fig. S3). The regions where this proportion was the highest (>99%) were Kédougou (Senegal), Northern (Sierra Leone), Ogooué-Ivindo (Gabon), Sahel (Burkina Faso), and Kasungu (Malawi). For candidates to be equipped with AST laboratories, the queries of the OSM database returned 354 candidates in Gabon, 944 in Sierra Leone, 1,129 in Senegal, 1,315 in Malawi, and 1,718 facilities in Burkina Faso (fig. S4). Of these candidates, more than 97% in each country were classified as hospitals.

Table 1. Yellow columns report the existing laboratories for antimicrobial susceptibility testing and the initial population living within 1 hour of travel time from these laboratories. Orange columns report the 5% additional health facilities to be equipped with laboratories and the supplementary population that this increase would cover. Green columns report the total number of laboratories after increasing their numbers and the total population covered in the country. Percentages next to the number of laboratories represent the proportion of health facilities in the country equipped for antimicrobial susceptibility testing.

Country	Existing laboratories	Initial population covered	Supplementary laboratories (5%)	Supplementary population covered	Total laboratories	Total population covered
Senegal	22 (1.98%)	10,947,489 (65.5%)	55 (4.96%)	3,346,472 (20%)	77 (6.94%)	14,293,961 (85.5%)
Sierra Leone	7 (0.74%)	2,213,817 (27.8%)	47 (4.99%)	3,493,234 (43.9%)	54 (5.73%)	5,707,051 (71.7%)
Gabon	17 (4.7%)	1,464,855 (66.1%)	18 (4.97%)	137,061 (6.2%)	35 (9.67%)	1,601,916 (72.3%)
Burkina Faso	23 (1.35%)	7,134,605 (34.1%)	85 (4.97%)	6,917,820 (33.2%)	108 (6.32%)	14,052,425 (67.3%)
Malawi	15 (1.13%)	10,925,221 (57.1%)	66 (4.96%)	7,111,551 (37.2%)	81 (6.09%)	18,036,772 (94.3%)

The analyses performed in AccessMod led to identifying the 5% of candidates to be equipped for AST covering the highest proportion of the population. The geographic distribution of these candidates varied among countries (Fig. 1A).

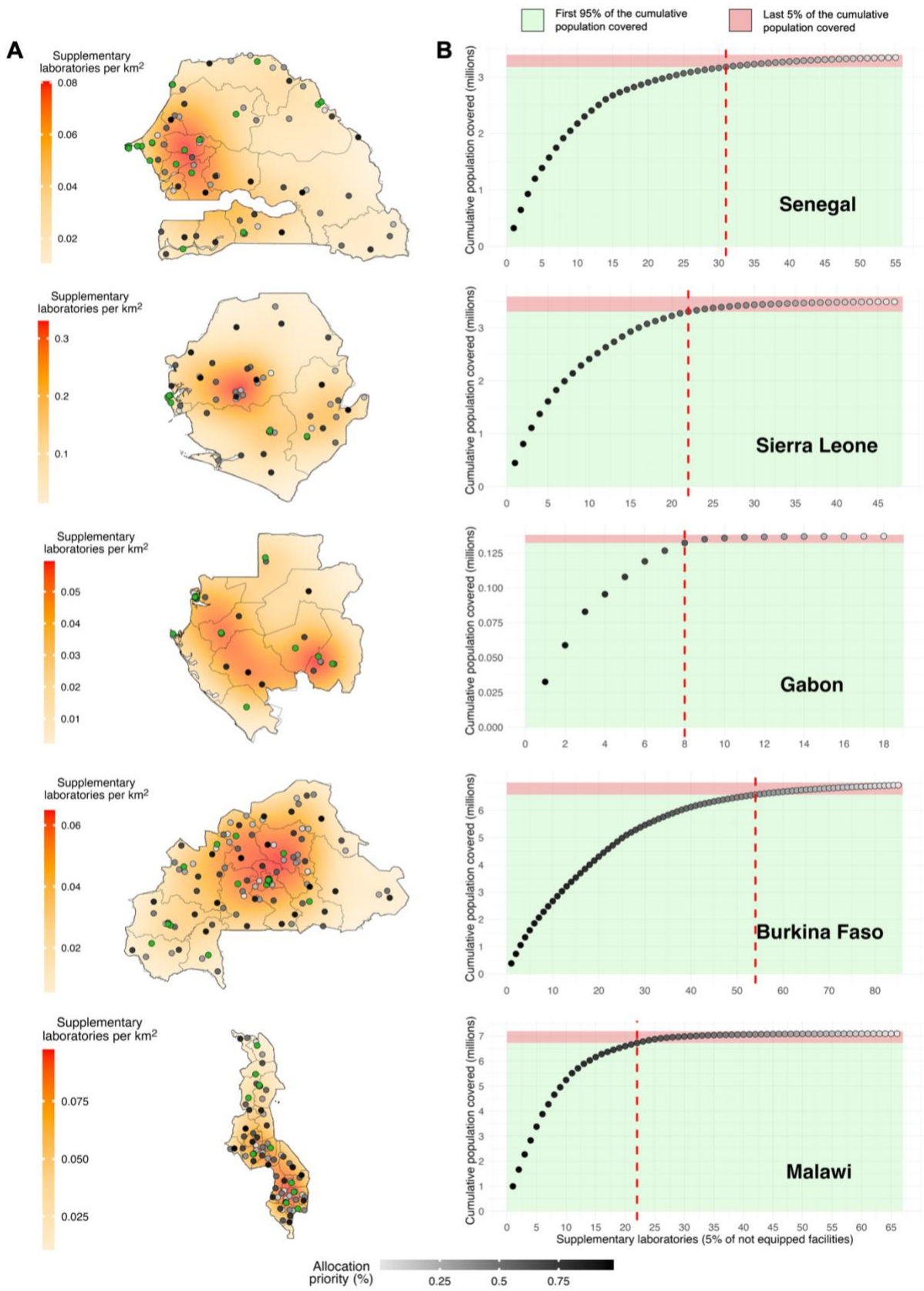

Fig. 1. Where to increase access to laboratories for antimicrobial susceptibility testing? (A) Locations of the existing health facilities with antimicrobial susceptibility testing laboratories (green dots) and the locations of the 5% of candidates (see "Methods") to be equipped for antimicrobial susceptibility testing to maximize access to them within 1 hour of

travel time (gray dots). In addition, the maps report the density per km^2 of all the laboratories (including the ones available before equipping additional candidates). (B) Cumulative population that obtained access to laboratories within 1 hour of travel time under a 5% increase in candidates to be equipped for antimicrobial susceptibility testing. The green area represents 95% of the population with access to laboratories within 1 hour of travel time after equipping 5% of candidates for antimicrobial susceptibility testing. The red line indicates the number of health facilities necessary to reach this target. The red area represents the remaining 5% of the population that obtained access to the laboratories.

In Senegal, the highest density of supplementary health facilities to be equipped for AST was identified in the Western part of the country. In Sierra Leone and Burkina Faso, the highest density was in the Central part of the countries, while for Malawi and Gabon in the Southern and Southwestern parts, respectively. The regions with the highest candidates to be equipped for AST were Saint-Louis and Tambacounda (Senegal), Northern and Eastern (Sierra Leone), Haut-Ogooué and Ngounié (Gabon), Boucle du Mouhoun and Centre-Nord (Burkina Faso), and Mzimba and Lilongwe (Malawi) (fig. S5). The regions with the highest population covered were Tambacounda (Senegal), Northern (Sierra Leone), Ngounié (Gabon), Boucle du Mouhoun (Burkina Faso), and Kasungu (Malawi).

Overall, a 5% increase in candidates to be equipped for AST increased the population covered by 43.9% in Sierra Leone, 37.2% in Malawi, 33.2% in Burkina Faso, 20% in Senegal, and 6.2% in Gabon (Table 1). In absolute terms, this increase ranges from the ~130,000 people covered in Gabon to the ~7 million covered in Malawi. Across all countries, the resulting population covered was ~21 million, comparable to the population of Zambia.
In addition, we tried to understand if selecting fewer candidates than the national 5% could cover at least 95% of the population covered with the whole 5% of candidates. This translates into saving resources when the gains in population covered with supplementary laboratories are small. For this reason, we inspected the cumulative population covered by each of the candidates selected by the REA. Specifically, increasing the proportion of candidates to be equipped for AST by 2.8% in Senegal, 2.4% in Sierra Leone, 2.3% in Gabon, 3.2% in Burkina Faso, and 1.7% in Malawi, was sufficient to cover 95% of the population covered using the whole 5% of supplementary candidates to be equipped for AST (Fig. 1B). For each country, this translated in saving an average of 26.8 candidates to be equipped for AST.

Veterinarians

The contiguity, the administrative, and the random allocation approaches were applied in the nine countries (Kenya, Panama, Ecuador, Liberia, Eritrea, Honduras, Nicaragua, Costa Rica, and Cambodia) and compared with the "optimal solution", the REA (Fig. 2). The contiguity approach performed best and reached >90% of the animal population covered by the REA when scaling up the number of veterinarians by 5%. In addition, the contiguity approach was, on average across each country, 20 times faster than the REA (Fig. 2 and Table S1). In contrast, the administrative and the random approach covered, on average, 35.1% and 42.5% of the animal population covered by the REA (average range of this value across all the MC

simulations for the random approach: 23.3%-63.4%). Furthermore, in the nine countries selected, the KDE maps values computed from the veterinarians allocated by each approach showed that the administrative and the random approach allocated veterinarians to candidates different from the ones selected by the REA. In contrast, the patterns of the veterinarians allocated through the contiguity approach corresponded to the ones of the REA (fig. S6). As a result, the average Pearson correlation coefficient computed between the values of each KDE map of the REA and the contiguity approach was 0.96 (95% bootstrapped CIs: [0.94, 0.97]). Therefore, the contiguity approach was applied to all countries with coldspots to generate a global map of supplementary veterinarians (fig. S7).

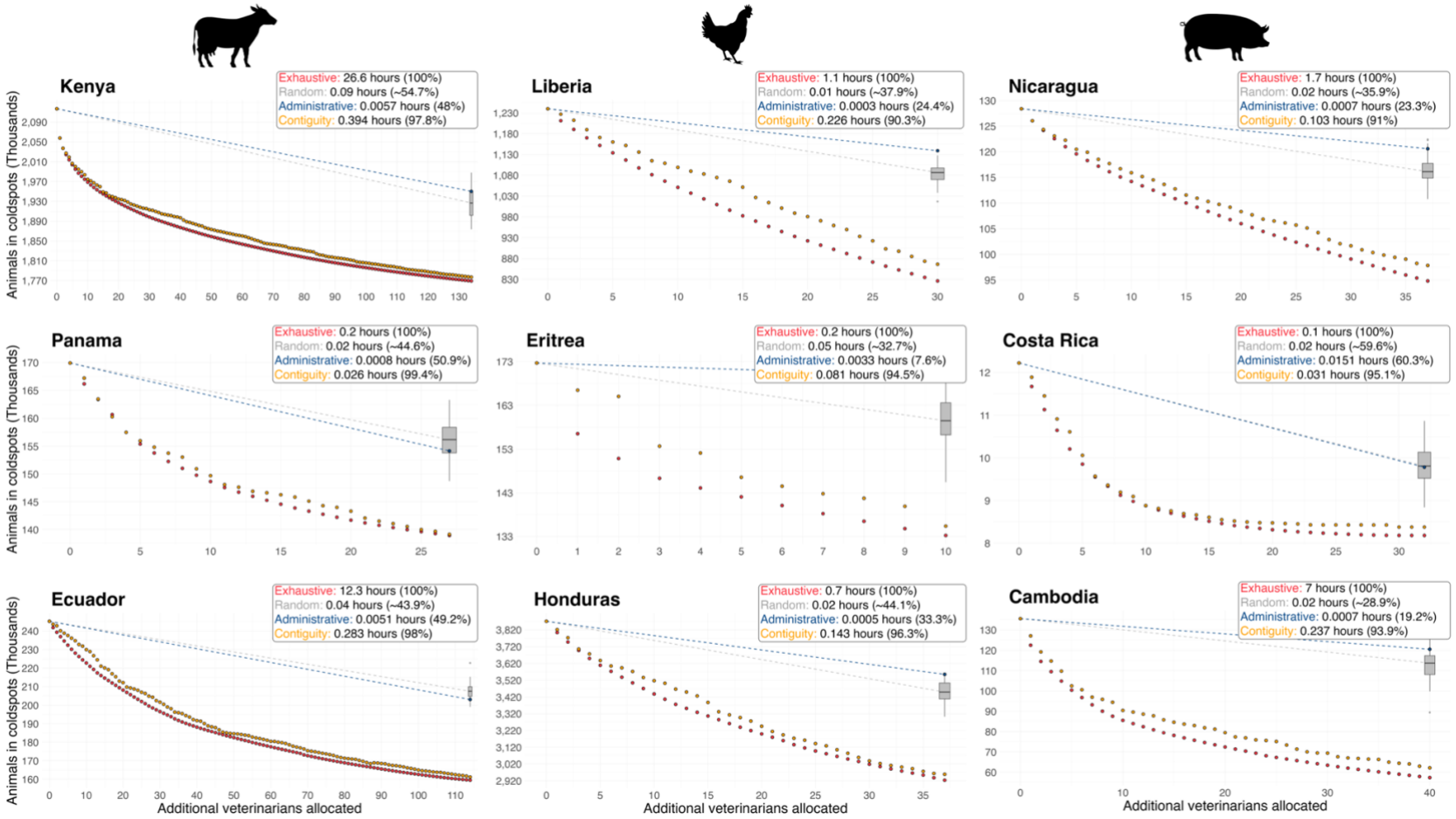

Fig. 2. Comparison of allocation approaches for veterinarians between different countries and food animal species. The curves represent the decrease of food animals living in coldspots for each additional veterinarian allocated to a country through the recursive exhaustive approach (red dots) and the contiguity approach (orange dots). Since the administrative and the random approaches are not iterative, the graphs report a single dot (administrative) and a boxplot (because of the 100 Monte Carlo simulations of the random approach) to represent their performance. The performance of each approach for improving access to care (IA_C, see "Methods") is reported in brackets, and it represents the percentage of food animals brought within 1 hour of travel time from a veterinarian out of the total brought by the recursive exhaustive approach.

Globally, a 5% increase in the number of veterinarians through the contiguity approach could reduce the coldspots area by 6,600,000 km^2 (comparable to 85% of the size of Australia, fig. S8) and reduce the number of animals in coldspots by 26.9% for cattle, 34% for chickens, and 44% for pigs (fig. S9). Together, these food animals correspond to 89% of the LSUs of Mexico. The continents where the contiguity approach removed the highest number of food animals from coldspots were Asia (11%) and Latin America (8.6%). Furthermore, for each species, all the allocations performed at the national level were merged into a unique dataset and the coordinates of the top 1,000 veterinarians that covered the highest animal population were extracted. The geographic pattern of these supplementary veterinarians showed the areas where to prioritize the increase of veterinary capacities (Fig. 3A).

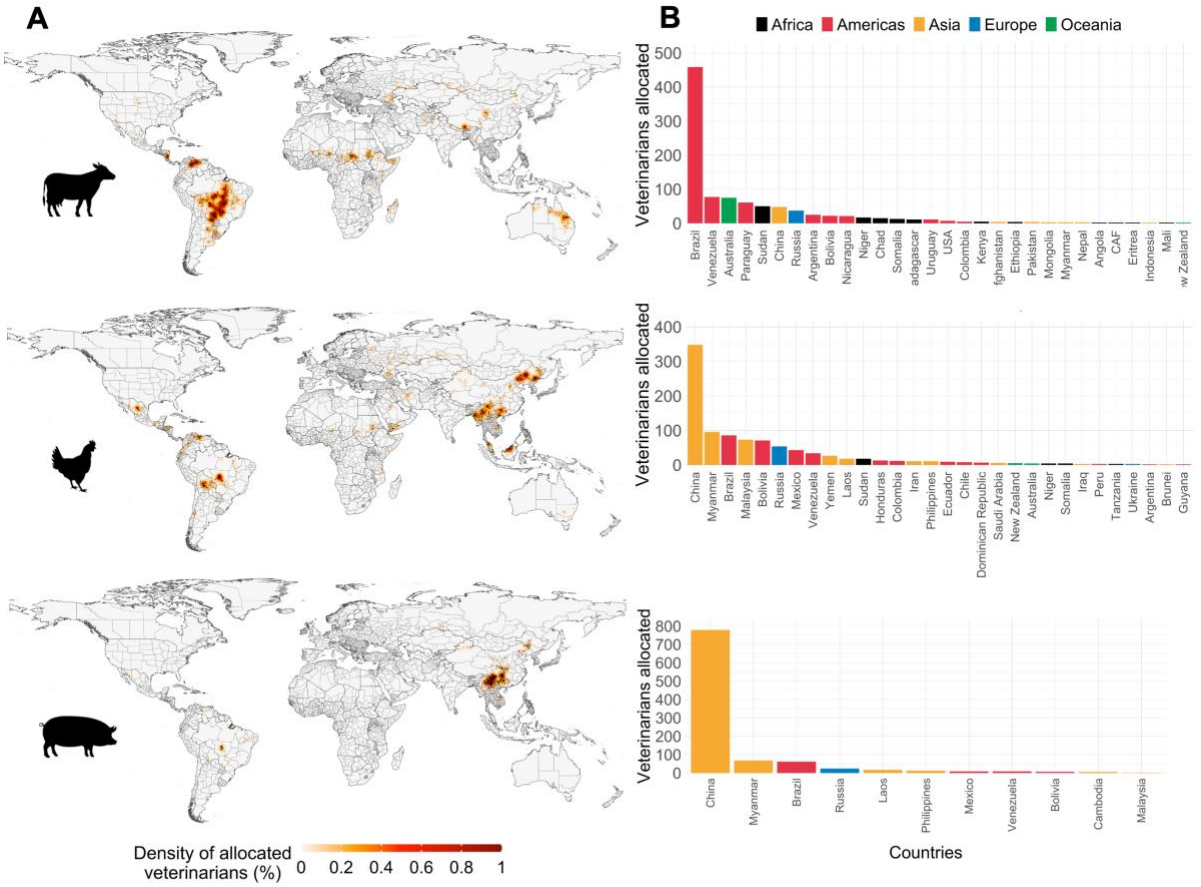

Fig. 3. Where to increase veterinary capacities. (A) Global maps of the regions that concentrate a 5% global increase in veterinarians to reduce coldspots of veterinary care (see "Methods") for cattle, chickens, and pigs. (B) Countries where veterinarians could be added to maximize the number of animals living within 1 hour of veterinary care, grouped by continent.

Countries that would concentrate the increase (5%) in the global number of veterinarians if these were targeted geographically were China (39.2%), Brazil (20.2%), Myanmar (5.6%), Venezuela (4%), and Russia (3.8%) (Fig. 3B). At the sub-national level, the regions that would concentrate the increase in the global number of veterinarians for cattle (fig. S10) were Mato Grosso, Rio Grande do Sul, Pará, and Goiás (Brazil), Xizang, Qinghai, and Yunnan (China), Mahajanga (Madagascar), Central Darfur (Sudan), Zinder (Niger), Kalmyk and Orenburg

(Russia), and Queensland and Northern Territory (Australia). For chickens (fig. S11), the regions identified were Mato Grosso (Brazil), Cochabamba and Santa Cruz (Bolivia), Durango (Mexico), Jilin, Yunnan, Nei Mongol, and Guangxi (China), Antsiranana (Madagascar), Northern Cape (South Africa), Zinder (Niger), Stavropol and Leningrad (Russia), New South Wales (Australia) and West Coast (New Zealand). For pigs (fig. S12) the regions were Mato Grosso and Parà (Brazil), Sonora (Mexico), Yunnan, Guizhou, Sichuan, and Guangxi (China), Gaza and Tete (Mozambique), and Omsk and Tatarstan (Russia).

Discussion

In this study, we estimated travel times to AST laboratories and veterinarians. The aim was to calculate the impact of a hypothetical 5% increase in the national health facilities to be equipped for AST and in the national veterinary capacities on the population living within 1 hour of travel time from these health services. First, the 5% threshold aimed at evaluating the positive impact of a modest increase in resources (a common situation in LMICs) on the population brought within reach of health services. Second, considering a 1-hour timeframe was a convenient benchmark for its easy interpretability and its comparability with the "golden hour" used in human medicine. However, while a 1-hour timeframe has been used in the context of trauma care for humans, it is important to note that this is not a universally recognized standard. The optimal timeframe for intervention can vary significantly based on the specific medical condition (*59*). For example, it is suggested that patients with severe hemorrhage require surgical intervention within 20 minutes from the event causing the hemorrhage (*60*), while a 2-hour timeframe is a commonly acknowledged critical time for obstetric emergencies (*61*). Another example is represented by people in need of antivenoms for snakebites. In their study, Ochoa and colloeagues used travel time intervals to access facilities with antivenoms of 0-30 minutes, 30-60 minutes, and >60 minutes if the neurotoxic effects of snakes' venom have, respectively, severe, moderate, and mild risk of mortality (*20*). Our choice of a 1-hour timeframe is therefore an initial attempt to evaluate the lack of access to health services, which might differ for specific health issues of the populations considered in this study.

Travel time scenarios can also change based on the directionality of movements, influenced by the topography of the areas considered for the accessibility analyses. In our case studies, we focused on people traveling toward AST laboratories and veterinarians traveling toward animals, assuming in both cases an isotropic movement. This implies that we did not incorporate DEM maps into our accessibility models in either case. This choice depends on the nature of the friction surfaces used in our study, which assume walking as the primary mode of transportation only up to the nearest roads where motorized vehicles are available, without time delays in accessing such vehicles (*20*). Therefore, for motorized traveling, variations in slopes caused by natural and artificial hills, as well as other physical barriers, are assumed to have no impact on the overall travel time when heading towards a health service or the population to be served. Further investigations, including DEM models for areas where

walking is the sole movement option for long distances, could enhance the precision of our travel time maps.

In addition, the population coverage could differ based on the choice of the friction surface. For instance, in the study for scaling up veterinary capacities, we used a global friction surface at the resolution of 1 km^2 (*48*) to estimate travel time between veterinarians and food animals. An area of 1 km^2 could include more than one road with different traveling speeds that are consequently merged to obtain a single value of time cost to cross each pixel. Generally, this cost is associated with the one of the paths that can be traveled the fastest (*62*). However, a limitation of this approach is the underestimation of the time required to reach populations reachable only through roads with traveling speeds lower than the one assigned to a friction surface pixel containing also high-speed roads. For this reason, different research groups produce their friction surfaces for the specific objectives of their accessibility analyses, for example at the resolutions of 100 m^2, 30 m^2, and 20 m^2 (*20*, *63*, *64*). Friction surfaces with these resolutions can produce catchment areas whose shape matches the one of the road networks, leading to more precise travel time maps. Similarly, high-resolution friction surfaces can better represent areas where walking is the only traveling option, and also adapt walking scenarios for a specific analysis, as shown by Watmough and colleagues for parents walking with children that are slower than the average walking pace (5 km/h) (*63*). Since the case study about AST laboratories was focused on a national-level, we could use the high-resolution friction surfaces available for the countries investigated. However, to the best of our knowledge, high-resolution global friction surfaces to be used in the case study about veterinarians are not yet available, and they might be associated with prohibitive computational costs if they did exist on a global scale. Future research efforts could focus both on producing high-resolution friction surfaces for every country and speeding-up the calculation of travel times to produce coldspots maps associated with detailed catchment areas.

For identifying the health facilities to be equipped for AST, we used the software AccessMod to test every candidate and then select the ones covering the largest population (REA, see "Methods"). This approach was instrumental to understanding when the marginal benefits of increasing access to AST laboratories plateaued (Fig. 1B). In every country analyzed (Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi) the REA showed that, on average, increasing the number of health facilities to be equipped for AST by 2.5% would account for 95% of the population covered under a 5% increase in health facilities equipped for AST (Fig. 1B). This could have important implications in countries where resources are limited, and help decision-makers to allocate them in high-priority areas (*65*). Nevertheless, the REA can be computationally expensive when the scale of the study increases. For increasing health facilities to be equipped for AST, only 5,460 candidates across five countries were tested.

In contrast, scaling up of veterinary capacities required to evaluate >250,000 candidates, ~42,000 of which were just in China and Brazil. For this reason, we developed an approximation of the REA to increase the number of veterinarians in each country with coldspots. As with the REA, the contiguity approach prioritizes candidates where a supplementary veterinarian covers the highest animal population. However, this approach

samples information about food animals that are lacking access to care directly in areas in the proximity of each candidate. For this reason, we could avoid calculating the catchment areas for each candidate and the population they can serve, making the scaling up through the contiguity approach faster than the scaling up using the REA.
Although this approach could be helpful to support the scaling-up of veterinarians during time-sensitive health issues, veterinarians are rarely involved in emergency medicine where travel time is critical for the survival of food animals. While travel times lower than 1 hour can be used as a proxy to evaluate the degree of access to care, it is worth mentioning that most veterinary interventions are of preventive nature. Therefore, understanding the broader context of veterinary care and its predominantly preventive focus is essential in guiding the implementation of scaling-up strategies.

As a result, the contiguity approach covered a similar animal population of the REA, but it was, on average, 20 times faster across the countries where we compared the two approaches (Fig. 2 and Table S1). Furthermore, scaling up approaches that are geographically targeted as the REA and the contiguity approach showed that small increases in the national health services can result in a high proportion of the population covered (Table 1). This finding can help identify regions that could carry a high "return on investment" to limit the number of people living far from AST laboratories and food animals living far from veterinarians. In addition, both case studies provided evidence that increasing health services evenly by administrative division can lead to less population covered than using geographic approaches targeting directly the population not covered (*66*). For example, in Kasungu (Malawi), only 2 candidates (out of 66 available) were chosen to be equipped for AST. However, they were sufficient for covering the highest population among all country regions. A similar situation was observed in Tambacounda (Senegal). This result was also confirmed in the case study about veterinarians, where the administrative approach always covered fewer food animals than the contiguity approach.

Investigation into individual animal species showed that the contiguity approach removed 26.9% of the world's cattle, 34% of chickens, and 44% of pigs living in veterinary coldspots. This result is a low-hanging fruit to improve the livelihood of the 1.3 billion people who rely on food animals for sustenance (*67*), improve animal welfare (*68*), and strengthen surveillance against potentially pandemic pathogens. East and Southeast Asia are the regions that would benefit the most from increased access to veterinary care for chickens and pigs. Such regions were prioritized by the contiguity approach because, according to the GLW4, they are among the ones with the highest density of chickens and pigs in the world (*50*). For cattle, Brazil resulted to be the country with the highest priority for supplementary veterinarians. However, a lack of access to care for the animals in this country could also depend on an overall lack of incentives for health workers to move to rural areas (*69*), rather than a shortage of veterinary capacities.

Unlike for veterinarians, the increase in AST laboratories showed that the distribution of the initial supplementary laboratories (Fig. 1, black dots) was not concentrated in specific regions. This underscores the widespread shortage of AST laboratories in many African countries (*13*,

70). Therefore, the first health facilities selected by the REA to be equipped for AST established a first robust network of laboratories, rather than simply enhancing an existing one. This is confirmed also by the population covered in some regions (fig. S3 and S5). For instance, in Burkina Faso, the Est region has ~500,000 more inhabitants than the Sahel region. However, in the Est region, fewer candidates were equipped for AST, and fewer people were covered than in the Sahel region, suggesting that the Est region has fewer candidates to evaluate for the scaling up analysis. Although beyond the scope of our study, these results suggest that efforts of international funders should also be directed to strengthening the overall network of health facilities available in Africa before equipping them for AST (*7*). When dealing with accessibility, this means considering also other aspects of the multifaceted challenge of access to care. In this context, travel time represents only one of its dimensions, which are represented by the costs to equip new laboratories, the availability of trained personnel, and regulatory hurdles.

Limitations

As for any modeling study, our analysis comes with limitations. First, due to restrictions on the granularity of the data shared by the African Society for Laboratory Medicine, the number of AST tests conducted by each laboratory could not be considered to assign them different "treatment capacities". Similarly, for veterinarians, we could not access the number of veterinarians sharing the same practice location and being able to travel to multiple coldspots. Second, addresses of veterinarians working exclusively for government or private companies might not be available on online platforms, which could result in an underestimation of the true number of veterinarians, and hence a different distribution of veterinary coldspots. In addition, all locations of veterinarians used to calculate coldspots have been predicted through geospatial models, and hence such predictions are associated with a degree of uncertainty (see Chapter 3). Third, the need for supplementary veterinarians for cattle in Brazil (Fig. 3) might depend on their limited presence on online registries that we used to outline their distribution (see Chapter 3). This could depend on the fact that veterinarians who have as sole clients a single large farm – a common business structure in Brazil (*71*) – may have little incentive to register on online platforms. Finally, the population covered by increasing the number of AST laboratories may vary depending on the source used for the human population map. In such a context, Hierink and colleagues showed that differences in accessibility could exceed 70% in large and sparsely populated administrative units (*43*). Therefore, future research efforts could compare the accessibility to AST laboratories when using population maps provided by services different from WorldPop, like LandScan (*72*) and Gridded Population of the World (*73*). Similarly, for countries where high-resolution friction surfaces (e.g., $\leq$100 m^2) are available, comparing the travel time maps that can be calculated through such friction surfaces could highlight differences in terms of population coverage and support the development of more accurate and harmonized friction surfaces.

References

1. United Nations. Transforming our world: The 2030 Agenda for sustainable development. (2015). doi:10.1201/b20466-7
2. F. Hierink, E. A. Okiro, A. Flahault, N. Ray. The winding road to health: A systematic scoping review on the effect of geographical accessibility to health care on infectious diseases in low- And middle-income countries. *PLoS ONE* **16**, 1–15 (2021). doi:10.1371/journal.pone.0244921
3. S. N. Grief, J. P. Miller. Infectious Disease Issues in Underserved Populations. *Physician Assistant Clinics* **4**, 107–125 (2019). doi:10.1016/j.cpha.2018.08.006
4. E. N. Hulland, K. E. Wiens, S. Shirude, J. D. Morgan, A. Bertozzi-Villa, T. H. Farag, N. Fullman, M. U. G. Kraemer, M. K. Miller-Petrie, V. Gupta, et al. Travel time to health facilities in areas of outbreak potential: Maps for guiding local preparedness and response. *BMC Medicine* **17**, 1–16 (2019). doi:10.1186/s12916-019-1459-6
5. P. W. Gething, F. A. Johnson, F. Frempong-Ainguah, P. Nyarko, A. Baschieri, P. Aboagye, J. Falkingham, Z. Matthews, P. M. Atkinson. Geographical access to care at birth in Ghana: A barrier to safe motherhood. *BMC Public Health* **12**, 1 (2012). doi:10.1186/1471-2458-12-991
6. International Road Federation. Low income countries account for only 3% of global road networks – A constraint for economic growth (2018). Available at: https://irfnet.ch/2018/12/17/low-income-countries-account-for-only-3-of-global-road-networks-a-constraint-for-economic-growth/. Accessed: 04/11/2023.
7. G. Falchetta, A. T. Hammad, S. Shayegh. Planning universal accessibility to public health care in sub-Saharan Africa. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 31760–31769 (2020). doi:10.1073/pnas.2009172117
8. C. Kelly, C. Hulme, T. Farragher, G. Clarke. Are differences in travel time or distance to healthcare for adults in global north countries associated with an impact on health outcomes? A systematic review. *BMJ Open* **6**, 1–9 (2016). doi:10.1136/bmjopen-2016-013059
9. P. P. Simões, R. M. V. R. Almeida. Maternal mortality and accessibility to health services by means of transit-network estimated traveled distances. *Maternal and Child Health Journal* **18**, 1506–1511 (2014). doi:10.1007/s10995-013-1391-x
10. T. Eryando, D. Lasut, B. Rachmat, G. Putro, M. Widiyanti. Influence of travel time on infant mortality rate in Bali province, Indonesia. *Journal of Southwest Jiaotong University* **58**, 532–542 (2023). doi:10.35741/issn.0258-2724.58.3.45
11. J. A. Ayukekbong, M. Ntemgwa, A. N. Atabe. The threat of antimicrobial resistance in developing countries: Causes and control strategies. *Antimicrobial Resistance and Infection Control* **6**, 1–8 (2017). doi:10.1186/s13756-017-0208-x
12. C. J. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022). doi:10.1016/S0140-6736(21)02724-0
13. Mapping Antimicrobial Resistance and Antimicrobial Use Partnership. The crisis within the crisis (2023). Available at: https://aslm.org/wp-content/uploads/2022/09/ASLM_MAAP-Policy-Brief_Embargoed-until-15-Sept-6AM-GMT.pdf?x26552. Accessed: 03/11/2023.
14. Food and Agriculture Organization. Livestock sector investment and policy toolkit (lsipt), making responsible decisions (2019). Available at: https://www.fao.org/3/ca6335en/CA6335EN.pdf. Accessed: 04/11/2023.

15. M. Herrero, D. Grace, J. Njuki, N. Johnson, D. Enahoro, S. Silvestri, M. C. Rufino. The roles of livestock in developing countries. *Animal* **7**, 3–18 (2013). doi:10.1017/S1751731112001954
16. World Health Organization. World Bank and WHO: Half the world lacks access to essential health services, 100 million still pushed into extreme poverty because of health expenses (2017).
17. A. T. Murray. Maximal Coverage Location Problem: Impacts, Significance, and Evolution. *International Regional Science Review* **39**, 5–27 (2016). doi:10.1177/0160017615600222
18. ArcMap. Location-allocation analysis (2021). Available at: https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm. Accessed: 31/10/2023.
19. H. Smith, D. Cakebread, M. Battarra, B. Shelbourne, N. Cassim, L. Coetzee. Location of a hierarchy of HIV/AIDS test laboratories in an inbound hub network: Case study in South Africa. *Journal of the Operational Research Society* **68**, 1068–1081 (2017). doi:10.1057/s41274-017-0240-5
20. C. Ochoa, M. Rai, S. Babo Martins, G. Alcoba, I. Bolon, R. Ruiz de Castañeda, S. K. Sharma, F. Chappuis, N. Ray. Vulnerability to snakebite envenoming and access to healthcare in the Terai region of Nepal: a geospatial analysis. *The Lancet Regional Health - Southeast Asia* **9**, 100103 (2023). doi:10.1016/j.lansea.2022.100103
21. Wikipedia. Optimal facility location (2023). Available at: https://en.wikipedia.org/wiki/Optimal_facility_location. Accessed: 15/11/2023.
22. R. Wei. Coverage Location Models: Alternatives, Approximation, and Uncertainty. *International Regional Science Review* **39**, 48–76 (2016). doi:10.1177/0160017615571588
23. H. Pirkul, D. A. Schilling. The Maximal Covering Location Problem with Capacities on Total Workload. *Management Science* **37**, 233–248 (1991). doi:10.1287/mnsc.37.2.233
24. World Health Organization. AccessMod 5 - Supporting Universal Health Coverage by modelling physical accessibility to health care (2023). Available at: https://www.accessmod.org. Accessed: 08/11/2023.
25. One Health Trust. Mapping Antimicrobial Resistance and Antimicrobial Use Partnership (2023). Available at: https://onehealthtrust.org/projects/mapping-antimicrobial-resistance-and-antimicrobial-use-partnership-maap/. Accessed: 02/11/2023.
26. Mapping AMR and AMU Partnership. National situation of antimicrobial resistance and consumption analysis from 2016-2018 - Senegal (2023). Available at: https://aslm.org/flip-books/SENEGAL-ENGLISH/REPORT-SENEGAL.html. Accessed: 02/11/2023.
27. Mapping AMR and AMU Partnership. National situation of antimicrobial resistance and consumption analysis from 2016-2018 - Sierra Leone (2023). Available at: https://aslm.org/flip-books/SIERRA LEONE/REPORT- SIERRA-LEONE-PRINT.html. Accessed: 02/11/2023.
28. Mapping AMR and AMU Partnership. National situation of antimicrobial resistance and consumption analysis from 2016-2018 - Gabon (2023). Available at: https://aslm.org/flip-books/GABON/REPORT-GABON-PRINT.html. Accessed: 02/11/2023.
29. Mapping AMR and AMU Partnership. National situation of antimicrobial resistance and consumption analysis from 2016-2018 - Burkina Faso (2023). Available at: https://aslm.org/flip-books/BURKINA FASO/REPORT-BURKINA-FASO.html.

Accessed: 02/11/2023.
30. Mapping AMR and AMU Partnership. National situation of antimicrobial resistance and consumption analysis from 2016-2018 - Malawi (2023). Available at: https://aslm.org/flip-books/MALAWI/REPORT-MALAWI-PRINT.html. Accessed: 02/11/2023.
31. African Society for Laboratory Medicine. MAAP Country Reports (2023). Available at: https://aslm.org/what-we-do/maap/maap-country-reports/. Accessed: 01/11/2023.
32. D. Kahle, H. Wickham. ggmap: Spatial visualization with ggplot2. *The R Journal* **5**, 144–161 (2013). doi:10.32614/rj-2013-014
33. healthsites.io - Building an open data commons of health facility data with OpenStreetMap (2023). Available at: https://healthsites.io. Accessed: 18/05/2023.
34. M. Padgham, R. Lovelace, M. Salmon, B. Rudis. Osmdata. *The Journal of Open Source Software* **2**, 305 (2017). doi:10.21105/joss.00305
35. Testing. Where Lab Tests Are Performed (2021). Available at: https://www.testing.com/articles/where-lab-tests-are-performed/. Accessed: 01/11/2023.
36. Database of Global Administrative Areas (GADM). Available at: https://gadm.org. Accessed: 01/02/2022. (2018).
37. USGS. USGS EROS Archive - Digital Elevation - Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global (2018). Available at: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1. Accessed: 03/11/2023.
38. AccessMod 5. 5.3. Using the 'Projects' module: creating and managing projects in AccessMod (2023). Available at: https://doc-accessmod.unepgrid.ch/pages/viewpage.action?pageId=819813. Accessed: 08/11/2023.
39. Buchhorn, M. Copernicus Global Land Service: Land Cover 100m, epoch "2019", Globe (2020). Available at: https://lcviewer.vito.be/download. Accessed: 08/11/2023.
40. Université de Genève. GeoHealth group (2023). Available at: https://www.unige.ch/environnement/en/hubs/digital-sciences-hub/research-groups/geohealth/. Accessed: 02/11/2023.
41. OpenStreetMap Foundation. OpenStreetMap (2023). Available at: https://www.openstreetmap.org. Accessed: 01/12/2021.
42. D. J. Weiss, A. Nelson, C. A. Vargas-Ruiz, K. Gligorić, S. Bavadekar, E. Gabrilovich, A. Bertozzi-Villa, J. Rozier, H. S. Gibson, T. Shekel, et al. Global maps of travel time to healthcare facilities. *Nature Medicine* **26**, 1835–1838 (2020). doi:10.1038/s41591-020-1059-1
43. F. Hierink, G. Boo, P. M. Macharia, P. O. Ouma, P. Timoner, M. Levy, K. Tschirhart, S. Leyk, N. Oliphant, A. J. Tatem, et al. Differences between gridded population data impact measures of geographic access to healthcare in sub-Saharan Africa. *Communications Medicine* **2**, 1–13 (2022). doi:10.1038/s43856-022-00179-4
44. University of Southampton. WorldPop - Open Spatial Demographic Data and Research (2023). Available at: https://www.worldpop.org. Accessed: 03/11/2023.
45. American College of Surgeon. ATLS - Advanced Trauma Life Support Program for Doctors. *American College of Surgeons* (2008).
46. M. Marconcini, A. Metz-Marconcini, S. Üreyen, D. Palacios-Lopez, W. Hanke, F. Bachofer, J. Zeidler, T. Esch, N. Gorelick, A. Kakarla, et al. Outlining where humans live, the World Settlement Footprint 2015. *Scientific Data* **7**, 1–14 (2020). doi:10.1038/s41597-020-00580-5
47. D. J. Weiss, A. Nelson, K. Gligorić, S. Bavadekar, E. Gabrilovich, J. Rozier, H. S.

Gibson, T. Shekel, C. Kamath, A. Lieber, et al. Global maps of travel time to healthcare facilities. *Nature Medicine* **26**, (2020). doi:10.1038/s41591-020-1059-1
48. The Malaria Atlas Project. Accessibility to Healthcare - Motorized friction surface (2020). Available at: https://malariaatlas.org/project-resources/accessibility-to-healthcare/. Accessed: 01/10/2022.
49. M. Gilbert, G. Nicolas, G. Cinardi, T. P. Van Boeckel, S. O. Vanwambeke, G. R. W. Wint, T. P. Robinson. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data* **5**, 1–11 (2018). doi:10.1038/sdata.2018.227
50. G. Cinardi, D. Da Re, M. Gilber, T. P. Robinson, W. G. R. Wint. Gridded Livestock of the World - 2015 (GLW4). *Harvard Dataverse* (2022). doi:10.7910/DVN/LHBICE
51. N. Ray, S. Ebener. AccessMod 3.0: Computing geographic coverage and accessibility to health care services using anisotropic movemen of patients. *International Journal of Health Geographics* **7**, 1–17 (2008). doi:10.1186/1476-072X-7-63
52. P. Timoner, F. Hierking, L. Baecher. inAccessMod: Get and prepare the required inputs for AccessMod. (2023).
53. M. Buchhorn, L. Smets, B. Bertels, B. De Roo, M. Lesiv, N. E. Tsendbazar, L. Linlin, A. Tarko. Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015-2019: Product User Manual. *Copernicus Global Land Operations* 1–93 (2020). doi:10.5281/zenodo.3938963.PU
54. E. W. Dijkstra. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik* **1**, 269–271 (1959). doi:10.1007/BF01386390
55. F. Iyun. Hospital service areas in Ibadan city. *Social Science and Medicine* **17**, 601–616 (1983). doi:10.1016/0277-9536(83)90304-0
56. P. M. Macharia, N. Ray, E. Giorgi, E. A. Okiro, R. W. Snow. Defining service catchment areas in low-resource settings. *BMJ Global Health* **6**, 4–7 (2021). doi:10.1136/bmjgh-2021-006381
57. T. Duong. Statistical visualisation for tidy and geospatial data in R via kernel smoothing methods in the eks package. *arXiv* 1–26 (2023).
58. J. van Etten. R Package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software* **76**, (2017). doi:10.18637/jss.v076.i13
59. W. M. Jang, J. Lee, S. J. Eun, J. Yim, Y. Kim, M. Y. Kwak. Travel time to emergency care not by geographic time, but by optimal time: A nationwide cross-sectional study for establishing optimal hospital access time to emergency medical care in South Korea. *PLoS ONE* **16**, 1–11 (2021). doi:10.1371/journal.pone.0251116
60. E. B. Lerner, R. M. Moscati. The golden hour: Scientific fact or medical 'urban legend'? *Academic Emergency Medicine* **8**, 758–760 (2001). doi:10.1111/j.1553-2712.2001.tb00201.x
61. World Health Organization. Monitoring emergency obstetric care - A handbook. 1–164 (2007).
62. D. J. Weiss, A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, E. Poyart, S. Belchior, N. Fullman, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018). doi:10.1038/nature25181
63. G. Watmough, M. Hagdorn, J. Brumhead. Tanzania friction surface, 2016-2020 [dataset]. *Data for Children Collaborative with UNICEF and University of Edinburgh, School of Geosciences* (2021). doi:10.7488/ds/3089
64. G. Carrasco-Escobar, E. Manrique, K. Tello-Lizarraga, J. J. Miranda. Travel Time to Health Facilities as a Marker of Geographical Accessibility Across Heterogeneous Land Coverage in Peru. *Frontiers in Public Health* **8**, 1–10 (2020).

doi:10.3389/fpubh.2020.00498
65. C. Ham. Priority setting in health care: Learning from international experience. *Health Policy* **42**, 49–66 (1997). doi:10.1016/S0168-8510(97)00054-7
66. S. S. R. Shariff, N. H. Moin, M. Omar. Location allocation modeling for healthcare facility planning in Malaysia. *Computers and Industrial Engineering* **62**, 1000–1010 (2012). doi:10.1016/j.cie.2011.12.026
67. WOAH. GBADs - The Global Burden of Animal Diseases (2023). Available at: https://gbads.woah.org. Accessed: 12/10/2023.
68. A. Radyowijati, H. Haak. Improving antibiotic use in low-income countries: An overview of evidence on determinants. *Social Science and Medicine* **57**, 733–744 (2003). doi:10.1016/S0277-9536(02)00422-7
69. P. F. A. da Silva, H. E. Shimizu, M. N. Sanchez, M. C. Ramos. Analysis of the distribution of the health workforce in Brazil. *Research, Society and Development* **11**, 1–12 (2022). doi:10.33448/rsd-v11i8.30992
70. World Health Organization. The African Health Monitor: Key Determinants for Health in the African Region. **1**, 1–194 (2013).
71. A. Sen, M. Chander. Privatization of veterinary services in developing countries: A review. *Tropical Animal Health and Production* **35**, 223–236 (2003). doi:10.1023/A:1023343429498
72. J. E. Dobsonn, E. A. Bright, P. R. Coleman, R. C. Durfee, B. A. Worley. LandScan: A Global Population Databbase for Estimating Population at Risk. *PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING* **66**, 849–857 (2000). doi:10.1201/9781482264678-24
73. Center for International Earth Science Information Network (CIESIN). Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11 (2018). Available at: https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11. Accessed: 12/11/2023.
74. B. Y. Chen, X. P. Cheng, M. P. Kwan, T. Schwanen. Evaluating spatial accessibility to healthcare services under travel time uncertainty: A reliability-based floating catchment area approach. *Journal of Transport Geography* **87**, 102794 (2020). doi:10.1016/j.jtrangeo.2020.102794
75. P. M. Macharia, N. Ray, C. W. Gitonga, R. W. Snow, E. Giorgi. Combining school-catchment area models with geostatistical models for analysing school survey data from low-resource settings: Inferential benefits and limitations. *Spatial Statistics* **51**, 100679 (2022). doi:10.1016/j.spasta.2022.100679

Supplementary materials

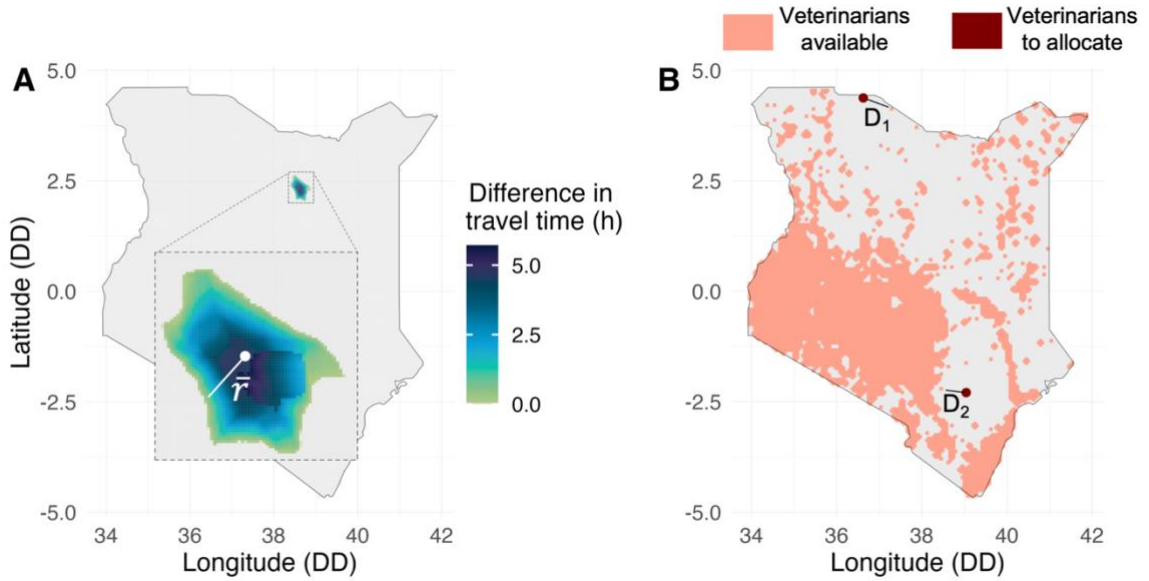

Fig. S1. Intermediate steps of the contiguity approach. (A) We defined a catchment area (*74*, *75*) by subtracting from the initial map of coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, the updated map of coldspots obtained after the allocation of 1 supplementary veterinarian (white dot). The blue area left shows the extension of the catchment area affected by a reduction in travel time. By considering the external perimeter of the catchment area we calculated its average radius ($\bar{r}$, white segment) and repeated this operation for 100 catchment areas defined from randomly selected candidates.

(B) The map of existing veterinarians reports two supplementary veterinarians that can be selected for a new allocation through the contiguity approach. Besides prioritizing areas with a high number of food animals living in coldspots, the approach also evaluates which veterinarian to allocate is the farthest from veterinarians already present in the country. For this reason, the approach calculates the minimum distance of each candidate from predicted veterinarians and then gives priority to the one for which this distance is the highest. In our example, the minimum distance D_1 is bigger than D_2 and hence the veterinarian in the North of Kenya (D_1) will have allocation priority on the one in the South.

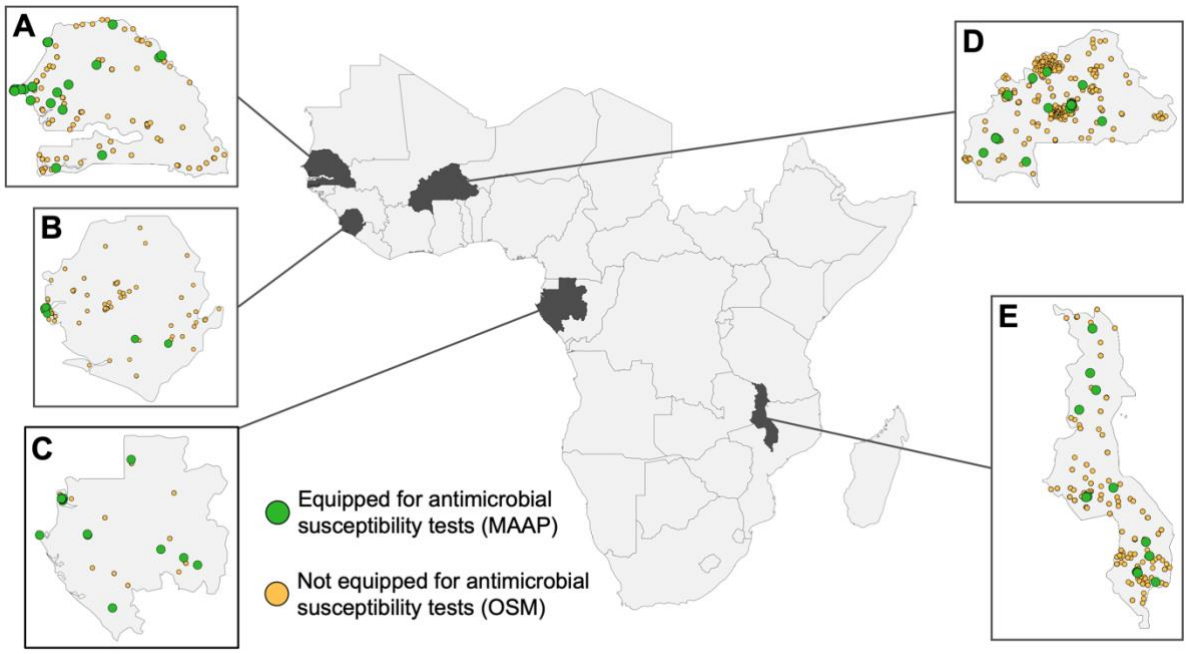

Fig. S2. Distribution of health facilities in 5 African countries. The dots represent health facilities (hospitals and diagnostic laboratories) equipped (green dots) and not equipped (yellow dots) for perform antimicrobial susceptibility testing in (A) Senegal, (B) Sierra Leone, (C) Gabon, (D) Burkina Faso, and (E) Malawi. The coordinates of the equipped health facilities were downloaded from the 2023 country reports produced within the "Mapping Antimicrobial Resistance and Antimicrobial Use" (MAAP) project (*25*), while the coordinates of not equipped health facilities were downloaded from OpenStreetMap (OSM) (*41*).

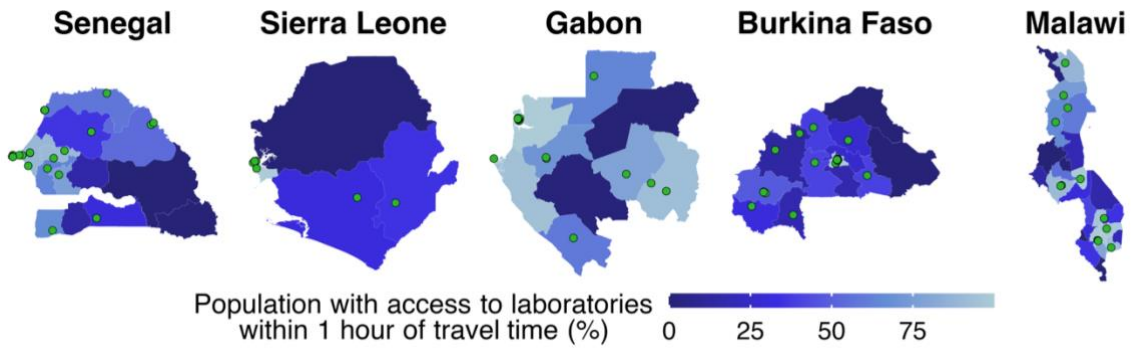

Fig. **S3**. **Population with access to existing laboratories**. Green dots represent the health facilities with laboratories for antimicrobial susceptibility testing in Senegal, Sierra Leone, Gabon, Burkina Faso, and Malawi, as reported by the 2023 country reports of the "Mapping Antimicrobial Resistance and Antimicrobial Use Partnership" project (*25*). The blue shades represent the proportion of the human population, grouped by region, with access to such laboratories within 1 hour of travel time.

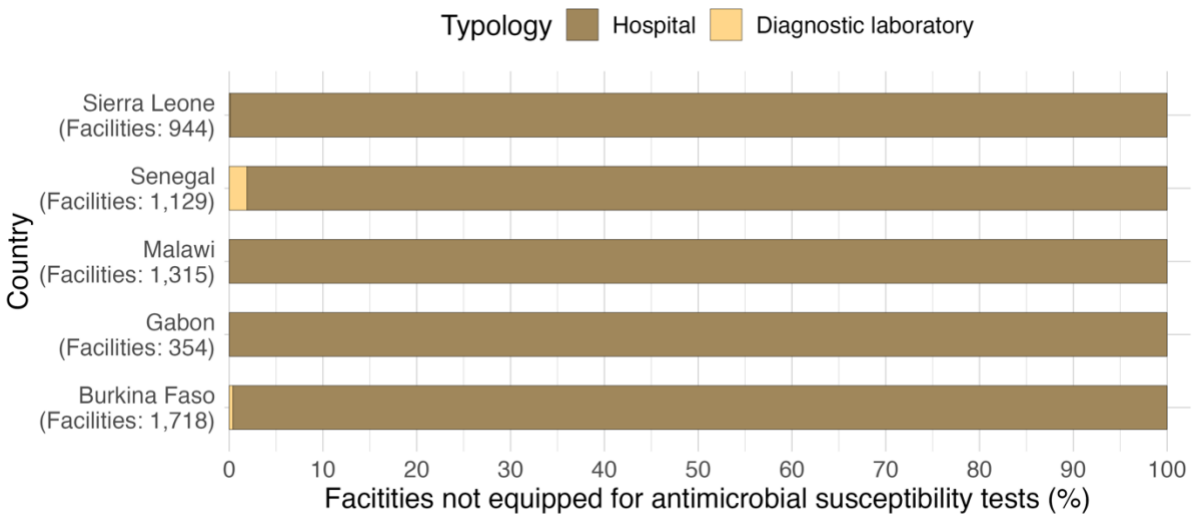

Fig. **S4**. **Typology of health facilities**. Typologies of health facilities that can be equipped with laboratories for antimicrobial susceptibility testing (i.e., hospitals and diagnostic laboratories) downloaded from OpenStreetMap (*41*).

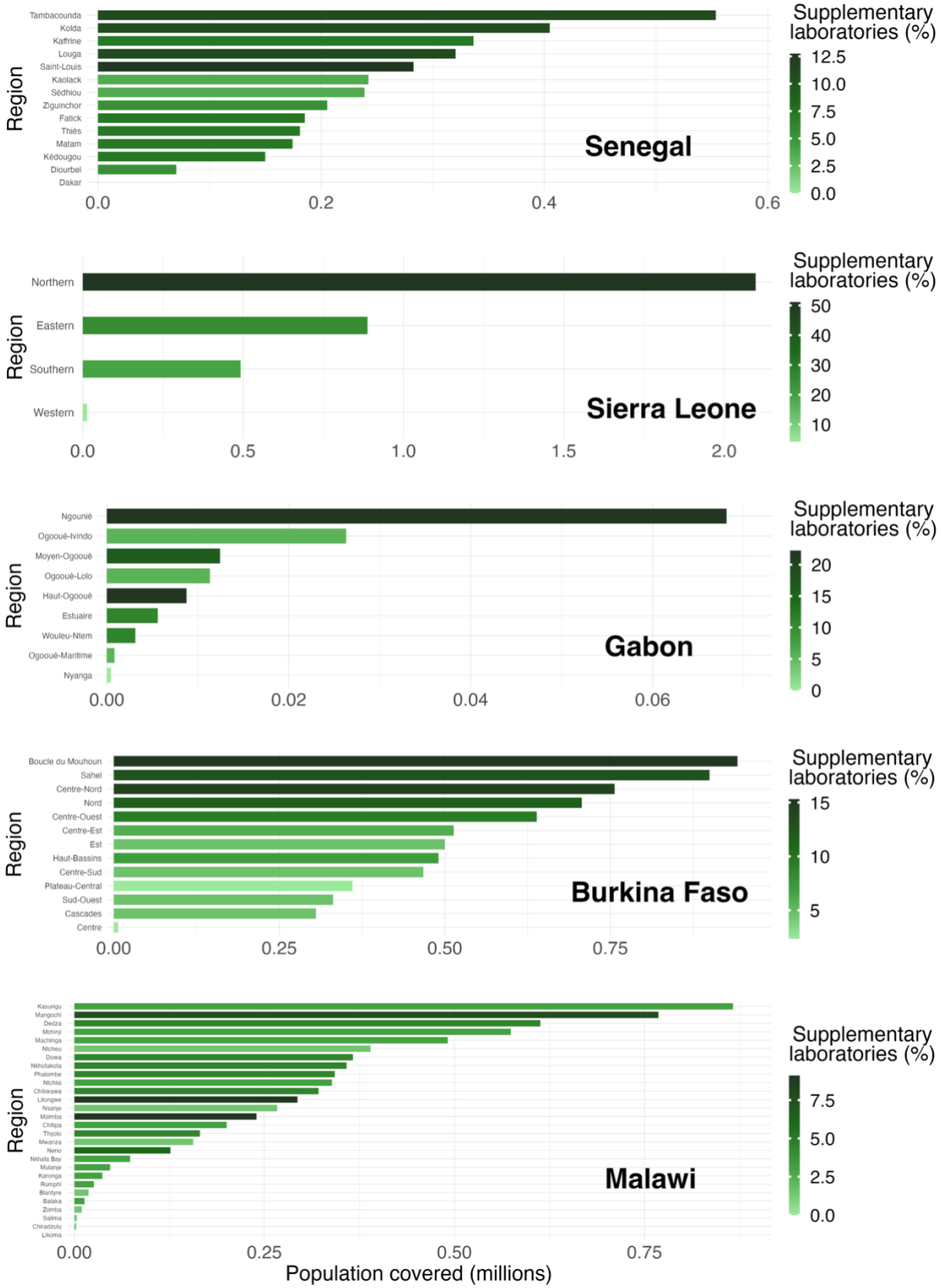

Fig. S5. Supplementary laboratories and population covered per region. Shades of green correspond to the proportion of health facilities equipped for antimicrobial susceptibility testing out of the 5% supplementary health facilities to be equipped in each country.

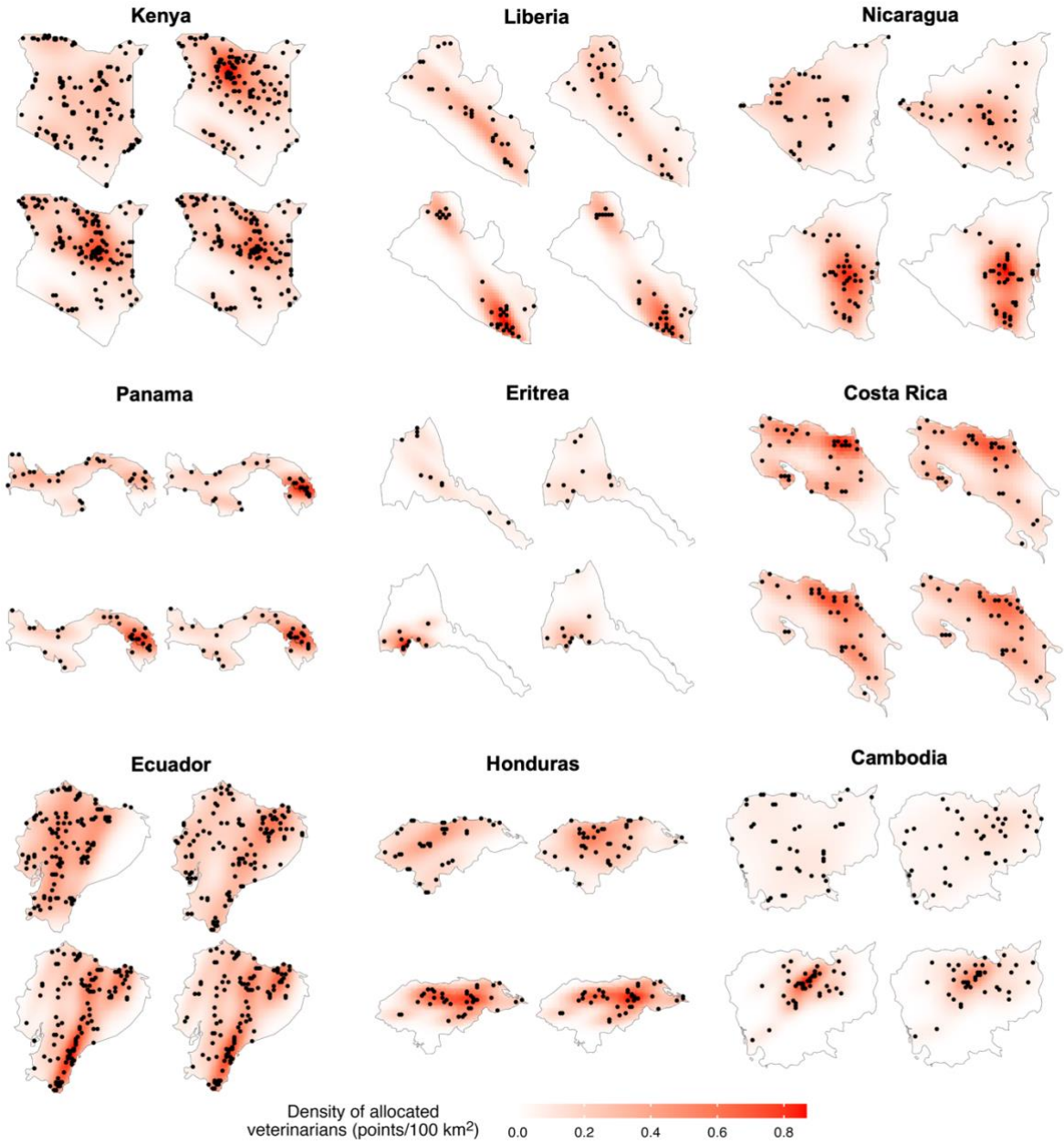

Fig. S6. Kernel Density Estimation maps of the allocated veterinarians.
The maps show the coordinates of the veterinarians allocated in the countries reported in Fig. 2. For each country-level panel, the dots represent the 5% of the supplementary veterinarians present in a country and allocated through the administrative approach (upper-left), random approach (upper-right), contiguity approach (bottom-left), and exhaustive approach (bottom-right). For the definition of each approach, see "Methods". The red patterns represent the density of supplementary veterinarians at each grid cell of ~10x10 km^2. The sets of dots chosen for the allocations performed with the random approach correspond to the Monte Carlo simulations that returned the median value of the proportion of food animals brought within 1 hour of travel time from a veterinarian with every simulation.

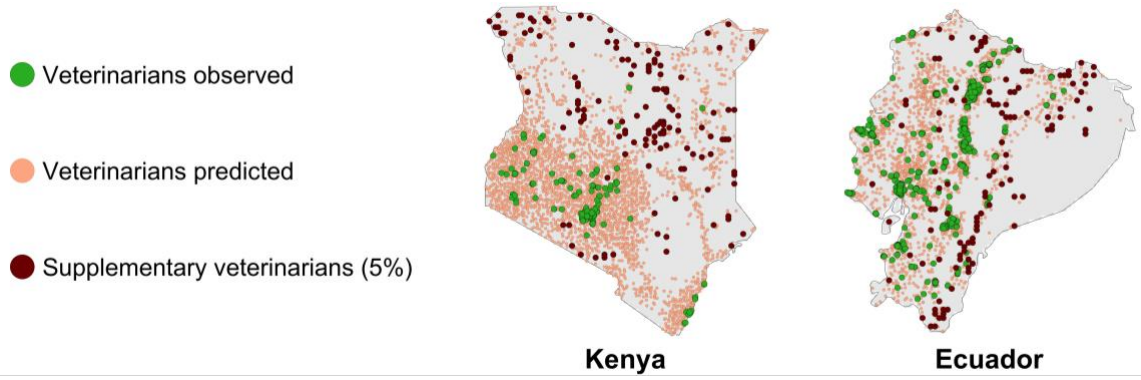

Fig. S7. Veterinarians allocated through the contiguity approach. Example of two countries with coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, where we compared the point pattern of veterinarians observed (green, see Chapter 3), veterinarians predicted (pink, see Chapter 3), and the 5% of supplementary veterinarians allocated through the contiguity approach (dark red, see "Methods").

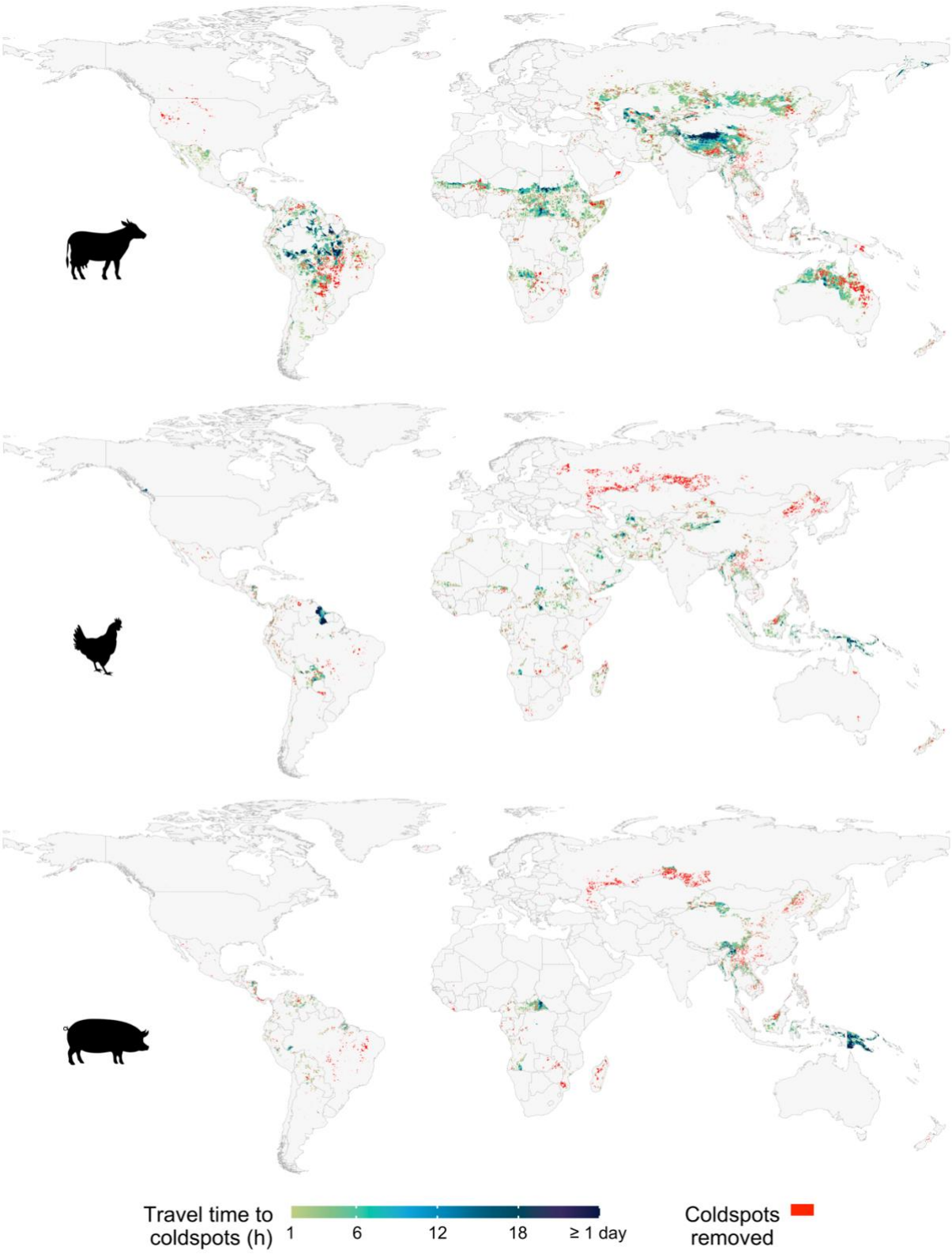

Fig. S8. Change of global coldspots upon the allocation of veterinarians. The maps report the change of countries' coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian (see Chapter 3 for the initial coldspots maps). Red areas represent the coldspots removed by a 5% scaling up of the veterinarians available per country through the contiguity approach (see "Methods").

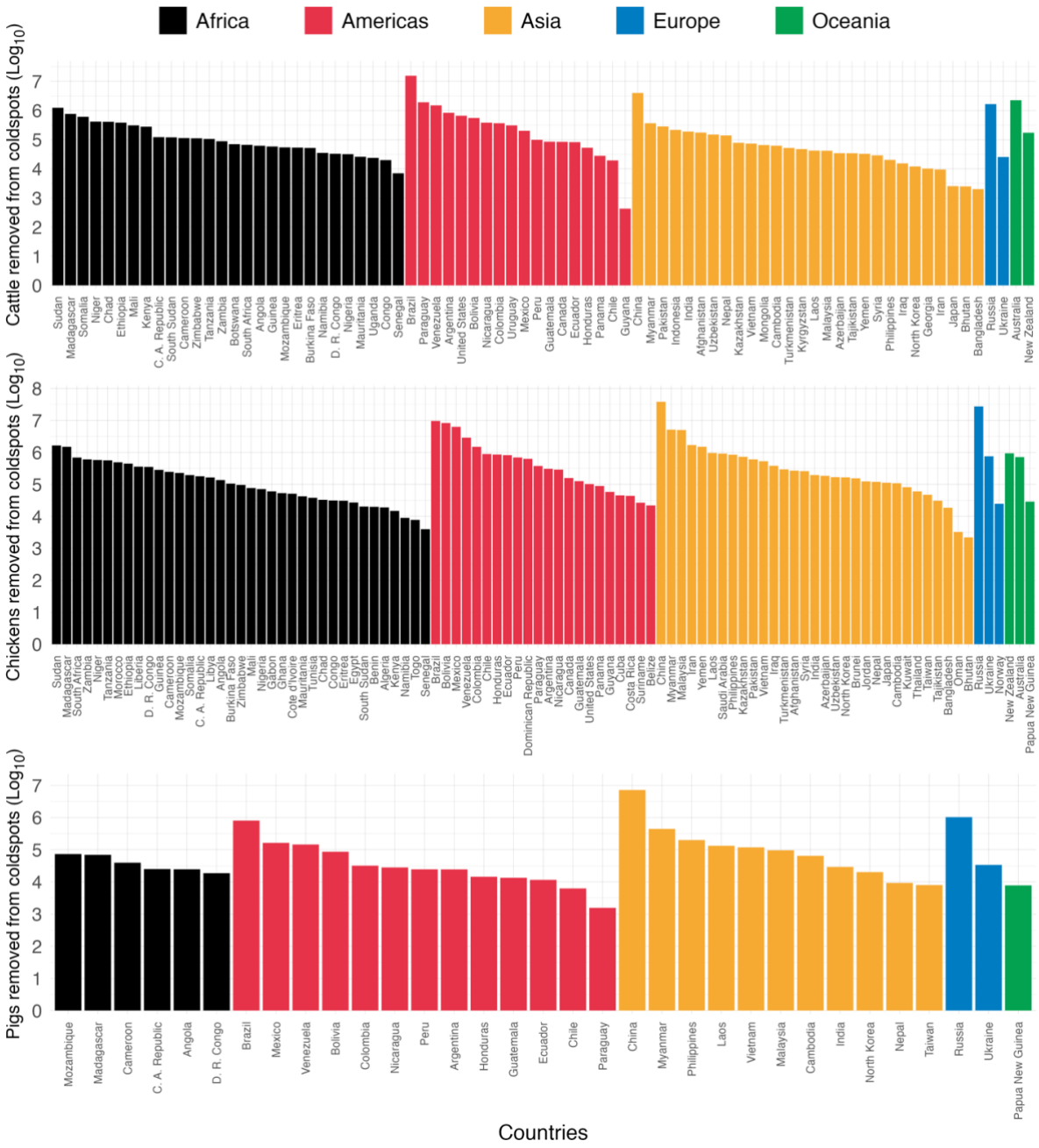

Fig. S9. Food animals removed from coldspots through the contiguity approach. The bars correspond to the Log10 number of cattle, chickens, and pigs removed from coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, by allocating a supplementary 5% of the available veterinarians per country through the contiguity approach (see "Methods")

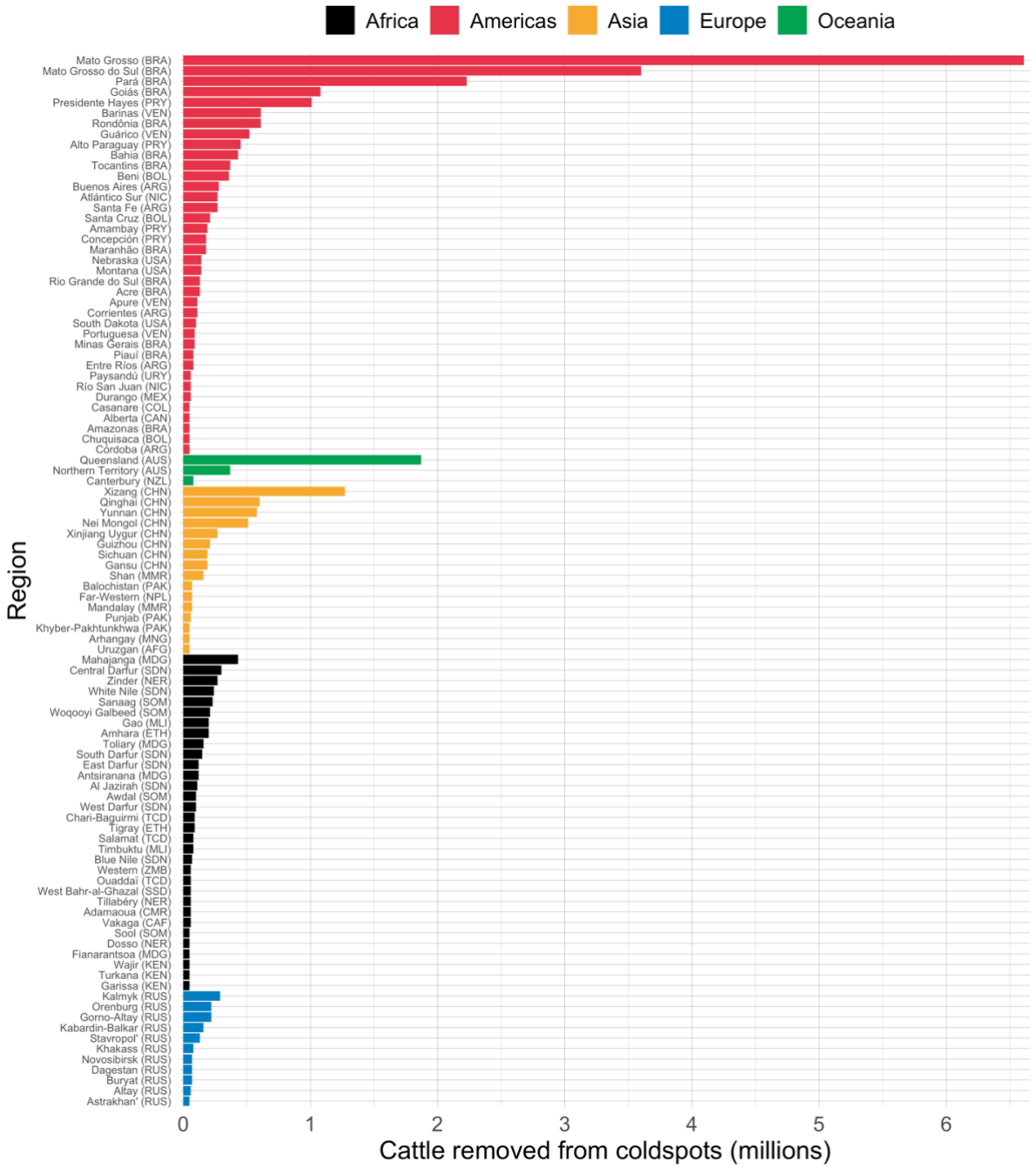

Fig. S10. Top 100 regions with the highest number of cattle removed from coldspots. The bars represent the top 100 regions, grouped by continent, with the highest number of cattle removed from coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, after the allocation of supplementary veterinarians through the contiguity approach (see "Methods"). The ISO3 of the countries to which each region belongs is reported in brackets.

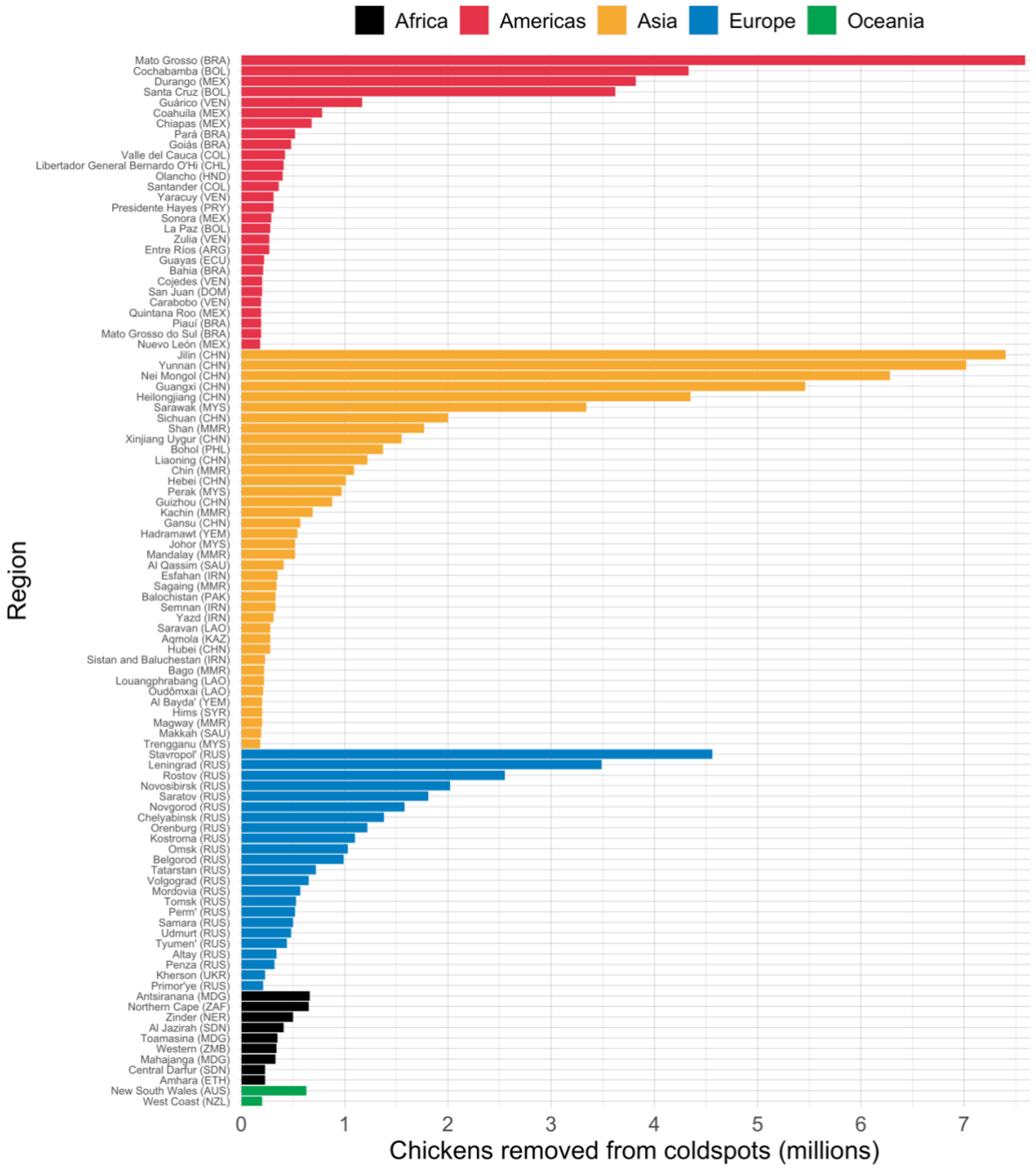

Fig. S11. Top 100 regions with the highest number of chickens removed from coldspots. The bars represent the top 100 regions, grouped by continent, with the highest number of chickens removed from coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, after the allocation of supplementary veterinarians through the contiguity approach (see "Methods"). The ISO3 of the countries to which each region belongs is reported in brackets.

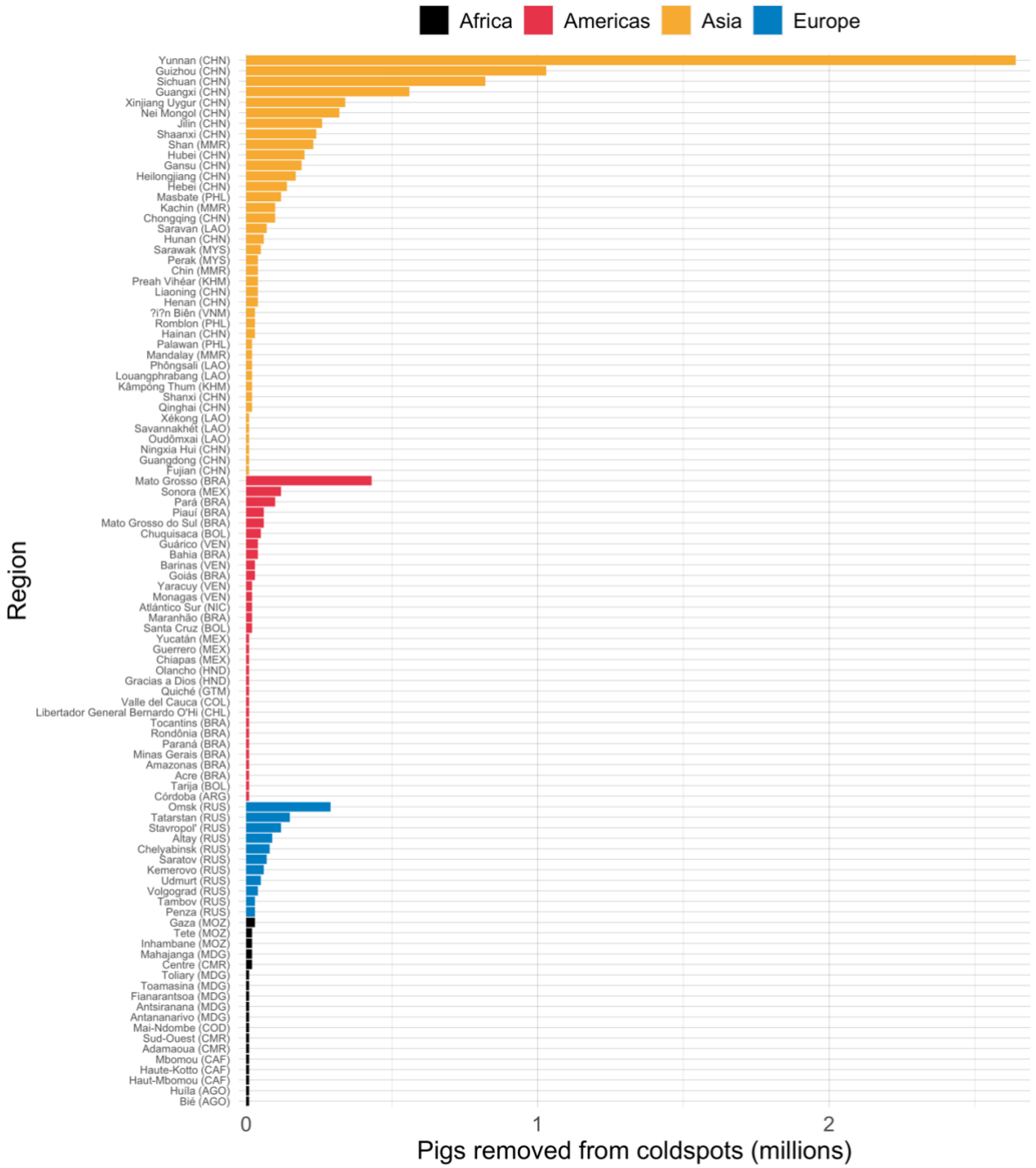

Fig. S12. Top 100 regions with the highest number of pigs removed from coldspots. The bars represent the top 100 regions, grouped by continent, with the highest number of pigs removed from coldspots, i.e., 10x10 km^2 pixels where food animals were farther than 1 hour of travel time from the nearest veterinarian, after the allocation of supplementary veterinarians through the contiguity approach (see "Methods"). The ISO3 of the countries to which each region belongs is reported in brackets.

Table S1. **Run times of the approaches for scaling up veterinarians**. Run time, in hours, of the scaling up approaches defined for the case study about the accessibility to veterinarians. Each approach was compared in nine countries (see "Methods" and Fig. 2).

Country	REA	Contiguity	Administrative	Random	REA/Contiguity
Kenya	26.6	0.394	0.0057	0.009	67.5
Ecuador	12.3	0.283	0.0051	0.04	43.5
Cambodia	7	0.237	0.0007	0.02	29.5
Nicaragua	1.7	0.103	0.0007	0.02	16.5
Liberia	1.1	0.226	0.0003	0.01	4.9
Honduras	0.7	0.143	0.0005	0.02	4.9
Panama	0.2	0.026	0.0008	0.02	7.7
Eritrea	0.2	0.081	0.0033	0.05	2.5
Costa Rica	0.1	0.031	0.0151	0.02	3.2
Average	5.5	1.7	0.003	0.03	20

Chapter 5

Open-access approaches during the COVID-19 pandemic: mapping healthcare resources and hotspots of infections

During the COVID-19 pandemic, it became evident that ensuring open-access to health data was crucial to support scientific advice for decision-making. As a member of the Health Geography and Policy Group (HEGEP), with a primary focus on mapping infectious diseases at a global scale, my work took a significant shift during this period. In recognition of the urgency and importance of the situation, from March 2020 until November 2020 I halted my ongoing projects to contribute to scientific outputs to inform decisions makers in Switzerland. The following paragraphs explain my contribution to the two scientific papers presented in this chapter.

Online platform to forecast intensive care units occupancy

One of my initial contributions involved collaborating with my research group to develop the open-access platform *icumonitoring.ch*. The objective of the platform was to map the intensive care unit (ICU) occupancy in Switzerland at the regional-, cantonal-, and hospital-level. It was programmed in the R language, using the functions of the *shiny* package (*1*), which translates the R code into HyperText Markup language (HTML) to be displayed on every internet browser. JavaScript and CSS languages were used to curate the aesthetics of the platform and format it into a dashboard. *icumonitoring.ch* was conceived to automatically display the regular updates of statistical models forecasting ICU occupancy, COVID-19-related deaths, positive COVID-19 cases, hospitalizations, and availability of ventilators. These outputs were displayed through geographical choropleth using the functions of the R package *leaflet* (*2*) for an instant overview of hospitals, cantons, and regions with a saturation of ICUs.

Since all these outputs had to be updated on a bi-weekly basis, the architecture of the platform was developed to automatically access databases stored remotely on cloud services. Specifically, *icumonitoring.ch* constantly communicated with two cloud services for accessing and storing data on remote servers:

1. polybox. This cloud service is maintained by the IT department of the ETH Zürich. Access to polybox is available (and encrypted) for every member of the university. During the pandemic, access was granted also to the members of the Information and Operation System (IES) that inventoried restricted-access datasets (updated twice a day) containing information about the ICU occupancy of every Swiss hospital. In addition, polybox was also used to store open-access data updated on the online repository OpenZH (*3*), which provided information on COVID-19 cases, deaths, hospitalizations, and availability of ventilators at the cantonal level. These datasets

were then used as inputs for the statistical models developed by other HEGEP members to produce bi-weekly forecasts of these same variables.

2. Amazon storage service (S3). The cloud service of Amazon can communicate with platforms developed using the R programming language through the R package *aws.s3* (*4*). The S3 service was used to store all the information that could be displayed online, as the tables and figures produced by the statistical models used for the forecasts. These outputs were automatically updated on S3 and accessed each time *icumonitoing.ch* was launched by the users. In this way, the platform could operate autonomously, loading only the most updated tables and figures available inside S3.

In addition, *icumonitoring.ch* was equipped with an authentication system to keep the hospital-level data provided by IES confidential. This measure ensured that the relevant authorities retained control over patient flow management in each hospital, preventing platform users from choosing where to be hospitalized, and potentially contribute to the imbalance in ICU occupancy. For this reason, in the user interface of *icumonitoring.ch*, we included a panel to input username and password. Upon the input of credentials, this information unlocked the download of hospital-level data from the encrypted remote storage service (S3) and displayed them in the user interface of the platform.

Once operational, *icumonitoring.ch* was updated on the cloud service provided by *shinyapps.io*, which is one of the products available within the R environment. This service allowed us to customize settings for the expected performance and scalability of the platform. Specifically, we defined settings to provide simultaneous access to the platform and enable navigation up to >2,500 users (at peak, the platform received 7,818 visitors in one day). After the development stage, *icumonitoring.ch* was supported by armasuisse, the procurement branch of the Swiss Armed Forces (*5*), and it was used daily for managing ICU occupancy in Switzerland until March 2022. The platform in presented in the first publication of this chapter.

Daily geocodings of COVID-19 hotspots

At the request of the FOPH, I was involved in another project to produce daily hotspots maps of COVID-19 cases in Switzerland. For this task, I used the R programming language to develop a script for geocoding new daily cases of COVID-19 in the country using Google's geocoding Application Programming Interface (API) available from the R package *ggmap* (*6*). As for the *icumonitoring.ch* project, a database of COVID-19 cases was updated daily inside the polybox storage service by members of the FOPH. Then, using the *cronjobs* functionalities of a Linux-based operative system (*7*), the R script was scheduled to run on our local server twice per day. At each run, the script automatically mixed the database with addresses of new COVID-19 cases with a database of random addresses of shops and restaurants sampled all over Europe through OpenStreetMap. This step, known as "anonymization by dilution" was crucial to prevent potential hacks in the Google API framework to rebuild the databases of COVID-19 patients in Switzerland. This daily routine enabled us to obtain precise coordinates

(longitude and latitude expressed in decimal degrees) of COVID-19 positive cases. The database was a foundational component in the creation of daily national maps to identify COVID-19 hotspots across Switzerland. This information, summarized in reports that my supervisor (Prof. Thomas Van Boeckel) disseminated with members at the FOPH (*8*), was used to monitor and identify areas at high risk of new infections, thereby facilitating targeted public health interventions and resource allocation (if deemed necessary by the competent authorities).

My work in geocoding COVID-19 cases did not stop at the identification of hotspots of COVID-19 infections. All the geocodings we performed were included in a database of >2.5 million COVID-19 tests performed in Switzerland between March 2020 and February 2021. This database was used in a broader study that assessed the number of hospitalizations, ICU admissions, deaths, and the number of COVID-19 tests performed across different socioeconomic strata of the Swiss population (Swiss-SEP) (*9*). This analysis led by Dr. Julien Riou and Prof. Matthias Egger at the University of Bern highlighted socioeconomic disparities in healthcare outcomes during the pandemic, underscoring the pressing need to enhance them in the poorest neighborhoods of the country. The findings of this analysis, to which I contributed geocoded COVID-19 cases, is presented in the second paper of this chapter.

Contribution remarks

My involvement in these two projects was mostly dedicated to operational support in developing and maintaining a framework to openly share outputs that would support decision-makers during the pandemic. Although these projects were conceived to run automatically, their architecture had to be constantly updated and reviewed throughout the whole pandemic to address new incoming requests from the FOPH, the Swiss Armed Forces, and the Swiss National COVID-19 Task Force. For example, the formats of the databases used as inputs for both *icumonitoring.ch* and the geocoding routine of new COVID-19 cases were changed over time. This implied a parallel and rapid adaptation of all the scripts used to produce the outputs described in this chapter.

However, the continuous updates made to the branches of these projects were essential in ensuring the relevance of the information provided to public health authorities throughout the pandemic. As a result, the daily forecasts provided by *icumonitoring.ch* and the maps of COVID-19 hotspots were widely disseminated, reaching not only the panel of scientific experts and decision-makers but also various media channels, such as the Swiss Radio and Television (SRF) (*10*, *11*), and regional (*12*) and national newspapers (*13*), to inform a broader audience.

References

1. W. Chang, J. Cheng, J. Allaire, Y. Xie, J. McPherson. shiny: Web Application Framework for R. R package version 1.3.2. https://CRAN.R-project.org/package=shiny. (2019).
2. J. Cheng, B. Karambelkar, X. Yihui. leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.0.3. https://CRAN.R-project.org/package=leaflet. (2019).
3. Kanton Zürich. SARS-CoV-2 open government data reported by the Swiss Cantons and the Principality of Liechtenstein (20239. Available at: https://github.com/openZH/covid_19. Accessed: 31/10/2023.
4. T. J. Leeper. aws.s3: AWS S3 Client Package. (2020).
5. FOPH. armasuisse (2023). Available at: https://www.ar.admin.ch. Accessed: 28/10/2023.
6. D. Kahle, H. Wickham. ggmap: Spatial visualization with ggplot2. *The R Journal* **5**, 144–161 (2013). doi:10.32614/rj-2013-014
7. opensource.com. How to use cron on Linux (2021). Available at: https://opensource.com/article/21/7/cron-linux. Accessed: 28/10/2023.
8. T. P. Van Boeckel, N. G. Criscuolo. COVID-19 in Switzerland: a geographic approach to targeted testing. 1–3 (2020).
9. J. Riou, R. Panczak, C. L. Althaus, C. Junker, D. Perisa, K. Schneider, N. G. Criscuolo, N. Low, M. Egger. Socioeconomic position and the COVID-19 care cascade from testing to mortality in Switzerland: a population-based analysis. *The Lancet Public Health* **6**, e683–e691 (2021). doi:10.1016/S2468-2667(21)00160-2
10. SRF. ETH-Forscher: «Schweizer Spitäler droht nächste Woche die Überlastung» (2020). Available at: https://www.srf.ch/play/tv/tagesschau/video/eth-forscher-schweizer-spitaeler-droht-naechste-woche-die-ueberlastung?urn=urn:srf:video:20508509-cde8-4c0f-95f0-0069c3ce4a31. Accessed: 28/10/2023.
11. SRF. Warten auf die Operation: Wie ungleich Corona die Spitäler bremst (2021). Available at: https://www.srf.ch/wissen/corona/knappe-intensivbetten-warten-auf-die-operation-wie-ungleich-corona-die-spitaeler-bremst. Accessed: 28/10/2020.
12. Blick. Intensivstationen im Kanton Solothurn bereits voll (2020). Available at: https://www.blick.ch/schweiz/eth-zahlen-zeigen-spitalbelegung-intensivstationen-im-kanton-solothurn-bereits-voll-id16155717.html. Accessed: 28/10/2023.
13. swissinfo.ch. Intensive care beds could run out in a week, study predicts (2020). Available at: https://www.swissinfo.ch/eng/coronavirus_intensive-care-beds-could-run-out-on-thursday--study-predicts/45647658. Accessed: 28/10/2023.

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.

†These authors contributed equally.

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

Published in:
Swiss Medical Weekly 150:w20277 (2020). DOI: 10.4414/smw.2020.20277

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 prevalence estimates [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 [7–9] while others focused on capturing and forecasting the dynamics of COVID-19 in individual Cantons [6]. 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
k	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%
ι	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

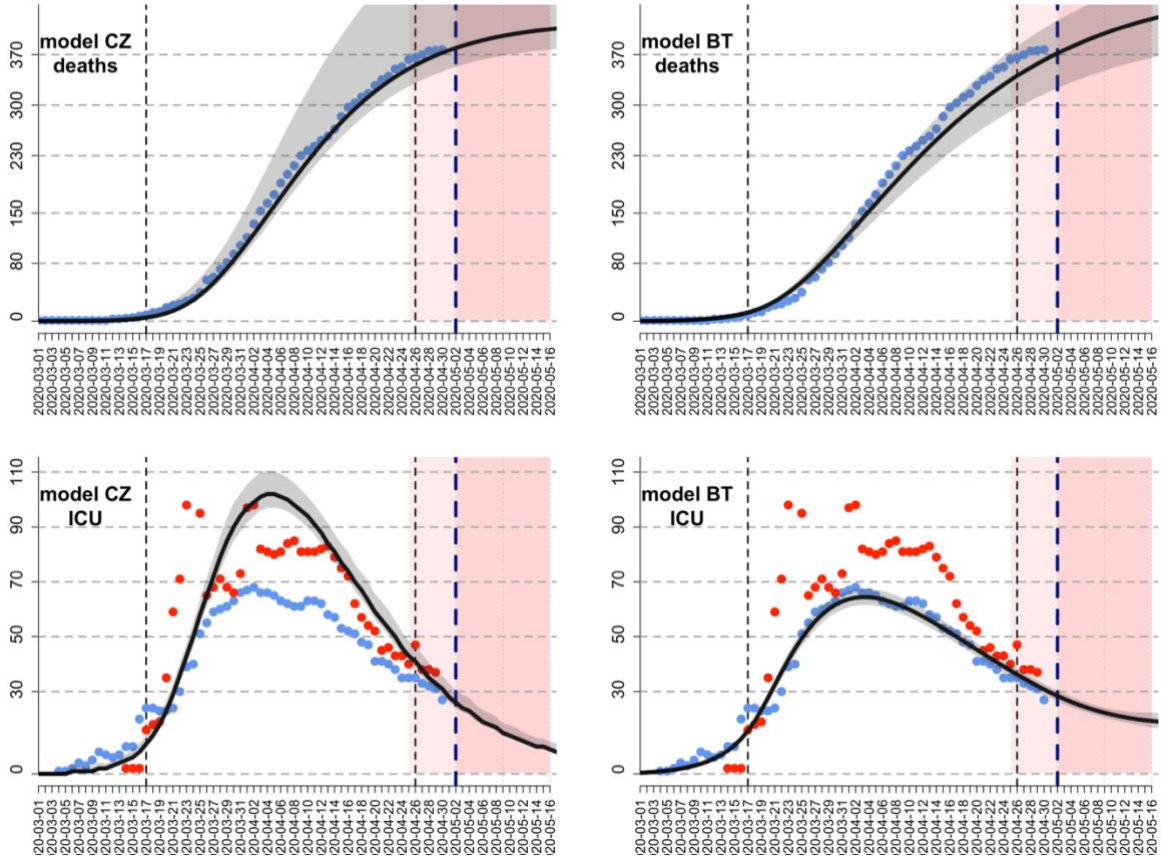

Fig. 1. Canton-level fit of epidemic model CZ and BT in Vaud on 03.05.2020. Black shades describe 95% confidence interval, red dots are ICU occupancy as reported in the IES systems,

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: $\alpha \times LOS_{deaths} + (1-\alpha) \times LOS_{recovered}$, 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: $\log10(\text{ICU}_{beds})/(\text{travel time}_{pixel->H})$, 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.

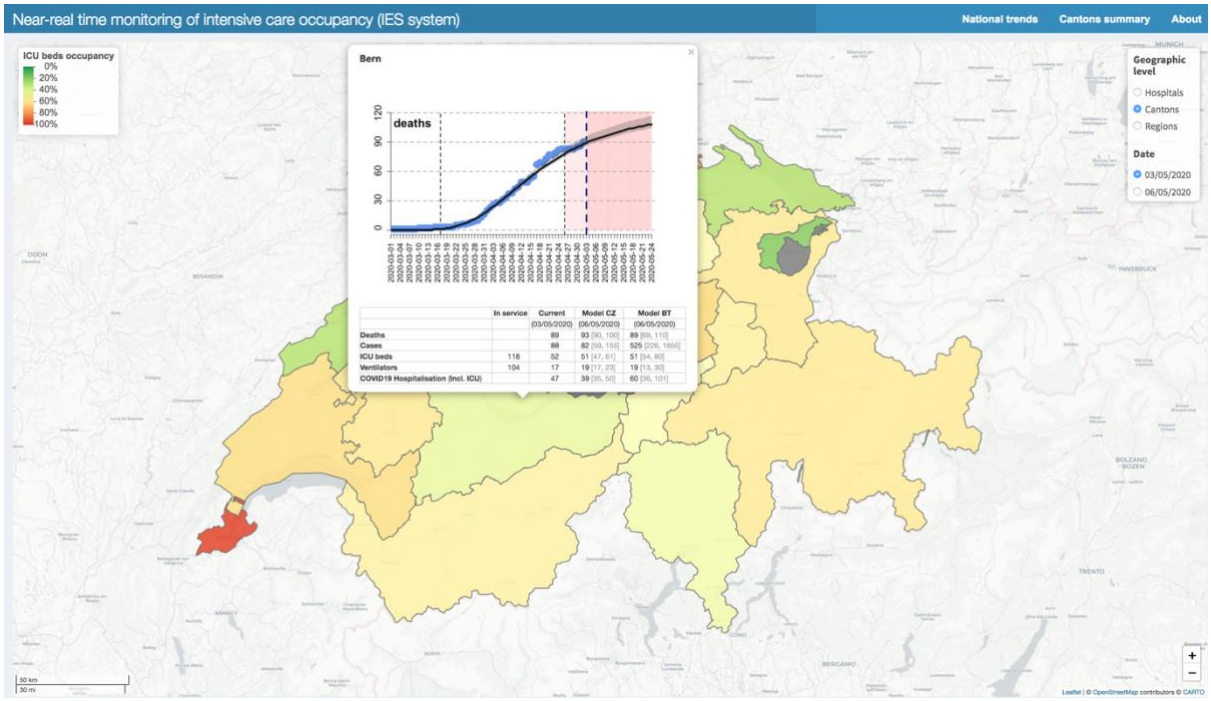

Fig 2. **Canton-level ICU occupancy, colors in other Cantons indicate ICU bed occupancy compared to the number of beds in service**. Pop-up windows indicate the situation in the Canton of Bern as reported in *icumonitoring.ch* for 03.05.2020 and projected for 06.05.2020.

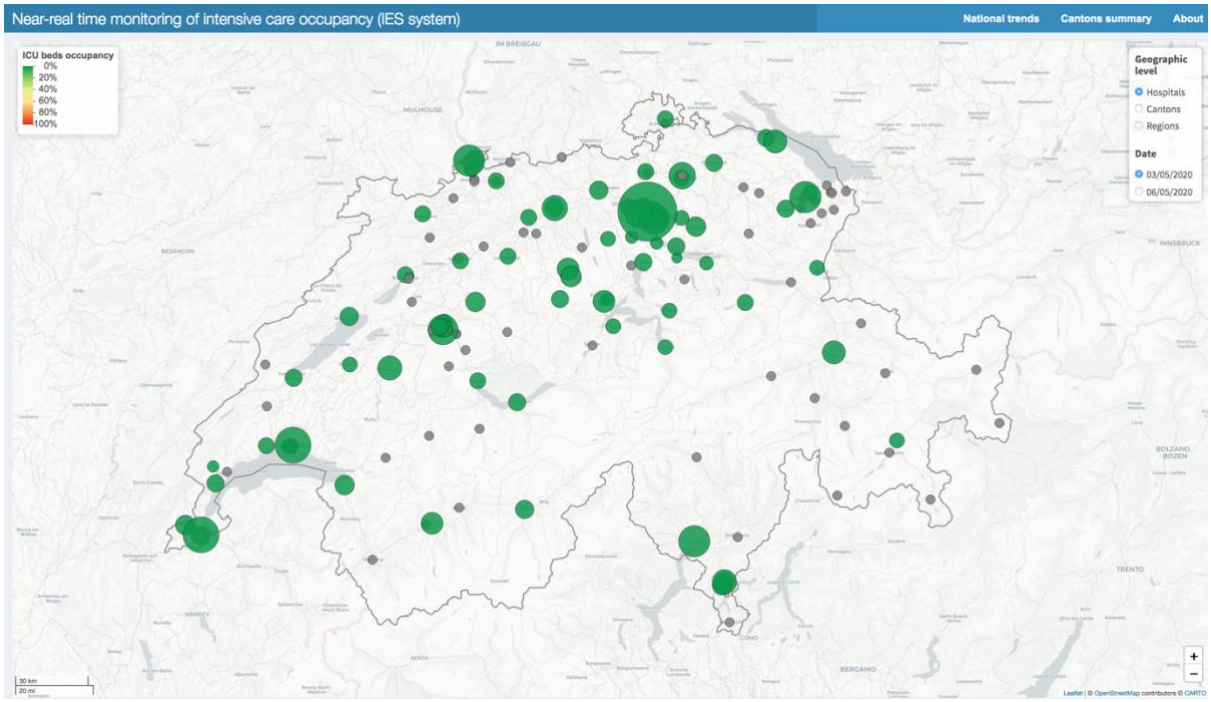

Fig 3. **Hospitals with ICU** (yellow), **and without ICU** (grey). bed occupancy in individual hospitals set to 0 artificially but available upon request to the corresponding author (bottom).

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.

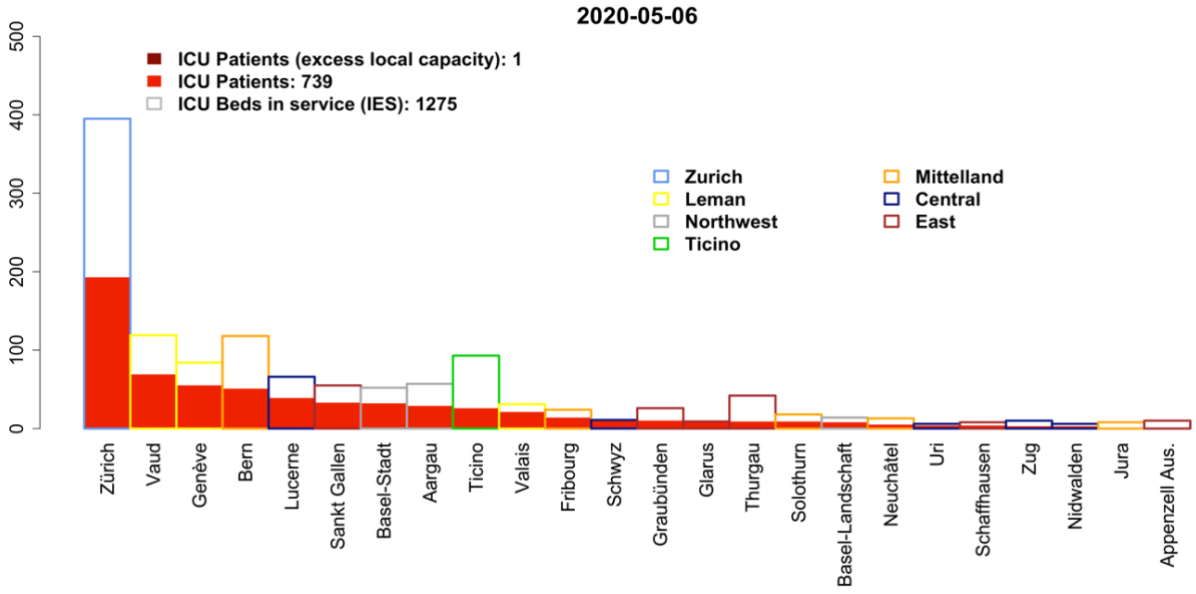

Fig 4. **Projection for ICU bed occupancy on 06.05.2020 ranked by Regions and Cantons** (model CZ).

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.

Authors contributions

Thomas Van Boeckel conceived the work, cleaned, and analyzed the data on ICU occupancy, and wrote the first version of the manuscript. Ms Cheng Zhao, and Dr Burcu Tepekule produced epidemic projections. Mr Nicola Criscuolo developed the Shiny web application. Pedro Wendel Garcia, and Matthias Hilty facilitated access to and pre-treated data from the RISC-19-ICU registry. Thierry Fumeaux facilitated access to data from SSIM and informed the model development. All authors critically revised and edited the manuscript.

Funding

This work is supported by the Vice-president for research of ETH, the Branco Weiss Foundation, and Project Horizon 2020 MOOD (MOnitoring Outbreak events for Disease surveillance in a data science context). We declare no conflicts of interest.

Acknowledgements

We thank Hans Ulrich Rothen for providing data on length of stay in ICU of patients with ARDS in the preparatory phase of this analysis on behalf of the Swiss Society of Intensive Care Medicine. We thank Dani Angst for providing access to ETH's Shiny server, as well as Katie Tiseo and Corinne Hertaeg for language editing.

RISC-19-ICU Investigators for Switzerland. Institute of Intensive Care Medicine, University Hospital Zurich, Zurich (Reto Schüpbach, MD; Philipp Bühler, MD; Silvio Brugger, MD, PhD; Jan Bartussek, PhD); Soins intensifs, Groupement Hospitalier de l'Ouest Lémanique - Hopital de Nyon, Nyon (Mallory Moret-Bochatay, MD); Interdisziplinaere Intensivstation, Spital Buelach, Buelach (Bernd Yuen, MD; Thomas Hillermann, MD); Soins Intensifs, Hopital cantonal de Fribourg, Fribourg (Hatem Ksouri, MD, PhD; Govind Oliver Sridharan, MD); Departement for intensive care medicine, Kantonsspital Nidwalden, Stans (Anette Ristic, MD; Michael Sepulcri, MD); Departement of Anesthesiology and Intensive Care Medicine, Cantonal Hospital St. Gallen, St. Gallen (Miodrag Filipovic, MD; Urs Pietsch, MD); Intensivstation, Regionalspital Emmental AG, Burgdorf (Petra Salomon, MD; Iris Drvaric, MD); Institut fuer Anesthaesie und Intensivmedizin, Zuger Kantonsspital AG, Baar (Peter Schott, MD; Severin Urech, MD); Intensivmedizin, St. Claraspital, Basel (Adriana Lambert, MD; Lukas Merki, MD); Department Intensive Care Medicine, Spitalzentrum Biel, Biel (Marcus Laube, MD); Intensivmedizin, Kantonsspital Graubünden, Chur (Frank Hillgaertner, MD; Marianne Sieber); Institut fuer Anaesthesie und Intensivmedizin, Spital Thurgau, Frauenfeld (Alexander Dullenkopf, MD; Lina Petersen, MD); Division of Neonatal and Pediatric Intensive Care, Geneva University Hospitals, Geneva (Serge Grazioli, MD; Peter C. Rimensberger, MD); Soins Intensifs, Hirslanden Clinique Cecil, Lausanne (Isabelle Fleisch, MD; Jerome Lavanchy, MD); Interdisziplinaere Intensivstation, Spital Maennedorf AG, Maennedorf (Katharina Marquardt, MD; Karim Shaikh, MD); Intensivmedizin, Schweizer Paraplegikerzentrum Nottwil, Nottwil (Hermann Redecker, MD); Intensivmedizin, Spital Oberengadin, Samedan (Michael Stephan, MD; Jan Brem, MD); Paediatric Intensive Care

Unit, Children's Hospital of Eastern Switzerland, St. Gallen (Bjarte Rogdo, MD; Andre Birkenmaier, MD); Klinik für Anaesthesie und Intensivmedizin, Spitalzentrum Oberwallis, Visp (Friederike Meyer zu Bentrup, MD, MBA); Interdisziplinaere Intensivstation, Stadtspital Triemli, Zurich (Patricia Fodor, MD; Pascal Locher, MD); Department Intensivmedizin, Universitaetsspital Basel, Basel (Martin Siegemund, MD; Nuria Zellweger); Department of Intensive Care Medicine, University Hospital Bern - Inselspital, Bern (Marie-Madlen Jeitziner, RN, PhD; Beatrice Jenni-Moser, RN, MSc); Intensivstation, Spital Grabs, Grabs (Christian Bürkle, MD); Medical ICU, Cantonal Hospital St.Gallen, St. Gallen (Gian-Reto Kleger, MD); Service d'Anesthesiologie, EHNV, Yverdon- les-Bains (Marilene Franchitti Laurent, MD; Jean-Christophe Laurent, MD); Abteilung für Anaesthesiologie und Intensivmedizin, Hirslanden Klinik Im Park, Zürich (Tomislav Gaspert, MD; Marija Jovic, MD); Intensivmedizin & Intermediate Care, Kantonsspital Olten, Olten (Michael Studhalter, MD); Institut für Anaesthesiologie und Intensivmedizin, Klinik Hirslanden, Zurich (Christoph Haberthuer, MD; Roger F. Lussman, MD); Anaesthesie Intensivmedizin Schmerzmedizin, Spital Schwyz, Schwyz (Daniela Selz, MD; Didier Naon, MD); Dipartimento Area Critica, Clinica Luganese Moncucco, Lugano (Romano Mauri, MD; Samuele Ceruti, MD); Institut für Anaesthesiologie Intensivmedizin & Rettungsmedizin, See-Spital Horgen & Kilchberg, Horgen (Julien Marrel, MD; Mirko Brenni, MD); Klinik für Operative Intensivmedizin, Kantonsspital Aarau, Aarau (Rolf Ensner, MD); Intensivstation, Kantonsspital Schaffhausen, Schaffhausen (Nadine Gehring, MD); Intensivstation, Spital Simmental-Thun-Saanenland AG, Thun (Antje Heise, MD), Division of Intensive Care, University Hospitals of Geneva, Geneva (Sara Cereghetti, MD; Filippo Boroli, MD; Jerome Pugin, MD, PhD).

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

Socioeconomic position and the cascade from SARS-CoV-2 testing to COVID-19 mortality: Analysis of nationwide surveillance data

Authors:
Julien Riou[1,2,†], Radoslaw Panczak[1,†], Christian L. Althaus[1], Christoph Junker[2], Damir Perisa[2], Katrin Schneider[2], **Nicola G. Criscuolo** [3], Nicola Low,[1] Matthias Egger [1,4,5*]

Affiliations:
[1]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
[2]Federal Office of Public Health, Liebefeld, Switzerland
[3]Department of Environmental Systems Science, ETH Zürich, Zurich, Switzerland
[4]Population Health Sciences, Bristol Medical School, University of Bristol, UK
[5]Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa

†These authors contributed equally

*Correspondence to: matthias.egger@ispm.unibe.ch

Published in:
The Lancet Public Health 6:e683–91 (2021). DOI: 10.1016/ S2468-2667(21)00160-2

Supplementary Material:
https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(21)00160-2/fulltext#supplementaryMaterial

Abstract

The inverse care law states that disadvantaged populations need more health care than advantaged populations but receive less. Gaps in COVID-19 related health care and infection control are not well understood at present. We examined inequalities in health and health care from testing for SARS-CoV-2 infection to COVID-19-related hospitalisation, ICU admission and death in Switzerland, a wealthy country that was strongly affected by the pandemic.

We used surveillance data reported to the Federal Office of Public Health from 1 March 2020 to 4 February 2021. We geocoded residential addresses of notifications to determine the Swiss neighbourhood index of socioeconomic position (Swiss-SEP). The index describes 1.27 million small neighbourhoods of about 50 households, based on rent per square meter, education and occupation of household heads, and crowding. We used negative binomial regression models to calculate incidence rate ratios (IRR) with 95% credible intervals [CrI] of the association between ten groups of the Swiss-SEP index defined by deciles (1=lowest, 10=highest) and outcomes. Models were adjusted for sex, age, canton, and wave of the epidemic (before or after June 8, 2020). We used three different denominators: the general population, the number of tests, and the number of positive tests.

Analyses were based on 2,548,638 tests, 423,656 positive tests, 17,762 hospitalisations, 1,785 ICU admissions and 6,060 deaths. Comparing the highest with the lowest Swiss-SEP group, and using the general population as the denominator, more tests were performed among people living in neighbourhoods of the highest socioeconomic position (adjusted IRR 1.21; 95%CrI 1.05-1.40). Among tested people, test positivity was lower (adjusted IRR 0.77; 0.71-0.84) in neighbourhoods of the highest socioeconomic position. Among people testing positive, the corresponding IRRs for hospitalisation, ICU admission and death were 0.67 (0.61-0.74), 0.50 (0.38-0.66) and 0.84 (0.71-1.01). The associations between neighbourhood socioeconomic position and outcomes were stronger in younger age groups, and there was heterogeneity between cantons.

The inverse care law and socioeconomic inequalities are evident in Switzerland throughout the care cascade. People living in neighbourhoods of low socioeconomic position were less likely to be tested, but more likely to test positive, be hospitalised, or die. It is essential to continue to monitor testing for SARS-CoV-2 infection, access and uptake of COVID-19 vaccination and outcomes of COVID-19. The Government and health care systems should take measures to reduce health inequalities in response to the SARS-CoV-2 pandemic.

Introduction

The pandemic of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infections has created unprecedented challenges for society and healthcare systems worldwide. With almost 40 million cases and over 1 million deaths as of mid-April 2021, Europe has been heavily affected by the pandemic.[1] Compared with neighbouring countries, Switzerland had the highest rate of confirmed COVID-19 cases, higher than France, Austria and Italy and over double the rate in Germany.[2] Similarly, there was substantial excess mortality in Switzerland

during the first wave and the highest excess mortality among neighbouring countries during the second wave[3].

Published in 1971, the inverse care law states that "the availability of good medical care tends to vary with the need for it in the population served".[4,5] Inequalities in health have been a concern for over 50 years in many regions, including in Europe.[6–8] In Switzerland, life expectancy varies between neighbourhoods, depending on the neighbourhood's socioeconomic position.[9] Health inequalities and inequities may also influence the outcomes of the COVID-19 pandemic.[10] In the UK Biobank cohort, testing positive for SARS-CoV-2 infection was related to area-level socioeconomic deprivation, lower educational level and non-white ethnicity.[11] The REal-time Assessment of Community Transmission-2 (REACT-2) study showed a higher prevalence of people with SARS-CoV-2 antibodies in neighbourhoods with high levels of social disadvantage and among minority ethnic communities.[12] Studies in the USA showed that patients from neighbourhoods or counties with lower median income or higher deprivation were more likely to require intensive care, and more likely to die.[13 14]

Inequalities and inequities in health care and infection control should be described and documented at the population level along the COVID-19 cascade, i.e. from testing and testing positive to medical care and clinical outcomes. We analysed nationwide surveillance data from the Swiss Federal Office of Public Health (SFOPH) to examine the impact of neighbourhood socioeconomic position on testing for SARS-CoV-2 infection, testing positive, hospitalisation, intensive care unit (ICU) admission and death.

Methods

Data sources

We used mandatory notifications for SARS-CoV-2 and COVID-19, received at the SFOPH between Mar 1, 2020, and Feb 4 2021.[15] The data consisted of individual-level information about tests and test results, hospitalisations, ICU admissions and deaths, together with age, sex and residential addresses. Data on negative tests were systematically collected from May 23, 2020, onwards. We excluded notifications with missing or invalid information on the residential address. We obtained geocoded general population data from the Swiss Statistical Office (STATPOP, 2018 edition).[16] We used the most recent (2018) directory of retirement and nursing homes to identify individuals living in such institutions.[17]

Index of neighbourhood socioeconomic position

The Swiss neighbourhood index of socioeconomic position (Swiss-SEP) is based on the national house-to-house census of the year 2000.[18] It includes 1.27 million neighbourhoods of about 50 households centred on the corresponding residential building, with overlapping boundaries. We used the median rent per square metre, the proportion of households headed by a person with primary education or less, the proportion headed by a person in manual or unskilled occupation and the mean number of persons per room (crowding) to characterise neighbourhoods. No data on household income are collected in the Swiss census. The index was constructed using principal component analysis and validated using independent data on

households' financial situation.[18] It was standardised to range from 0 (lowest socioeconomic position) to 100 (highest position).

Geocoding and linkage to Swiss-SEP

Geocoding of the residential addresses was done using the publicly available geodata from the Swiss Federal Office of Topography. Swiss-SEP index values were aggregated into ten groups using deciles as cut-offs. Where a post code only was available, we used the Swiss-SEP value corresponding to the centroid of the area. Data were aggregated by canton (26 groups), sex (2 groups), age (9 groups, from 0-9 to 70-79 and 80 years and older), Swiss-SEP (10 groups) and epidemic wave (2 groups, before June 8, 2020, 14 weeks or after, 35 weeks) at the FOPH. June 8, 2020, was the nadir of case counts. The dataset consisted of aggregated data only. No ethical approval was required, in line with the Swiss Human Research Act.

Statistical analysis

We examined the association between Swiss-SEP group and counts of SARS-CoV-2 tests, positive tests, hospitalizations, ICU admissions, and deaths in negative binomial regression models to account for unknown overdispersion.[19] We considered three different denominators: the general population, the total number of tests and the number of positive tests. Denominators were included as offsets in each model. In a first univariate model, we estimated the incidence rate ratio (IRR) per increase in Swiss-SEP group for each outcome and denominator. The model assumes that the association with Swiss-SEP is linear on the logarithmic scale. We tested this assumption by comparing this model with one where each group was included separately. In a second model, we estimated the IRR adjusted for age group, sex, canton and epidemic wave. The adjustment for canton included a random intercept and slope by canton, allowing for interaction between Swiss-SEP group and canton. In a third model, we assessed two-way interactions between Swiss-SEP, and age group, sex and epidemic wave. We used the leave-one-out cross-validation information criterion (LOOIC) for model selection.[20]

In sensitivity analyses, we replicated the analyses after i) excluding all cases geocoded with the post code and not the full address and ii) excluding cases with an address corresponding to one of 1,586 retirement or nursing homes. All analyses were conducted using Stan[21] in R version 4.0.4,[22] with package 'rstanarm'.[23] We used weakly informative prior distributions for all model parameters.[20] The supplementary appendix provides more information.

Results

Notifications and geocoding

During the study period (May 23 2020 to Feb 4 2021 for the number of reported tests, Mar 1 2020 to Feb 4 2021 for all other outcomes), the following were notified (Fig. 1): 2,895,139 SARS-CoV-2 test results, 488,531 positive SARS-CoV-2 tests (16.9% of all tests), 20,070 COVID-19 hospitalisations (9.9% of positive cases), 1,983 COVID-19 ICU admissions (0.4% of positive cases) and 7,628 deaths from COVID-19 (for a case fatality rate of 1.6%). Geocoding was successful for 2,548,638 (88.0%) tests, 423,656 (86.7%) positive tests, 17,762 (88.5%) hospitalisations, 1,785 (90.0%) ICU admissions and 6,060 (79.4%) deaths. In over

90% of geocoded notifications, geocoding was based on the exact address (Table S1). Few geocodes corresponded to retirement or nursing homes, ranging from 1.5% for people admitted to ICU to 3.8% for people admitted to hospital. For deaths, 1,864 notifications (30.8%) were associated with such institutions (Table S1).

Table 1 shows the observed distribution of geocoded notifications and the size of the 2018 general population across age, sex, epidemic wave and Swiss-SEP group. About 40% of the Swiss population is aged 50 years or older. This age group accounted for a third of all SARS-CoV-2 tests (32.7%) and for 39.5% of positive tests, but for 90.0% of hospitalisations, 92.3% of ICU admissions and 99.4% of deaths. Women contributed more tests (53.1% of all tests) and more positive tests (52.5%) than men. Men accounted for most hospitalisations (57.9%), ICU admissions (72.7%) and deaths (55.8%). The group of lowest socioeconomic position accounted for 13.4% of positive tests, 15.2% of hospitalisations, 15.7% of ICU admissions and 13.2% of deaths and the highest for 7.4%, 5.8%, 5.1% and 6.3%, respectively. In the general population, the number of people living in neighbourhoods of lower socioeconomic position is higher than in neighbourhoods of higher position. Most of the data were from the second wave, which lasted longer and was more severe than the first: 81.7% of deaths were from the second wave (Table 1).

The rates of SARS-CoV-2 tests per population increased with Swiss-SEP group (Fig. 2), whereas they decreased for positive tests, hospitalisations, ICU admissions and deaths. The slopes for positive tests, hospitalisations and ICU admissions were steeper when rates were calculated per test rather than by population, and somewhat less steep when expressed per positive test.

Model fit

Modelling Swiss-SEP groups as a continuous variable led to a similar or better model fit compared to discrete variables (Table S2). Adjusting for age, sex, epidemic wave, and canton improved the fit further. Visual comparison of model predictions and observed data from the latter model illustrates the good fit, with most observed data points within the 95% credible interval of model estimates (Fig. S1). One exception was the higher number of tests among people living in neighbourhoods in the highest Swiss-SEP group, which was not captured well. Also, several data points were outside the credible interval for positive tests per population. The fit improved for positive tests when stratifying the data by epidemic wave (Fig. S8).

Analyses of Swiss-SEP groups, age, and sex

Both in unadjusted and adjusted analyses, each increase in Swiss-SEP group was associated with an increase in SARS-CoV-2 testing per population (Fig. 3). The adjusted rate ratio (IRR) was 1.02 (95% CrI 1.01-1.04) per one group increase, corresponding to 21% (95% CrI 5-40%) more tests in the highest compared to the lowest group (Table 2). There was no association with positive tests. A decreasing trend in positive tests was found with increasing Swiss-SEP group when expressing rates per number of tests rather than by population (adjusted IRR 0.97, 95% CrI 0.96-0.98), corresponding to a 23% (95% CrI 16-29) lower test positivity in the highest compared to the lowest socioeconomic group. The greater uptake of testing in neighbourhoods of higher socioeconomic position masked the higher number of positive tests among individuals from lower socioeconomic position neighbourhoods. Rates of

hospitalisations and ICU admissions decreased with higher socioeconomic position. Estimates were similar in the unadjusted and adjusted analyses, and similar with different denominators (Table 2, Fig. 3).

COVID-19 related mortality declined with increasing socioeconomic position of neighbourhoods (Table 2). The association became stronger when excluding residents of retirement or nursing homes (Fig. 4A). After excluding residents, the adjusted IRRs per increase in Swiss-SEP group were 0.96 (95% CrI 0.93-0.99) for COVID-19 deaths per population, 0.94 (95% CrI 0.92-0.97) for deaths among those tested and 0.98 (95% CrI 0.96-1.0) for deaths among those testing positive. These estimates translated into 34% (95% CrI 10-50%), 41% (95% CrI 24-52%) and 19% (95% CrI 3-31%) lower mortality, respectively, comparing the highest with the lowest group.

Testing intensity, positive tests and the clinical outcomes were also associated with age and sex (Fig. S2). Testing per 100,000 population was less intense, and positive tests less frequent in children 0 to 9 years. The risk of hospitalisation increased with age, and ICU admissions and mortality increased from age 50 years onwards. Testing and positive tests were about as frequent among men and women, but the rates of hospitalisations, ICU admissions, and mortality were all lower in women than in men. For all outcomes, there was heterogeneity across cantons (Fig. S3 and S4).

Interactions

We examined two-way interactions between Swiss-SEP group and age, sex, epidemic wave, and canton. The associations between socioeconomic position and outcomes became weaker with increasing age (Fig. S7, Table S8). The interaction with age is illustrated in Fig. 4B for mortality in those testing positive. The association with neighbourhood socioeconomic position becomes weaker moving from age group 0-49 years to older age groups and disappears in age group 80 years and older. There was little evidence of interactions with sex or epidemic wave (Fig. S7). There was also heterogeneity across cantons, particularly for testing and positive tests. The canton of Geneva was an outlier, with a stronger positive association of Swiss-SEP group with testing and a stronger negative association with test positivity. Associations with testing and test positivity were also somewhat stronger for the cantons of Bern, Obwalden, and Uri, and weaker or absent for other cantons (Fig. S5 and S6, Table S9).

Discussion

In this analysis of the Swiss COVID-19 surveillance data we found that people living in areas with higher socioeconomic position were more likely to get tested for SARS-CoV-2 infection but less likely to test positive, be admitted to hospital or the ICU, and less likely to die. Associations with neighbourhood socioeconomic position were similar during the first and second wave, but they were more pronounced in some cantons than others. Our analysis also showed that testing was less intense and positive tests less frequent in children. The risk of hospitalisation increased continuously with age, and ICU admissions and mortality increased from age 50 years onwards, in line with previous studies.[24] Testing and positive tests were

about as frequent among men and women, but the rates of hospitalisations, ICU admissions, and mortality were all higher in men than in women, again confirming previous findings.[25,26] Using national data reflecting one country and its health system, our analysis covered the entire cascade from testing for SARS-CoV-2 infection to testing positive, the need for hospital care and death. The coverage of both the first and second wave of the COVID-19 epidemic. Another strength is the availability of the Swiss-SEP index, which has criterion validity, with mean household income continuously increasing from the lowest to highest socioecnomic position group,[18] based on data from more than one million small neighbourhoods centred on individuals' residences ('ego-centred neighbourhoods')[18].

The national datasets allowed us to examine the association with neighbourhood socioeconomic position in three different populations, the whole of the general population, the population tested for SARS-CoV-2 infection, and the group with positive test results. Associations with socioeconomic position were consistent across the three populations, except for the rate of positive tests in the population. The greater uptake of testing in higher SEP neighbourhoods masked the higher number of positive tests among individuals of lower socioeconomic position. There are also weaknesses of the data that limit interpretation. The data about tests are limited by the absence of complete data on reasons for testing. For example, the lower test positivity among children under ten years of age could reflect that children were more likely to be tested within the context of infection control measures rather than because of symptoms. Alternatively, the lower test positivity could indicate a lower susceptibility to SARS-CoV-2 infection in this age group. Not all notifications could be geocoded due to incomplete addresses, and selection bias may thus have been introduced. Also, the Swiss-SEP index of retirement and nursing homes might not reflect the neighbourhood where residents spent most of their lives, thus misclassifying their socioeconomic position. This might explain why the association with the index of socioeconomic position became stronger when excluding deaths in residents of these institutions.

Strengths and weaknesses in relation to other studies

Data on indicators of socioeconomic position are often not collected in clinical studies or routine surveillance systems. Khalatbari-Soltani et al.[10] observed that, up to April 2020, no study about COVID-19 had reported data on socioeconomic indicators such as educational level, income or housing conditions.[10] Since then, several studies have found associations between area-level deprivation and SARS-CoV-2 infection, more severe COVID-19 disease, and mortality.[11–14] In common with these, our study used a small area-based measure of socioeconomic position. Area-based measures are more readily available than individual measures and have the advantage of capturing effects at the level of the individual and the place. A seroprevalence study in the canton of Geneva, Switzerland did not find strong associations with individual level indicators of social position, other than becoming unemployed. However, the study was not representative of the population of Geneva, including many more individuals with tertiary education, fewer with mandatory school only, and fewer non-Swiss nationals.[27,28] A Swedish study used individual-level data from population-based registers to show that people with lower income and level of education, and immigrants from low- or middle-income countries were at higher risk of death from COVID-19.[29] Data on

immigrant status are not recorded in the Swiss surveillance system. A unique strength of the present study is that it covers the entire cascade from testing to mortality at the national level. Furthermore, previous studies were often based on population surveys, with unequal participation across socioeconomic and ethnic groups,[11,12,27,28] excluded children,[11,12] or were restricted to selected hospitals or cities[13,30].

Meaning and implications

Fifty years ago, Tudor Hart proposed the "inverse care law", which stipulates that "the availability of good medical care tends to vary inversely with the need for it in the population served." [4] Similarly, the "inverse equity hypothesis" proposed in 2000 by Victora and colleagues[31] states that newly introduced health interventions would be initially adopted by the wealthier segments of a population, who likely had the least need for such interventions. Both hypotheses were borne out in the present study, in the unique setting of a pandemic and of infection control. Early diagnosis of SARS-CoV-2 infections and adequate initial management may improve the prognosis of COVID-19, whereas prognosis is worse in patients diagnosed late, with low oxygen saturation and signs of pneumonia. Rapid diagnosis and rapid isolation are the key to preventing transmission: communities with higher testing levels will benefit from lower rates of transmission. The SARS-CoV-2 tests were a new technology and testing capacity was limited in Switzerland, particularly during the first wave of the pandemic. In both waves, testing was less intense in neighbourhoods of lower socioeconomic position. It is possible that people living in areas of lower socioeconomic position had less access to test centres, because of poorer access to private transport or inability to take time off work. Greater availability of testing and conditions that eased uptake in these areas could have improved outcomes and reduced transmission.

The higher rate of positive tests in neighbourhoods of lower socioeconomic position might reflect higher risks of SARS-CoV-2 infection at work and at home. People in manual occupations are unable to work from home and might have more unprotected contacts with others, on building sites or in factories. At home, living conditions might also be more crowded. A study from the USA used mobile phone data to show that the adoption of social distancing was lower in counties with higher proportions of people below the poverty level.[32] Detailed maps of the socioeconomic position of Swiss neighbourhoods have been published.[18,33] Health policy measures should consider the vulnerability of different communities and prevent inequities in health and infection control. The Swiss National COVID-19 Science Taskforce recommended that, in this unpredictable crisis, the state should assume the role of insurer and cushion the negative effects with appropriate economic and social policies.[34] Without such support, it is understandable that those affected will not be in favour of control measures that threaten their livelihoods.

Switzerland is one of the wealthiest countries globally,[35] with wealth more unequally distributed than in other European countries: the Gini coefficient is estimated at 0.86 based on 2015 tax data.[36] Switzerland has a well-developed health care system and universally mandated health insurance, which in principle guarantees access to care for all. The Swiss health care system has been described as providing a good balance between individual responsibility and

community solidarity,[37] but there is evidence that high out of pocket payments, including co-payments and deductibles, may prompt some not to seek care. A survey in the canton of Geneva[38] showed that depending on income, up to 31% of respondents reported having foregone healthcare for economic reasons. In our study, Geneva was the canton with the strongest association between neighbourhood socioeconomic position and testing. It is also the canton with the highest Gini index for wealth (0.92).

In conclusion, this nationwide study found that people living in neighbourhoods of higher socioeconomic position are more likely to be tested in Switzerland but less likely to test positive, be hospitalised, or die. The higher incidence of SARS-CoV-2 infections, combined with a higher prevalence of co-morbidities in neighbourhoods of lower socioeconomic position will have contributed to worse outcomes,[39,40] including the higher risk of hospitalisation and death.[41,42] At the time of writing, vaccination coverage was still low in Switzerland, but the government has gradually been easing measures.[43] It is essential to continue to monitor testing for SARS-CoV-2 infection, access and uptake of COVID-19 vaccination and outcomes of COVID-19. Government and health care systems should take measures to reduce health inequalities in their response to the SARS-CoV-2 pandemic.

Table 1. Distribution of study data across age, sex, and neighbourhood index of socioeconomic position (SEP).

Variable	Total tests* (%)	Positive tests (%)	Hospitalisations (%)	ICU admissions (%)	Deaths (%)	Population (%)
Total	2,548,638*	423,656	17,762	1,785	6,060	8,225,085
Age (years)						
0-9	70,557 (2.8%)	4,918 (1.2%)	184 (1.0%)	9 (0.5%)	7 (0.1%)	850,653 (10.3%)
10-19	281,535 (11.0%)	36,614 (8.6%)	104 (0.6%)	4 (0.2%)	0 (0.0%)	816,438 (9.9%)
20-29	489,337 (19.2%)	74,586 (17.6%)	274 (1.5%)	12 (0.7%)	2 (0.0%)	1,017,953 (12.4%)
30-39	484,283 (19.0%)	72,577 (17.1%)	462 (2.6%)	35 (2.0%)	6 (0.1%)	1,181,590 (14.4%)
40-49	388,318 (15.2%)	67,572 (15.9%)	930 (5.2%)	77 (4.3%)	22 (0.4%)	1,170,915 (14.2%)
50-59	352,170 (13.8%)	71,328 (16.8%)	2,063 (11.6%)	229 (12.8%)	103 (1.7%)	1,240,120 (15.1%)
60-69	208,943 (8.2%)	40,947 (9.7%)	3,082 (17.4%)	462 (25.9%)	398 (6.6%)	899,148 (10.9%)
70-79	137,956 (5.4%)	26,631 (6.3%)	4,584 (25.8%)	642 (36.0%)	1,232 (20.3%)	675,678 (8.2%)
80+	135,539 (5.3%)	28,483 (6.7%)	6,079 (34.2%)	315 (17.6%)	4,290 (70.8%)	372,590 (4.5%)
Sex						
Men	1,194,479 (46.9%)	201,270 (47.5%)	10,293 (57.9%)	1,297 (72.7%)	3,383 (55.8%)	4,083,429 (49.6%)
Women	1,354,159 (53.1%)	222,386 (52.5%)	7,469 (42.1%)	488 (27.3%)	2,677 (44.2%)	4,141,656 (50.4%)
COVID-19 wave						
First wave	32,660 (1.3%)	22,742 (5.4%)	2,856 (16.1%)	386 (21.6%)	1,110 (18.3%)	-
Second wave	2,515,978 (98.7%)	400,914 (94.6%)	14,906 (83.9%)	1,399 (78.4%)	4,950 (81.7%)	-
Neighbourhood index of						

SEP (group)						
1 (lowest)	273,223 (10.7%)*	56,603 (13.4%)	2,691 (15.2%)	280 (15.7%)	799 (13.2%)	983,462 (12.0%)
2	265,271 (10.4%)*	48,692 (11.5%)	2,261 (12.7%)	251 (14.1%)	717 (11.8%)	885,738 (10.8%)
3	250,009 (9.8%)*	45,297 (10.7%)	2,031 (11.4%)	224 (12.5%)	681 (11.2%)	843,178 (10.3%)
4	246,696 (9.7%)*	42,620 (10.1%)	1,904 (10.7%)	190 (10.6%)	614 (10.1%)	827,815 (10.1%)
5	252,454 (9.9%)*	42,255 (10.0%)	1,799 (10.1%)	178 (10.0%)	662 (10.9%)	825,729 (10.0%)
6	245,782 (9.6%)*	41,730 (9.8%)	1,741 (9.8%)	162 (9.1%)	656 (10.8%)	813,863 (9.9%)
7	254,675 (10.0%)*	39,794 (9.4%)	1,555 (8.8%)	141 (7.9%)	553 (9.1%)	809,535 (9.8%)
8	247,772 (9.7%)*	38,146 (9.0%)	1,431 (8.1%)	138 (7.7%)	551 (9.1%)	792,051 (9.6%)
9	246,265 (9.7%)*	37,238 (8.8%)	1,317 (7.4%)	130 (7.3%)	447 (7.4%)	759,312 (9.2%)
10 (highest)	266,491 (10.5%)*	31,281 (7.4%)	1,032 (5.8%)	91 (5.1%)	380 (6.3%)	684,402 (8.3%)

SEP, socioeconomic position; first wave, before Jun 8, 2020; second wave, from Jun 8, 2020.
* Data on total tests relate to the period May 23, 2020, to Feb 4, 2021, rather than the full study period from Mar 1, 2020 to Feb 4, 2021.

Table 2. Association of group of neighbourhood index of socioeconomic position (SEP) with five outcomes related to SARS-CoV-2 surveillance and care: total tests, positive tests, hospitalisations, intensive care unit (ICU) admissions and deaths. Three denominators are considered accordingly: population, total tests and positive tests.

Outcome	**Denominator**	**Unadjusted IRR per Swiss-SEP group increase (95% CrI)**	**Adjusted IRR per Swiss-SEP group increase (95% CrI)**	**Adjusted IRR comparing highest and lowest Swiss-SEP group (95%CrI)**
Total tests*	Per population	1.02 (1.01-1.03)	1.02 (1.01-1.04)	1.21 (1.05-1.40)
Positive tests	Per population	0.99 (0.98-1.00)	1.00 (0.99-1.02)	1.03 (0.88-1.20)
	Per total test*	0.97 (0.97-0.98)	0.97 (0.96-0.98)	0.77 (0.71-0.84)
Hospitalizations	Per population	0.94 (0.93-0.96)	0.94 (0.93-0.96)	0.59 (0.50-0.71)
	Per total test*	0.94 (0.92-0.96)	0.92 (0.91-0.94)	0.49 (0.43-0.57)
	Per positive test	0.97 (0.95-0.98)	0.96 (0.95-0.97)	0.67 (0.61-0.74)
ICU admissions	Per population	0.91 (0.89-0.94)	0.90 (0.86-0.94)	0.39 (0.27-0.57)
	Per total test*	0.90 (0.87-0.93)	0.89 (0.87-0.93)	0.37 (0.27-0.50)
	Per positive test	0.92 (0.90-0.95)	0.93 (0.90-0.96)	0.50 (0.38-0.66)
Deaths	Per population	0.97 (0.94-1.00)	0.97 (0.92-1.02)	0.75 (0.49-1.17)
	Per total test*	0.97 (0.94-1.01)	0.95 (0.93-0.98)	0.66 (0.54-0.85)
	Per positive test	0.98 (0.95-1.01)	0.98 (0.96-1.00)	0.84 (0.71-1.01)

* Data on total tests relate to the period May 23, 2020, to Feb 4, 2021, rather than the full study period from Mar 1, 2020, to Feb 4, 2021.

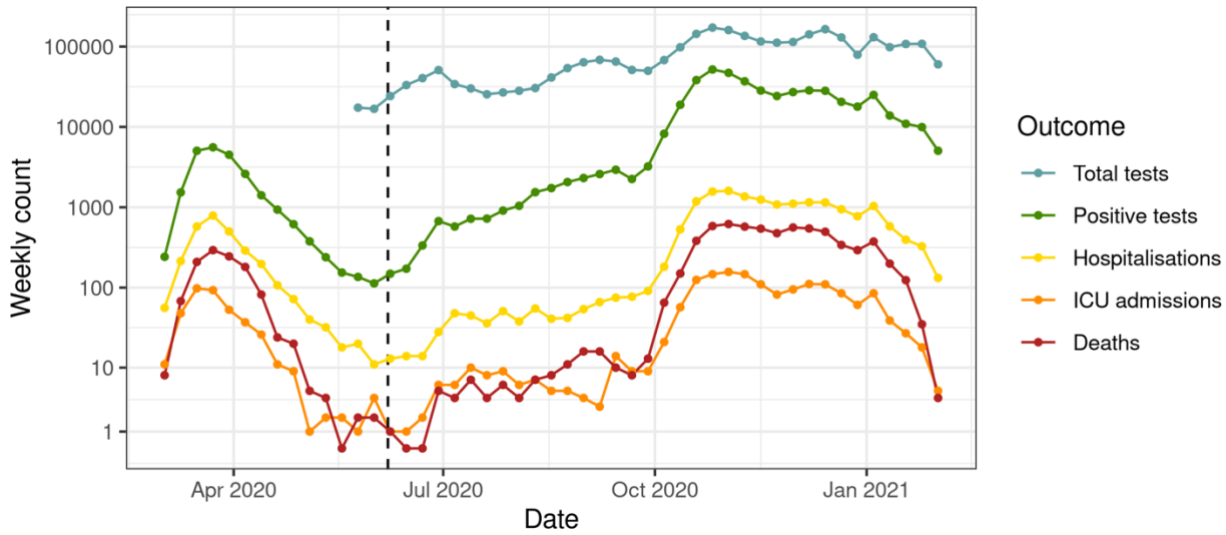

Fig. **1**. From testing to mortality: evolution of notifications to the Federal Office of Public Health during the COVID-19 pandemic in Switzerland from Mar 1 2020 to Feb 4 2021. The counts of total tests were available only from May 23 2020. The dashed line shows the date chosen for the separation between the first and second waves, Jun 8 2020.

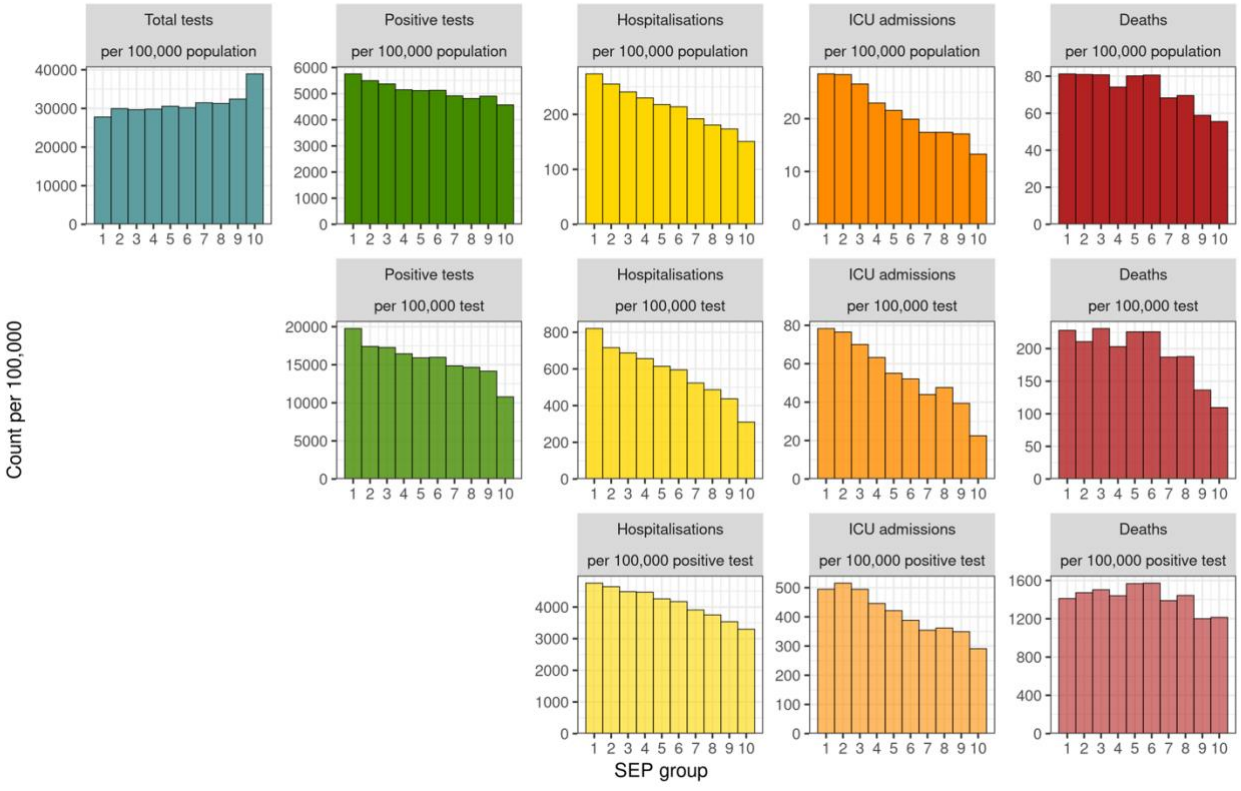

Fig. **2**. Counts of notified SARS-CoV-2 tests, positive tests, hospitalisations, intensive care unit (ICU) admissions and deaths across groups of socioeconomic position (SEP) per 100,000 population, tests, or positive tests. Higher SEP groups correspond to neighbourhoods of higher SEP. The study period was Mar 1, 2020, to Feb 4, 2021, except for total tests which only covered the period May 23, 2020, to Feb 4, 2021.

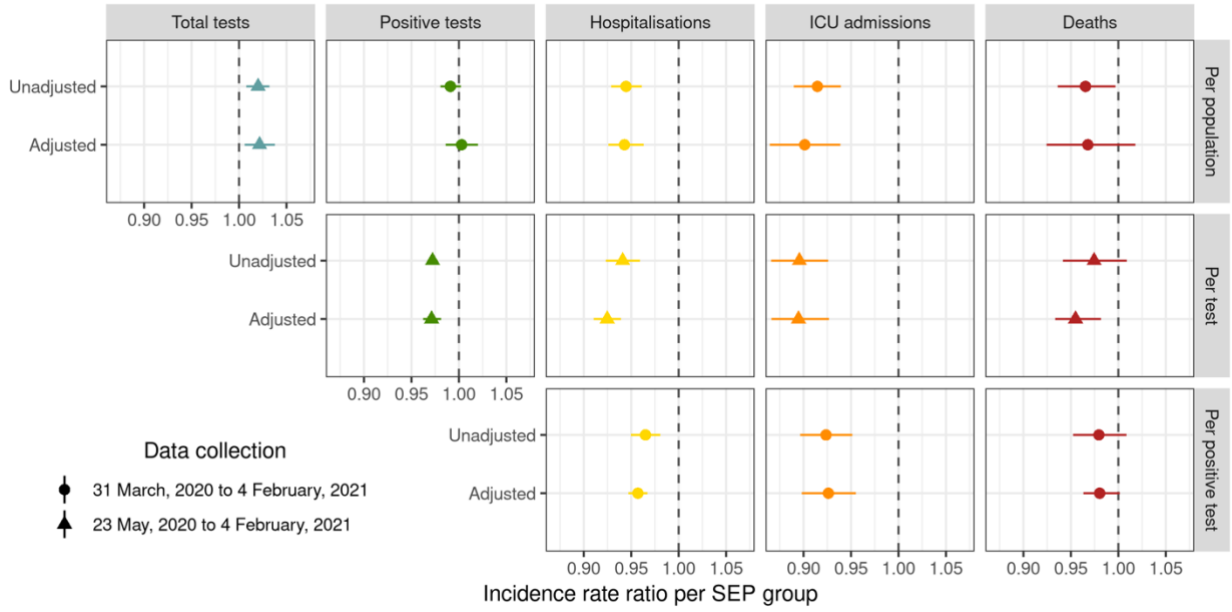

Fig. 3. Unadjusted and adjusted incidence rate ratios (IRR) per increase in group of neighbourhood socioeconomic position (SEP) for the counts of SARS-CoV-2 tests, positive tests, hospitalisations, ICU admissions and mortality per population, tests, or positive tests. Median posteriors and 95% credible interval are shown in each case. IRR estimates higher than 1 correspond to a positive association with Swiss-SEP groups, estimates lower than 1 correspond to a negative association. Adjusted estimates are adjusted for age, sex, canton, and epidemic wave. The study period was Mar 1, 2020, to Feb 4, 2021, except for total tests which covered May 23, 2020, to 4 February 2021.

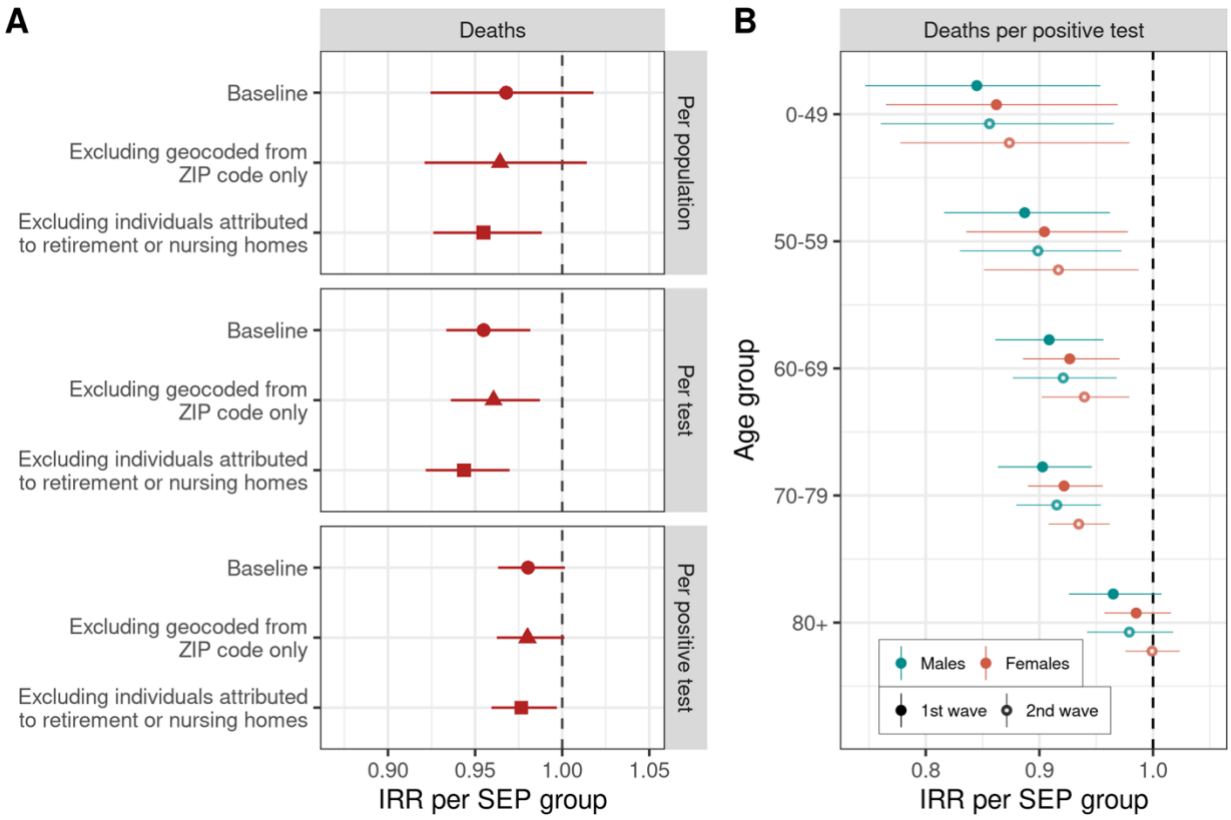

Fig. 4. (A) Adjusted incidence rate ratio (IRR) per increase in group of neighbourhood socioeconomic position (SEP) for COVID-19 deaths per population, per test or per positive test in the baseline analysis or in sensitivity analyses: 1) excluding all cases geocoded from the ZIP code only and 2) excluding cases with a residential address corresponding to retirement or nursing homes. (B) Adjusted IRR for COVID-19 deaths per positive test by age group, sex and epidemic wave.

Contributors

ME, JR, RP, NL and CLA conceived the study. JR, RP, NL, and ME drafted the first version of the manuscript. JR and RP did all statistical analyses, DP and KP managed the data at the Federal Office of Public Health. NC programmed the geocoding routine and RP assigned the values of the neighbourhood index of socio-economic position. All authors contributed to the interpretation of data and read and approved the final manuscript.

Funding

Federal Office of Public Health, Swiss National Science Foundation (grant 189498).

Acknowledgements

This study would not have been possible without the extraordinary efforts of the infectious disease surveillance team at the Federal Office of Public Health. We are also grateful to Thomas Van Boeckel and Carole Dupont for helpful comments on an earlier draft of this paper. This study was funded by the FOPH and the Swiss National Science Foundation (grant 189498).

Data sharing

The data are accessible to researchers upon reasonable request for data sharing to the corresponding author. Requests for data need to be approved by the Federal Office of Public Health. The code is available from https://github.com/ jriou/covid-sep-ch.

Conflict of interest statement

The authors declare no competing interests.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/ (accessed Feb 7, 2021).
2. Coronavirus Update (Live): 106,578,121 Cases and 2,324,793 Deaths from COVID-19 Virus Pandemic - Worldometer. https://www.worldometers.info/coronavirus/ (accessed Feb 7, 2021).
3. Eurostat. 580 000 excess deaths between March and December 2020. https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210312-2 (accessed March 17, 2021).
4. Tudor Hart J. THE INVERSE CARE LAW. The Lancet 1971; 297: 405–12.
5. The Lancet. 50 years of the inverse care law. The Lancet 2021; 397: 767.
6. GBD 2019 Diseases and Injuries Collaborators. 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. Lancet 2020; 396: 1204–22.
7. Mackenbach JP, Stirbu I, Roskam AJ, et al. Socioeconomic inequalities in health in 22 European countries. N Engl J Med 2008; 358: 2468–81.
8. Marmot M. Social determinants of health inequalities. Lancet 2005; 365: 1099–104.
9. Moser A, Panczak R, Zwahlen M, et al. What does your neighbourhood say about you? A study of life expectancy in 1.3 million Swiss neighbourhoods. Journal of Epidemiology & Community Health 2014; 68: 1125-32. doi: 10.1136/jech-2014-204352.
10. Khalatbari-Soltani S, Cumming RC, Delpierre C, Kelly-Irving M. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. J Epidemiol Community Health 2020; 74: 620–3.
11. McQueenie R, Foster HME, Jani BD, et al. Multimorbidity, polypharmacy, and COVID-19 infection within the UK Biobank cohort. PLOS ONE 2020; 15: e0238091.
12. Ward H, Atchison C, Whitaker M, et al. SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. Nat Commun 2021; 12: 905.
13. Quan D, Luna Wong L, Shallal A, et al. Impact of Race and Socioeconomic Status on Outcomes in Patients Hospitalized with COVID-19. J Gen Intern Med 2021; published online Jan 27. DOI:10.1007/s11606-020-06527-1.
14. Ossimetha A, Ossimetha A, Kosar CM, Rahman M. Socioeconomic Disparities in Community Mobility Reduction and COVID-19 Growth. Mayo Clin Proc 2021; 96: 78–85.
15. Federal Office of Public Health. Infectious diseases requiring notification – key facts. 2020. https://www.bag.admin.ch/bag/en/home/krankheiten/infektionskrankheiten-bekaempfen/meldesysteme-infektionskrankheiten/meldepflichtige-ik.html (accessed Dec 17, 2020).
16. Federal Office of Statistics. Population and households from 2010. https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/gebaeude-wohnungen-haushalte-personen/bevoelkerung-haushalte-ab-2010.html (accessed March 17, 2021).

17. Federal Office of Statistics. Retirement and Nursing Homes. https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/gesundheitswesen/alters-pflegeheime.html (accessed March 17, 2021).
18. Panczak R, Galobardes B, Voorpostel M, Spoerri A, Zwahlen M, Egger M. A Swiss neighbourhood index of socioeconomic position: development and association with mortality. J Epidemiol Community Health 2012; 66: 1129–36.
19. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis, 3. ed. Boca Raton: CRC Press/Chapman & Hall, 2014.
20. Gelman A, Hill J, Vehtari A. Regression and Other Stories, 1st edition. Cambridge University Press, 2020.
21. Carpenter B, Gelman A, Hoffman MD, et al. Stan : A Probabilistic Programming Language. Journal of Statistical Software 2017; 76. DOI:https://doi.org/10.18637/jss.v076.i01.
22. R Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2020 https://www.R-project.org/.
23. Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via Stan. R package version 2018; 2: 1758.
24. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020; 395: 1054–62.
25. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. Nat Commun 2020; 11: 6317.
26. Wenham C, Smith J, Morgan R. COVID-19: the gendered impacts of the outbreak. Lancet 2020; 395: 846–8.
27. Richard A, Wisniak A, Perez-Saez J, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies, risk factors for infection and associated symptoms in Geneva, Switzerland: a population-based study. medRxiv 2020; : 2020.12.16.20248180.
28. Stringhini S, Wisniak A, Piumatti G, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. Lancet 2020; 396: 313–9.
29. Drefahl S, Wallace M, Mussino E, et al. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. Nat Commun 2020; 11: 5097.
30. Kim B, Rundle AG, Goodwin ATS, et al. COVID-19 testing, case, and death rates and spatial socio-demographics in New York City: An ecological analysis as of June 2020. Health Place 2021; 68: 102539.
31. Victora CG, Vaughan JP, Barros FC, Silva AC, Tomasi E. Explaining trends in inequities: evidence from Brazilian child health studies. Lancet 2000; 356: 1093–8.
32. Garnier R, Benetka JR, Kraemer J, Bansal S. Socioeconomic Disparities in Social Distancing During the COVID-19 Pandemic in the United States: Observational Study. J Med Internet Res 2021; 23: e24591.
33. Panczak R, Spörri A, Zwahlen M, Egger M. Featured graphic. The socioeconomic position of Swiss neighbourhoods. ENVIRONMENT AND PLANNING A 2013; 45: 751–2.

34. Swiss National COVID-19 Science Taskforce. Warum aus gesamtwirtschaftlicher Sicht weitgehende gesundheitspolitische Massnahmen in der aktuellen Lage sinnvoll sind. [Why far-reaching health policy measures make sense from a macroeconomic perspective in the current situation]. https://sciencetaskforce.ch/policy-brief/warum-aus-gesamtwirtschaftlicher-sicht-weitgehende-gesundheitspolitische-massnahmen-in-der-aktuellen-lage-sinnvoll-sind/ (accessed March 31, 2021).
35. Rohner U. The Global wealth report 2020. Zürich: Credit Suisse, 2020.
36. Swiss Feder Tax Administration. L'évolution de la richesse en Suisse de 2003 à 2015. Bern: SFTA, 2019.
37. Biller-Andorno N, Zeltner T. Individual Responsibility and Community Solidarity - The Swiss Health Care System. N Engl J Med 2015; 373: 2193–7.
38. Guessous I, Gaspoz JM, Theler JM, Wolff H. High prevalence of forgoing healthcare for economic reasons in Switzerland: A population-based study in a region with universal health insurance coverage. Prev Med 2012; 55: 521–7.
39. Panczak R, Zwahlen M, Woitek U, Rühli FJ, Staub K. Socioeconomic, temporal and regional variation in body mass index among 188,537 Swiss male conscripts born between 1986 and 1992. PLoS One 2014; 9: e96721.
40. de Mestral C, Stringhini S, Guessous I, Jornayvaz FR. Thirteen-year trends in the prevalence of diabetes according to socioeconomic condition and cardiovascular risk factors in a Swiss population. BMJ Open Diabetes Res Care 2020; 8. DOI:10.1136/bmjdrc-2020-001273.
41. Singh AK, Gillies CL, Singh R, et al. Prevalence of co-morbidities and their association with mortality in patients withCOVID-19: A systematic review and meta-analysis. Diabetes Obes Metab 2020; 22: 1915–24.
42. Reilev M, Kristensen KB, Pottegard A, et al. Characteristics and predictors of hospitalization and death in the first 11 122 cases with a positive RT-PCR test for SARS-CoV-2 in Denmark: a nationwide cohort. Int J Epidemiol 2020; 49: 1468–81.
43. Coronavirus: the situation in Switzerland. SWI swissinfo.ch. https://www.swissinfo.ch/eng/society/covid-19_coronavirus--the-situation-in-switzerland/45592192 (accessed March 15, 2021).

Chapter 6

Discussion

Summary of findings

In this doctoral thesis, we investigated the means to centralize, disseminate, and use open-access data to generate outputs that help decision-makers in defining policies to control infectious diseases. Specifically, we focused on furthering the evidence base on antimicrobial resistance (AMR) in food animals and global veterinary capacities, and on providing short-term projections to control the COVID-19 pandemic in Switzerland.

In Chapter 2, we presented the open-access platform *resistancebank.org*. When launched in September 2019, this platform was conceived as a repository to centralize PPS from LMICs to evaluate AMR prevalence estimates of four pathogens in food animals: *Escherichia coli*, non-typhoidal *Salmonella* spp., *Staphylococcus aureus*, and *Campylobacter* spp. Two years later, the platform was expanded to include PPS on AMR in fisheries and aquaculture. As of October 2023, *resistancebank.org* contains 2,045 PPS, making the platform – to the best of our knowledge – the largest open-access repository of PPS on AMR in food animals. Besides centralizing PPS, *resistancebank.org* is also the first platform providing freely available high-resolution maps to investigate the spatial granularity of AMR trends, as well as country-level reports available for download. Implementing these resources within an intuitive user interface was essential to allow people without technical knowledge of Geographic Information Systems (GIS) to easily navigate our platform and access its resources.

In Chapter 3, we presented a global address book of veterinarians, showing how data not primarily intended for public health can be used to identify areas where food animals lack access to veterinary care. Similarly for the database of PPS openly disseminated by *resistancebank.org* (Chapter 2), this database of veterinarians is – to the best of our knowledge – the first database centralizing >300,000 veterinary practices across 115 countries, reporting their geographic information at a finer scale than the national level. The experience acquired while working with the database of *resistancebank.org* was crucial for the management of such a large database and supported the development of the methods used for data curation.
Assembling a database of veterinarians as comprehensively as possible was essential to train and validate geospatial models to produce robust predictions of veterinary densities in areas where addresses could not be obtained. As a result, we produced the very first map of veterinarians' distribution at 10x10 km^2. We could show that >93% of the areas where food animals are further than 1 hour of motorized travel time from veterinarians (i.e., "veterinary coldspots") are in LMICs. The biomass of food animals in coldspots amounts to 189 million livestock units (LSUs), which is equivalent to 1.2 times the biomass of all animals raised for food in the U.S. In addition, by mapping veterinary coldspots at a fine-scale resolution, this study provides for the first time a detailed distribution of areas lacking veterinary capacities at a sub-regional level, rather than only at a national or regional one.

The content of Chapter 4 was strictly linked to that of Chapter 3. The increased proficiency in manipulating geographic objects and the acquired familiarity with algorithms for facility location analysis and for computing travel time maps formed the cornerstone of every aspect

presented in this chapter. Specifically, we investigated ways to improve accessibility to health services (HS) to support the surveillance of AMR and the health of food animals. We quantified the population brought within 1 hour of motorized travel time from HS (i.e., "population covered") under a hypothetical 5% increase in each country's national capacities. The first case study focused on a network of laboratories for antimicrobial susceptibility testing (AST) across five African countries. Using a geographically targeted approach we allocated new laboratories in areas where the largest population lacks access to care and showed that ~21 million people could be brought under 1 hour of travel time from AST laboratories under an increase of 271 laboratories across five countries. However, almost half of these laboratories (143) were sufficient to cover 95% of the whole population covered. The second case study considered the increase in the veterinary workforce for food animals in coldspots (see Chapter 3). For this, a scaling up approach that targeted areas with the highest population of food animals in coldspots was used to place a supplementary 5% of veterinarians in each country. However, the global scale of the study required the development of an approximation for the exact solution to the computational challenge. By targeting locations with a high number of food animals in coldspots, far from existing veterinarians, and with a low travel time, we developed an approach that was, on average across nine countries tested, 20 times faster than those that tested every location, reaching at least 90% of its animal population coverage. Once applied at the global scale, the 5% increase in national veterinarians ensured access to veterinary care for 26.9% of cattle, 34% of chickens, and 44% of pigs currently living in coldspots.

Finally, in Chapter 5, we presented two applications for using open-access data during the COVID-19 pandemic in Switzerland. First, the open-access platform *icumonitoring.ch* was presented. Its aims were focused on supporting decision-makers with bi-weekly forecasts of ICU occupancy, deaths, hospitalizations, and availability of ventilators in Swiss hospitals, to try to prevent hospital overflow (*1*, *2*). The unique contribution of *icumonitoring.ch* is that it used epidemic models to produce forecasts of ICU occupancy as a proxy for monitoring the evolution of the pandemic, given the absence of SARS-CoV-2 tests during its early stage. In addition to forecasts at the hospital-level (with restricted access), forecasts were also aggregated at the cantonal- and regional-level to provide a more general picture of the ICU capacity available in different administrative areas of Switzerland. All forecasts from *icumonitoring.ch* were regularly communicated directly to the Swiss Armed Forces and the Federal Office of Public Health (FOPH) to plan interventions and prevent maximum occupancy of ICUs. It is worth noting that much of the expertise used in developing *icumonitoring.ch* was acquired during the development of *resistancebank.org*. Without the work described in Chapter 2, it would have been impossible to develop *icumonitoring.ch* in a timeframe of almost one month. The knowledge gained from *resistancebank.org*, especially familiarizing with services for remote data storage, significantly contributed to the effective design and implementation of *icumonitoring.ch*, highlighting how important it is the transferability of skills across different projects.

A second contribution we made during the pandemic was the daily geocoding of COVID-19 cases to map new hotspots of infections. These maps were meant to inform decision-makers on where to prioritize interventions for containing the virus' spread. Furthermore, these geocodings were included in a large database of >2.5 million COVID-19 tests performed in

Switzerland between March 2020 and February 2021. This database, along with that of hospitalizations, ICU admissions, and deaths, was useful to emphasize that the Swiss population living in neighborhoods of a higher socioeconomic position (Swiss-SEP) performed more COVID-19 tests, had a lower positivity to such tests, and had lower hospitalization and death rate than the population in the poorest neighborhoods of Switzerland.

Open-access data for enhancing existing surveillance

Infectious diseases present health challenges that transcend conventional boundaries. The focus of this thesis on AMR, veterinarians, and COVID-19 underscores some of the threats affecting both human and animal health. Therefore, effective measures against infectious diseases require collaborative strategies across each public sector. Along with the initiatives we presented in this work, other existing strategies using open-access data have been launched, each targeting the unique challenges within its respective sector.

In addressing the critical issue of antimicrobial resistance (AMR), the World Health Organization (WHO) has taken a significant step with the introduction of the Global Antimicrobial Resistance and Use Surveillance System (GLASS), providing open-access data on AMR in humans at the country-level (*3*). While such efforts exist for human health, a comparable initiative in the context of food animals, where the majority of globally sold antibiotics are used (*4*), has not yet been developed by international organizations focusing on food animals like the Food and Agriculture Organization. Recognizing this gap, the creation of *resistancebank.org* emerged as an essential response to potential deficiencies in AMR surveillance within the domain of animal health in LMICs. Our platform not only addresses this gap but also distinguishes itself by offering – to the best of our knowledge – a resource currently unavailable on other web-based platforms focused on AMR in humans: spatial granularity. *resistancebank.org* stands out by providing 10x10 km^2 resolution maps that allow users to investigate AMR trends at the sub-regional level. The integration of such maps within systems provided with functions to zoom in and navigate these maps empowers researchers and individuals unfamiliar with GIS tools to promptly identify specific areas for planning interventions against AMR.

Our work about veterinarians arises from the need to get insights into their capacities at a finer scale than a national or regional one. Specifically, this work can supplement existing efforts monitoring the capacities of each actor involved in the veterinary sector. The most notable of these efforts is led by the World Organisation for Animal Health (WOAH) (*5*). In countries where it has not been possible for these organizations to obtain detailed numbers of veterinarians – for example, due to the lack of field officers – our national estimations, albeit incomplete, are some of the first that are available about veterinary capacities. In addition, WOAH has established guidelines to evaluate the national performances in veterinary services (PVS) (*6*); to date, more than 140 countries have engaged in the PVS pathway. This consists of a standardized global methodology that every country can use to i) evaluate the status of its national veterinary services, ii) identify its strengths and weaknesses, and iii) plan workforce

development. In this context, the methodology underlying the PVS pathway could be supplemented with additional criteria for the evaluation. One example could be the abundance, within a country, of the areas we classified as veterinary coldspots.

Similarly, additional outcomes of our studies can help inform other initiatives of international organizations. For example, one of the main objectives of the Fleming Fund – the United Kingdom program to improve AMR surveillance in LMICs – is to equip countries to collect and use data on drug resistance (*7*). Using our accessibility maps produced for the scaling up of AST laboratories, they can be supported in deciding which health facility has the priority to be equipped for AST.
Furthermore, *resistancebank.org* could serve as a valuable resource for funding organizations, aiding them in determining which research groups working on AMR should have priority for fund allocation based on the possibility of performing more PPS in a specific area. For example, a region displayed on our maps that shows limited PPS indicates heightened uncertainty in the spatial predictions of AMR prevalence estimates in that region. Consequently, research groups that have conducted PPS in these regions should be given priority in funding assignments to increase the number of PPS conducted in such areas, share their results on *resistancebank.org*, and ultimately help us reduce the uncertainty of our maps.

Another way to enhance existing surveillance is to use veterinary capacities predicted at the fine-scale resolution in combination with existing country-level capacities systematically inventoried in high-income countries (HICs). First, fine-scale predictions could be aggregated at the country-level and compared – as shown in Chapter 3 – with national estimates as an additional means to validate predictions of our geospatial models. Second, country-level capacities collected systematically – a common situation in HICs – are more reliable than the addresses of veterinarians to inform about the total national capacities available in a country. However, the addresses bring spatial information that country-level estimates lack. Therefore, such estimates can be disaggregated at the fine-scale resolution and incorporated within the geographic pattern represented by the addresses to get precise information about the national capacities available, but also strong predictions about their spatial availability. Third, national estimations collected in different years could be used to associate a temporal trend with the distribution of veterinarians. In this way, models could be improved to predict in which sub-national areas a possible growth (or shortage) of veterinarians is likely to happen.

As concerns the surveillance of the COVID-19 pandemic, several user-friendly platforms were developed to track disease spread. At the global-level, Johns Hopkins University developed the "COVID-19 dashboard" to track cases in real-time (*8*), displaying locations of new infections through an interactive web map. In addition, Google and UNESCO created similar versions of these platforms, aiming to disseminate COVID-19 information to a wider audience (*9*, *10*). In Switzerland, *icumonitoring.ch* was one of the first open-access platforms developed, thanks also to the support of the Swiss Armed Forces which needed a resource to understand how to distribute ventilators and health workforce upon potential requests of the cantons.
In addition, these efforts pushed the development of other platforms focusing on the monitoring of COVID-19 cases in Switzerland, like *corona-data.ch* (*11*) and the official platform provided

by the FOPH (*12*). However, in contrast to these platforms, *icumonitoring.ch* used ICU occupancy as a proxy for disease surveillance, providing information not only at the regional or national levels but also at the hospital-level.

In addition to enhancing existing surveillance methods, the projects presented in this thesis share a common goal: complementing broader One Health surveillance initiatives. One Health represents an integrated approach aimed at optimizing the health of humans, animals, and the environment by fostering collaborations of experts within these interconnected fields. Recognizing the close links among such fields can help develop innovative surveillance and disease control methods (*13*). The WHO and its partner organizations are actively promoting the adoption of the One Health approach in national, regional, and international health policies. This involves strengthening country capacities and monitoring risks, as well as preparing for early detection and response to emerging pathogens.
In the case of AMR, the maps of *resistancebank.org* can be incorporated within the GLASS framework and highlight areas where antibiotics are losing effectiveness in food animals. This can support the funding of alternatives to antibiotics, which is essential for limiting the resources necessary for developing new antibiotics and preventing resistant bacteria from carrying AMR genes within ecosystems. Similarly, the study on veterinarians exemplifies a surveillance approach aligning with the principles of One Health. Including maps of deforestation in our models of accessibility to veterinarians could highlight areas requiring increased surveillance, particularly those where recent agricultural expansion has occurred. This approach could help prevent potential cross-species transmission of infectious diseases in areas where interactions between wild animals that lost their habitat and food animals could be more frequent (*14*).

Policy implications

Enhancing global collaboration in animal health

Our work with open-access data and platforms allowed us to identify geographic regions where interventions could be targeted. These include hotspots of AMR in food animals, coldspots of veterinary capacities, and coldspots of AST laboratories. While this could guide decision-makers on how to optimize public health spending, many LMICs may lack the necessary resources for such investment. Particularly, the COVID-19 pandemic caused increases in hunger and poverty (*15*), further straining limited resources. The resources provided by our work can help in reducing the costs associated with disease surveillance through targeted approaches. For example, maps on *resistancebank.org* could guide researchers to perform new PPS only in areas where they haven't been performed yet, as it was investigated in China by Zhao and colleagues (*16*).

Nevertheless, the need for international assistance from high-income countries (HICs) to support LMICs in strengthening their capacities against infectious diseases should remain a priority. According to the Organisation for Economic Cooperation and Development, in 2018,

half of international health aid went to purchasing medicines from industries in the donor country, and over half of the aid was targeted at reducing HIV/AIDS, malaria, and tuberculosis (*17*). For sustained progress in public health within LMICs, global funders should also increase their investments in improving health systems, finance training programs to create a stronger workforce of veterinarians prepared to work in remote areas (*18*), and use our maps to guide the strategic enforcement of HS in underserved areas, like when members of the Fleming Fund approached our group to consult the AMR maps displayed on *resistancebank.org*.

Informing strategies to improve accessibility to health services

In this doctoral thesis, we also explored the drivers behind the geographic patterns of HS. The first important information that was evident from our findings is that geographic approaches aiming at increasing HS capacities need to be targeted towards areas with the highest population lacking prompt access to care. Specifically, what our findings clearly show is that scaling up HS equally among administrative divisions is far from optimal because the distribution of animal and human populations is generally heterogeneous within a country. Therefore, geographically approaches that focus primarily on populations lacking access to care can target specifically this heterogeneity to maximize the HS coverage.

An additional insight refers to the creation of incentives to improve the coverage of HS. For example, from the global distribution of veterinarians, we could identify that a low travel time to cities was an important driver for a high presence of veterinarians. This enforces previous findings showing that veterinarians prefer to work in urban areas, where there is a higher pool of patients, a higher possibility to share a studio (thereby saving funds for rents and essential veterinary equipment), and likely more appealing social and living conditions than in remote areas (*19*, *20*). Therefore, decision-makers could use our maps primarily to identify areas to prioritize for increasing the veterinary workforce, and then for improving the benefits for veterinarians working in remote areas. For example, incentives could be i) competitive salaries and bonuses, ii) financial support for continuing education and training, iii) job placement assistance for partners and family members, and iv) programs for engagement and integration in rural communities.

Raising public awareness

Platforms sharing open-access data present an intuitive way to let the public quickly grasp the severity of health issues like AMR or COVID-19. This is reflected, for instance, in the number of users that accessed *icumonitoring.ch* between May 2020 and August 2021. In this period, >175,000 unique users visited the platform (>10,000 users per month). Therefore, a significant portion of the platform's user base likely comprised ordinary citizens who sought to stay informed about the progression of the pandemic, rather than members of authorities directly involved in COVID-19 management. The heightened visibility of the platform can be attributed to extensive coverage by Swiss press outlets such as SRF, swissinfo.ch, Blick, and The Local (*21–23*). Additionally, its inclusion in SRF's evening news broadcasts on February 15th, 2021, and October 24th, 2021, contributed substantially to the increase in the number of users.

While this is crucial for increased dissemination of data that can serve for disease surveillance, it is also fundamental to enhance public awareness. When health-related data are summarized into outputs easily understandable by the public, they can become incentives to follow and respect preventive measures aimed at limiting the spread of diseases. For example, press articles sharing *icumonitoring.ch* forecasts that pointed at reaching the maximum capacity of ICUs (*24*) can represent incentives for the public to follow hygiene measures, social distancing, and the use of masks to avoid potential hospitalizations.

Promoting scientific collaboration

An implication of working with open-access data is the improvement of scientific collaboration among different sectors. Firstly, providing access to centralized databases of PPS and health services such as veterinarians can support other research groups in using these resources to produce further scientific knowledge on topics of global importance for human and animal health. This can have important impacts, especially for researchers in LMICs, where publication and subscription fees for scientific journals are barriers to accessing up-to-date research findings. Another example is represented by the country reports available from *resistancebank.org*. These outputs synthesize scientific findings that are easily interpretable by decision-makers, thereby promoting swift communication between the scientific and governmental sectors.

Secondly, the open dissemination of our resources can represent a way to encourage data sharing from the government and/or private agencies. For instance, one of the strategies used to build the address book of veterinarians (see Chapter 3) was to directly request databases of addresses by sending emails to veterinary associations and governmental agencies. The first outcome of this process was a low response rate (16.8%). Second, in countries such as Brazil, we received negative answers due to privacy policies. In such a context, three different scenarios could manifest after we produced maps of veterinarians using open-access data. Specifically, organizations that did not collaborate with our study could 1) acknowledge our work but show no interest in being involved in it, 2) recognize that our estimations of national veterinary capacities differ from the ones they have recorded, and 3) show interest in our work and decide to supplement our open-access data with theirs, hence providing further detailed information about veterinary capacities to strengthen our findings.
Scenario n° 1 would not change the amount of open-access data about veterinary capacities that we centralized and proposed as a proxy of animal health surveillance. This could also be the case for Scenario n° 2 unless these organizations decide to show that their data lead to different estimations of capacities than the ones we provided, thereby disseminating previously restricted data. In Scenario n° 3, their restricted data will be openly available to the public after the dissemination of our work. Therefore, the chances of a potential dissemination of these data are higher than in the case that no study using open-access data is conducted.

An example that can represent the situation described in Scenario n° 3 refers to the integration of data available from the Information and Operation Systems (IES) of Switzerland in the

forecasts of ICU occupancies during the pandemic of COVID-19. These data were restricted, and upon our initial interest in using open-access data provided by the platform OpenZH, the Swiss Armed Forces decided to establish a collaboration with our research group to provide us with detailed data about ICU occupancies. Such data were useful to supplement data from OpenZH used in the epidemic models defined to forecast COVID-19 cases.

A way to increase situations like those described in Scenario n° 3 is asking for support from international organizations involved in health capacity building. For example, although a formal collaboration has not yet been established, we have initiated connections with members of WOAH by presenting our work on veterinarians at scientific conferences. Additional efforts to reach more organizations are represented by publishing our findings in open-access journals and reaching organizations to request possible meetings and webinars to introduce our work. If such organizations agree to disseminate our findings within their extensive network of national collaborators who may possess restricted data about HS capacities, they could facilitate the establishment of new partnerships and the potential release of restricted data.

Limitations and future directions

The projects explored in this doctoral thesis come with limitations and possibilities for future improvements. For *resistancebank.org*, although the platform was originally conceived to collect AMR data as well as share them, it has so far received <10 PPS submitted directly by researchers. This could be due to two reasons. Firstly, researchers working on AMR might be reluctant to share their surveys because they do not see being exposed to our platform as a short-term advantage. As an incentive, we ensured that their contribution would have been available on the user interface of the platform, with the possibility for their work to gain increased attention in the AMR community. However, additional incentives might be required. One possibility could be represented by using a monthly newsletter sharing updates and information about *resistancebank.org*, informing subscribers about new PPS uploaded, thus improving the dissemination of other people's work, and sending reminders underlying the importance of increasing the existing database of PPS we centralized. Secondly, it could be that *resistancebank.org* is not yet recognized as the focal point for AMR data in food animals. This would mean increasing the visibility of the platform by advertising it at future scientific conferences and establishing collaboration with other research groups working on AMR. Also, it might be important to hire press agencies specialized in advertising our work to people active in this field. Another option would be to submit requests to organizations funding AMR research (e.g., Fleming Fund, Wellcome Trust, Bill & Melissa Gates Foundation, etc.) and establish partnerships with journal editors to encourage researchers performing phenotypic AMR surveys to upload their data on open-access repositories like ours. This system would be similar to the ones requiring researchers performing genomic surveys to upload complete genomic data into a public repository (*25*).

Another current limitation of *resistancebank.org* is the limited number of visitors from LMICs. From September 2019 to October 2023, 10,271 visitors (206 per month) accessed the platforms

from 146 countries (Fig. 1). However, only 28.4% of these visitors were from LMICs. This could be explained by a limited knowledge of *resistacebank.org* in these countries, which could depend on factors like poor internet connectivity and low availability of computers. Therefore, it could be necessary to promote the use of the platform by establishing partnerships with universities and research centers in LMICs, hiring national officers to travel to these institutes, and informing researchers and decision-makers about *resistancebank.org*. I would have personally embarked on some of these field trips after the platform was published, but COVID-19 disrupted these plans.

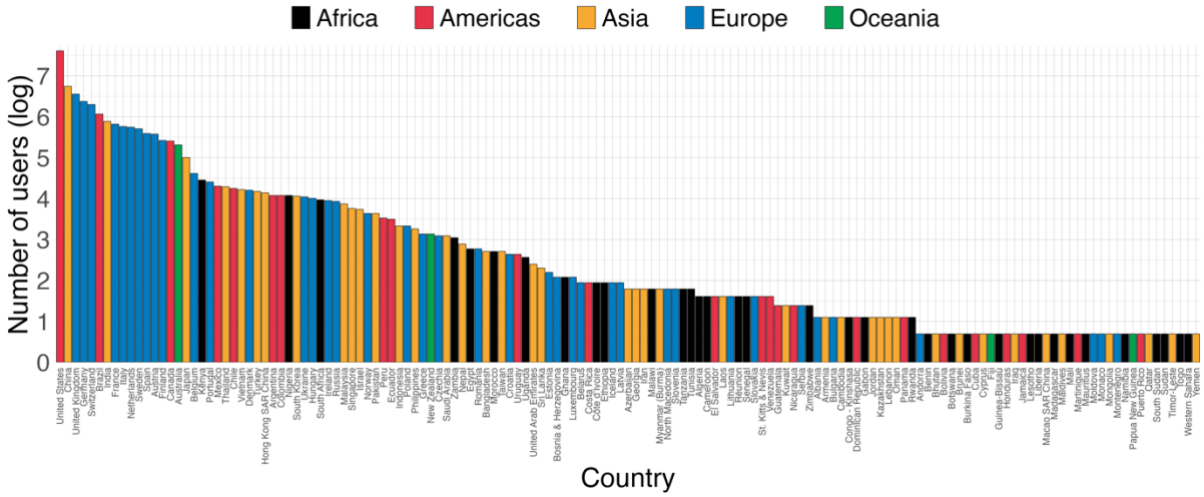

Fig. 1. Users of *resistancebank.org* by country from September 2019 to October 2023.

Additional limitations concern the accessibility to HS such as AST laboratories and veterinarians. For our analyses, we considered only one dimension of accessibility: travel time, using 1 hour as a threshold value. This definition provides a clear interpretability of the definition of "coldspots" of HS. However, there are other important aspects to consider. For example, 1 hour of traveling does not mean the same actual expenses for traveling among different countries/regions. Traveling to remote areas by different means might be more expensive for certain categories of people than for others. Therefore, our accessibility models could be expanded to incorporate covariates as the gross domestic product to produce travel time maps weighted for a measure of income related to the financial costs of traveling. Similarly, additional covariates to incorporate in these models could also assess the HS performance such as i) their capacity in terms of services provided per day/week, ii) the level of training of their personnel, and iii) the quality of care provided based on standards certified by competent agencies.

A common limitation that applies to multiple case studies is represented by the difference in the rates of open-access data reported within countries. This is related, for instance, to the PPS available in the literature or the number of veterinarians' addresses (and platforms) used to sample them. One solution to leverage these differences relies on the use of geospatial models in combination with spatial covariates, training them in data-rich regions to make reasonable predictions in data-poor regions. However, an additional solution could be necessary to increase the amount of data available to train such models and reduce uncertainty. This could

be represented by establishing a network of scientific collaborators spanning several countries to overcome language and cultural barriers that could prevent foreigners from identifying useful literature and/or online resources containing additional open-access data.

Another avenue to improve our maps and get additional insights into accessibility to HS is to consider different travel time thresholds. For example, a sensitivity analysis could be conducted on the threshold value of 1 hour to understand how the human and animal population served by HS changes when considering longer times of traveling. In addition, these thresholds could be adjusted according to specific medical and veterinary emergencies. Considering human obstetric emergencies, a 2-hour travel time threshold is a commonly acknowledged critical time from postpartum hemorrhage to death in the absence of medical intervention (*26*). Therefore, accessibility maps could be calculated when considering potential animals in danger after giving birth. Similarly, this travel time threshold could be applied also to AST laboratories since the collection of bacteriological samples is not a life-threatening emergency. The results of these sensitivity analyses could help in assigning to each health facility a standard travel time limit that represents the urgency in providing their unique service to a population. This could be helpful for decision-makers when planning interventions of capacity building.
In addition, although our maps suggested locations where a scale-up in HS could carry the highest benefits for the human and animal populations, we did not explore mechanisms on how this could be achieved in practice. Therefore, field surveys could be conducted to get additional information about the specific needs of farmers and people in remote areas who travel to HS.

In conclusion, this doctoral thesis focused on the potential of using open-access data to limit the burden of infectious diseases. The focal point of our studies was to show how data created for public health purposes and data not primarily intended for public health can both be used to define a similar resource: proxies of surveillance that LMICs can adopt when official systems are still in development. However, it is challenging to provide short-term evaluations of the impacts that our work could have on enhancing surveillance systems. As a temporary proxy, we can affirm that the number of users that our platforms have received is encouraging. A high number of users (including the >30 still accessing *resistancebank.org* daily) is giving us hope, because it could mean that researchers and organizations all over the world are potentially building new AMR knowledge based on the resources we openly disseminated, in a constant effort to understand more about AMR and try to limit its future burden. Similarly, as concerns the impacts of our studies on the accessibility of HS such as veterinarians, we can so far rely only on the positive feedback received by our colleagues during workshops and conferences. What is certain, is that we contributed to centralizing open-access data with a great potential for application in public health that was not previously considered for this purpose. In this way, we can at least affirm that we produced a new resource highlighting potential challenges for food animals in LMICs that require the attention of international funders and decision-makers.

In the context of a shortage of resources, the public health challenges we explored in this thesis pose competing priorities for investments. However, before identifying such priorities, there is one key message that emerges from these years of work: it is essential to create and support

initiatives aimed at increasing the availability of open-access data. The more data that will be openly accessible in the future (with appropriate anonymization when applicable), the more accurate evidence-based surveillance systems can become, increasing the amount of knowledge available for reducing the negative health outcomes of infectious diseases.

References

1. VICE. Coronavirus Has Northern Italy's Hospitals on the Brink of Collapse (2020). Available at: https://www.vice.com/en/article/k7ex4a/coronavirus-has-northern-italys-hospitals-on-the-brink-of-collapse. Accessed: 13/11/2023.
2. Bermuda Real. COVID-19: Horrifying Images Of Hospital Patients In Madrid Lying On The Floor As City Is Overrun With Cases (2020). Available at: https://www.bermudareal.com/covid-19-horrifying-images-of-hospital-patients-in-madrid-lying-on-the-floor-as-city-is-overrun-with-cases/. Accessed: 13/11/2023.
3. World Health Organization. Global Antimicrobial Resistance Surveillance System (GLASS) Report - Early implementation 2017-2018. 1–268 (2017).
4. *Working Party on Agricultural Policies and Markets* (2015).
5. World Organisation for Animal Health. Home (2023). Available at: https://www.woah.org/en/home/. Accessed: 14/11/2023.
6. World Organisation for Animal Health. Performance of Veterinary Service (PVS) Pathway (2022). Available at: https://www.woah.org/en/what-we-offer/improving-veterinary-services/pvs-pathway/#:~:text=The PVS Pathway empowers national,inefficiencies and opportunities for innovation. Accessed: 20/11/2022.
7. Department of Health and Social Care. The Fleming Fund - About Us (2023). Available at: https://www.flemingfund.org/about-us/. Accessed: 16/11/2023.
8. E. Dong, H. Du, L. Gardner. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533–534 (2020). doi:10.1016/S1473-3099(20)30120-1
9. Google. COVID-19 Open Data Repository (2022). Available at: https://health.google.com/covid-19/open-data/. Accessed: 05/10/2023.
10. UNESCO. Open access to facilitate research and information on COVID-19 (2021). Available at: https://en.unesco.org/covid19/communicationinformationresponse/opensolutions. Accessed: 05/10/2023.
11. Daniel Probst. COVID-19 Info Switzerland (2023). Available at: https://corona-data.ch. Accessed: 14/11/2023.
12. Swiss Federal Office of Public Health. COVID-19 Switzerland (2023). Available at: https://www.covid19.admin.ch/it/overview. Accessed: 14/11/2023.
13. World Health Organization. One Health (2023). Available at: https://www.who.int/news-room/fact-sheets/detail/one-health. Accessed: 23/12/2023.
14. J. H. Ellwanger, B. Kulmann-Leal, V. L. Kaminski, J. M. Valverde-Villegas, A. B. G. DA VEIGA, F. R. Spilki, P. M. Fearnside, L. Caesar, L. L. Giatti, G. L. Wallau, et al. Beyond diversity loss and climate change: Impacts of Amazon deforestation on infectious diseases and public health. *Anais da Academia Brasileira de Ciencias* **92**, 1–33 (2020). doi:10.1590/0001-3765202020191375
15. H. Kakaei, H. Nourmoradi, S. Bakhtiyari, M. Jalilian, A. Mirzaei. Effect of COVID-19 on food security, hunger, and food crisis. *COVID-19 and the Sustainable Development Goals* 3–29 (2022). doi:10.1016/B978-0-323-91307-2.00005-5
16. C. Zhao, Y. Wang, K. Tiseo, J. Pires, N. G. Criscuolo, T. P. Van Boeckel. Geographically targeted surveillance of livestock could help prioritize intervention against antimicrobial resistance in China. *Nature Food* **2**, 596–602 (2021). doi:10.1038/s43016-021-00320-x
17. Development initiatives. Aid spent on health: ODA data on donors, sectors, recipients (2023). Available at: https://devinit.org/resources/aid-spent-health-oda-data-donors-

sectors-recipients/. Accessed: 13/11/2023.
18. T. W. Graham, J. Turk, J. McDermott, C. Brown. Preparing veterinarians for work in resource-poor settings. *JAVMA* **243**, 1523–1528 (2023).
19. J. M. Richards. An Ecological Analysis of the Geographic Distribution of Veterinarians in the United States. *Journal of Vocational Behavior* **11**, 216–231 (1977). doi:10.1016/0001-8791(77)90008-2
20. S. Truchet, N. Mauhe, M. Herve. Veterinarian shortage areas: what determines the location of new graduates? *Review of Agricultural, Food and Environmental Studies* **98**, 255–282 (2017). doi:10.1007/s41130-018-0066-9
21. SRF. Warten auf die Operation: Wie ungleich Corona die Spitäler bremst (2021). Available at: https://www.srf.ch/wissen/corona/knappe-intensivbetten-warten-auf-die-operation-wie-ungleich-corona-die-spitaeler-bremst. Accessed: 28/10/2020.
22. Blick. Intensivstationen im Kanton Solothurn bereits voll (2020). Available at: https://www.blick.ch/schweiz/eth-zahlen-zeigen-spitalbelegung-intensivstationen-im-kanton-solothurn-bereits-voll-id16155717.html. Accessed: 28/10/2023.
23. swissinfo.ch. Intensive care beds could run out in a week, study predicts (2020). Available at: https://www.swissinfo.ch/eng/coronavirus_intensive-care-beds-could-run-out-on-thursday--study-predicts/45647658. Accessed: 28/10/2023.
24. SRF. ETH-Forscher: «Schweizer Spitäler droht nächste Woche die Überlastung» (2020). Available at: https://www.srf.ch/play/tv/tagesschau/video/eth-forscher-schweizer-spitaeler-droht-naechste-woche-die-ueberlastung?urn=urn:srf:video:20508509-cde8-4c0f-95f0-0069c3ce4a31. Accessed: 28/10/2023.
25. R. Van Noorden. US agency updates rules on sharing genomic data. *Nature* 1–2 (2014). doi:10.1038/nature.2014.15800
26. A. J. Van Duinen, H. A. Adde, O. Fredin, H. Holmer, L. Hagander, A. P. Koroma, M. M. Koroma, A. J. M. Leather, A. Wibe, H. A. Bolkan. Travel time and perinatal mortality after emergency caesarean sections: An evaluation of the 2-hour proximity indicator in Sierra Leone. *BMJ Global Health* **5**, (2020). doi:10.1136/bmjgh-2020-003943