"Everything comes from somewhere"

Exploring the potential role of the environment in transmission of foodborne pathogens

Annemieke Christine Mulder

The publication of this thesis was financially supported by the National Institute for Public Health and the Environment (RIVM).

ISBN: 978-94-6483-015-6
DOI: https://doi.org/10.33540/824
Cover design and lay-out: Publiss | www.publiss.nl
Print: Ridderprint | www.ridderprint.nl

"Everything comes from somewhere"

Exploring the potential role of the environment in transmission of foodborne pathogens

"Everything comes from somewhere"

Onderzoek naar de mogelijke rol van het milieu in de transmissie van door voedsel overgedragen ziekteverwekkers

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 22 mei 2023 des middags te 2.15 uur

door

Annemieke Christine Mulder

geboren op 31 augustus 1991
te 's-Hertogenbosch

Promotor:
Prof. dr. D.J.J. Heederik

Copromotoren:
Dr. E. Franz
Dr. L. Mughini Gras

Table of content

1

CHAPTER 1

General introduction

The coronavirus disease 2019 (COVID-19) pandemic that started in December 2019 in China,[1] has embedded words like "epidemiology" and "zoonoses" in normal daily conversations worldwide. On the one hand, this facilitates the explanation of some key concepts of our field of work, but does it also facilitate the writing of a thesis about the eco-epidemiology of zoonotic infectious diseases and does it make acceptance of research findings more straightforward?

A short answer would be no.

Indeed, the emergence and spread of zoonotic infectious diseases are complex and unique for each disease-causing micro-organism or virus. This complexity is evident in the definition of zoonoses itself, i.e., any disease or infection that is naturally transmissible from vertebrate animals to humans.[2] As soon as a zoonotic pathogen spills over from vertebrate animals to humans, opportunities arise to adapt and transmit from human to human as well (e.g. COVID19, Monkeypox). When zoonotic pathogens move to this next evolutionary stage, they pose an even bigger public health threat than before.[2] Besides, zoonoses can be caused by several microorganisms, such as bacteria, viruses, parasites or fungi. These microorganisms can infect susceptible individuals in a population via various transmission pathways depending on the pathogen in question, such as through direct contact with animals and/or feces, or via food, water, soil, air or any other matrix or substrate that can be contaminated with infectious forms of that pathogen.[2] All these different options for transmission lead to many questions as to where a person acquired a zoonosis. Common questions are, for example, how did a given zoonotic pathogen infect a person in the first place? What was the source of the infection and how did the pathogen spread from that source to humans? Was it the food eaten? Did someone sneeze too close at work? Or was it a contaminated environment, including a farm nearby? Etc. These are important questions, because: "*Everything comes from somewhere*" [3], and zoonoses are no exception.

Zoonoses explained

Pathogens need a suitable reservoir where they can multiply and maintain themselves for spread to new hosts. In the context of this thesis, these reservoirs are not humans but vertebrate animals. Zoonotic pathogens need those animal reservoirs as they provide the necessary principal habitat for them to survive and multiply in sufficient numbers, evolve, and be transmitted to a susceptible person or another living animal.[4] When a zoonotic pathogen is transmitted from a vertebrate animal to a human, a zoonotic spillover event occurs.[5] If this spillover is caused by previously unknown agents, or known microorganisms manifesting themselves in places or in species in

which the disease was apparently unknown, the related zoonosis can be defined as an emerging zoonosis.[6] Spillover events can also be caused by endemic zoonoses, which are constantly maintained at a baseline level in a specific geographic area and an animal population.[6]

When an animal or human host harbors a zoonotic pathogen, the source usually can be found at a specific time and place and can be identified as another animal, individual, or an object (e.g. fomite), as well as food, feed, water, soil. etc. Thus, the source is present in the direct, external environment of the host.[4] Identifying the source of infection is important to prevent others from acquiring the infection. However, identifying one source may lead to another source being involved.[4] The pathway in which a pathogen travels from one source to another one up to the final host is the "transmission route".[4] Tracing back the sources of infection along the chain of transmission ends with revealing the original reservoir, where sources can be a part of.[4] But what are the mechanisms behind the transmission of zoonotic pathogens?

Zoonoses can be spread in essentially three ways, either directly, indirectly or through airborne transmission.[4] Direct transmission is defined as the immediate transfer of pathogens to a susceptible individual via direct contact with animals or droplet spread (particles larger than 5 μm[7]).[4] However, droplet spread could be included in the definition of airborne transmission instead, as the used definitions are not conclusive. Overall, direct transmission routes are often linked to behavior.[4] Most preventive measures, therefore, focus on reducing risk behavior, such as washing hands after contact with animals.[4]

Indirect transmission is the transmission of a zoonotic pathogen from an animal through fomites, such as inanimate objects and surfaces, or vehicles like food, water, or vectors like arthropods (e.g. ticks, mosquitoes, etc.).[4] Control measures for endemic zoonoses often focus on reducing the risk of transmission by enhancing the safety of procedures (for example, hygiene during slaughtering for meat production) or controlling the size of the vector population.[4]

With airborne transmission, pathogens are suspended in aerosols (fluid or solid) for prolonged periods of time (as opposite to droplets, that are too large in diameter).[4] However, droplets can remain airborne for prolonged periods of time as well, depending on particle movement and size which are affected by conditions like wind speed, humidity, and temperature, especially in open air[8]. Therefore, one could argue to include droplet spread here. Airborne transmission is particularly efficient for viruses[4]. When a zoonosis is able to transmit through airborne transmission from human to human, this can result in a fast increase in number of cases in an outbreak which is difficult to contain. A recent example of such a rapid developing outbreak

resulting in a pandemic is COVID-19. Measures to mitigate airborne transmission are, for instance, the provision of sufficient and effective ventilation, and avoidance of overcrowding. However, the evidence base for such measures is still limited and there is hardly any consensus on the effectiveness of those measures.[9]

In case of zoonotic outbreaks related to livestock, it is important to prevent farm-to-farm transmission, for instance by limiting or banning animal transport in a predefined area surrounding the farm harboring infected animals.[10] Other options to prevent further spread are vaccinating animals or improving hygiene measures on farms and during transport.[10] More drastic measures include, for example, culling all livestock (including healthy animals) from an infected farm and sometimes even its neighboring farms.[4,10,11] Besides reducing farm-to-farm transmission, those measures also reduce the potential for spillover to humans and causing outbreaks of zoonotic diseases in the human population.

The origins of zoonotic infectious diseases

Before the 19th century, the origins of diseases, including zoonoses, were difficult to grasp. There were different theories about the origin of infectious diseases. One example of such a theory was born from the societal fear for a supernatural origin of the plague.[12] Another one was that infection with the plague was caused by disease carrying vapours (miasmas) emanated from corpses.[13] It took scientists up to the 19th century to discover that the plague was caused by the bacterium *Yersinea pestis* and that it originated from rodents as reservoirs and fleas as vectors, while the plague has been one of the first known emerging zoonoses in human history.[14] It caused three human pandemics from the 6th century onwards, of which the "Black Death" pandemic (1347-1351) is considered as one of the most lethal in history of humankind, with 17-28 million people killed in Europe alone (~30-40% of the population).[14]

At the end of the 19th century, the German physician and pathologist Rudolf Virchow introduced a new term to describe human diseases shared with animals: "zoonoses", which is derived from the ancient Greek (*zoon*: animals, and *noson*: disease).[6,15] This was an important step forward in recognizing the need for a more comprehensive view on animal and human medicine. As he said himself:

"Between animal and human medicine, there is no dividing line, nor should there be".
Rudolf Virchow (1821-1902)

To date, zoonoses are responsible for a significant and growing burden of human disease worldwide due to their number, their frequency and their severity in relation

to human health.[16] In fact, 58% of the known human infectious diseases have a non-human animal source and up to 75% of the new diseases affecting people over the past decade originated from animals.[17] Note that some of the original zoonoses have evolved and gradually adapted to human-to-human transmission over time and are transmissible between humans nowadays (e.g. human tuberculosis, COVID-19).[18] Some zoonotic pathogens even became purely human pathogens, such as the human immunodeficiency virus (HIV), which mutated into human-only strains, while the disease originally was a zoonosis.[17]

Worldwide, there are over 150 zoonotic diseases, of which only 13 are responsible for 2.2 million deaths per year.[19] Moreover, the majority of recent epidemics and pandemics have been caused by zoonotic pathogens.[16] Between 2003 and 2020, six Public Health Emergencies of International Concern have been declared by the World Health Organization (WHO).[20] Five of them had a zoonotic origin: the H1N1 influenza pandemic (2009), the West African Ebola virus disease epidemic (2013–2016), the Democratic Republic of Congo Ebola virus disease epidemic (2018–2020), the Zika virus disease pandemic (2015–2016) and the COVID-19 pandemic (2020–present).[20] This emphasizes the need to prevent the emergence and spread of infectious agents. To this end, it is important to gain insights into the origins, sources and transmission routes of the pathogens in question.

Zoonoses in the Netherlands

Zoonoses can cause different diseases in humans. These diseases and their accompanying symptoms can either be gastrointestinal (e.g. campylobacteriosis or salmonellosis), respiratory (e.g. Q-fever) or systemic (e.g. hantavirus infection). The most frequently reported human zoonoses in the Netherlands (excluding vector-borne zoonoses) are campylobacteriosis, salmonellosis, and infection with Shiga toxin-producing *Escherichia coli* (STEC), which are endemic worldwide.[21] In the Netherland, these three pathogens had a mean estimated number of incident cases in the general human population (x 1,000) of respectively 73.0 (9.5-198), 26.0 (2.4-81) and 2.1 (0.2-8.8) in 2019.[21] Together, these pathogens cause over 4,000 human disability-adjusted life years (DALYs) in the Netherlands each year, with corresponding expenses of ~€87 million/year.[21]

The gastrointestinal bacteria *Campylobacter*, *Salmonella* and STEC O157 have different animal sources. In the Netherlands, for example, the main animal sources of human infection with *Campylobacter* are broilers and cattle.[22] Pigs and laying hens are the main sources of *Salmonella*[23] and domestic ruminants (cattle, sheep and goats) are the main sources of human STEC O157 infection[24]. Human infections are predominantly acquired through the fecal-oral route via contaminated food or water.[25] All these three infections

can result in gastroenteritis (GE), which can be defined as having three or more loose stools within 24 hours or any clinically relevant vomiting (i.e. vomiting events other than regurgitation, vomiting due to motion sickness/vertigo, traumatic event, nauseous event, or drug/alcohol abuse).[26] Those infections can have serious and sometimes long-term sequelae beyond gastroenteritis, such as Irritable Bowel Syndrome (IBS) following salmonellosis or campylobacteriosis, Guillain-Barré syndrome after campylobacteriosis or hemolytic-uremic syndrome (HUS) following infection with STEC.[27-29]

As animals can be infected with the aforementioned zoonotic pathogens as well, pathogen presence in livestock is also monitored. 44% of the poultry flocks were (highly) contaminated with *Campylobacter* in 2019.[30] Besides, 22 of the 198 sampled broiler farms (11%) reported the presence of *Salmonella* in their samples, but the prevalence of *Salmonella (S.)* Typhimurium and *S.* Enteritidis was low (0.1%).[30] Additional analyses revealed the presence of other *Salmonella*-serotypes: *S.* Paratyphi B var. Java (10x), *S.* Infantis (9x).[30] *S.* Agona, *S.* Goldcoast and *S.* Saint Paul each were present at one farm.[30] Furthermore, 27 laying hen farms reported 45 suspicious stables in 2019, which potentially harbored either *S.* Enteritidis or (monophasic) *S.* Typhimurium infected laying hens.[30] Verification-studies of 26 stables confirmed the contamination of 18 stables.[30] Positive samples of pigs in slaughter houses were dominated by *S.* Typhimurium (48 of the 133 positive samples in 2019).[30] STEC was only found in one meat poultry farm, but was present in 59% of the fecal samples taken at pig farms in 2019 and detected at almost all dairy goat and sheep farms that were included in the Dutch surveillance of zoonoses in 2016.[30,31] Besides in small ruminants, STEC is also common in dairy cows: 21% of the dairy cow farms participating in the Dutch surveillance of zoonoses in dairy cows in 2021 were positive.[32]

The manifestation of these zoonotic infections in animals can differ from its manifestation in humans. *Salmonella*, for instance, does not generally cause clinical disease in pigs or poultry.[33] The likelihood that it causes disease in animals depends, as always, on both bacteria- and host-related factors, such as the bacterial subtype and virulence, the animal species, its age and health status, among others. If clinical symptoms are observed in poultry upon *Salmonella* infection, it is mainly in young chickens up to two weeks of age, showing weakness, loss of appetite, watery diarrhea and poor growth, which may eventually lead to death.[33] Also STEC infection in ruminants is usually asymptomatic,[34] as *Campylobacter* infection is in chickens.[35] However, there are changes in the intestinal barrier function in *Campylobacter*-infected chickens,[35] and these changes are associated with decreased growth in otherwise asymptomatic birds.[35] The fact that these endemic zoonoses usually do not cause symptoms in their livestock sources makes detection, as well as the implementation of measures in the transmission chain, more difficult when compared to animals that do exhibit symptoms (e.g. avian influenza A viruses).

Drivers of change

The number of (re-)emerging zoonoses is increasing worldwide, as is the risk for both public and animal health.[36,37] The explanation for this increase can be found in factors that either directly or indirectly influence the emergence of zoonotic diseases. Those are "drivers of change" and can be subdivided into three main categories: globalization and environment (e.g. climate, the human-made environment, global travel and animal transport), sociodemographic factors (e.g. demographic, intensive livestock farming and social inequality), and public health systems (e.g. healthcare system, animal health and food and water quality).[37]

The first category of the aforementioned "drivers of change", globalization and environment, contributed to 61% of individual infectious disease threat events (IDTEs; outbreaks affecting > 5 persons) in Europe alone in the period 2008-2013.[37] Europe, therefore, is one of the hotspots for infectious disease (re-)emergence.[36,37] This directly relates to the major change in dynamic global trends in recent years, which led to an increased number of opportunities for emerging infectious diseases (EIDs) to occur and expand apace.[37] Not only travel and tourism contributed to the increased number of IDTEs, also food and water quality, the human-made and natural environment and climate change were in the top five of most important drivers of IDTEs in Europe.[37] Besides, also the increased transport of animals over larger distances worldwide has to be considered. During such transports, there can be contact between different animals from various countries, increasing the risk of potential infections and increased opportunities for global spread of the disease.[10]

Most IDTEs were the result of a combination of two or more drivers in Europe.[37] One of the reasons that the Netherlands is considered to be a high risk country for zoonotic disease (re-)emergence, is its high and increasing human population density (>500 inhabitants/km^2) in combination with the close proximity of residents to intensive livestock farms, which affect the environment in several ways. For instance, intensive livestock farming leads to air pollution (fine and coarse dust, endotoxins and ammonia), but also water and soil quality might be affected depending on local conditions, manure treatment methods, and environmental regulations.[38] Those farm pollutants not only affect the environment. Farm emissions, for instance, can have potential health effects, such as zoonotic infections, infections with antimicrobial-resistant bacteria, and respiratory disorders.[38] Therefore, agricultural pollutants can be of relevance for public health.[38]

The threats, often caused by the (re-)emergence of (zoonotic) infectious diseases, can become major public health risks and even emergencies as soon as they lead to many human cases and/or casualties.[5] But before a new disease can emerge, there needs

to be alignment of several drivers and determinants.[5] In other words, the zoonotic pathogen needs to overcome a series of barriers in a specific order to be able to spillover to a human host.[5] Examples of such barriers are the ecological, epidemiological and behavioral determinants of pathogen exposure and within human factors affecting susceptibility to infection.

One Health

The complexity of zoonoses requires a multilevel and multidisciplinary approach in order to study the influence of all different determinants, drivers and the relationships between them at the human-animal-environment interface (HAEI).[39,40] Examples of multilevel thinking are present in every era of epidemiological research.[41] One of the best known, and first spatial-epidemiological study, is the work of John Snow of 1854 on Cholera and its link to contaminated drinking water from wells in the London city center and thus the human-made environment.[41,42] Despite a focus shift towards non-communicable (chronic) diseases in the mid-twentieth century, multilevel thinking re-emerged at the 2000 International AIDS Conference.[41] The overarching theme of this conference was to recognize causes of AIDS and its transmission on multiple levels.[41] In 2004, this perspective was adopted and included in the "Manhattan Principles", which were developed during the "Building Interdisciplinary Bridges to Health in a 'Globalized World' symposium.[43] The outcome included 12 priorities to combat health threats to human and animal health and called for an international, interdisciplinary approach to prevent disease.[43] This established the foundation for the "One Health, One World™" concept.[43]

To move the concept of One Health forward, specific action steps were defined during the "Stone Mountain Meeting" in 2010.[43] During this meeting, experts of the Centre for Disease Control and Prevention (CDC), the World Organization for Animal Health (WOAH), the Food and Agriculture Organization of the United Nations (FAO), and the World Health Organization (WHO) defined seven key activities to advance the One Health agenda.[43] They identified not only the need for greater incorporation of animal health, but also the need of environmental support to human health and the need for greater incorporation of environmental (contamination with toxic chemicals, agricultural intensification, population growth etc.), spatial (e.g. distance to poultry farm) and ecosystem (e.g. habitat of hosts or vectors) drivers into disease burden assessments and interventions.[44,45]

The aforementioned is summarized in the following definition of One Health:

"One Health is a collaborative, multisectoral, and transdisciplinary approach – working at the local, regional, national, and global levels – with the goal of achieving optimal health outcomes recognizing the interconnection between people, animals, plants, and their shared environment."

CDC (2022)

This view of One Health can be presented as a disease triangle[46], with the pathogenic microorganism, often zoonotic, in its center and human, animals and the environment each as one corner (Figure 1). This captures the interplay within the HAEI and each determinant and driver of disease can be assigned to one of the corners (humans, animals or the environment), which are all interacting with each other and the disease causing pathogen in question.

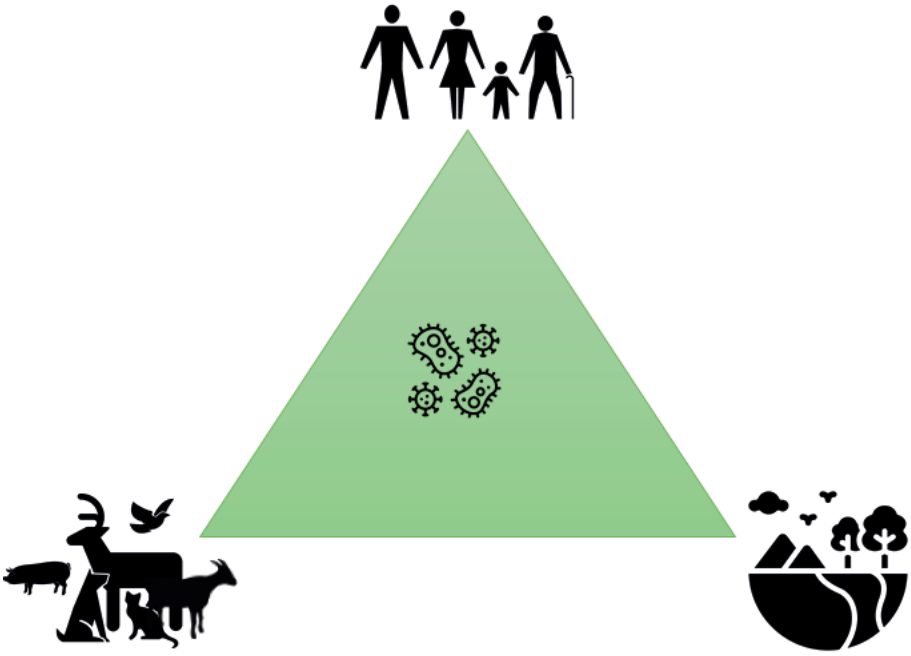


Figure 1. A visual representation of the definition of One Health, summarized as a disease triangle with the pathogenic microorganism in the center, interacting with all corners of the HAEI (human-animal-environment interface). *Icons from flaticon.com*

Eco-epidemiology

Previous epidemiological studies tried to incorporate the effects of the environment on human health by focusing on the interrelation between human disease and the environment using environmental epidemiology methods by studying population's exposure to contaminants in water, air and soil. Results are the smoke-free legislation, control of particulate air pollution and the introduction of proper sanitation. Current efforts in the Netherlands to reduce exposure to air pollution include the introduction of areas where only clean vehicles are allowed, the mandatory use of particulate filters in vehicles and the introduction of clean public transport.[47,48]

Although the above-described studies achieved remarkable gains, they did not include the effects of ecological (e.g. habitat type) and socio-demographic (e.g. address/region) drivers on public health, and infectious diseases were scarcely part of the equation. Besides, it remains complex to unravel all drivers playing a role in human health and diseases from those (re-)emerging pathogens transmitted through multiple routes, including exposure to animals, food, water and contaminated environments.[39] Studying zoonotic infectious disease eco-epidemiology covers this field and seeks to understand the environmental, spatial, ecological, and socio-demographic drivers of infectious disease emergence and transmission, including those of outbreaks, by involving the health of humans, animals and ecosystems, but also environmental sustainability and socioeconomic stability.[44,45,49-51] This is derived from a suggestion of David Waltner-Toews, which states that a sustainable environment and thus sustainable and healthy ecosystems will eventually lead to sustainable human and animal health.[50,52] Both human and animal health cannot be sustained in an environment where resources are depleted and polluted.[50,53] In general, this discipline incorporates the interrelations of all living things and could be described as human and animal ecology, or the part of human and animal ecology relating to states of health.[41,54] Health, therefore, can be viewed in terms of dynamic states influenced by factors on multiple levels, such as the cellular, the individual, the community and the population.

There are several definitions of eco-epidemiology, as it is a developing field, but they all have an 'ecological' perspective.[41] The definition of eco-epidemiology adopted within this thesis is the following:

"Eco-epidemiology (or ecological epidemiology) is the study of the ecology of infectious diseases. It includes population and community level studies of the interactions between hosts and there pathogens and parasites, and covers diseases of both humans and wildlife."
Nature (2022)

This definition clarifies that eco-epidemiology is a way to study complex interactions in the field of zoonotic infectious diseases and that it can play a major role in finding answers about the drivers, transmission routes and sources of zoonotic infectious diseases.

Knowledge gaps

In theory, the One Health approach has the potential to improve our understanding of infectious disease emergence, spread, and control, especially for zoonoses, including their interrelations with humans, animals and the environment. However, each pathogen behaves in a different way, with specific reservoirs, hosts and transmission routes, which are all interrelated within the HAEI. Besides, the (re-)emergence of pathogenic micro-organisms continuously creates new fields to study. Although associations are being gradually unraveled by researchers worldwide, several knowledge gaps remain to be filled.

Climate change as driver for foodborne diseases

Climate is one of the drivers of change in the (re-)emergence of infectious diseases. This driver is part of the overall category "globalization and environment", which contributes most (61%) to IDTEs in Europe and affects exposure pathways of foodborne and waterborne diseases amongst others.[37] Recent developments in climate change steer changes in temperature, humidity, wind and rainfall worldwide.[55] In the Netherlands, for instance, extreme rainfall events (usually > 30 mm rainfall/hour for > 1 hour) are expected to increase in frequency, intensity and duration.[56] As the urban sewage drainage system capacities (~20 mm of rainfall per hour) are not (yet) adapted to those amounts of water in a relatively short timeframe, this can lead to urban pluvial flooding.[56,57] During such flooding events, pluvial water mixes with sewage water, thereby contaminating floodwater with fecal material and its related pathogens.[58] Most of those pathogens are recognized causative agents of gastrointestinal and respiratory infections, such as noroviruses, enteroviruses or *Campylobacter*.[58] Thus, the floodwater becomes a vehicle for several pathogens. This can lead to human exposure to contaminated floodwater and creates a risk for public health, through for example post-flooding cleaning or recreational activities, or through accidental exposure when passing through (e.g. walking or cycling) or splash exposure. The possible risk for the development of gastrointestinal, influenza-like illness and dermatological complaints through floodwater exposure is already described in a previous Dutch study.[56,58] However, research on health risks, such as the development of acute gastroenteritis (AGE) or acute respiratory infection (ARI), associated with exposure to urban pluvial flooding is limited.

The human-made environment as driver for foodborne disease and gut microbiota

Spatial distribution of foodborne pathogens from livestock

Another group of drivers from the overall category "globalization and environment" cover the human-made environment. Examples are rapid urbanization, built environment and intensive agriculture and livestock farming, which can enable propagation and dissemination of pathogens.[37] For instance, the Netherlands is a densely populated country with over 500 inhabitants/km^2 of which ~ 93% currently (2022) lives in urban areas.[38] This high degree of urbanization, in combination with the presence of residential areas in close proximity to intensive livestock farms, but also to natural areas and wildlife, creates a high potential for the (re-)emergence of zoonotic pathogens[38], as is also shown by the IDTE heat map of the world.[59] Indeed, several micro-organisms have recently emerged in the Netherlands. Examples are West Nile virus, COVID-19, and Seoul virus.[60] The close proximity of the human and animal population also leads to outbreaks of specific endemic diseases, such as the avian influenza outbreak in poultry in 2003, smaller outbreaks of avian influenza in 2021-2022, and the Q-fever epidemic among Dutch residents in 2007-2010.[60] Since October 2021, there are 37 findings of highly pathogenic avian influenza at commercial poultry farms reported in the Netherlands. This ongoing outbreak started with the finding of two infected wild swans, but also farm-to-farm transmission plays a role. The continuously changing epidemiology of avian influenza makes it complicated to stop this particular outbreak, and prevent recurrent outbreaks of avian influenza from happening. It also raises concerns about potential transmission from animals to humans, related human infection and further evolutionary adaptations of the pathogen in question towards having the potential of human-to-human transmission. Especially because the epidemiology of avian influenza is not yet completely understood.

The most commonly reported zoonotic agents among human cases in the Netherlands (*Campylobacter*, non-typhoid *Salmonella*, and STEC, particularly STEC O157) are endemic in this country and worldwide. Because these pathogens are considered to be typically foodborne pathogens[61,62], most interventions have been focused on reducing the pathogen load and potential spread throughout the food production chain. However, at least for campylobacteriosis, there has been no significant decrease in the occurrence of human cases in many European countries so far, while the incidence of human salmonellosis has stabilized in several European countries after decades of significant decline.[30] Also for human STEC O157 infections, there are multiple sporadic cases reported without a link to food sources each year.[63] These unexplained, sporadic cases and the stabilization of the number of reported cases, together with the

presence of the pathogens in livestock and wild animals, the high human population density and the proximity of Dutch residents to livestock farms, raises the question as to whether environment-mediated transmission pathways are important as well. The potential role of the environment in transmission of pathogens from livestock to humans can be illustrated by the shear amount of pathogen that is shed into the environment by the entire livestock population, the so-called "reservoir output load". As an example, we used cattle and STEC O157 in the Netherlands. Each positive dairy cow sheds 1,000,000 CFU kg^{-1} feces into its direct surroundings.[64] As one cow produces 24,983 kg feces per year[65] and there are 46,197 positive milk cows in the Netherlands each year[64,66], the reservoir output load of STEC O157 from dairy cows is estimated to be $1.15*10^{15}$ CFU each year for this pathogen only. Thus, the effects of livestock farming as potential driver for the (re-)emergence of zoonoses, especially in close proximity to the human population, should not be underestimated and are in need of further scientific exploration.

Livestock-associated Campylobacter in surface water

A first step to study the possibility of environment-mediated transmission routes from livestock farms to humans, can be to gather knowledge about which type of environment acts as vehicle for transmission for a pathogen. A step further would be to determine from which source(s) the pathogens originate and which source(s) contribute most to environmental contamination. In case of *Campylobacter*, it is important to involve the aquatic environment as possible transmission route, as *Campylobacter* is commonly found in surface water contaminated with animal feces, sewage effluent and agricultural runoff.[67] A previous source attribution study attempted to quantify the relative contributions of different animal sources to surface water contamination with *Campylobacter*.[68] They found that the main contributors to surface water contamination with both *Campylobacter* species of public health significance, *C. jejuni* and *C. coli*, were poultry and wild birds in the Netherlands and Luxembourg, respectively.[68] However, the extensive use of non-local isolates (i.e. animal isolates were not all from the same countries as the surface water isolates) and non-recent isolates (i.e. animal isolates were not all from the same years as the surface water isolates), as well as retail food data and a coarse spatial resolution of the analyses, complicated interpretation of results.[68] Moreover, it remained unclear to which extent surface water contamination by *Campylobacter* was determined by fecal pollution from different types of livestock and wildlife. Therefore, more information is required to further support the implementation of measures related to the prevention of pathogen spread (*Campylobacter* in this case) from farm animals into the environment.

Livestock-associated spatial risk factors for human disease

Studying causal relationships in terms of environment-mediated transmission routes of zoonotic pathogens is complex. Not only because of all potential determinants involved in transmission of zoonoses, but also because of all other issues involved, including person-hours for performing the study, study-equipment, related expenses (covering men-hours and costs of study-equipment purchase and maintenance), timing, ethics and many other factors. In short, determining whether there is a causal effect and obtaining an accurate estimate of the effect requires rigorous and time-consuming research.[69]

To start testing the aforementioned hypothesis, it is possible to look for potential increased infection risk linked to specific transmission routes, for example by looking at associations between human infection and exposure to (the pathogens via) specific risk factors to get a first impression of the factors influencing disease occurrence risk. Examples are epidemiological studies like case-control and cohort studies including relevant variables to determine the importance of particular types of exposure.[70] A way to study specifically the association between human infection and exposure to certain risk factors suggestive of environmental transmission is the use of spatial regression analysis within an ecological design to determine the probability of exposure.[71-73] The latter methodology generally requires less resources and, therefore, can be a preferred way of exploring new ideas with possibly existing data. However, it should be noted that especially in the field of zoonotic infectious diseases, the interpretation of such studies is difficult due to the complex dynamics of disease spread. Frequently, ecological studies depend on aggregated data only, which can suffer from the problem of ecological fallacy leading to false negative or positive finding.[69] Here, incorrect interpretations occur when inferences about individuals are reasoned from inferences about the group which an individual belongs to.[69] Furthermore, potential confounding can have a major effect, as on aggregate level there are multiple factors that can vary and even co-vary between geographical entities and that cannot be fully and adequately controlled for.[69]

Despite these disadvantages, the application of spatial regression analysis with the necessary precautions can be a first step to gain more insights into environment-mediated transmission pathways in otherwise foodborne pathogens (STEC O157, *Campylobacter*, *Salmonella*), to be able to study associations with their animal reservoirs (livestock) in the Netherlands. As the aforementioned three zoonotic pathogens have limited human-to-human transmission, their transmission is relatively less complex, which can be seen as a prerequisite to use a more classical epidemiological design.

The number of studies focusing on spatial associations between human zoonotic infections and the probability of exposure to livestock is limited.[71-73] Besides, existing studies into the spatial association between, for example, STEC O157 infections and livestock only include one domestic ruminant species (either cattle, sheep or goat) at a time in the analysis, while ignoring other important animal reservoirs.[72,73] This might impact the outcome of those studies, especially in countries like the Netherlands, where high numbers of different types of livestock are present and mixed on relatively small geographical scales. Moreover, the probability of exposure should be a combination between the number of residents and the number of animals in a certain area, while past studies determined exposure as the number of animals in a certain area solely.[71-75] Therefore, it is necessary to develop new methods to study the association between human zoonotic infections and livestock density, preferably using exemplar zoonotic pathogens that have been studied previously in the same area for the sake of comparison, i.e. STEC O157. Ideally, these new methods should be applicable to the study of other pathogens as well, such as *Campylobacter* and *Salmonella*. This is especially of interest for *Campylobacter,* as in the Netherlands there was an observed drop in campylobacteriosis incidence after the massive culling operations following the H7N7 avian influenza epidemic in 2003.[11] This led to the formulation of the hypothesis that those culling operations caused a drop in human cases due to a reduction in environmental contamination with *Campylobacter* from the temporally closed (because of the epidemic) poultry farms and slaughterhouses. However, this hypothesis has never been studied further. Furthermore, a Dutch study showed that it is possible to quantify airborne concentrations of livestock-related bacteria such as *C. jejuni* at residential levels amongst others, which was associated with the presence of poultry farms in the neighborhood.[70] As there is growing evidence that foodborne pathogens can also be transmitted via the environment, the question raises whether there also is a spatial association between *Campylobacter* and *Salmonella* infection and livestock density in the Netherlands.

The urban-rural gradient and gut microbial diversity and composition

The human living environment is not only a driver of infectious diseases, but it also interacts with our microbiome. This microbiome is unique for every individual, just like a fingerprint. It consists of trillions of different microorganisms together with their joint genomes, living in a complex ecosystem inside the human gut, with a unique set of strains of each species present in each individual. This ecosystem plays an important role in terms of both physical and mental human health. Although it strives towards a stable equilibrium, its composition and diversity can be influenced by many factors.[76,77] A limited number of studies investigated and showed the potential health effects of the farming environment on local residents beyond farmers and their families. For example, living in close proximity to livestock farms was associated

with protection against asthma and atopy.[78,79] When focusing more specifically on the effects of environmental microbial exposure on the diversity and composition of the human gut microbiome, differences in both airway and gut microbiotas between infants in urban and rural living environments have been observed.[80] Besides, an elevated risk of asthma and atopic traits with urbanization-related changes in the microbiota have been shown.[80] As environmental microbial exposure affects the human microbiome throughout life[81,82], it would be interesting to study whether human gut microbiome diversity and composition in adults is affected by the urban-rural gradient throughout the Netherlands.

Thesis aims and outline

In this thesis, we aim to fill the aforementioned knowledge gaps using mainly existing data sets from different surveillance and research activities. Those knowledge gaps were formulated based on questions that arose from previous studies and observations but remained largely unanswered so far. By answering these questions, we aim to provide novel insights into the potential role of environmental transmission into the epidemiology of infectious agents, namely selected exemplary zoonotic pathogens that are typically considered as foodborne pathogens (i.e. *Campylobacter*, *Salmonella* and STEC O157), in the Netherlands. This was done by studying the potential human health effects linked to extreme rainfall events, the pathogen characteristics in the environment (for *Campylobacter*), and the potential spatial association between human infection with all three foodborne pathogens in question and the local livestock density, using newly developed methods within the concept of eco-epidemiology. In addition, this thesis explored the spatial association between the urban-rural gradient and gut microbial diversity and composition. Broadly, the aims of the different chapters focused on corresponding drivers of change fitting in the category "globalization and the environment".[37] These drivers were climate change and the human-made environment. Together, the themes created a flow throughout this thesis, from syndrome-level to pathogen-specific, as well as more holistic approaches focused on the potential effects of the environment on human infectious diseases and gut microbiota. The first theme focusses on climate change as driver for foodborne infectious diseases and aims specifically at studying the associations between health risks associated with exposure to urban pluvial flooding. Here, health risks are defined as the possible risk for the development of acute gastroenteritis (AGE) and acute respiratory infection (ARI) after such exposure. The subsequent theme covered four chapters of this thesis related to the human-made environment. Three of these chapters focused on the role of potential environmental transmission of the pathogens *Campylobacter*, STEC O157 and *Salmonella* from livestock farming. Particularly, the first chapter aimed to study the prevalence and genotype diversity of

Campylobacter in surface water, as well as the relative contributions of several putative animal sources to surface water contamination with *Campylobacter*. Additionally, potential effects of local livestock density, type of surface water, and season were assessed. The next two chapters focused on studying the spatial associations between human STEC O157, campylobacteriosis and salmonellosis incidence and livestock exposure using a novel exposure measure, i.e. the population-weighted number of farm animals. The last chapter covered a more holistic approach where we did not focus on a specific infectious disease, but rather on the human gut microbiome as readout. Here, we studied the potential (indirect) effects of the urban-rural gradient on the human gut microbiome, which is generally assumed to play an important role in the overall health status of individuals.

Climate change as driver for foodborne disease

Chapter 2 determines the risks of AGE and ARI associated with exposure to pluvial floodwater in the Netherlands. Furthermore, specific risk factors for AGE and ARI in pluvial flood-ravaged urban areas are identified. Here, we performed a retrospective cross-sectional study for this purpose in locations in the Netherlands with reported pluvial flooding. To do so, we used data from questionnaires with information on self-reported AGE and ARI symptoms after floodwater exposure.

The human-made environment as driver for foodborne disease and gut microbiota

Spatial distribution of foodborne pathogens from livestock

Livestock-associated Campylobacter in surface water

In **Chapter 3**, *Campylobacter* prevalence and genotype diversity in surface water is quantified, as well as the relative contributions of several putative animal sources to surface water contamination with *C. jejuni* and *C. coli* in the Netherlands. For this purpose, we used local and recent source data with a high spatial resolution and whole-genome sequencing (WGS) is used for tracing the (domestic and wild) animal sources of *Campylobacter* isolates in surface water. Additionally, potential effects of local livestock density, type of surface water and season are assessed.

Livestock-associated spatial risk factors for human disease

Chapter 4 describes a study that assesses the spatial association between sporadic human STEC O157 infections and the combined exposures to cattle, small ruminants, poultry and pigs in the Netherlands, thereby including the main animal reservoirs of human STEC O157 infection. Different state-of-the-art methods are

used that include the population-weighted numbers of animals in the calculation of the probability of exposure to livestock.

Chapter 5 directly builds on Chapter 4 as the newly developed method is applied to two other foodborne zoonotic pathogens: *Campylobacter* and *Salmonella*. It thus assesses whether human salmonellosis and campylobacteriosis incidence are spatially associated with local density of small ruminants, dairy cows, veal calves, laying hens, broilers and pigs in the Netherlands. Additionally, here we accounted for geographical coverage of the diagnostic catchment areas, as this was not necessary for STEC O157. It is possible that repeated exposure to the pathogen through the environment can potentially hinder the analyses based on the incidence of reported cases. Therefore, serological data is used as well to look into possible effects of acquired immunity.

The urban-rural gradient and gut microbial diversity and composition

In **Chapter 6**, data on gut microbiome and corresponding metadata from participants questionnaires of the PIENTER-III cohort is used to explore the potential (indirect) effects of the urban-rural gradient on the diversity and composition of the human gut microbiome in the Netherlands.

References

1 Worobey, M. et al. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. Science, 8715, doi:10.1126/science.abp8715.

2 Joint FAO/WHO Expert Committee on Zoonoses (World Health Organization & Food and Agriculture Organization of the United Nations). Joint WHO/FAO Expert Committee on Zoonoses [meeting held in Stockholm from 11 to 16 August 1958] : second report. Report No. 924120169X, 83 (World Health Organization, Geneva, 1959).

3 Quammen, D. Spillover - Animal infections and the next human pandemic. 1st edn, (W. W. Norton & Company, Incl., 2012).

4 Wormser, G. P. & Colebunders, R. L. Control of Communicable Diseases Manual. 21st edn, Vol. 49 (APHA Press, 2009).

5 Plowright, R. K. et al. Pathways to zoonotic spillover. Nature Reviews Microbiology 15, 502-510, doi:10.1038/nrmicro.2017.45 (2017).

6 Chomel, B. B. Encyclopedia of Microbiology (Third Edition)(ed Moselio Schaechter) 820-829 (Academic Press, 2009).

7 World Health Organization. Infection Prevention and Control of Epidemic- and Pandemic-Prone Acute Respiratory Infections in Health Care. WHO Guidelines Approved by the Guidelines Review Committee, (World Health Organization, 2014).

8 Yang, X., Yang, H., Ou, C., Luo, Z. & Hang, J. Airborne transmission of pathogen-laden expiratory droplets in open outdoor space. Science of The Total Environment 773, 145537, doi:https://doi.org/10.1016/j.scitotenv.2021.145537 (2021).

9 Morawska, L. & Milton, D. K. It Is Time to Address Airborne Transmission of Coronavirus Disease 2019 (COVID-19). Clinical Infectious Diseases 71, 2311-2313, doi:10.1093/cid/ciaa939 (2020).

10 Bekedam, H. et al. Zoönosen in het vizier. (2021).

11 Friesema, I. H. M., Havelaar, A. H., Westra, P. P., Wagenaar, J. A. & van Pelt, W. Poultry culling and Campylobacteriosis reduction among humans, the Netherlands. Emerging infectious diseases 18, 466-468, doi:10.3201/eid1803.111024 (2012).

12 Riva, M. A., Benedetti, M. & Cesana, G. Pandemic fear and literature: observations from Jack London's The Scarlet Plague. Emerg Infect Dis 20, 1753-1757, doi:10.3201/eid2010.130278 (2014).

13 Cipolla, C. M. Miasmas and disease: public health and the environment in the pre-industrial age. (Yale University Press, 1992).

14 Parkhill, J. et al. Genome sequence of Yersinia pestis, the causative agent of plague. Nature 413, 523-527, doi:10.1038/35097083 (2001).

15 Leal Filho, W., Ternova, L., Parasnis, S. A., Kovaleva, M. & Nagy, G. J. Climate Change and Zoonoses: A Review of Concepts, Definitions, and Bibliometrics. International Journal of Environmental Research and Public Health 19, 893 (2022).

16 Judson, S. D. & Rabinowitz, P. M. Zoonoses and global epidemics. Current Opinion in Infectious Diseases 34, 385-392, doi:10.1097/QCO.0000000000000749 (2021).

17 Woolhouse, M. E. & Gowtage-Sequeria, S. Host range and emerging and reemerging pathogens. Emerg Infect Dis 11, 1842-1847, doi:10.3201/eid1112.050997 (2005).

18 Slingenbergh, J., Gilbert, M., Balogh, K. d. & Wint, W. Ecological sources of zoonotic diseases. Revue scientifique et technique-Office international des épizooties 23, 467-484 (2004).

19 Wageningen University & Research. Zoonoses, <https://www.wur.nl/en/Research-Results/Research-Institutes/Bioveterinary-Research/In-the-spotlight/Zoonoses.htm> (2022).

20 Wilder-Smith, A. & Osman, S. Public health emergencies of international concern: a historic overview. J Travel Med 27, 1-13, doi:10.1093/jtm/taaa227 (2020).

21 Lagerweij, G., Pijnacker, R., Friesema, I., Mughini Gras, L. & Franz, E. Disease burden of food-related pathogens in the Netherlands, 2019. 1-51 (Rijksinstituut voor Volksgezondheid en Milieu RIVM, 2020).

22 Mughini-Gras, L. et al. Sources and transmission routes of campylobacteriosis: A combined analysis of genome and exposure data. J Infect 82, 216-226, doi:10.1016/j.jinf.2020.09.039 (2021).

23 Mughini-Gras, L. et al. Risk Factors for Human Salmonellosis Originating from Pigs, Cattle, Broiler Chickens and Egg Laying Hens: A Combined Case-Control and Source Attribution Analysis. PLOS ONE 9, e87933, doi:10.1371/journal.pone.0087933 (2014).

24 Mughini-Gras, L. et al. Attribution of human infections with Shiga toxin-producing Escherichia coli (STEC) to livestock sources and identification of source-specific risk factors, The Netherlands (2010–2014). Zoonoses and Public Health 65, e8-e22, doi:10.1111/zph.12403 (2018).

25 Musher, D. M. & Musher, B. L. Contagious Acute Gastrointestinal Infections. New England Journal of Medicine 351, 2417-2427, doi:10.1056/NEJMra041837 (2004).

26 Majowicz, S. E. et al. A common, symptom-based case definition for gastroenteritis. Epidemiology and Infection 136, 886-894, doi:10.1017/S0950268807009375 (2008).

27 Baumgart, D. C. & Sandborn, W. J. Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet 369, 1641-1657, doi:10.1016/s0140-6736(07)60751-x (2007).

28 Scheiring, J., Andreoli, S. P. & Zimmerhackl, L. B. Treatment and outcome of Shiga-toxin-associated hemolytic uremic syndrome (HUS). Pediatr Nephrol 23, 1749-1760, doi:10.1007/s00467-008-0935-6 (2008).

29 Willison, H. J., Jacobs, B. C. & van Doorn, P. A. Guillain-Barré syndrome. The Lancet 388, 717-727, doi:10.1016/S0140-6736(16)00339-1 (2016).

30 Vlaanderen, F. et al. in State of Zoonoses 2019 (Rijksinstituut voor Volksgezondheid en Milieu RIVM, 2020).

31 Opsteegh, M. et al. Surveillance zoönosen in de melkgeiten-en melkschapenhouderij in 2016. (2018).

32 Cuperus, T. et al. in Surveillance into zoonoses in dairy cattle 2021 (Rijksinstituut voor Volksgezondheid en Milieu RIVM, 2022).

33 De Busser, E. V., De Zutter, L., Dewulf, J., Houf, K. & Maes, D. Salmonella control in live pigs and at slaughter. The Veterinary Journal 196, 20-27, doi:https://doi.org/10.1016/j.tvjl.2013.01.002 (2013).

34 Nastasijevic, I. STEC O157 in the beef chain - risk assessment and management. Vol. 2011 (CABI International, 2011).

35 Awad, W. A. et al. Campylobacter infection in chickens modulates the intestinal epithelial barrier function. Innate Immun 21, 151-160, doi:10.1177/1753425914521648 (2015).

36 Jones, K. E. et al. Global trends in emerging infectious diseases. Nature 451, 990-993, doi:10.1038/nature06536 (2008).

37 Semenza, J. et al. Determinants and Drivers of Infectious Disease Threat Events in Europe. Emerging Infectious Disease journal 22, 581, doi:10.3201/eid2204.151073 (2016).

38 Smit, L. A. M. & Heederik, D. Impacts of Intensive Livestock Production on Human Health in Densely Populated Regions. GeoHealth 1, 272-277, doi:10.1002/2017GH000103 (2017).

39 Li, A. M. L. Ecological determinants of health: food and environment on human health. Environmental Science and Pollution Research 24, 9002-9015, doi:10.1007/s11356-015-5707-9 (2017).

40 Mackenzie, J. S., Jeggo, M., Daszak, P. & Richt, J. A. One Health: The Human-Animal-Environment Interfaces in Emerging Infectious Diseases: Food Safety and Security, and International and National Plans for Implementation of One Health Activities. (Springer Berlin Heidelberg, 2013).

41 March, D. & Susser, E. The eco- in eco-epidemiology. International Journal of Epidemiology 35, 1379-1383, doi:10.1093/ije/dyl249 (2006).

42 Shephard, David A. E. & Warner, David O. Cholera, Chloroform, and the Science of Medicine: A Life of John Snow. Anesthesiology 100, 1633-1634, doi:10.1097/00000542-200406000-00066 (2004).

43 Centers for Disease Control and Prevention. One Health Basics - History, <https://www.cdc.gov/onehealth/basics/history/index.html> (2022).

44 Eisenberg, J. N. et al. Environmental determinants of infectious disease: a framework for tracking causal links and guiding public health research. Environ Health Perspect 115, 1216-1223, doi:10.1289/ehp.9806 (2007).

45 Rabinowitz, P. M. et al. Toward proof of concept of a one health approach to disease prediction and control. Emerg Infect Dis 19, 1-8, doi:10.3201/eid1912.130265 (2013).

46 Scholthof, K.-B. G. The disease triangle: pathogens, the environment and society. Nature Reviews Microbiology 5, 152-156, doi:10.1038/nrmicro1596 (2007).

47 Pekkanen, J. & Pearce, N. Environmental epidemiology: challenges and opportunities. Environ Health Perspect 109, 1-5, doi:10.1289/ehp.011091 (2001).

48 Zijlema, W. L. et al. The LifeLines Cohort Study: a resource providing new opportunities for environmental epidemiology. Archives of Public Health 74, 32, doi:10.1186/s13690-016-0144-x (2016).

49 Lau, C. L. et al. Unravelling infectious disease eco-epidemiology using Bayesian networks and scenario analysis: A case study of leptospirosis in Fiji. Environmental Modelling & Software 97, 271-286, doi:https://doi.org/10.1016/j.envsoft.2017.08.004 (2017).

50 Lerner, H. & Berg, C. A Comparison of Three Holistic Approaches to Health: One Health, EcoHealth, and Planetary Health. Frontiers in Veterinary Science 4, doi:10.3389/fvets.2017.00163 (2017).

51 Mills, J. N., Gage, K. L. & Khan, A. S. Potential influence of climate change on vector-borne and zoonotic diseases: a review and proposed research plan. Environ Health Perspect 118, 1507-1514, doi:10.1289/ehp.0901389 (2010).

52 Waltner-Toews, D. Eco-Health: a primer for veterinarians. Can Vet J 50, 519-521 (2009).

53 Lerner, H. & Berg, C. The concept of health in One Health and some practical implications for research and education: what is One Health? Infection Ecology & Epidemiology 5, 25300, doi:10.3402/iee.v5.25300 (2015).

54 Susser, M. Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology. (Oxford University Press, 1973).

55 Wuebbles, D. J. et al. Climate Science Special Report: Fourth National Climate Assessment, Wuebbles DJ, Fahey DW, Hibbard KA, Dokken DJ, Stewart BC, Maycock TK, eds.Vol. I Exec. Sum., 12-34 (U.S. Global Change Research Program, 2017).

56 De Man, H. et al. Gastrointestinal, influenza-like illness and dermatological complaints following exposure to floodwater: a cross-sectional survey in The Netherlands. Epidemiology and Infection 144, 1445-1454, doi:10.1017/S0950268815002654. (2015).

57 Ten Veldhuis, J., Clemens, F., Sterk, G. & Berends, B. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water research 44, 2910-2918, doi:10.1016/j.watres.2010.02.009. (2010).

58 De Man, H. et al. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. Water research 48, 90-99, 10.1016/j.watres.2013.1009.1022. (2014).

59 Allen, T. et al. Global hotspots and correlates of emerging zoonotic diseases. Nature Communications 8, 1124, doi:10.1038/s41467-017-00923-8 (2017).

60 Vlaanderen, F. et al. in State of Zoonoses 2020 (Rijksinstituut voor Volksgezondheid en Milieu RIVM, 2021).

61 Friesema, I. H. M., Slegers-Fitz-James, I. A., Wit, B. & Franz, E. Surveillance and characteristics of food-borne outbreaks in the Netherlands, 2006 to 2019. Eurosurveillance 27, doi:10.2807/1560-7917.ES.2022.27.3.2100071 (2022).

62 Strachan, N. J. C., Dunn, G. M., Locking, M. E., Reid, T. M. S. & Ogden, I. D. Escherichia coli O157: Burger bug or environmental pathogen? International Journal of Food Microbiology 112, 129-137, doi:10.1016/j.ijfoodmicro.2006.06.021 (2006).

63 Franz, E. et al. Phylogeographic analysis reveals multiple international transmission events have driven the global emergence of Escherichia coli O157: H7. Clinical Infectious Diseases 69, 428-437, doi:10.1093/cid/ciy919 (2018).

64 Franz, E., Semenov, A. V. & Van Bruggen, A. H. C. Modelling the contamination of lettuce with Escherichia coli O157:H7 from manure-amended soil and the effect of intervention strategies. Journal of Applied Microbiology 105, 1569-1584, doi:10.1111/j.1365-2672.2008.03915.x (2008).

65 Nennich, T. D. et al. Prediction of Manure and Nutrient Excretion from Dairy Cattle. Journal of Dairy Science 88, 3721-3733, doi:10.3168/jds.S0022-0302(05)73058-7 (2005).

66 Vlaanderen, F. et al. Staat van Zoönosen 2020. (Rijksinstituut voor Volksgezondheid en Milieu RIVM, 2021).

67 Jones, K. Campylobacter in water, sewage and the environment. Applied Microbiology 90, 68S-79S, doi:10.1046/j.1365-2672.2001.01355.x (2001).

68 Mughini-Gras, L. et al. Quantifying potential sources of surface water contamination with Campylobacter jejuni and Campylobacter coli. Water Res 101, 36-45, doi:10.1016/j.watres.2016.05.069 (2016).

69 Heederik, D. J. J., Smit, L. A. M. & Vermeulen, R. C. H. Go slow to go fast: a plea for sustained scientific rigour in air pollution research during the COVID-19 pandemic. Eur Respir J 56, doi:10.1183/13993003.01361-2020 (2020).

70 de Rooij, M. M. T. et al. Insights into Livestock-Related Microbial Concentrations in Air at Residential Level in a Livestock Dense Area. Environ Sci Technol 53, 7746-7758, doi:10.1021/acs.est.8b07029 (2019).

71 Elson, R. et al. A spatial and temporal analysis of risk factors associated with sporadic Shiga toxin-producing Escherichia coli O157 infection in England between 2009 and 2015. Epidemiol Infect 146, 1928-1939, doi:10.1017/S095026881800256X (2018).

72 Friesema, I., Van De Kassteele, J., De Jager, C. M., Heuvelink, A. E. & Van Pelt, W. Geographical association between livestock density and human Shiga toxin-producing Escherichia coli O157 infections. Epidemiol. Infect. 139, 1081-1087, doi:10.1017/S0950268810002050 (2011).

73 ÓHaiseadha, C., Hynds, P. D., Fallon, U. B. & O'Dwyer, J. A geostatistical investigation of agricultural and infrastructural risk factors associated with primary verotoxigenic E. coli (VTEC) infection in the Republic of Ireland, 2008–2013. Epidemiology and Infection 145, 95-105, doi:10.1017/S095026881600193X (2017).

74 Hallisey, E. et al. Transforming geographic scale: a comparison of combined population and areal weighting to other interpolation methods. International Journal of Health Geographics 16, 29, doi:10.1186/s12942-017-0102-z (2017).

75 Mulder, A. C., Snabilie, M. & Braks, M. A. Ecology and prevention of Lyme borreliosis. Vol. 4 373-387 (Wageningen Academic Publishers, 2016).

76 Spector, T., Wolf, J. & Hadjigeorgiou, G. Rethinking the future of nutrition - Using large-scale biological data, AI and novel digital technologies to improve human health [White paper]. Chapter 5 - The gut microbiome plays an important role in our health. 25-30 (ZOE, 2020).

77 Stege, P. B. et al. Impact of long-term dietary habits on the human gut resistome in the Dutch population. Sci Rep 12, 1892, doi:10.1038/s41598-022-05817-4 (2022).

78 Borlée, F. et al. Residential proximity to livestock farms is associated with a lower prevalence of atopy. Occupational and Environmental Medicine 75, 453, doi:10.1136/oemed-2017-104769 (2018).

79 Müller-Rompa, S. E. K. et al. An approach to the asthma-protective farm effect by geocoding: Good farms and better farms. Pediatr Allergy Immunol 29, 275-282, doi:10.1111/pai.12861 (2018).

80 Lehtimäki, J. et al. Urbanized microbiota in infants, immune constitution, and later risk of atopic diseases. Journal of Allergy and Clinical Immunology 148, 234-243, doi:10.1016/j.jaci.2020.12.621 (2021).

81 Caporaso, J. G. et al. Moving pictures of the human microbiome. Genome Biol 12, R50, doi:10.1186/gb-2011-12-5-r50 (2011).

82 Gacesa, R. et al. Environmental factors shaping the gut microbiome in a Dutch population. Nature 604, 732-739, doi:10.1038/s41586-022-04567-7 (2022).

CHAPTER 2

"Sickenin' in the rain" – increased risk of gastrointestinal and respiratory infections after urban pluvial flooding in a population-based cross-sectional study in the Netherlands

Annemieke C. Mulder
Roan Pijnacker
Heleen de Man
Jan van de Kassteele
Wilfrid van Pelt
Lapo Mughini-Gras
Eelco Franz

Published in BMC Infectious Diseases, 2019

Abstract

Background

Climate change is expected to increase the chance of extreme rainfall events in the Northern Hemisphere and herewith, there is an increased chance of urban pluvial flooding. Urban pluvial flooding often consists of street flooding and/or flooding of combined sewerage systems, leading to contamination of the floodwater with several gastrointestinal and/or respiratory pathogens. An increase in flooding events therefore pose a health risk to those exposed to urban floodwater. We studied the association between exposure to pluvial floodwater and acute gastroenteritis (AGE) and acute respiratory infection (ARI).

Methods

We performed a retrospective, cross-sectional survey during the summer of 2015 in 60 locations in the Netherlands with reported flooding. Two weeks after the flooding, questionnaires were sent to households in these locations, collecting data on self-reported AGE and ARI and information on floodwater exposure in the previous 2 weeks. Multivariable generalized estimating equations (GEE) regression models, accounting for the clustered data structure, were used to identify risk factors for AGE and ARI.

Results

In total, 699 households with 1,656 participants (response rate 21%) returned the questionnaire. Contact with floodwater was significantly associated with AGE (aOR 4.2, 95%CI 2.1–8.4) and ARI (aOR 3.3, 95%CI 2.0–5.4). Risk factors for AGE were skin contact with floodwater (aOR 4.0, 95%CI 1.8-9.0), performing post-flooding cleaning operations (aOR 8.6, 95%CI 3.5-20.9) and cycling through floodwater (aOR 2.3, 95%CI 1.0-5.0). Skin contact with floodwater (aOR 3.6, 95%CI 1.9-6.9) and performing post-flooding cleaning operations (aOR 5.5, 95%CI 3.0-10.3) were identified as risk factors for ARI.

Conclusions

Results suggest an association between direct exposure to pluvial floodwater and AGE and ARI. As it is predicted that the frequency of pluvial flooding events will increase in the future, there is a need for flood-proof solutions in urban development and increased awareness among stakeholders and the public about the potential health risks. Future prospective studies are recommended to confirm our results.

Introduction

Climate change leads to extreme weather events that can have devastating effects on human society and the environment [1]. Climate change models predict an almost exponential increase in atmospheric water-holding capacity, with increasing temperature and subsequent rise in atmospheric water content leading to an increase in extreme rainfall events in the Northern Hemisphere [2]. As extreme rainfall events are expected to increase in frequency, intensity and duration, pluvial flooding events are also expected to increase in urban settings [3]. Indeed, in the last few decades, countries like the Netherlands have observed an increasing trend in extreme rainfall events, causing recurrent pluvial flooding, especially in urban areas [4].

Most urban sewage drainage systems in the Netherlands can only support an intensity of ~20 mm of rainfall per hour. Extreme rainfall events (usually > 30 mm rainfall/hour for > 1 h) may overwhelm this drainage capacity, leading to 10-50 cm of pluvial floodwater to accumulate on the surface [3, 5]. Urban pluvial flooding often entails street flooding and/or flooding of combined sewerage systems, where rainwater mixes with sewage water, thereby heavily contaminating floodwater with fecal material. In this way, floodwater becomes a possible vehicle of several pathogens such as noroviruses, enteroviruses, or *Campylobacter* [6], many of which are recognized causative agents of gastrointestinal and respiratory infections.

People will inevitably be exposed to this floodwater, especially in urban settings and when engaged in post-flooding cleaning or recreational activities (e.g. swimming, playing in water, etc.). It is also possible that people are accidentally exposed to floodwater when passing through it (e.g. walking, cycling, etc.) or because of splash exposure. Increased flooding events may therefore pose a threat for public health, particularly for the development of acute gastroenteritis (AGE) or acute respiratory infection (ARI) via ingestion and inhalation of, and dermal contact with, pathogens in floodwater as is shown by the study of De Man, et al. [6]. However, little research is available on health risks associated with urban pluvial flooding.

With a focus on a high-income country like the Netherlands, the aim of this study was to quantify the AGE and ARI risks associated with exposure to pluvial floodwater, as well as to identify specific risk factors for AGE and ARI, in pluvial flood-ravaged urban areas.

Methods

Study design

To collect data on self-reported AGE and ARI after urban pluvial flooding, a retrospective cross-sectional survey was conducted during the summer of 2015 using the methodology described in De Man, et al. [3]. In total, 3,382 households that were surrounded by floodwater after extreme rainfall events (extreme rainfall events are rainfall events for which streets are flooded) in 60 locations in the Netherlands received a self-administered questionnaire by regular post. The questionnaire was send 2 weeks after the start of the flooding event and participants were asked to answer questions considering a recall period of 2 weeks, i.e. from the start of the flooding until reception of the questionnaire. Flood locations were identified by monitoring possible flood events on the internet, such as social media and press releases, which generally receive much media attention in the Netherlands. The national website of the Dutch fire brigade was the main source of information to determine which households were surrounded by floodwater (http://www.112meldingen.nl/). This website provides the causes of a given alert, which the Dutch fire brigade receives from the citizens themselves, corroborates on-site, and disseminates to the public, specifying the affected areas (postal codes). With this information, the addresses of the affected areas to which the questionnaires were sent were derived from publicly available mapping tools like googlemaps.com [3].

Data collection

The questionnaire was developed by adapting the questionnaire used in the study of De Man, et al. [3] to collect epidemiologically relevant information for each household member (individual participant). Information was collected on basic demographic characteristics (i.e. age, sex, number of household members). Furthermore, information was gathered about the occurrence of gastrointestinal and respiratory complaints during the 2 weeks before questionnaire completion, underlying chronic diseases, type of exposure (i.e. contact) to floodwater, and type of activity leading to contact with floodwater, duration of flooding in minutes and the magnitude of exposure (height of floodwater in cm). Participants reporting one or more types of contact or types of activities leading to contact with floodwater are hereafter referred to as being exposed to floodwater.

Questions could be answered for up to five household members (most Dutch households are composed of ≤ 4 people according to Netherlands Statistics, https://www.cbs.nl/, [7]), and the participant was asked to report information on all household members.

Ethical considerations

This study involved collection and analysis of fully anonymized data, so no ethical approval was necessary according to Dutch regulations. [8, 9] People gave consent upon anonymously completing and returning the questionnaire. Questionnaires were received in de-identified form, containing only data on postal code, sex, and age of the participants. Therefore, names and addresses could not be linked to the questionnaire responses, guaranteeing anonymity of the respondents. Participants were informed that the data they provided were to be analyzed for scientific purposes, and that by returning the questionnaires, they gave consent to do so. No written consent was therefore necessary. Parents were asked to complete the questionnaire and give consent on behalf of their children (family members < 16 years old).

Case definitions

Based on the symptoms reported in the questionnaire, participants were defined as cases of AGE or ARI when experiencing the following complaints in the 2 weeks before completing the questionnaire, and as non-cases otherwise:

- AGE: any individual experiencing ≥ 3 diarrheal discharges per day or any clinically relevant vomiting (i.e. vomiting events other than regurgitation, vomiting due to motion sickness/vertigo, traumatic event, nauseous event, or drug/alcohol abuse) [10].
- ARI: acute onset of symptoms (within 2 weeks before completion of the questionnaire) and common cold, shortness of breath, coughing, sore throat, wheezing, chest pain or sneezing [11].

Data analysis

Generalized estimating equations (GEEs) were used to compare AGE and ARI in cases with the non-cases, with regard to their type of contact and type of activity leading to contact with floodwater. As our study design led to clustering of data at the household-level, we corrected for dependence of observations deriving from individuals living in the same household [12]. An exchangeable working correlation structure was chosen for the GEE models, meaning that the correlation between individuals within a cluster was the same for all clusters. An overall model was built for all age categories (categorical: 0-5, 6-15, 16-25, 26-45, 46-65 and > 65 years), as well as separate models for adults (categorical: 16-25, 26-45, 46-65 and > 65 years) and for children (< 16 years, continuous), which were further subdivided in a model for type of exposure and a model for type of activity leading to exposure.

Variables with a *p*-value ≤ 0.20 in the univariate analyses were selected for inclusion in multivariable GEE logistic regression models. Potential confounders that were always adjusted for in multivariable analyses were age category, sex (male, female, missing) and summer (summer: June, July, August; not-summer: other months; missing: NAs). Additionally, models for AGE were corrected for chronic disorders: disease of the gastrointestinal tract (yes/no), reflux (yes/no), food allergy (yes/no), or pregnancy with vomiting during the study period. Models for ARI were corrected for hay fever (yes/no), and lung anomalies (e.g. asthma, chronic obstructive pulmonary disease – COPD, etc.; yes/no).

A manual backward-forward selection procedure was used to retain only those variables with p ≤ 0.05. Collinear variables were selected based on improved model fit as revealed by the Quasi-likelihood under the Independence model Criterion (QIC) statistic [13]. Biologically plausible interactions between independent, correlated variables were also assessed. Interactions were first tested in the univariate analyses (only interaction terms were included, confounders and other possible risk factors were added later in the multivariate analyses), and those with a *p*-value ≤ 0.20 were included in the multivariable model. We present the adjusted odds ratios (aOR) and 95% confidence intervals from the final multivariable models. For results from the univariate analyses, see Supplementary material, Table S1 and Table S2.

We used R (version 3.4.3) [14] and the gee package (version 4.8) [15] for data analyses.

Results

Sample description

In total, 3,382 households were invited and 699 households (response 21%; 1,656 individual participants) completed the questionnaire after three extreme rainfall events caused urban pluvial flooding on respectively July 5th, July 27th and August 15th, 2015, in 60 municipalities in the Netherlands (Figure 1). On average, the participants returned the questionnaire 9 days after sending (median = 6 days).

Twenty-one individual participants (1.3%) were excluded, because information reported on exposure was contradictory. Information about complaints and exposure was incomplete for 477 individual participants (28.8%). Therefore, they were also eliminated from the analysis (Supplementary material, Figure S1). The remaining 1,158 individual participants (582 households) represented the study population of which we had information for both cases and non-cases (Figure 2). There were no pregnant participants, who also reported vomiting during the recall period (Supplementary material, Figure S1). Table 1 gives an overview of the descriptive statistics.

277 participants (24%) have an underlying chronic disease, 690 participants (60%) were exposed to floodwater (i.e. contact) and 966 participants (83%) performed a type of activity leading to contact with floodwater.

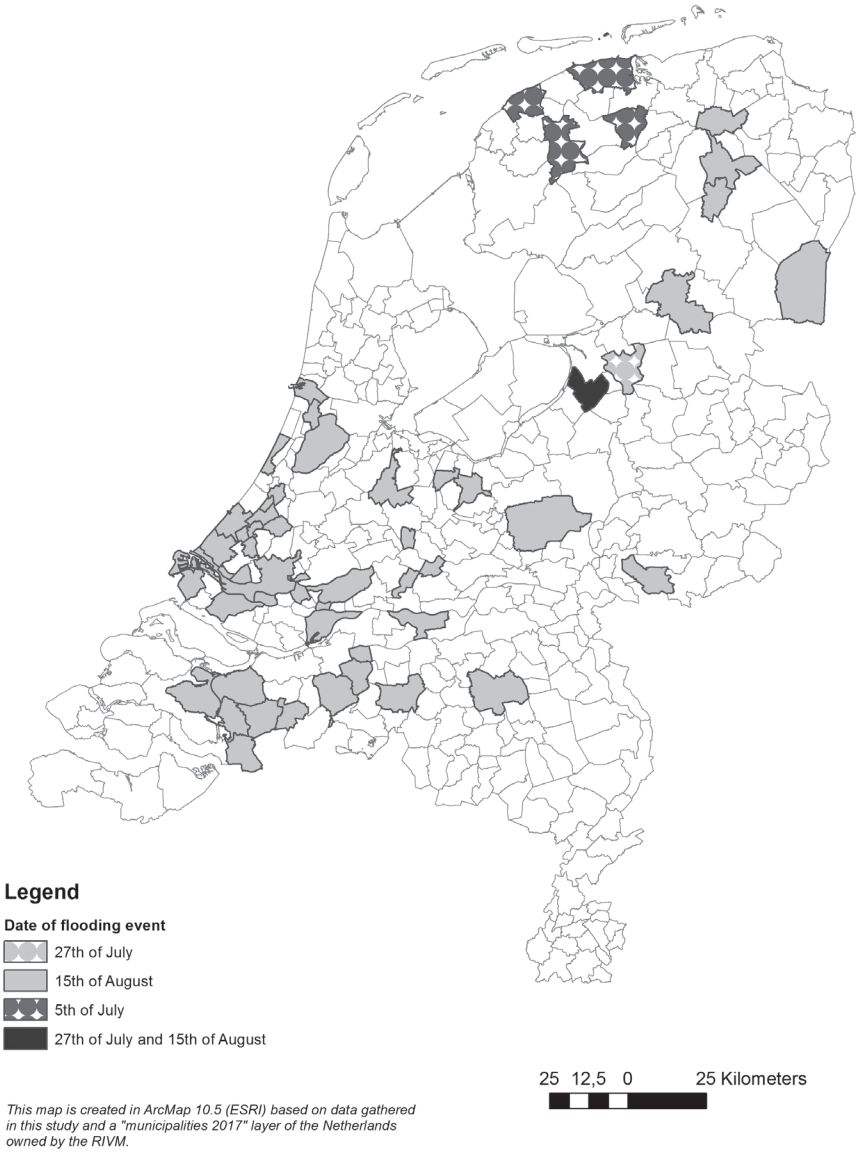
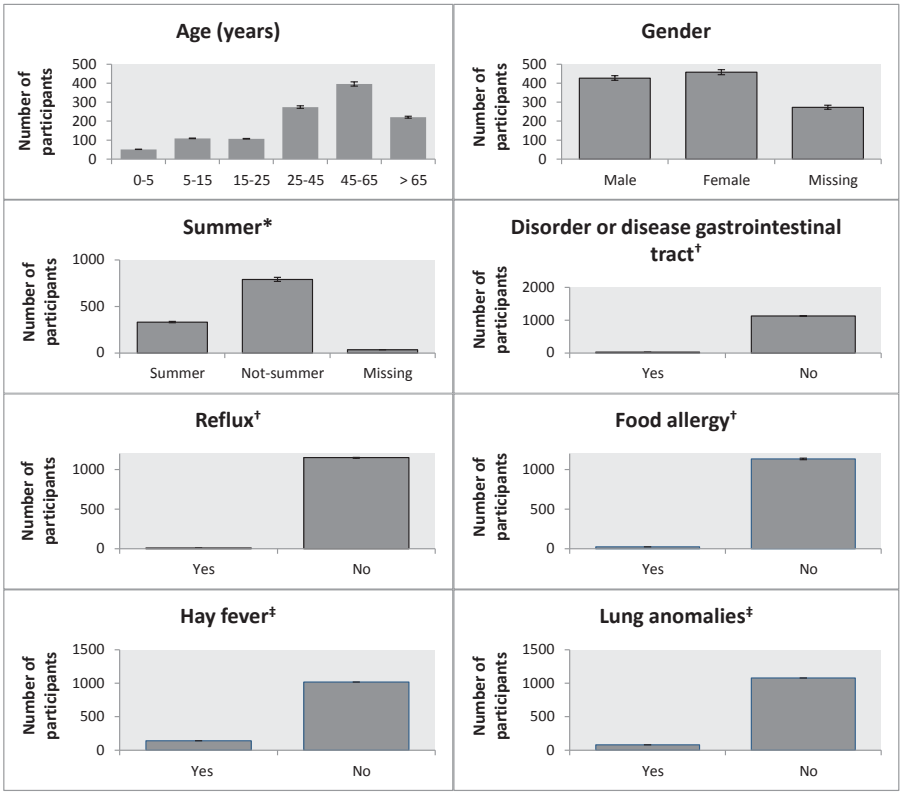


Figure 1. Map showing the municipalities that were affected by pluvial flooding at 05-07-2015, 27-07-2015 and 15-08-2017.

In the study population, 40% were women, 37% were male, and of 23% of the participants the sex was unknown. The median age was 47 years (25-75% percentile P_{25-75} 27 – 61). Compared to the Dutch general population in 2015 (16,900,726 inhabitants [16]), the sample contained slightly more women (52% vs. 51%) and was slightly older (average age 43 *vs.* 41 years). Most households (281;42%) reported two individual participants and 28 households (4%) reported five individual participants (Figure 3).



*Not used in analyses of AGE and type of exposure, because of a too small number of outcome events to perform analyses
†Adjusted for in models for AGE
‡Adjusted for in models for ARI

Figure 2. General characteristics of the study participants.

Prevalence of AGE and ARI

During the 2-week period after flooding, 75 (12.3%) participants in 51 households reported AGE, and 128 (21.1%) participants in 114 households reported ARI (Table 2), after being exposed. Table 2 also shows the results of the univariate and multivariate analyses. In the multivariate analyses, exposure to floodwater was significantly associated with increased odds of AGE (aOR 4.2, 95% CI 2.1 – 8.4). Furthermore, exposure to floodwater was associated with increased odds of reporting ARI (aOR 3.3, 95% CI 2.0 – 5.4). The analyses on the association between floodwater exposure and AGE in children < 16 years could not be performed given the small number of outcome events. Indeed, there were no children without contact with floodwater that also had AGE. Furthermore, Table 2 shows that there was no significant association between floodwater exposure and ARI in children (OR 1.6, 95% CI 0.6 – 4.0) in the univariate analysis. In adults, floodwater exposure was associated with both AGE and ARI (aOR 3.6, 95% CI 1.8 – 7.3, and aOR 4.0, 95% CI 2.3 – 6.8, respectively).

Table 1. Descriptive statistics of the risk factors of developing health complaints after contact with floodwater.

Risk factors	Total number of participants N (%)	Total number of adults N (%)	Total number of children N (%)
Chronic diseases			
Chronic disorder or disease of the gastrointestinal tract	27 (2)	27 (3)	0 (0)
Reflux	8 (1)	7 (1)	1 (1)
Food allergy	23 (2)	18 (2)	5 (3)
Pregnancy with vomiting	0 (0)	0 (0)	0 (0)
Hay fever	140 (12)	124 (12)	16 (9)
Lung anomalies	79 (7)	68 (7)	11 (6)
No chronic diseases	881 (76)	738 (75)	143 (81)
Type of exposure			
Skin contact	527 (46)	457 (47)	70 (40)
Droplets of floodwater in the mouth	19 (2)	15 (2)	4 (2)
Gulp of floodwater in the mouth	139 (12)	113 (11)	26 (15)
Head submerged in floodwater	5 (0)	4 (0)	1 (0)
Not exposed	468 (40)	393 (40)	75 (43)
Type of activity			
Cleaning inside	320 (28)	311 (32)	9 (5)
Cleaning outside	213 (18)	208 (21)	5 (3)
Swum in floodwater	3 (0)	0 (0)	3 (2)

Risk factors	Total number of participants N (%)	Total number of adults N (%)	Total number of children N (%)
Used rubber boat	7 (1)	4 (0)	3 (2)
Walked through floodwater	198 (17)	158 (16)	40 (23)
Cycled through floodwater	86 (7)	71 (7)	15 (8)
Driven through floodwater by car	139 (12)	135 (14)	4 (2)
No activities	192 (17)	95 (10)	97 (55)
Location of flooding			
Toilet overflow	250 (22)	209 (21)	41 (23)
No toilet overflow	908 (78)	773 (79)	135 (77)
Street flooding	779 (67)	650 (66)	129 (73)
No street flooding	379 (33)	332 (34)	47 (27)

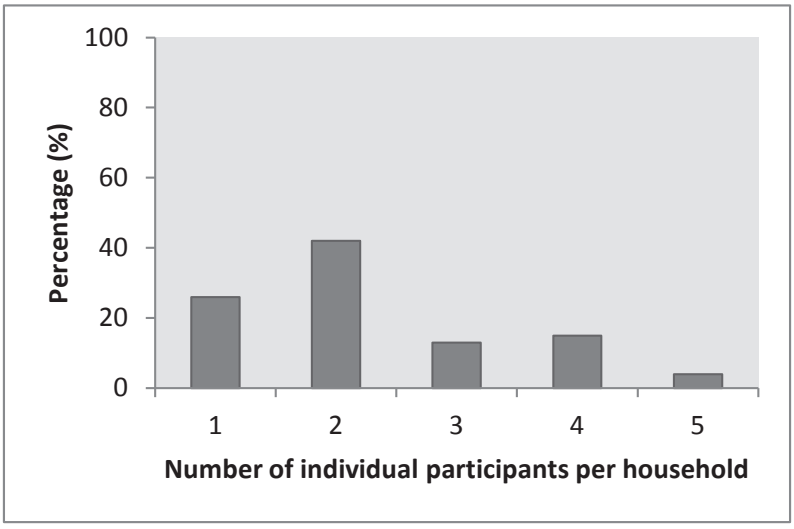


Figure 3. Percentage of participating households with 1, 2, 3, 4 or 5 individual participants.

Table 2. Health complaints of individual participants and the association with exposure to pluvial floodwater.

Without exposure	**With exposure**			**Children without exposure**	**Children with exposure**			**Adults without exposure**	**Adults with exposure**		
N **= 550** N (%)	***N*** **= 608** N (%)	**OR** (95% CI)	**aOR** (95% CI)	***N*** **= 95** N (%)	***N*** **= 81** N (%)	**OR** (95% CI)	**aOR** (95% CI)	***N*** **= 455** N (%)	***N*** **= 527** N (%)	**OR** (95% CI)	**aOR** (95% CI)
AGE 12 (2)	75 (12)	**4.9 (2.5, 9.6)**	**4.2[a] (2.1, 8.4)**	0 (0)	12 (15)	-	-	12 (3)	63 (12)	**4.2 (2.2, 8.2)**	**3.6[a] (1.8, 7.3)**
ARI 38 (7)	128 (21)	**3.5 (2.2, 5.7)**	**3.3[b] (2.0, 5.4)**	10 (11)	14 (17)	1.6 (0.6, 4.0)	-	28 (6)	114 (22)	**4.1 (2.5, 6.7)**	**4.0[b] (2.3, 6.8)**

Note: There were no children without contact with the floodwater, but with AGE, leading to a too small number of outcome events to perform the univariate analysis. Furthermore, there were not enough children to prevent the same from happening in the multivariate analyses for contact with the floodwater and ARI

[a]Adjusted for age, sex, summer, having a chronic disorder or disease of the gastrointestinal tract, having reflux, having a food allergy or pregnancy with vomiting during the study period

[b]Adjusted for age, sex, summer, having hay fever and having lung anomalies (e.g. asthma, chronic obstructive pulmonary disease - COPD, etc.) Factors with a p-value <0.05 are expressed in bold

Multivariate analyses

Results from the final multivariable models for AGE and ARI are presented in Tables 3 and 4 respectively. For results from the univariate models, see Supplementary material, Table S1 and Table S2.

In the models for type of activity and AGE or ARI, an interaction term was included between contact with floodwater and cleaning floodwater inside or outside the house. The strata included in this interaction term are: no contact with floodwater at all (reference category), contact with floodwater but no involvement in post-flooding cleaning operations, contact with floodwater and involvement in indoor cleaning operations, contact with floodwater and involvement in outdoor cleaning operations, and contact with floodwater and involvement in both indoor and outdoor cleaning operations. The rationale is that people with the least exposure to floodwater have the lowest chance to get AGE or ARI (water contact without cleaning), while people with the largest exposure to floodwater (water contact with cleaning inside and outside the house) have the highest chance to get AGE or ARI.

AGE

All age categories

In the multivariate model assessing AGE risk in relation to the type of exposure to floodwater for all age categories (overall model), having had skin contact with floodwater was significantly associated with AGE (aOR 4.0, 95% CI 1.8 – 9.0; Table 3). Regarding the type of activity, the aORs in all strata of the interaction term that included contact with floodwater were significantly increased. Indeed, the model showed increasing aORs from having had contact with floodwater without having been engaged in any cleaning operation (aOR 2.9, 95% CI 1.3 – 6.7) to having had contact with floodwater and having been engaged in cleaning operations both inside and outside the household (aOR 8.6, 95% CI 3.5 – 20.9). Cycling in floodwater was also a significant risk factor for AGE (aOR 2.3, 95% CI 1.0 – 5.0).

Table 3. Results of multivariate analyses for AGE in all age categories (overall), children (<16years) and adults.

Model & Covariates	Overall without AGE N (%)	Overall with AGE N (%)	aOR (95% CI) [a]	Children without AGE N (%)	Children with AGE N (%)	aOR (95% CI) [a]	Adults without AGE N (%)	Adults with AGE N (%)	aOR (95% CI) [a]
Type of exposure									
Skin contact	456 (70)	71 (93)	**4.0 (1.8, 9.0)**	59 (65)	11 (100)	NA	397 (71)	60 (92)	**3.8 (1.5, 9.4)**
Type of activity									
Water contact									
No (ref)			ref			ref			ref
Yes, no cleaning	190 (18)	22 (25)	**2.9 (1.3, 6.7)**	NA	NA	NA	128 (14)	12 (16)	1.5 (0.5, 4.3)
Yes, cleaning inside	173 (16)	17 (20)	**3.7 (1.4, 9.6)**	NA	NA	NA	167 (18)	17 (23)	**3.3 (1.3, 8.4)**
Yes, cleaning outside	69 (6)	14 (16)	**6.9 (2.5, 18.7)**	NA	NA	NA	69 (8)	12 (16)	**5.3 (1.9, 14.7)**
Yes, cleaning in & out	106 (10)	24 (28)	**8.6 (3.5, 20.9)**	NA	NA	NA	103 (11)	24 (32)	**7.6 (3.1, 18.6)**
Cycled	62 (6)	24 (28)	**2.3 (1.0, 5.0)**	10 (6)	5 (42)	NA	52 (6)	19 (25)	**2.4 (1.0, 5.7)**
Do not know	5 (0)	2 (2)	NA	2 (1)	0 (0)	NA	3 (0)	2 (3)	**27 (4.7, 158)**

Note*: aOR* adjusted odds ratio, *CI* confidence interval, *NA* not applicable, overall analyses of exposure vs AGE could not be performed for children, *ref* reference category
[a]Adjusted for age, sex, summer, having a chronic disorder or disease of the gastrointestinal tract, having reflux, having a food allergy or pregnancy with vomiting during the study period
Factors with a *p*-value <0.05 are expressed in bold

Table 4. Results of multivariate analyses for ARI in all age categories (overall), children (<16years) and adults.

Model & Covariates	Overall without ARI N (%)	Overall with ARI N (%)	aOR (95% CI) [a]	Children without ARI N (%)	Children with ARI N (%)	aOR (95% CI) [a]	Adults without ARI N (%)	Adults with ARI N (%)	aOR (95% CI) [a]
Type of exposure									
Skin contact	406 (68)	121 (88)	**3.6 (1.9, 6.9)**	54 (64)	16 (89)	NA	352 (69)	105 (88)	**3.7 (1.8, 7.7)**
Type of activity									
Water contact									
No (ref)			ref			ref			ref
Yes, no cleaning	168 (17)	44 (27)	**3.0 (1.6, 5.7)**	NA	NA	NA	112 (13)	28 (20)	**3.8 (1.9, 7.6)**
Yes, cleaning inside	155 (16)	35 (21)	**4.0 (2.2, 7.4)**	NA	NA	NA	149 (18)	35 (25)	**4.3 (2.2, 8.4)**
Yes, cleaning outside	61 (6)	22 (13)	**4.4 (2.2, 8.7)**	NA	NA	NA	59 (7)	22 (15)	**5.2 (2.5, 10.7)**
Yes, cleaning in & out	98 (10)	32 (19)	**5.5 (3.0, 10.3)**	NA	NA	NA	95 (11)	32 (23)	**5.7 (2.9, 11.3)**
Cycled	61 (6)	25 (15)	NA	12 (8)	3 (13)	NA	49 (6)	22 (15)	NA
Do not know	2 (0)	5 (3)	**9.4 (1.4, 64.9)**	0 (0)	2 (8)	NA	2 (0)	3 (2)	NA

Note: *aOR* adjusted odds ratio, *CI* confidence interval, *NA* not applicable (overall analyses of exposure vs AGE could not be performed for children), *ref* reference category
[a]Adjusted for age, sex, summer, having hay fever and having lung anomalies (e.g. asthma, chronic obstructive pulmonary disease - COPD, etc.) Factors with a *p*-value <0.05 are expressed in bold

Adults
In the multivariate model assessing AGE risk in relation to the type of exposure to floodwater among adults, having had skin contact with floodwater was significantly associated with AGE (aOR 3.8, 95% CI 1.5 – 9.4; Table 3). Regarding the type of activity, the aORs in the strata of the interaction term that included contact with floodwater and cleaning were significantly increased. Indeed, the model showed increasing aORs from having had contact with floodwater and having been engaged in indoor cleaning operations (aOR 3.3, 95% CI 1.3 – 8.4) to having had contact with floodwater and having been engaged in both indoor and outdoor cleaning operations (aOR 7.6, 95% CI 3.1 – 18.6). Cycling in floodwater was also a risk factor for AGE (aOR 2.4, 95% CI 1.0 – 5.7).

Children
No multivariate model for AGE in children could be built because there were not enough children with AGE to perform the analysis.

ARI

All age categories
In the multivariate model assessing ARI risk in relation to the type of exposure to floodwater for all age categories (overall model), having had skin contact with floodwater was significantly associated with ARI (aOR 3.6, 95% CI 1.9 – 6.9; Table 4). Regarding the type of activity, the aORs in all strata of the interaction term that included contact with floodwater were significantly increased. Indeed, the model showed increasing aORs from having had contact with floodwater without having been engaged in any cleaning operation (aOR 3.0, 95% CI 1.6 – 5.7) to having had contact with floodwater and having been engaged in both indoor and outdoor cleaning operations (aOR 5.5, 95% CI 3.0 – 10.3).

Adults
In the multivariate model assessing ARI risk in relation to the type of exposure to floodwater among adults, having had skin contact with floodwater was significantly associated with ARI (aOR 3.7, 95% CI 1.8 – 7.7; Table 4). Regarding the type of activity, the aORs in all strata of the interaction term that included contact with floodwater were significantly increased, with increasing aORs from having had contact with floodwater but no involvement in cleaning operations (aOR 3.8, 95% CI 1.9 – 7.6), to having had contact with floodwater and having been involved in cleaning operations both inside and outside the household (aOR 5.7, 95% CI 2.9 – 11.3).

Children
There was no significant association between exposure to floodwater and ARI for children.

Magnitude of exposure

The average height of the water (magnitude of exposure) in the streets was 19 cm (median = 15 cm, height $_{min}$ = 2 cm, height $_{max}$ = 150 cm). The average height of the water after toilet overflow was 5 cm (median = 2 cm, height $_{min}$ = 1 cm, height $_{max}$ = 30 cm). Additional univariate analyses showed that the magnitude of exposure in the streets was a risk factor for both AGE (OR 1.0; 95% CI 1.0 – 1.1) and ARI (toilet overflow, OR 1.1; 95% CI 1.0 – 1.2).

Discussion

This study suggests an association between direct exposure to urban pluvial floodwater and the occurrence of both AGE and ARI in a high-income country like the Netherlands. Identified risk factors were contact with floodwater, such as skin contact with floodwater (for both AGE and ARI), post-flooding cleaning operations (for both AGE and ARI) and cycling through floodwater (for AGE).

This study included 1,656 individual participants, resulting in a response rate of 21%. It is unknown how representative our sample is with regard to the target population (i.e. the people who experienced flooding), as we did not know the demographics of the people invited to participate in the study. Addressee-unknown invitations were sent to all house numbers of streets in which flooding had been reported. Therefore, it was impossible for us to know who the invitees were. It could be argued that representativeness of the sample could be assessed based on the demographics of the Dutch general population. However, this is not an optimal solution, because the 'target population', i.e. the people who experienced flooding, does not necessarily mirror the general population.

A response rate of 20-30% is commonly reported in this type of retrospective studies where self-reported health complaints are investigated [17, 18], and some studies report even lower response rates [19]. A low response rate could render these studies prone to, for example, selection bias [18]. The type of selection bias that might have played a role in this study is self-selection bias, by which the group of participants who were exposed and diseased is overrepresented, as people who experienced flooding and health complaints are particularly motivated to complete and return the questionnaire [18]. Another limitation that is inherent to retrospective studies that

use self-reported data is recall bias, where people might have forgotten (mild) AGE and ARI episodes [18]. On the other hand, an overestimation of the incidence might also have occurred, because of 'telescoping' (i.e. when people remember episodes as being more recent than they actually are) [18]. Overall, retrospective studies with self-reported data like ours tend to produce incidence estimates that overestimate the true incidence [20, 21]. A major advantage of these type of studies is that they allow for collection of data about ARI and AGE cases that are not reported to the General Practitioner (GP), i.e. cases that do not require medical attention.

Previous population-based studies on AGE in the Netherlands showed a baseline incidence of 0.95 episodes/person-year. In our study targeting pluvial flood-ravaged areas, the incidence of AGE was estimated at 1.69 episodes/person-year, which is almost twice as higher. This could suggest that flooding events increase AGE risk in the affected population also as compared to the baseline AGE incidence in the whole country [19]. Likewise, the incidence of ARI derived from our study was 3.74 episodes/person-year, which is more than twice as higher than the incidence of influenza-like illness (ILI) in the general Dutch population (1.72 episodes/person-year) [18]. Although ARI is not the same as ILI, they both include respiratory diseases and give an indication of the number of cases with respiratory disease in the country. However, direct comparison with the ARI/AGE incidence of the general population is hampered by the fact that we used a shorter recall period (2 vs. 4 weeks), which generally produces higher incidence estimates [20, 21]. Moreover, our study might have been subject to reporting and selection bias, as is described above.

De Man, et al. [3] performed a comparable study in the Netherlands at a much smaller scale (149 households) and lacked sufficient numbers of outcome events to use well-defined (standardized) AGE or ARI syndrome definitions. The larger scale of our study (699 households) allowed for estimates that are more precise and for the use of standardized definitions for both AGE and ARI. In agreement with De Man, et al. [3], we found that participants exposed to pluvial floodwater were more likely to develop gastrointestinal and respiratory complaints (Table 2).

We also investigated differences in risk factors for AGE and ARI in children and adults, but the number of children enrolled in the study was rather low, so the analyses for this age category were underpowered (Table 2). This may be reason as to why the association between ARI and exposure to floodwater was positive but not statistically significant. Low statistical power did also not allow De Man, et al. [3] to study differences in risk factors for AGE and ARI between adults and children. However, it is evident that children may display certain risk factors (e.g. playing in/with floodwater) more often than adults (e.g. post-cleaning operations). It was

also shown by De Man, et al. [6] that children are more likely to ingest floodwater compared to adults (1.7 ml vs. 0.016 ml), because they play in or around floodwater, and therefore have a higher risk of AGE (33% vs. 3.9%).

Sanitary sewer overflow (SSOs) events were shown to be associated with gastrointestinal illness (GI; emergency room (ER) visits with a primary diagnosis of GI) [22]. However, specific risk factors leading to exposure and eventually GI were not identified. Furthermore, SSOs probably entail different risk factors (swimming, contaminated drinking water) compared to flooding of combined sewerage systems/ street flooding as studied in this paper.

For ARI and AGE, skin contact was identified as a risk factor, which is probably a proxy for ingestion or inhalation of contaminated floodwater (Tables 3 and 4). This may happen, for example, when people do not wash their hands after floodwater contact or people splash aerosolized water particles in their face, while only reporting skin contact [23]. This could possible also explain why for example ingesting droplets of floodwater was not identified as risk factor, as people only report the most evident risk factors, i.e. skin contact.

Post-flooding cleaning operations were associated with increased risk for both ARI and AGE (Tables 3 and 4), which was only associated with ARI in De Man, et al. [3]. It could be that we were able to identify it as a risk factor for AGE due to our larger sample size. Our finding is supported by previous literature describing that cleaning leads to aerosolization of contaminated water droplets and their inhalation/ingestion, explaining why it could be a risk factor for AGE [23]. Similarly, cycling through floodwater was a significant exposure for AGE (Table 3), as water droplets could splash into their face and mouth. As reported in De Man, et al. [3], these associations may reflect, to some extent, the primary transmission routes of the pathogens in question, i.e. inhalation for ARI and ingestion for AGE. Indeed, several are typically associated with flood-ravaged settings in developed countries. Sales-Ortells and Medema [24] found, for example, *Campylobacter* in all water plaza samples in which people recreate, leading to a risk of developing AGE for those people.

A limitation of this study is that there were no water samples obtained from pluvial floodwater as well as no faecal samples from the participants. Therefore, the causative agents remain unknown. However, pluvial floodwater in the Netherlands was shown to always be contaminated with faeces, as was demonstrated by the presence of *E. coli*, intestinal enterococci, and enteropathogens such as enterovirus, norovirus and *Campylobacter* in water samples of pluvial floodwater [6].

Furthermore, there was no data on sex for 23% of the participants, primarily children, which was because the questionnaire only collected information regarding sex of the person filling in the questionnaire, but not for other household members. This explains most of the missings and is unlikely to affect our results because sex effects on AGE and ARI are less likely to be seen in children [19, 25]. Moreover, we do not know whether participants used a measure tape to measure the height of the floodwater in cm to answer the question. Although the magnitude of exposure came out as a risk factor in the analysis for AGE and ARI, those data could potentially be erroneous as most of the collected data are self-reported estimates of the true height of the floodwater.

Chronic diseases, which the models were corrected for, may be effect modifiers of the association between floodwater exposure and AGE/ARI. This was checked by running the models with and without chronically diseased cases therein: results can be found in Supplementary material, Table S3. They show that chronic diseases were not significant effect modifiers. However, they were associated with the outcome, so the models were always corrected for underlying chronic disease.

Future prospective studies are needed to confirm the associations found in this study. An example of such a study could be a longitudinal study that monitors the rainfall pattern in a given area. As soon as an extreme rainfall event occurs (> 30 mm rainfall/hour for > 1 h), it should be checked whether flooding has also occurred in that specific area. A representative selection of the inhabitants of the flooded area (which have been actively enrolled at the beginning of the study, perhaps upon financial incentive) should use health diaries to record their symptoms and their exposure to floodwater. This would reduce both recall and selection bias. In order to further confirm causality, floodwater and patient faeces samples should also be taken to identify and characterize potential pathogens.

Due to climate change, extreme rainfall events will occur more frequently, leading to more people being exposed to pluvial floodwater in urban areas in countries like the Netherlands [2]. As this study suggests, such events are likely to lead to increased risk of health effects. Generally, people in developed countries tend to not perceive the associated health risks. The Netherlands is a highly populated country (~400 people/km^2) [3], so urban flooding will not pass easily unnoticed. Recently, awareness has been arisen, with Dutch residents demanding municipal services to improve drainage systems as to prevent pluvial flooding [26]. In response, governmental authorities are promoting greenness in urban areas to facilitate natural drainage of the water in the soil [26]. This study adds another perspective to this debate and shows that it is necessary to take proper care of water drainage systems/sewage systems in terms of their drainage capacity to mitigate health risks in urban areas.

Conclusion

In conclusion, this study suggests an association between different non-mutually exclusive types of direct exposure and activities leading to exposure to floodwater and AGE and ARI, which is a possible reflection of the transmission routes of the pathogens in question. However, future prospective studies are needed to confirm this association. Since pluvial flooding events will increase in the future, proving the causal link would reinforce the need for flood-proof solutions in urban development and increased awareness among stakeholders and the public about the associated health risks.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

The study was conceived by RP, WP and HM. HM and RP collected the data and ACM analyzed those data with help of JK. ACM and RP interpreted findings, with feedback from EF and LMG. ACM drafted the first version of the manuscript. All authors provided critical edits and revisions to the paper and have reviewed and approved the final version of the paper.

Ethics approval and consent to participate

This study involved collection and analysis of fully anonymized data, so no ethical approval was necessary according to Dutch regulations [8, 9]. People gave consent upon anonymously completing and returning the questionnaire. Questionnaires were received in de-identified form, containing only data on postal code, sex, and age of the participants. Therefore, names and addresses could not be linked to the questionnaire responses, guaranteeing anonymity of the respondents. Participants were informed that the data they provided were to be analyzed for scientific purposes, and that by returning the questionnaires, they gave consent to do so. No written consent was therefore necessary. Parents were asked to complete the questionnaire and give consent on behalf of their children (family members < 16 years old).

Competing interests

The authors declare that they have no competing interests.

References

1. Wuebbles DJ, Fahey DW, Hibbard KA, DeAngelo B, Doherty S, Hayhoe K, Horton R, Kossin JP, Taylor PC, Waple AM, Weaver CP. Executive summary. In: Wuebbles DJ, Fahey DW, Hibbard KA, Dokken DJ, Stewart BC, Maycock TK, editors. Climate Science Special Report: Fourth National Climate Assessment, vol. I. Washington, DC: U.S. Global Change Research Program; 2017. p. 12–34.
2. Min S-K, Zhang X, Zwiers FWandHegerl GC. Human contribution to more- intense precipitation extremes. Nature. 2011;470(7334):378.
3. De Man H, Mughini Gras L, Schimmer B, Friesema IHM, De R. Husman AMandVan Pelt W: Gastrointestinal, influenza-like illness and dermatological complaints following exposure to floodwater: a cross-sectional survey in The Netherlands. Epidemiol Infect. 2015;144(7):1445–54.
4. Roth M, Buishand TA, Jongbloed G. Trends in moderate rainfall extremes: A regional monotone regression approach. J Climate. 2015;28(22):8760–9.
5. Ten Veldhuis J, Clemens F, Sterk GandBerends B. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water Res. 2010;44(9):2910–8.
6. De Man H, Van Den Berg H, Leenen E, Schijven J, Schets F, Van der Vliet J, van Knapen F, de Roda Husman AM. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. Water Res. 2014;48:90–9. https://doi.org/10.1016/j.watres.2013.09.022.
7. Particuliere huishoudens naar samenstelling en grootte, 1 januari 2015. https://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37975&D1=0-19,21-26&D2=0&D3=0,5,10,15-20&HDR=T&STB=G1,G2&VW=T. Accessed 20 Mar 2018.
8. Central committee on Research Involving Human Subjects (CCMO) https:// english.ccmo.nl/. Accessed 18 Oct 2018.
9. Autoriteit Persoonsgegevens https://www.autoriteitpersoonsgegevens.nl/nl/ onderwerpen/avg-europese-privacywetgeving/algemene-informatie-avg Assessed 18 Oct 2018.
10. Majowicz SE, Hall G, Scallan E, Adak GK, Gauci C, Jones TF, O'Brien S, Henao OandSockett PN. A common, symptom-based case definition for gastroenteritis. Epidemiol Infect. 2007;136(7):886–94.
11. Teirlinck A, Van Asten L, Brandsema P, Dijkstra F, Donker G, Van Gageldonk- Lafeber A, Hooiveld M, de Lange M, Marbus Sand Meijer A. Annual report Surveillance of influenza and other respiratory infections in the Netherlands: Winter 2016/2017. Utrecht: RIVM; 2017.
12. Hardin JW. Generalized Estimating Equations (GEE). In: Encyclopedia of Statistics in Behavioral Science. edn. Wiley Online Library; 2005.
13. Pan W. Akaike's Information Criterion in Generalized Estimating Equations. Biometrics. 2001;57(1):120–5.
14. RCT, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna. https://www.r-project.org/. Accessed 18 Jan 2018.
15. gee: Generalized Estimation Solver. R package version 4.8. https://cran.r- project.org/web/packages/gee/index.html. Accessed 19 July 2018.
16. Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari. https:// opendata.cbs.nl/statline/#/CBS/nl/dataset/7461bev/table?dl=5052. Accessed 20 Mar 2018.
17. Doorduyn Y, Van Pelt WandHavelaar AH. The burden of infectious intestinal disease (IID) in the community: a survey of self-reported IID in The Netherlands. Epidemiol Infect. 2012;140(7):1185–92.
18. Mughini-Gras L, Pijnacker R, Enserink R, Heusinkveld M. van der Hoek Wandvan Pelt W: Influenza-like illness in households with children of preschool age. Pediatr Infect Dis J. 2016;35(3):242–8.

19. Mughini-Gras L, Pijnacker R, Heusinkveld M, Enserink R, Zuidema R, Duizer E, Kortbeek Tandvan Pelt W. Societal Burden and Correlates of Acute Gastroenteritis in Families with Preschool Children. Sci Rep. 2016;6:22144.

20. Cantwell LB, Henao OL, Hoekstra RMandScallan E. The effect of different recall periods on estimates of acute gastroenteritis in the United States, FoodNet Population Survey 2006-2007. Foodborne Pathog Dis. 2010; 7(10):1225–8.

21. Viviani L, van der Es M, Irvine L, Tam CC, Rodrigues LC, Jackson KA, O'Brien SJ, Hunter PR, Committee IIDSE. Estimating the Incidence of Acute Infectious Intestinal Disease in the Community in the UK: A Retrospective Telephone Survey. PLOS ONE. 2016;11(1):e0146171.

22. Jagai JS, DeFlorio-Barker S, Lin CJ, Hilborn ED, Wade TJ. Sanitary Sewer Overflows and Emergency Room Visits for Gastrointestinal Illness: Analysis of Massachusetts Data, 2006–2007. Environ Health Perspect. 2017;125(11):–117007.

23. Morawska L. Droplet fate in indoor environments, or can we prevent the spread of infection? Indoor Air. 2006;16(5):335–47.

24. Sales-Ortells H, Medema G. Microbial health risks associated with exposure to stormwater in a water plaza. Water Res. 2015;74:34–46.

25. Ezeonwu B, Ibeneme C, Aneke F, Oguonu T. Clinical features of acute gastroenteritis in children at university of Nigeria teaching hospital, ituku- ozalla, enugu. Ann Med Health Sci Res. 2013;3(3):361–4.

26. Meindertsma B and Van der Parre H, 2018. Burgers met natte voeten wijzen vooral naar de overheid. https://nos.nl/artikel/2230728-burgers-met-natte- voeten-kijken-naar-overheid-maar-willen-niet-meer-betalen.html. Accessed 8 May 2018.

Supplementary material

Table S1. Results of univariate analyses for AGE in all age categories (overall), children (<16years) and adults (factors with a p-value <0.20 are expressed in bold).

Model & Covariates	Overall without AGE N (%)	Overall with AGE N (%)	OR (95% CI)	Children without AGE N (%)	Children with AGE N (%)	OR (95% CI)	Adults without AGE N (%)	Adults with AGE N (%)	OR (95% CI)
Type of exposure									
Skin contact	456 (70)	71 (93)	**5.5 (2.1, 14.8)**	59 (65)	11 (100)	NA	397 (71)	60 (92)	**4.9 (1.7, 13.7)**
Droplets of water in the mouth	16 (2)	3 (4)	**2.4 (0.8, 7.0)**	4 (4)	0 (0)	NA	12 (2)	3 (5)	**2.5 (0.8, 8.1)**
Gulp of water in the mouth	119 (18)	20 (26)	**2.0 (1.1, 3.9)**	25 (27)	1 (8)	NA	94 (17)	19 (29)	**1.9 (1.0, 3.5)**
Head submerged	4 (1)	1 (1)	3.3 (0.5, 23.4)	0 (0)	1 (8)	NA	4 (1)	0 (0)	NA
Type of activity									
Cleaning inside	279 (26)	41 (47)	**2.3 (1.2, 4.4)**	9 (5)	0 (0)	NA	270 (30)	41 (55)	**2.7 (1.5, 4.9)**
Cleaning outside	175 (16)	38 (44)	**3.4 (2.0, 5.8)**	3 (2)	2 (17)	NA	172 (19)	36 (48)	**3.7 (2.2, 6.3)**
Played/run/splashed	72 (7)	17 (20)	1.5 (0.7, 3.4)	46 (28)	10 (83)	NA	26 (3)	7 (9)	**2.4 (0.6, 8.8)**
Swum	2 (0)	1 (1)	1.3 (0.2, 9.8)	2 (1)	1 (8)	NA	0 (0)	0 (0)	NA
Used rubber boat	5 (0)	2 (2)	**2.8 (0.8, 9.5)**	2 (1)	1 (8)	NA	3 (0)	1 (1)	**3.0 (0.8, 11.7)**
Walked	165 (15)	33 (38)	**2.8 (1.6, 4.7)**	35 (21)	5 (42)	NA	130 (14)	28 (37)	**3.0 (1.7, 5.3)**
Cycled	62 (6)	24 (28)	**4.5 (2.2, 8.9)**	10 (6)	5 (42)	NA	52 (6)	19 (25)	**4.7 (2.3, 9.4)**
Driven	117 (11)	22 (25)	**1.5 (0.9, 2.7)**	4 (2)	0 (0)	NA	113 (12)	22 (29)	**1.8 (1.0, 3.4)**
Do not know	5 (0)	2 (2)	1.6 (0.0, 230)	2 (1)	0 (0)	NA	3 (0)	2 (3)	**7.7 (1.4, 43.1)**

Note: OR, odds ratio; CI, confidence interval; NA, not applicable (overall analyses exposure vs AGE was not applicable for children, so the separate univariate analyses, multivariate model and the interaction term for children were not created)

Table S2. Results of univariate analyses for ARI in all age categories (overall), children (<16years) and adults (factors with a p-value <0.20 are expressed in bold).

Model & Covariates	Overall without ARI N (%)	Overall with ARI N (%)	OR (95% CI)	Children without ARI N (%)	Children with ARI N (%)	OR (95% CI)	Adults without ARI N (%)	Adults with ARI N (%)	OR (95% CI)
Type of exposure									
Skin contact	406 (68)	121 (88)	**3.9 (2.0, 7.4)**	54 (64)	16 (89)	NA	352 (69)	105 (88)	**3.9 (1.9, 7.9)**
Droplets of water in the mouth	12 (2)	7 (5)	1.6 (0.5, 4.9)	4 (5)	0 (0)	NA	8 (2)	7 (6)	**2.9 (1.0, 8.1)**
Gulp of water in the mouth	106 (18)	33 (24)	**1.7 (1.0, 2.8)**	24 (29)	2 (11)	NA	82 (16)	31 (26)	**1.8 (1.1, 3.0)**
Head submerged	3 (1)	2 (1)	1.4 (0.5, 4.3)	0 (0)	1 (6)	NA	3 (1)	1 (1)	1.1 (0.2, 6.7)
Type of activity									
Cleaning inside	253 (26)	67 (40)	**2.2 (1.5, 3.2)**	9 (6)	0 (0)	NA	244 (29)	67 (47)	**2.2 (1.5, 3.3)**
Cleaning outside	159 (16)	54 (33)	**2.2 (1.5, 3.4)**	5 (3)	0 (0)	NA	154 (18)	54 (38)	**2.4 (1.6, 3.7)**
Played/run/splashed	71 (7)	18 (11)	0.8 (0.4, 1.8)	46 (30)	10 (42)	NA	25 (3)	8 (6)	1.3 (0.4, 4.2)
Swum	1 (0)	2 (2)	2.8 (1.0, 8.2)	1 (1)	2 (8)	NA	0 (0)	0 (0)	NA
Used rubber boat	3 (0)	4 (2)	0.9 (0.0, 201)	1 (1)	2 (8)	NA	2 (0)	2 (1)	**4.2 (0.6, 30.8)**
Walked	149 (15)	49 (30)	**2.4 (1.5, 3.8)**	33 (22)	7 (29)	NA	116 (14)	42 (30)	**2.6 (1.7, 4.1)**
Cycled	61 (6)	25 (15)	**1.9 (1.0, 3.5)**	12 (8)	3 (13)	NA	49 (6)	22 (15)	**2.4 (1.2, 4.4)**
Driven	104 (10)	35 (21)	**1.8 (1.1, 3.0)**	4 (3)	0 (0)	NA	100 (12)	35 (25)	**1.8 (1.1, 3.1)**
Do not know	2 (0)	5 (3)	**11.4 (2.2, 59.7)**	0 (0)	2 (8)	NA	2 (0)	3 (2)	**5.7 (1.3, 25.7)**

Note: OR, odds ratio; CI, confidence interval; NA, not applicable (overall analyses exposure vs ARI was not significant for children, so the separate univariate analyses, multivariate model and the interaction term for children were not created)

Table S3. Additional analysis of AGE and ARI with regard to their risk factors with and without individuals having chronic diseases. (Factors with a p-value <0.05 are expressed in bold)

Model & Covariates	AGE with chronic disease aOR (95% CI) [a]	AGE without chronic disease aOR (95% CI) [a]	ARI with chronic disease aOR (95% CI) [b]	ARI without chronic disease aOR (95% CI) [b]
Type of exposure				
Skin contact	**3.9 (1.7-8.9)**	**5.0 (1.8-13.5)**	**3.4 (1.8-6.5)**	**3.2 (1.6-6.1)**
Type of activity				
Water contact				
No (ref)	Ref	Ref	Ref	Ref
Yes, no cleaning	**2.9 (1.3-6.4)**	**3.2 (1.4-7.6)**	**3.2 (1.7-5.9)**	**3.0 (1.5-6.1)**
Yes, cleaning inside	**3.7 (1.5-9.4)**	**3.8 (1.4-10.7)**	**3.9 (2.2-7.1)**	**3.7 (1.9-7.3)**
Yes, cleaning outside	**6.8 (2.6-17.9)**	**8.2 (3.0-22.9)**	**4.3 (2.2-8.4)**	**5.1 (2.5-10.5)**
Yes, cleaning in & out	**8.6 (3.7-20.3)**	**8.8 (3.5-22.3)**	**5.7 (3.1-10.7)**	**4.0 (2.0-8.4)**
Cycled	**2.2 (1.0-4.9)**	**2.7 (1.2-6.4)**	NA	NA
Do not know	NA	NA	**8.8 (1.5-50.7)**	**7.4 (0.6-87.4)**

Note: aOR, adjusted odds ratio; CI, confidence interval; ref, reference category; NA, not applicable
[a] Adjusted for age, sex
[b] Adjusted for age, sex and summer

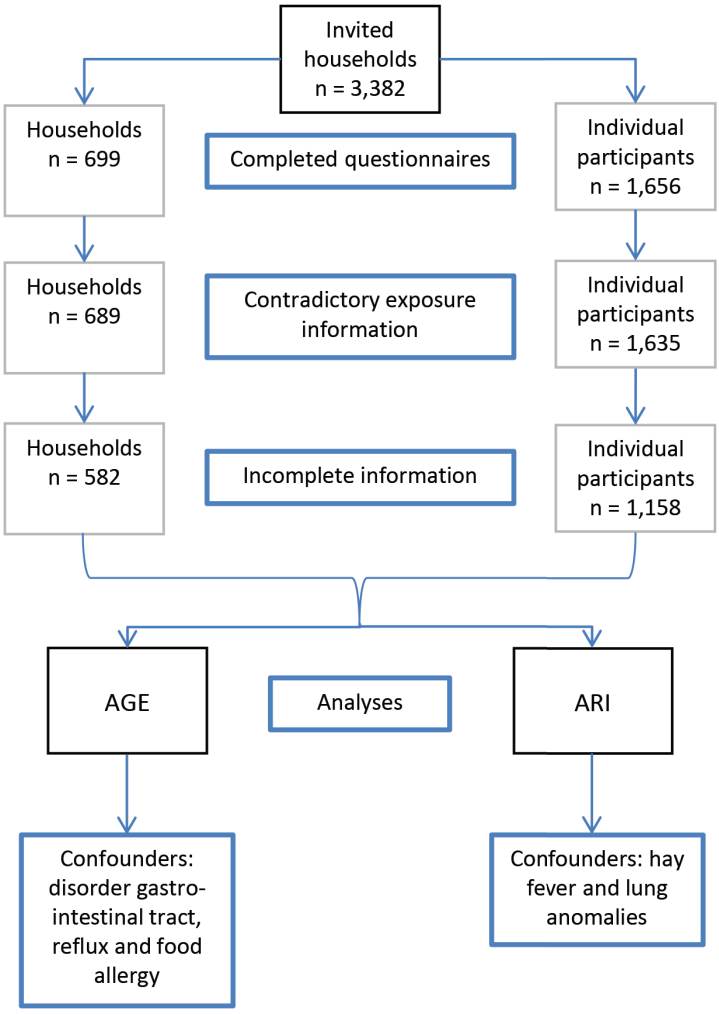


Figure S1. Flowchart discarded information.

2

CHAPTER 3

Tracing the animal sources of surface water contamination with *Campylobacter jejuni* and *Campylobacter coli*

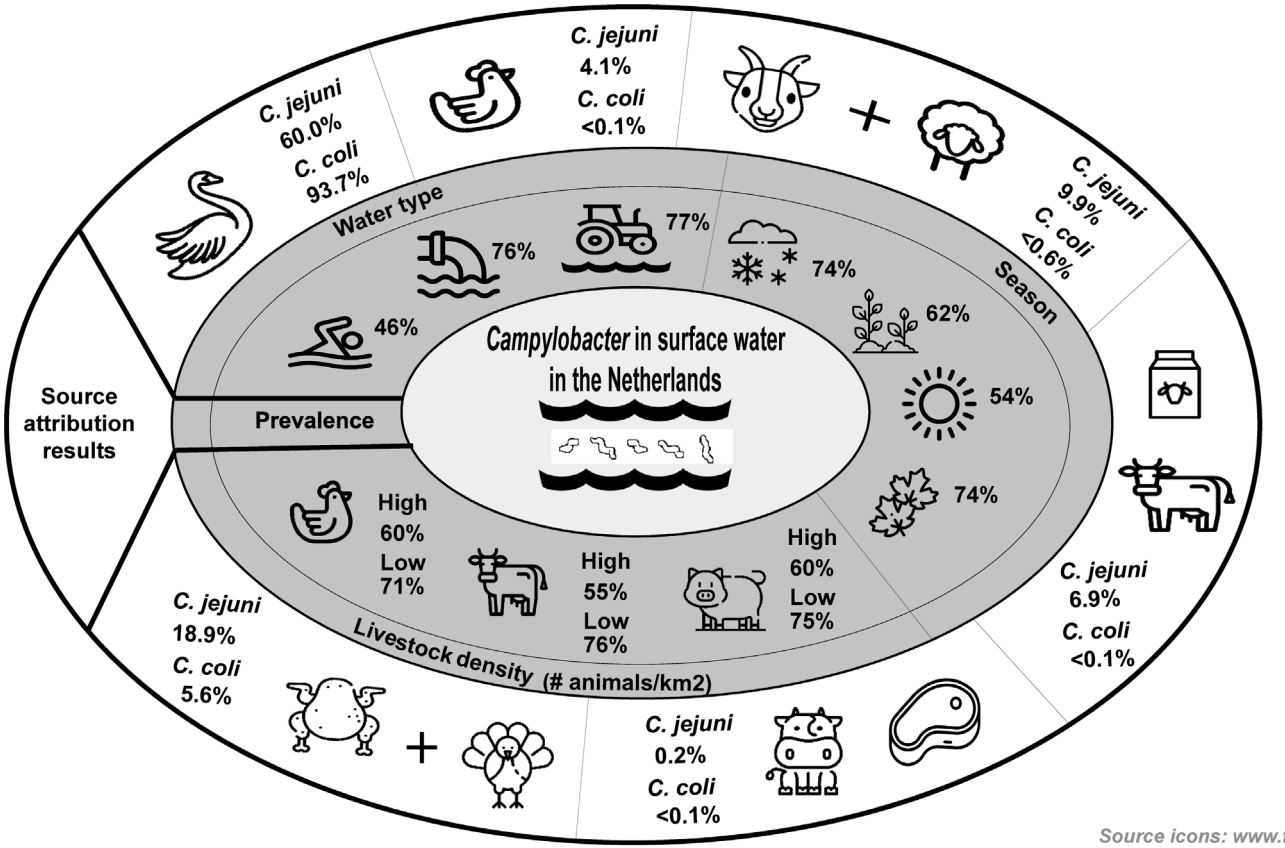
Annemieke C. Mulder
Eelco Franz
Sharona de Rijk
Moyke A.J. Versluis
Claudia Coipan
Ralph Buij
Gerard Müskens
Miriam Koene
Roan Pijnacker
Birgitta Duim
Linda van der Graaf-van Bloois
Kees Veldman
Jaap A. Wagenaar
Aldert L. Zomer
Franciska M. Schets
Hetty Blaak
Lapo Mughini-Gras

Published in Water Research, 2020

Abstract

Campylobacter jejuni and *C. coli*, the primary agents of human bacterial gastroenteritis worldwide, are widespread in surface water. Several animal sources contribute to surface water contamination with *Campylobacter*, but their relative contributions thus far remained unclear. Here, the prevalence, genotype diversity, and potential animal sources of *C. jejuni* and *C. coli* strains in surface water in the Netherlands were investigated. It was also assessed whether the contribution of the different animal sources varied according to surface water type (i.e. agricultural water, surface water at discharge points of wastewater treatment plants [WWTPs], and official recreational water), season, and local livestock (poultry, pig, ruminant) density. For each surface water type, 30 locations spread over six areas with either high or low density of poultry, ruminants, or pigs, were sampled once every season in 2018-2019. *Campylobac ter* prevalence was highest in agricultural waters (77%), and in autumn and winter (74%), and lowest in recreational waters (46%) and in summer (54%). In total, 76 *C. jejuni* and 177 *C. coli* water isolates were whole-genome sequenced. Most *C. coli* water isolates (78.5%) belonged to hitherto unidentified clones when using the seven-locus sequence type (ST) scheme, while only 11.8% of the *C. jejuni* isolates had unidentified STs. The origin of these isolates, as defined by core-genome multi-locus sequence typing (cgMLST), was inferred by comparison with *Campylobacter* strain collections from meat-producing poultry, laying hens, adult cattle, veal calves, small ruminants, pigs, and wild birds. Water isolates were mainly attributed to wild birds (*C. jejuni*: 60.0%; *C. coli*: 93.7%) and meat-producing poultry (*C. jejuni*: 18.9%; *C. coli*: 5.6%). Wild bird contribution was high among isolates from recreational waters and WWTP discharge points, and in areas with low poultry (*C. coli*) or high ruminant (*C. jejuni*) densities. The contribution of meat-producing poultry was high in areas with high density of poultry, springtime, agricultural waters and WWTP discharge points. While wild birds and poultry were the main contributors to *Campylobacter* contamination in surface water, their contribution differed significantly by water type, season, and local poultry and ruminant densities.

Graphical abstract



Source icons: www.flaticon.com

Introduction

Campylobacteriosis is the most frequently reported zoonosis in Europe, with an estimated 70 thousand cases annually in the Netherlands alone (~17 million inhabitants) (Pijnacker et al., 2019). *Campylobacter jejuni* and *Campylobacter coli* are the two species of the *Campylobacter* genus that account together for over 90% of human campylobacteriosis cases in Europe (Centre of Disease Control and Prevention, 2017; European Food Safety and European Centre of Disease Prevention and Control, 2018). Besides gastroenteritis, a *Campylobacter* infection can result in more severe diseases, such as Guillain-Barré syndrome, reactive arthritis, and irritable bowel disease, which strongly contribute to the disease burden of campylobacteriosis (Halvorson et al., 2006; Nachamkin et al., 1998; Ternhag et al., 2008). Although up to 80% of all human campylobacteriosis cases can be attributed to the poultry reservoir, several epidemiological studies have shown that only 40% of these poultry-associated cases can be explained by the consumption of chicken meat (Friesema et al., 2012; Mughini Gras et al., 2012; Veilinga and van loock, 2002; Wagenaar et al., 2013; Wagenaar et al., 2015; Wilson et al., 2008). Most interventions to control *Campylobacter* infections have focused on spread through the food production chain, particularly poultry meat, with limited effects. Accordingly, there has been no appreciable decrease in the incidence of human campylobacteriosis so far (European Food Safety and European Centre of Disease Prevention and Control, 2018). This emphasizes the need to study transmission routes other than food (Sears et al., 2011; Stern et al., 2003), such as those involving the aquatic environment, as *Campylobacter* is commonly found in surface water contaminated with animal faeces, sewage effluent, and agricultural runoff (Jones, 2001).

Even though *Campylobacter* is believed to survive poorly outside the host, some specialist strains have been found to be successfully adapted to survival outside an animal host in certain sylvatic (Hepworth et al., 2011), farmland (French et al., 2005) and environmental niches (Colles et al., 2011; French et al., 2005; Sopwith et al., 2008). These strains are generally more resistant to physical stress than other strains (Sopwith et al., 2008). *Campylobacter* also has the ability to convert into a viable but nonculturable state, to advert conditions while being outside the host (Collins and Colwell, 1986; Murphy et al., 2006). These characteristics indicate that surface water serves more as a vehicle of transmission for *Campylobacter* among animals, from animals to humans and vice versa, rather than as an amplifying reservoir per se.

A previous source attribution study has shown that poultry and wild birds are the most important contributors to surface water contamination with *C. jejuni* and *C. coli* in the Netherlands and Luxembourg, followed by ruminants and pigs (Mughini-Gras et al., 2016). The relative contributions of wild birds and poultry seemed to

vary with season, water type, and the magnitude of the local poultry production, suggesting substantial dissemination of *Campylobacter* into the environment from poultry farms in poultry-rich regions. Although the aforementioned study quantified the contributions of different animal sources to *C. jejuni* and *C. coli* contamination in surface water, the authors acknowledged that the interpretation of their findings was limited by the extensive use of non-local and non-recent source data, retail food data, and coarse spatial resolution of the analyses (Mughini-Gras et al., 2016). Therefore, further testing of the previously formulated hypotheses by using more representative data in additional smaller-scale analyses was necessary.

Due to the paucity of epidemiological research on non-foodborne transmission routes of *Campylobacter*, innovative control measures to limit *Campylobacter* spread into the environment have not yet been developed. Although the aquatic environment seems to contribute to the transmission of *Campylobacter* to humans, the extent to which this is determined by faecal pollution from different types of livestock and wildlife remains unclear. This study aimed to quantify *Campylobacter* prevalence and genotype diversity in surface water, as well as the relative contributions of several putative animal sources to surface water contamination with *C. jejuni* and *C. coli* in the Netherlands using high-throughput genomic data derived from whole-genome sequencing (WGS). Additionally, potential effects of local livestock density, type of surface water, and season were assessed.

Materials and methods

Water samples

Study areas

Water samples were collected in areas which largely varied in densities of specific livestock groups in the Netherlands. Specifically, six areas were selected, with either a high or a low density of poultry (i.e. broiler chickens, laying hens and turkeys combined), pigs, or ruminants (i.e. cattle, sheep and goats combined). To this end, the livestock density per municipality (number of animals/km^2) was calculated per livestock group based on official agricultural census data and land surface per municipality available at the time of this study set-up (December 2018) from Statistics Netherlands (CBS, 2018a,b). Then, the first and last quintiles of the frequency distributions of the density of each livestock group were calculated, which thereby defined the high and low densities of each livestock group in question (high poultry density area: 19,677 poultry/km^2; low poultry density area: <1 poultry/km^2; high pig density area: 2,729 pigs/km^2; low pig density area: 4 pigs/km^2; high ruminant density area: 451 ruminants/km^2; low ruminant density area: 26 ruminants/km^2). The corresponding areas were geographically identified for each livestock group

separately (Figure 1) using ArcMap 10.5 (ESRI). However, those areas could contain multiple livestock types, as the different livestock types are widely spread and mixed throughout the Netherlands (Smit and Heederik, 2017). The selection was based on the following criteria: i) a municipality could only be included in one livestock density area; ii) the areas needed to contain enough surface water sampling sites to allow for data collection (see also *Sampling sites*); iii) the areas needed to be of comparable size.

Sampling sites

Three different types of surface freshwater were selected: i) recreational water at official bathing sites that have to comply with European bathing water legislation (EUR-Lex., 2006) and therefore generally have relatively low levels of faecal contamination; ii) surface water (e.g. drainage ditches, irrigation canals, etc.) in farmlands and pastures that are mainly faecally contaminated by agricultural activities (i.e. run-off from farms, grazing fields and crops fertilized with manure or accessed by free-ranging animals, etc.); and surface water at the discharge sites of effluents of wastewater treatment plants (WWTP), which are a source of contamination with faecal material mainly of human origin. Within each of the six study areas (Figure 1), 15 surface water sampling sites were selected, i.e. five sites for each of the three types of surface water. Each of the selected 90 sampling sites was sampled four times, once per season (summer: June to August; autumn: September to November; winter: December to February; spring: March to May), resulting in a total of 360 planned water samples.

Water sample collection and analysis

Sampling was performed between April 2018 and February 2019 by an accredited contractor (OMEGAM-Water B.V.). Water samples were taken according to the ISO 19458:2006 procedure and immediately cooled and transported to the laboratory at the Dutch National Institute for Public Health and the Environment (RIVM), where they were stored at 4°C and analyzed within 24 hours from sampling. Samples in a total volume of 1000 ml were filtered using 0.45 µm cellulose-based membrane filters (Milli- pore). The filters were placed in Preston broth and incubated under microaerobic conditions using CampyGen sachets (Oxoid) for 48 h at 37°C. Samples were then streaked (10 μl) on modified charcoal cefoperazone deoxycholate (mCCDA) agar and the plates were incubated under microaerobic conditions for 48 h at 41.5°C. From each sample, a maximum of five colonies was inspected by light microscopy for *Campylobacter* characteristics, and a maximum of five visually confirmed colonies per sample were analyzed using Matrix-Assisted Laser-Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS, Bruker Microflex LT, Germany) for species identification. Per individual sample, one *C. jejuni* isolate and one *C. coli* isolate was selected at random for whole-genome sequencing (WGS).

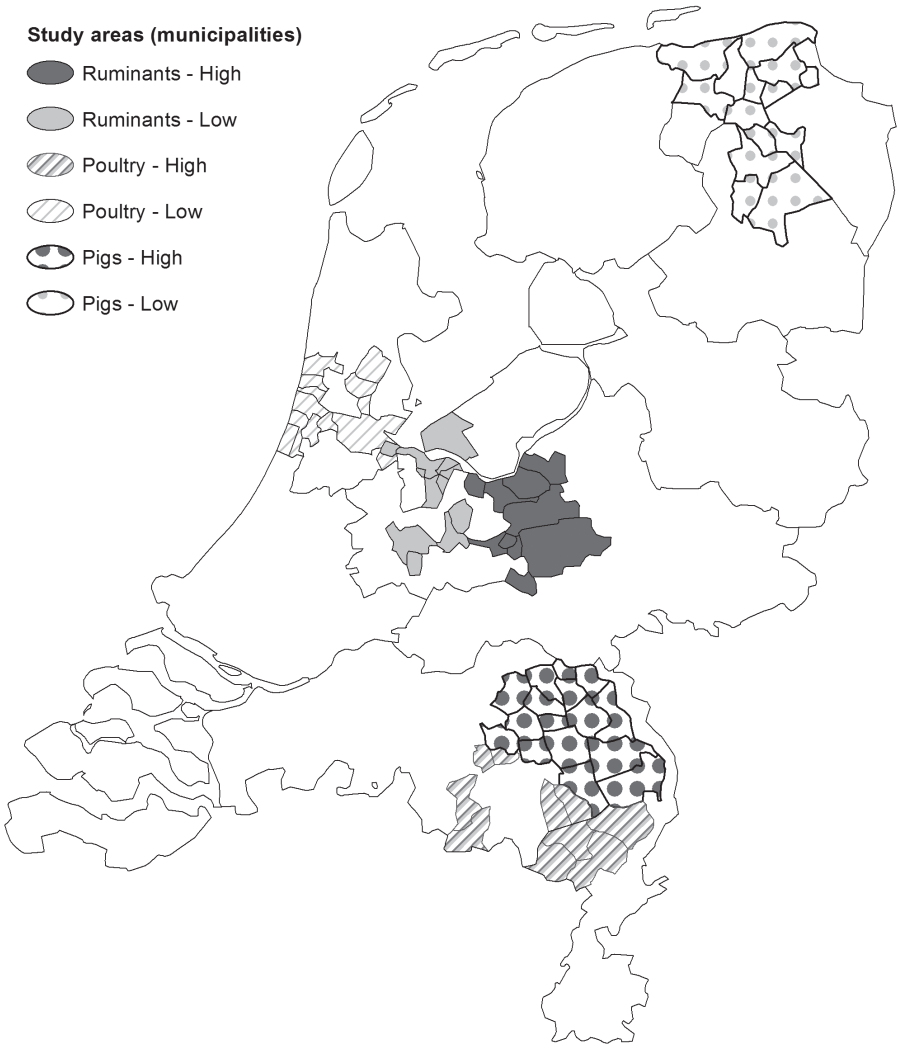


Figure 1. Representation of the six study areas for surface water sampling with high and low densities of poultry, pigs, or ruminants.

Animal data collection

Livestock

Livestock isolates of *C. jejuni* and *C. coli* from faecal samples and carcasses were collected at farms and slaughterhouses by Wageningen Bioveterinary Research (WBVR) and Wageningen Food Safety Research (WFSR), in collaboration with

the RIVM and the Netherlands Food and Consumer Product Safety Authority (NVWA). This was done within the framework of established and nationally representative surveillance programs for zoonotic agents (Opsteegh et al., 2018) and antimicrobial resistance (de Greeff et al., 2019) in food-producing animals, as well as routine inspection and testing activities by veterinary services, in the Netherlands during 2014-2019. Additional isolates from small ruminants (sheep and goats) were obtained through a small-scale internal project including *ad-hoc* sampling events at farms in the Netherlands conducted by engaging field veterinarians collaborating with the Veterinary Microbiological Diagnostic Centre (VMDC) of Utrecht University.

Isolates were obtained from faecal samples analysed without enrichment by direct streaking onto mCCDA (Oxoid) plates following the ISO 10272-1:2017 procedure, whereas carcass samples were analysed with an enrichment step in accordance with the same procedure. Species identification was performed using MALDI-TOF MS.

Wild birds

Fresh droppings or cloacal swabs of some of the most widespread species of waterfowl, pigeons and gulls in the Netherlands were collected by Wageningen Ecological Research (WER) in June and December 2018 to cover both the summer and winter seasons. Sampling was performed by convenience at different locations across the country selected based on previous studies (Lange et al., 2013), which did not include the water sampling sites. Both herbivorous (i.e. geese, mute swan, ducks, common wood pigeon) and omnivorous/piscivorous (gulls, great cormorant) bird species were sampled, as the latter species have higher *Campylobacter* concentrations in their faeces, whereas herbivorous birds produce more faeces per day (de Lange, 2013).

Whole-genome sequencing (WGS)

All gathered *Campylobacter* isolates from both surface water and animals were subject to WGS. DNA isolation was done using the UltraClean® Microbial DNA Isolation Kit (Qiagen, USA). WGS was performed on Illumina Hiseq and NextSeq platforms (Illumina, USA) using 2 × 150-bp reads. Genomes were assembled with SPAdes v3.10.1 (Bankevich, 2012) and checked for completeness and contamination using CheckM (Parks et al., 2015); genomes with >5% contamination or <95% completeness were excluded. The sequences were deposited in ENA Sequence Read Archive project PRJEB38253.

A standard core genome multilocus sequence typing (cgMLST) scheme for *Campylobacter* was applied (Cody et al., 2017) using Seemanns' MLST tool to scan contig files against traditional PubMLST typing schemes (https://github.com/

tseemann/mlst) modified for cgMLST schemes (https://github.com/aldertzomer/cgmlst). The cgMLST profile was assessed using the sequence definitions in BIGSdb (accessed at November 9th, 2019). Additional searches for missing genes were performed using the Basic Local Alignment Search Tool (BLAST) v2.5.0 (Altschul et al., 1990) on the assembled genomes. For the alleles not yet present in BIGSdb, multiple alignments of each locus were performed using MAFFT v7.407 (Katoh et al., 2002) and these were assigned unique identification numbers. All the loci for which none of these approaches provided unambiguous results were considered as missing. Loci with missing allele numbers in >5% of the isolates were excluded from the analysis. For description purposes, the sequence types (STs) based on the conventional 7-locus MLST scheme (Dingle, 2001) were also derived from the WGS data.

Data analysis

Prevalence and ST diversity in surface water

The prevalence of *C. jejuni* and *C. coli* (and their different STs) in surface water was calculated for the different types of surface water, seasons and livestock density areas. Simpson's diversity index was calculated to quantify the diversity of STs (Anandan et al., 2014). The structure of the *Campylobacter* population was visualized using both conventional MLST- and cgMLST-based minimum spanning trees (MST) to appreciate interrelationships and clusters among the *Campylobacter* isolates.

Analysis of Molecular Variance (AMOVA)

To test for genetic differentiation in STs between the sources, statistics were estimated using analysis of molecular variance (AMOVA) (Excoffier, 1992), an extension of the analysis of variance (ANOVA) that focusses on the (genetic) heterogeneity between groups. If the mutual heterogeneity between two sources was not significant, they were pooled into a new group. AMOVA was conducted using the R packages "poppr" (version 2.8.5) and "hierfstat" (version 0.04-22) in R (version 3.6.0) (RCT, 2015).

Source attribution analysis

The *C. jejuni* and *C. coli* isolates from surface water were attributed to the putative animal sources as defined by the AMOVA. cgMLST-based source attribution analysis using an established population genetics model, i.e. STRUCTURE (version 2.3.4), was performed to estimate relative probabilities for each *Campylobacter* strain found in surface water to originate from each of the animal sources (Hubisz et al., 2009; Pritchard et al., 2000). A model with no admixture and with the "USEPOPINFO" flag was used to determine the ancestry of the isolates to be attributed, i.e. the surface water isolates. Therefore, each animal population was considered as discrete

and the origin of each water isolate *i* was estimated under the assumption that the isolate comes directly from one of the *K* animal sources, with a prior probability for each source of 1/*K* (Porras-Hurtado et al., 2013; Pritchard et al., 2009). The USEPOPINFO flag was used to pre-specify the population of origin of the animal isolates as to assist inference of the origin for the water isolates, whose (animal) populations of origin were set as unknown (Pritchard et al., 2009). By pre-setting the populations of origin of the animal isolates based on the AMOVA results, the cluster structure corresponded to the pre-defined populations, which were in agreement with the genetic information and made the output more interpretable. The very few missing alleles (0.3% of the total), minimized by performing additional searches and blasting, as well as by excluding loci with considerable and systematic missingness over the isolates, were then handled with the default software function, which ignores missing data when updating parameters. The length of the burning period was set at 1,000, followed by 10,000 iterations, which were able to provide adequate convergence of parameter estimation. The overall proportion of surface water isolates attributed to a given source was then calculated as the sum of the relative probabilities for that source of the surface water isolates divided by the total number of surface water isolates. The 95% confidence intervals (CIs) for the attribution were derived in R with the "boot" (version 1.3-24) package to provide bootstrapped values of the average attributions per source with 1,000 replications.

Effects of livestock density, water type and season

Significance testing of the differences in attribution estimates (i.e. the source probabilities) for the surface water isolates between the livestock density areas, types of water, and seasons, was performed using multiple linear regression analysis in R. A logarithmic transformation of the outcome variable (i.e. source probabilities) was applied. For this analysis, the attribution estimates for meat-producing poultry and laying hens were combined into 'poultry', and those for adult cattle, veal calves and small ruminants into 'ruminants' in order to reflect the livestock groups used in the definition of the livestock density areas. Finally, the multivariate shared relationships of the variables type of water, season and livestock density, with the attributions of surface water isolates were explored using canonical correlation analysis (CCA).

Results

Isolate collection

In total, 360 water samples were planned to be taken during this study. However, due to a few sampling locations being temporarily inaccessible or being without water due to drought at a given sampling event, a total of 348 samples were eventually collected. In total, 411 isolates (304 *C. coli* and 107 *C. jejuni*) were obtained from

those surface water samples. From each individual sample, only one *C. jejuni* isolate and one *C. coli* isolate was selected at random for WGS. This resulted in a selection of 253 water isolates (177 *C. coli* and 76 *C. jejuni*). In total, 570 *C. jejuni* and 152 *C. coli* isolates were obtained from different livestock species (Table 1) and 47 *C. jejuni* and 15 *C. coli* isolates were obtained from wild birds (Supplementary Material, Table S1). This resulted in a total of 1,037 *Campylobacter* isolates (253 from surface water and 784 from animals) which were subjected to WGS. As was described in the materials and methods section (Whole-genome sequencing (WGS)), 88 loci with missing allele numbers in >5% of the isolates were excluded, resulting in 1,255 loci with 99.7% complete allele numbers in the whole dataset.

Table 1. Total number of *Campylobacter* isolates obtained from each livestock group.

Livestock	**Total number of isolates** (N)	***C. jejuni* isolates** (N)	***C. coli* isolates** (N)
Broiler chickens	200	186	14
Laying hens	56	55	1
Turkeys	38	37	1
Beef cattle	96	96	0
Dairy cattle	62	61	1
Veal calves	49	39	10
Small ruminants	111	86	25
Pigs	110	10	100

Prevalence and STs in surface water

Prevalence

The overall *Campylobacter* prevalence in surface water samples was 66% (Table 2). Prevalence was highest in agricultural waters (77%) and at WWTP discharge points (76%), and lowest in official recreational waters (46%). Prevalence was highest in autumn (74%) and winter (74%) and lowest in spring (62%) and summer (54%).
Prevalence was generally higher in the low livestock density areas (poultry 71%, pigs 75%, and ruminants 76%) as compared to the high livestock density areas (poultry 60%, pigs 60%, and ruminants 55%) (Table 3). However, in areas with high poultry and ruminant densities, *C. jejuni* prevalence was higher, but *C. coli* prevalence was lower, as compared to areas with low poultry and ruminant densities. In general, *C. coli* was more often isolated than *C. jejuni*, except for the high ruminant density area where the prevalence was the same for both *Campylobacter* species (33%). Also in the high poultry density area, the difference in prevalence between *C. coli* (40%) and *C. jejuni* (33%) was small compared to that in the other areas.

Table 2. Total number of surface water samples tested (N), number and percentage of *C. jejuni* and *C. coli* positive samples (Pos and %) per surface water type and season

	Spring			Summer			Autumn			Winter			Total		
	N	Pos	%	N	N	Pos	%	Pos	%	N	Pos	%	N	Pos	%
Agricultural waters	30	24	80	26	17	65	26	20	77	29	24	83	111	85	77
WWTP discharge points	30	22	73	30	21	70	30	23	77	30	25	83	120	91	76
Recreational waters	30	10	33	29	8	28	29	20	69	29	16	55	117	54	46
Total	90	56	62	85	46	54	85	63	74	88	65	74	348	230	66

Table 3. Total number of surface water samples tested (N) and percentage (%) of *C. jejuni* and *C. coli* positive samples per livestock density areas.

	Low poultry density	High poultry density	Low pig density	High pig density	Low ruminant density	High ruminant density
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Samples	59 (100)	52 (100)	60 (100)	58 (100)	58 (100)	60 (100)
Positive	42 (71)	31 (60)	45 (75)	35 (60)	44 (76)	33 (55)
C. coli [a]	37 (63)	21 (40)	39 (65)	27 (47)	36 (62)	20 (33)
C. jejuni [a]	7 (12)	17 (33)	11 (18)	10 (17)	13 (22)	20 (33)

[a] The sum of the number of positive samples of *C. jejuni* and *C. coli* is not equal to the number of positive samples, because each surface water sample can contain multiple isolates of *C. jejuni* and *C. coli*.

Sequence types

Overall, 105 (41.5%) sequenced *Campylobacter* isolates had known STs, whereas about 60% of the isolates (58.5%) had thus far unidentified STs. Most *C. coli* water isolates (78.5%) belonged to those unidentified clones, while only 11.8% of the *C. jejuni* isolates had an unknown ST. The four most prevalent STs in surface water were ST45 (n=10, 4.0%), ST1766 (n=5, 2.0%), ST137 (n=4, 1.6%) and ST2654 (n=4, 1.6%). The prevalence of STs differed considerably between the three different types of surface water. While only 15 of 54 isolates from recreational waters had known STs, which were each detected once, isolates from WWTP discharge points and agricultural waters had higher proportions of known STs: 41/91 and 49/85 isolates (i.e. 45,1% and 57,7%), respectively. At WWTP discharge points, the two most common STs were ST2654 (n=4, 4.4%) and ST137 (n=3, 3.3%), whereas the two most common STs in agricultural waters were ST45 (n=9, 10.6%) and ST1766 (n=4, 4.7%). An overview of all STs found in surface water and STs grouped by season, water type and livestock density area are reported in Figure S1, Supplementary Material.

Source heterogeneity

The AMOVA showed that there was significant genetic heterogeneity between most of the sources. Non-significant heterogeneity was observed between broilers and turkeys, and between dairy cattle and beef cattle. Therefore, these sources were combined into 'meat-producing poultry' (i.e. broilers and turkeys) and 'adult cattle' (i.e. dairy and beef cattle) for further analyses. The values and corresponding p-values for each pair of sources are reported in Table S2, Supplementary Material.

ST diversity among Campylobacter isolates from surface water

Simpson's diversity index based on the STs found in surface water was 0.96 for *C. jejuni* and 0.94 for *C. coli* (overall 0.97), indicating the probability that two isolates randomly selected from surface water belong to different STs. Diversity was lowest in areas with low pig density, at WWTP discharge sites, and in spring, and highest in areas with high ruminant density, recreational waters and in winter (Table 4).

Table 4. Simpson's index of diversity of *Campylobacter* STs from surface water.

Variable	Total	*C. jejuni*	*C. coli*
Overall	0.97	0.96	0.94
Livestock density area			
Low ruminant density	0.94	0.88	0.88
High ruminant density	0.95	0.95	0.00
Low poultry density	0.92	0.83	0.83
High poultry density	0.93	0.92	0.75
Low pig density	0.88	0.81	0.67
High pig density	0.91	0.74	0.88
Water type			
Agricultural waters	0.94	0.88	0.89
WWTP discharge points	0.93	0.93	0.91
Recreational waters	0.96	0.91	0.75
Season			
Spring	0.92	0.88	0.75
Summer	0.94	0.86	0.90
Autumn	0.94	0.91	0.86
Winter	0.96	0.94	0.91

Population structure

The cgMLST-based MSTs visualizing the population structure of the *C. jejuni* and *C. coli* isolates are shown in Figure 2 and Figure 3, respectively. For *C. jejuni* (Figure 2), surface water and wild bird isolates generally clustered together, but surface water isolates were also found among isolates from other sources, specifically meat-producing poultry and laying hens. The *C. coli* tree showed significant clustering of the surface water and wild bird isolates, which were clearly separated from the isolates of the other sources (Figure 3). A similar structure was appreciable in the MSTs based on conventional MLST (Figure S2 and Figure S3, Supplementary Material).

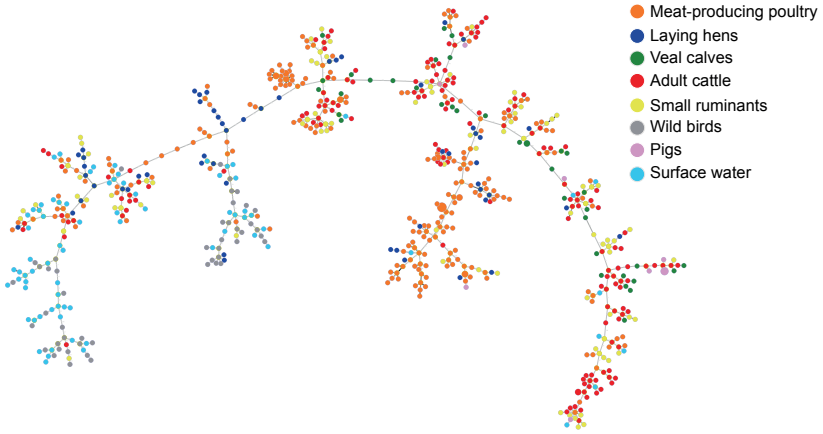


Figure 2. Core-genome MLST-based minimum spanning tree showing the population structure of the *C. jejuni* isolates from surface water and from the different animal sources.

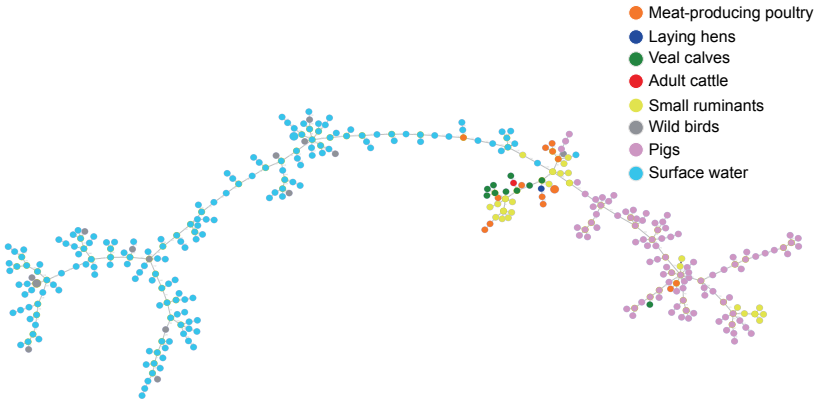


Figure 3. Core-genome MLST-based minimum spanning tree showing the population structure of the *C. coli* isolates from surface water and from the different animal sources.

Attribution of animal sources to surface water

Most *C. jejuni* isolates in surface water could be attributed to wild birds (60.0%, 95%CI: 48.9-71.3%), followed by meat-producing poultry (18.9%, 95%CI: 10.5-26.8%), small ruminants (9.9%, 95%CI: 4.4-16.4%), adult cattle (6.9%, 95%CI: 2.3-12.4%), laying hens (4.1%, 95%CI: 0.4-8.8%), and veal calves (0.2%, 95%CI: 0.0-0.5%) (Figure 4). The attribution of *C. jejuni* in surface water to pigs was too small to be sized. The vast majority of *C. coli* isolates in surface water could also be attributed to wild birds (93.7%, 95%CI: 90.4-96.4%), followed by meat-producing poultry (5.6%, 95%CI: 3.1-8.9%) and small ruminants (0.6%, 95%CI: 0.1-1.7%), while the other animal sources accounted altogether for <0.1% of *C. coli* isolates. When the attribution estimates were split by livestock density area, wild birds were again the predominant contributor to *C. jejuni* contamination in surface water in all areas (Table 5). The second contributor was meat-producing poultry in most areas, except for the area with low pig density where small ruminants where the second most important contributor. The contribution of wild birds to *C. jejuni* contamination in surface water was higher in the low poultry density area (99.4%) than in the high poultry density area (59.2%). The opposite was observed for meat-producing poultry: in the high poultry density area, the contribution of meat-producing poultry to *C. jejuni* contamination in surface water (24.9%) was higher than in the low poultry density area (0.3%). In both the low and high ruminant density areas, small ruminants and laying hens ranked respectively as third and fourth most important contributors to *C. jejuni* contamination in surface water, without large differences in the contributions between the two areas. In the low ruminant density area, however, the contribution of adult cattle (6.8%) was higher than in the high ruminant density area where there was no detectable contribution of adult cattle at all. Also for *C. coli*, the contribution of wild birds was higher in the low poultry density area (93.2%) than in the high poultry density area (87.0%), although the difference was smaller than for *C. jejuni*. The opposite was observed for meat-producing poultry (low poultry density area: 4.1%; high poultry density area: 12.8%). There were no sizeable contributions of other sources of *C. coli* water contamination.

Wild birds were the predominant contributor to *C. jejuni* and *C. coli* contamination in all three water types (*C. jejuni*: recreational waters 92.3%, agricultural waters 53.5%, and WWTP discharge points 51.8%, *C. coli*: recreational waters 98.9%; WWTP discharge points 95.0%; agricultural waters 88.4%). Meat-producing poultry was the second most important contributor in all three types of water for both *C. jejuni* and *C. coli*. Finally, wild birds were also the largest contributor to surface water contamination with both *C. jejuni* and *C. coli* in all seasons. For *C. jejuni*, the second most important contributor in spring was cattle, while meat-producing poultry was the second most important contributor in the other seasons. For *C. coli*, meat-producing poultry was the second most important contributor in all seasons.

Table 5. Source attribution estimates of the *C. jejuni* and *C. coli* isolates in surface water per livestock density area, water type, and season based on cgMLST.

Species	Variable	Category	Wild birds	Meat-producing poultry	Laying hens	Adult cattle	Small ruminants	Veal calves
C. jejuni	**Livestock**	Poultry high	59.2 (35.8-82.2)	24.9 (8.1-44.4)	0.3 (0.0-0.9)	15.4 (0.0-33.0)	0.2 (0.0-0.4)	0.0 (0.0-0.0)
	density	Poultry low	99.4 (99.0-99.7)	0.3 (0.2-0.5)	0.1 (0.0-0.2)	0.2 (0.0-0.6)	0.0 (0.0-0.0)	0.0 (0.0-0.0)
		Pigs high	40.6 (10.8-70.3)	23.2 (3.0-46.8)	10.1 (0.2-28.3)	5.6 (0.0-16.8)	20.6 (0.3-50.4)	0.0 (0.0-0.0)
		Pigs low	50.0 (12.6-87.3)	16.8 (0.0-0.4)	1.1 (0.0-3.2)	14.0 (0.0-37.6)	17.6 (0.1-47.5)	0.7 (0.0-2.0)
		Ruminants high	64.4 (44.3-84.5)	18.4 (3.8-33.9)	4.6 (0.0-13.9)	0.0 (0.0-0.0)	12.6 (0.5-27.5)	0.0 (0.0-0.0)
		Ruminants low	54.6 (28.9-78.5)	19.8 (3.4-39.8)	7.2 (0.0-21.4)	6.8 (0.0-1.9)	11.1 (0.0-28.5)	0.6 (0.0-1.9)
	Water	Agricultural waters	53.5 (35.4-69.3)	21.3 (8.7-34.6)	3.8 (0.3-10.2)	10.3 (2.2-19.8)	10.7 (1.4-21.9)	0.3 (0.0-0.8)
	type	Recreational waters	92.3 (78.0-99.7)	7.2 (0.3-20.5)	0.4 (0.0-1.3)	0.0 (0.0-0.0)	0.1 (0.0-0.3)	0.0 (0.0-0.0)
		WWTP discharge	51.8 (34.2-69.6)	21.8 (10.0-36.5)	6.1 (0.0-15.2)	6.5 (0.0-16.3)	13.6 (3.6-25.5)	0.18 (0.0-0.5)
	Season	Spring	54.5 (27.2-81.7)	12.3 (0.3-33.1)	0.0 (0.0-0.0)	23.9 (0.0-50.4)	9.3 (0.1-27.5)	0.0 (0.0-0.1)
		Summer	40.2 (16.1-66.1)	36.2 (15.4-57.6)	1.1 (0.0-2.6)	5.2 (0.0-13.2)	17.3 (0.2-38.5)	0.0 (0.0-0.0)
		Autumn	66.4 (47.2-85.5)	12.1 (0.6-26.1)	8.9 (0.0-22.4)	4.6 (0.1-13.4)	7.6 (0.1-19.0)	0.4 (0.0-1.2)
		Winter	66.7 (47.4-85.4)	18.0 (0.8-26.2)	3.5 (0.0-22.4)	3.2 (0.1-13.3)	8.4 (0.1-19.4)	0.2 (0.0-1.2)
C. coli	**Livestock**	Poultry high	87.0 (74.6-98.6)	12.8 (1.1-28.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
	density	Poultry low	93.2 (85.8-98.6)	4.1 (1.1-9.8)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	2.6 (0.0-7.6)	0.0 (0.0-0.0)
		Pigs high	90.9 (79.3-98.9)	9.0 (1.0-20.5)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
		Pigs low	99.1 (98.7-99.4)	0.9 (0.6-1.2)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)
		Ruminants high	93.3 (82.9-99.0)	6.5 (1.0-17.9)	0.1 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.3)	0.0 (0.0-0.0)
		Ruminants low	94.7 (88.6-98.8)	5.0 (1.0-11.2)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
	Water	Agricultural waters	88.4 (81.6-95.3)	11.4 (5.4-19.1)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.1)	0.0 (0.0-0.0)
	type	Recreational waters	98.9 (98.5-99.3)	1.0 (0.0-1.4)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)

Species	Variable	Category	Wild birds	Meat-producing poultry	Laying hens	Adult cattle	Small ruminants	Veal calves
		WWTP discharge	95.0 (90.7-98.4)	3.4 (1.3-7.3)	0.1 (0.0-0.1)	0.0 (0.0-0.0)	1.5 (0.1-4.2)	0.0 (0.0-0.0)
	Season	Spring	92.2 (83.6-98.6)	7.7 (1.1-16.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
		Summer	93.0 (85.8-98.7)	6.9 (1.3-13.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
		Autumn	94.7 (89.6-98.8)	5.0 (1.2-10.5)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.1 (0.0-0.2)	0.0 (0.0-0.0)
		Winter	94.4 (88.0-98.7)	3.5 (1.0-7.8)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	2.1 (0.0-6.0)	0.0 (0.0-0.0)

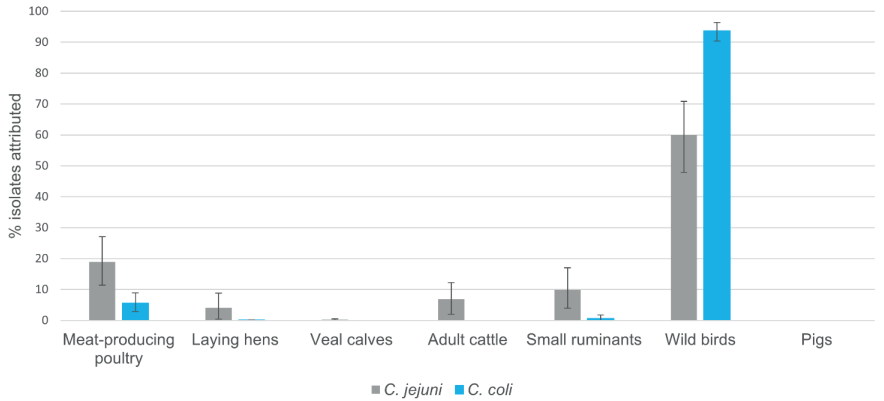


Figure 4. Overall attributions of *C. jejuni* and *C. coli* isolates from surface water to the animal sources based on cgMLST. Error bars indicate 95% confidence intervals.

Effects of livestock density, water type and season on the attribution estimates

The differences in attribution estimates for the *C. jejuni* and *C. coli* isolates from surface water (described for each source in the results section (Prevalence and STs in surface water) between livestock density areas, water types, and seasons were further tested for statistical significance using multiple linear regression. The significant differences are summarized in Table 6.

For *C. jejuni*, significantly higher attributions to wild birds were associated with high ruminant density (β=9.21, 95%CI 1.62;16.80), recreational waters (β=15.31, 95%CI 8.04;22.57), and agricultural waters (β=5.66, 95%CI 0.92;10.40). Furthermore, the attributions to ruminants were negatively associated with high ruminant density (β=-10.66, 95%CI -20.42;-0.79), and positively associated with agricultural waters (β=11.19, 95%CI 1.79;20.59) and WWTP discharge points (β=0.10, 95%CI 0.77;20.05), mostly during the warmer seasons. Attributions to poultry were positively associated with WWTP discharge points in winter (β=5.24, 95%CI 0.90;9.59). For *C. coli*, significantly higher attributions to poultry were associated with high poultry density (β=2.51, 95%CI 0.10;4.91) and agricultural waters (β=2.14, 95%CI 0.69;3.58). Furthermore, the attributions to wild birds were negatively associated with high poultry density (β=-2.63, 95%CI -5.04;-0.21), and positively associated with recreational waters (β=1.95, 95%CI 0.07;3.83) and WWTP discharge points (β=2.13, 95%CI 0.66;3.61).

Figure 5 shows the canonical correlation analysis (CCA) plots to visualize the results of the regression analyses for both *C. jejuni* and *C. coli.* The dots represent the attributions of the surface water isolates to the different sources and the arrows represent the different variables used in the linear regression analysis (i.e. livestock density area, water type or season) to test for differences in the attributions. The stronger the association of a variable with the attributions to a specific source, the longer the arrows. If an arrows points in the same direction as a particular source (dot), this means that there is a positive association between that source and the given variable.

Table 6. Significant associations of *C. jejuni* and *C. coli* attributions based on cgMLST with livestock density, water type, and season.

Species	Source	Variable	Season	Beta	95% CI	P-value	Ref. category water	Ref. category livestock
C. jejuni	**Poultry***	WWTP discharge points	Winter	5.24	0.90;9.59	0.02	Recreational waters	-
	Ruminants**	Ruminants high	Spring	-10.60	-20.42;-0.79	0.04	-	Ruminants low
		Agricultural waters	Spring	11.19	1.79;20.59	0.03	Recreational waters	-
		Recreational waters	Summer	-1.04	-20.05;-0.77	0.04	WWTP discharge points	-
		WWTP discharge points	Summer	0.10	0.77;20.05	0.04	Recreational waters	-
	Wild birds	Ruminants high	Spring	9.21	1.62;16.80	0.02	-	Ruminants low
		Agricultural waters	Winter	5.66	0.92;10.40	0.02	WWTP discharge points	-
		Recreational waters	Spring	15.31	8.04;22.57	0.00	Agricultural waters	-
		Recreational waters	Spring	10.93	3.67;18.20	0.01	WWTP discharge points	-
		Recreational waters	Winter	7.51	2.10;12.93	0.01	WWTP discharge points	-
C. coli	**Poultry**	Poultry high	Spring	2.51	0.10;4.91	0.04	-	Poultry low
		Agricultural waters	Autumn	2.14	0.69;3.58	0.00	WWTP discharge points	-
	Wild birds	Poultry high	Spring	-2.63	-5.04;-0.21	0.03	-	Poultry low
		Recreational water	Winter	1.95	0.07;3.83	0.04	WWTP discharge points	-
		WWTP discharge points	Autumn	2.13	0.66;3.61	0.01	Agricultural waters	-

*It includes the attributions of meat-producing poultry (broilers and turkeys) and laying hens.
**It includes the attributions of adult cattle, veal calves and small ruminants.

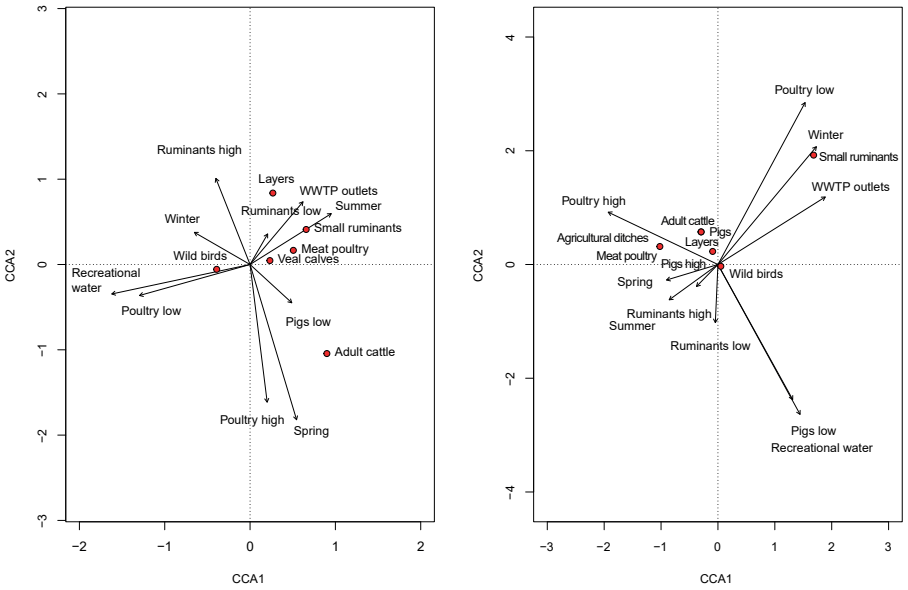


Fig. 5. Canonical correlation analysis plot of *C. jejuni* (left) and *C. coli* (right).

Discussion

In this study, the prevalence, genotype diversity and animal origin of *C. jejuni* and *C. coli* strains isolated from surface water in the Netherlands were investigated. Furthermore, it was assessed whether the estimated contributions of the different animal sources varied significantly with the type of surface water (i.e. recreational waters, agricultural waters, and WWTP discharge points), season, and local livestock density.

C. jejuni and/or *C. coli* strains were detected in 66% of surface water samples, demonstrating the widespread presence of these pathogens in surface water, which is an indication of faecal contamination. In contrast to most animal sources, surface water was mainly contaminated with *C. coli*, with a *C. coli* to *C. jejuni* isolation ratio of about 3:1. This finding agrees with previous European studies (Mughini-Gras et al., 2016; Rosef et al., 2001; Shrestha, 2019). Prevalence of both *C. coli* and *C. jejuni* in agricultural water and water at WWTP discharge points was higher compared to that in recreational water. The relatively low prevalence in recreational water was anticipated, as the microbiological water quality at these official EU bathing sites has to comply with European guidelines for faecal contamination. The higher prevalence in agricultural water was also expected, as these water bodies are usually closer to farms, grazing fields, or fields fertilized with manure where agriculture runoff is more

likely to occur. Furthermore, the similarly high *Campylobacter* prevalence at WWTP discharge points was foreseen, as regular wastewater treatment does not completely remove bacteria (Rechenburg, 2009).

Campylobacter prevalence was higher during autumn and winter compared to spring and especially summer. This finding is in agreement with previous studies showing that *Campylobacter* prevalence in surface water is lower when there are more hours of sunshine (Jones, 2001), probably due to higher ultraviolet radiation levels and temperatures, which eventually lead to reduced *Campylobacter* survival in aquatic environments. Indeed, the summer season in the Netherlands has more hours of sunshine in comparison to the spring (KNMI, 2020), and thus higher levels of ultraviolet radiation and ambient temperatures, which might contribute to decreased *Campylobacter* presence in surface water. In agreement with this finding, an increased risk for human campylobacteriosis associated with swimming in surface water in spring compared to swimming in the summer was previously reported (Mughini Gras et al., 2012). Despite the clear difference in *Campylobacter* prevalence between the warmer (spring and summer) and the colder (autumn and winter) seasons, the differences between spring and summer and between autumn and winter were less prominent or absent. A possible explanation is that the water samples representing spring and summer in this study were taken in 2018, a year characterized by extremely dry spring and summer seasons in the Netherlands, including a drought record in July (KNMI, 2018). From the autumn of 2018 onwards, there was a reduction in precipitation deficit, with a recovery towards normal levels in the winter of 2018/2019. This shows that the weather conditions during sampling were quite similar in terms of precipitation for both the two warmer seasons (spring and summer) and the two colder seasons (autumn and winter) and suggests that the role of the aquatic environment as exposure route for humans to *Campylobacter* varies with the seasons and their weather conditions. Also, a higher frequency of recreational activities in surface water during summer compared to those in winter contributes to this variability.

Strain diversity in surface water as reflected by STs, was very high. There were also high numbers of surface water isolates with novel STs, mainly among *C. coli*. This may be due to *C. coli* being generally under-represented in human and animal samples that have been studied and typed previously, but they were also under-represented in the animal sources that were explored in this study. In agreement with a previous study conducted in the Netherlands and Luxembourg (Mughini-Gras et al., 2016), the most prevalent ST in surface water was ST45, which was most often detected in agricultural waters. ST45 has been recognized to be ubiquitous and to be more frequently found in the environment than other STs that are common in humans

(French et al., 2005; Mughini- Gras et al., 2016; Sopwith et al., 2008). This has lent weight to the hypothesis that ST45 is a potential environmentally adapted ST that is able to survive advert conditions while being outside the host (Colles et al., 2011; French et al., 2005; Sopwith et al., 2008). Also the other STs prevalent in surface water, i.e. ST1766, ST137 and ST2654, have frequently been isolated from surface water and/or wild birds worldwide, as reported in the *Campylobacter* PubMLST database.

The population structure of both *C. coli* and *C. jejuni* from surface water showed predominance of wild bird-like strains compared to other sources. This was particularly the case for *C. coli*, where 94% of the water isolates were attributable to wild birds. The remaining *C. coli* were predominantly attributable to poultry, in particular in areas with high poultry density. Interestingly, while *C. coli* were relatively frequently detected in pigs as well, the cluster of *C. coli* isolates from pigs was clearly separated from the surface water isolates. For *C. jejuni*, wild birds explained about 60% of all isolates in water, and livestock sources, particularly poultry, significantly contributed to water contamination as well. This was supported by the finding that *C. jejuni* prevalence was higher in the high poultry and ruminant density areas compared to the low poultry and ruminant density areas, while for *C. coli* it was the other way around, suggesting a more prominent role of these livestock groups in contaminating surface water with *C. jejuni* relative to *C. coli*.

Source attribution analysis confirmed that wild birds were the likely source of the majority of strains found in surface water, followed by poultry (broilers, turkeys and layers combined) and ruminants (cattle, sheep and goats combined). Similar results were found in studies performed in Luxembourg for *C. jejuni* and *C. coli* (Mughini-Gras et al., 2016) and in New Zealand for *C. jejuni* (Shrestha, 2019) in which about 61% of the surface water isolates originated from wild birds. As the wild bird isolates in this study mainly comprised isolates from aquatic bird species and only from one terrestrial species (common wood pigeon - *Columba palumbus*), this finding is highly plausible. Of note is that when repeating the source attribution analysis with the common wood pigeon isolates as separate group, the attribution to the terrestrial bird species was about 9% (data not shown), showing that aquatic wild birds remain the most likely source of strains found in surface water. With a larger collection of wild bird isolates it would be interesting to focus future studies on how *Campylobacter* prevalence and its attributions to different bird species differ according to their habitats, migration patterns and roosting behaviors (Ito et al., 1988; Waldenström et al., 2002; Whelan et al., 1988).

It was previously reported that *C. jejuni* and *C. coli* isolates from surface water in the Netherlands were mainly attributable to poultry (52%), followed by wild birds (37%). However, that study was performed in poultry-rich regions and results may therefore be explained by a relative high environmental dissemination of *Campylobacter* strains from poultry farms (Mughini-Gras et al., 2016). In agreement with this, the linear regression results in the present study show that surface water strains attributable to poultry were significantly more likely to be found in high compared to low poultry density areas and in agricultural waters. This supports the previously postulated hypothesis that geographical variation in the relative contribution of poultry as a source of surface water contamination with *Campylobacter* is associated with local differences in the magnitude of poultry production (Mughini-Gras et al., 2016). This could also explain the observed decrease in human campylobacteriosis incidence in areas where poultry farms and slaughterhouses were emptied (i.e. culled), thoroughly disinfected and closed to control the devastating H7N7 avian influenza epidemic in 2003 in the Netherlands (Friesema et al., 2012). Indeed, it is possible that this is a reflection of reduced environmental *Campylobacter* load due to the temporary inactiveness of poultry farms.

C. jejuni strains attributable to ruminants were more likely to be isolated from surface water in low vs. high ruminant density areas, which is counterintuitive. A possible explanation could be that in the low ruminant density areas, farming operations are less intensive (and more extensive) in nature, with differences being related to farm size (CBS, 2018a,b), grazing opportunities (e.g. use of pasture lands, time animals spent in pastures) (Van Den Pol-Van Dasselaar et al., 2015), and likely management of manure and distance to surface waterways as well. However, the attribution results of *C. jejuni* are more uncertain then the attribution results of *C. coli* strains, which could also influence the results.

Besides that livestock densities influence the relative contributions of *Campylobacter* of different sources in surface water, it was also shown that there are seasonal and water type-dependent variations in those contributions. Those variations may reflect different conditions facilitating access to, contact with, and discharge of fecal material, into surface water. An example is *C. jejuni* contamination in water at WWTP discharge points. Although contamination from sewage is mainly of human origin, water at WWTP discharge points had a significantly higher contribution of poultry-associated strains than other types of surface water. As poultry is the primary source of human *Campylobacter* infections (Mughini Gras et al., 2012), the *C. jejuni* contamination in water at WWTP discharge points is likely to reflect a pattern more similar to that of the (main) sources of human infections, i.e. poultry, as observed previously (Mughini-Gras et al., 2016), than that of other animal sources.

A few methodological considerations are called for. We used a no admixture model, meaning that each water isolate was assumed to come 'as is' from one of the animal sources. This model was appropriate for this study as we aimed to quantify the fraction of isolates found in surface water that is directly attributable to each of the animal sources, thereby considering only the last transfer step of the (potentially longer and more complex) *Campylobacter* transmission chains among hosts and the environment, i.e. the transfer step from animals to surface water. Indeed, *Campylobacter* is not able to grow outside the host, so its presence in the environment is only a matter of die-off rather than growth. This means that the isolates found in the environment originate as such from (the faeces of) a specific host and are not generated in the environment itself. In the admixture model, on the other hand, the isolates are assumed to have mixed ancestry and this is modelled by saying that, for example, isolate i has inherited a given proportion of its genome from ancestors in population k (Porras-Hurtado et al., 2013; Pritchard et al., 2009). However, here we were interested in knowing the most likely animal origin of an isolate as a whole based on its genome and our goal was not to make evolutionary inference about the life history of strains.

The application of USEPOPINFO allowed for the inclusion of isolates of known origin (i.e. the animal isolates) as to attribute only the isolates of unknown origin (i.e. the water isolates) (Pritchard et al., 2009). Therefore, a potential bias derives from the assumption that the pre-defined (animal) populations are correct, while misclassification might occur. However, the use of AMOVA to (re-)define the groupings of animal isolates to be used as sources in the attribution analysis based on the genetic similarities of their isolates made it possible to consider the pre-defined populations as actual populations.

Conclusions

The results of this study led to the following conclusions:

- *C. coli* is the dominant *Campylobacter* species in surface water.
- *Campylobacter* prevalence is highest in agricultural waters and during the coldest months of the year and lowest in recreational waters and warmer months.
- Wild birds and meat-producing poultry are the main contributors to *Campylobacter* contamination of surface water, with water type, season, and local livestock (particularly poultry and ruminant) density being significant drivers of these constributions.
- Poultry-associated *Campylobacter* strains are mostly found in agricultural waters, water at WWTP discharge points, and in areas with high poultry density.

- Wild bird-associated *Campylobacter* strains are mostly found in areas with low poultry density, high ruminant density, recreational waters and WWTP discharge points.
- Ruminant-associated *Campylobacter* strains are mostly found in low ruminant density areas, agricultural waters and WWTP discharge points, mostly during the warmer seasons.

The above conclusions may have public health implications, because even if we can ensure that poultry meat is *Campylobacter*- free at the point of consumption, leading to a reduction in human campylobacteriosis cases, human exposure can also occur via environmental pathways and specifically the aquatic environment. This calls for interventions aimed at controlling environmental dissemination of *Campylobacter* at primary livestock production and WWTPs, provided that cost-benefit analyses show that the public health benefits outweigh the costs of such interventions. Conversely, virtually nothing can be done to control *Campylobacter* in wildlife. In this regard, the finding that >90% of *Campylobacter* strains from recreational waters are attributable to wild birds, and that the higher contribution of wild birds to recreational water contamination relative to other types of water is significant, implies that the risk of acquiring campylobacteriosis from, e.g., swimming in official recreational water sites in the Netherlands, is largely beyond human control.

Data statement

The livestock density data are available at Statistics Netherlands (https://www.cbs.nl/en-gb). The *Campylobacter* sequences are deposited in ENA Sequence Read Archive (project PRJEB38253).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Annemieke C. Mulder: Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Eelco Franz**: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing - review & editing. **Sharona de Rijk**: Formal analysis, Investigation, Methodology, Writing - review & editing. **Moyke A.J. Versluis**: Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Claudia Coipan**: Formal analysis, Formal analysis,

Investigation, Project administration, Software, Writing - review & editing. **Ralph Buij**: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing - review & editing. **Gerard Müskens**: Investigation, Writing - review & editing. **Miriam Koene**: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing. **Roan Pijnacker**: Formal analysis, Formal analysis, Investigation, Writing - review & editing. **Birgitta Duim**: Formal analysis, Investigation, Methodology, Writing - review & editing. **Linda van der Graaf-van Bloois**: Formal analysis, Investigation, Methodology, Writing - review & editing. **Kees Veldman**: Formal analysis, Investigation, Methodology, Writing - review & editing. **Jaap A. Wagenaar**: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing - review & editing. **Aldert L. Zomer**: Formal analysis, Investigation, Methodology, Software, Writing - review & editing. **Franciska M. Schets**: Conceptualization, Funding acquisition, Investigation, Resources, Writing - review & editing. **Hetty Blaak**: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Writing - review & editing. **Lapo Mughini-Gras**: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - original draft, Writing - review & editing.

Acknowledgements

The authors would like to thank Menno van der Voort, Greetje A.A. Castelijn, Joke van der Giessen, Marieke Opsteegh and Angela van Hoek for support in isolate collection and logistics, and Ben Bom for support in selecting the study areas. This study was financially supported by the Netherlands' Organization for Health Research and Development (ZonMw) with grant number 50-52200- 98-316 (project name: "DEPiCT - Discerning Environmental Pathways of Campylobacter Transmission") and the Dutch Ministry of Health, Welfare and Sport with grant number 9.2.09.E (project name: "Campylobacter source attribution"). The funding source had no involvement in the conduct of the research and/or preparation of the article. There are no conflicts of interest to declare for any of the authors.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of molecular biology 215 (3), 403–410. https://doi.org/ 10.1016/S0022- 2836(05)80360- 2.

Anandan, G., Thomas, A., Benickson, C., Chitra, D., Geethu, M., Augustine, J., Mithun, R., Shiva, R., Kavipriya, J., 2014. Estimation of Tree Species Diversity in Four Campuses of Roever Institutions using Simpsons Diversity Index. Jour- nal of Biodiversity & Endangered Species 2 (4), 1–3. https://doi.org/10.4172/ 2332-2543.1000135.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19 (5), 455–477. https://doi.org/10.1089/cmb.2012.0021.

CBS 2018a. Statline - Landbouw; gewassen, dieren en grondgebruik naar gemeente. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80781ned/table?ts= 1568709176121.

CBS 2018b. Statline - Regionale kerncijfers Nederland. https://opendata.cbs.nl/ statline/#/CBS/nl/dataset/70072ned/table?ts=1568709352253.

Centre of Disease Control and Prevention 2017 . Campylobacter (campylobacteriosis) for professionals. https://www.cdc.gov/campylobacter/technical.html.

Cody, A.J., Bray, J.E., Jolley, K.A., McCarthy, N.D., Maiden, M.C., 2017. Core genome multilocus sequence typing scheme for stable, comparative analyses of Campylobacter jejuni and C. coli human disease isolates.. Journal of clinical microbiology 55 (7), 2086–2097. https://doi.org/10.1128/JCM.00080-17.

Colles, F.M., Ali, J.S., Sheppard, S.K., McCarthy, N.D., Maiden, M.C., 2011. Campylobacter populations in wild and domesticated Mallard ducks (Anas platyrhynchos). Environ Microbiol Rep 3 (5), 574–580. https://doi.org/10.1111/j.1758-2229.2011. 00265.x.

Collins, D.M., Colwell, R.R., 1986. Viable but nonculturable stage of Campylobacter jejuni and its role in survival in the natural aquatic environment. American Society for Microbiology Journals 52 (3), 531–538.

de Greeff, S., Mouton, J., Schoffelen, A. and Verduin, C. 2019. NethMap 2019: Consumption of antimicrobial agents and antimicrobial resistance among medically important bacteria in the Netherlands / MARAN 2019: Monitoring of Antimicrobial Resistance and Antibiotic Usage in Animals in the Netherlands in 2018. https://doi.org/10.21945/rivm-2019-0038.

de Lange, H.J., Lammertsma, D.R., Keizer-Vlek, H., de Haan, M. 2013. De invloed van watervogels op de bacteriologishe zwemwaterkwaliteit.

Dingle, K.E., Colles, F.M., Wareing, D.R., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.J., Willems, R.J., Urwin, R., Maiden, M.C., 2001. Multilocus sequence typing system for Campylobacter jejuni. J Clin Microbiol 39 (1), 14–23. https://doi.org/10.1128/ JCM.39.1.14-23.2001.

EUR-Lex, 2006. DIRECTIVE 2006/7/EC OF THE EUROPEAN PARLIAMENT AND OF THECOUNCIL of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. Official Journal of the European Union 37–51.

European Food Safety and European Centre of Disease Prevention and Control 2018. The European Union summary report on trends and sources of zoonosis, zoonotic agents and food-borne outbreaks in 2017. https://doi.org/10.2903/j. efsa.2018.5500.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes application to human mitochondrial DNA restriction data. Genetics 131 (2), 479–491.

French, N., Barrigas, M., Brown, P., Ribiero, P., Williams, N., Leatherbarrow, H., Birtles, R., Bolton, E., Fearnhead, P., Fox, A., 2005. Spatial epidemiology and natural population structure of Campylobacter jejuni colonizing a farmland ecosystem. Environ Microbiol 7 (8), 1116–1126. https://doi.org/10.1111/j.1462-2920. 2005.00782.x.

Friesema, I.H.M., Havelaar, A.H., Westra, P.P., Wagenaar, J.A., van Pelt, W., 2012. Poultry Culling and Campylobacteriosis Reduction among Humans, the Netherlands. Emerg Infect Dis 18 (3), 466–468. https://doi.org/10.3201/eid1803.111024.

Halvorson, H.A., Schlett, C.D., Riddle, M.S., 2006. Postinfectious irritable bowel syndrome – A meta-analysis. Am J Gastroenterol 101 (8), 1942. https://doi.org/10.1111/j.1572-0241.2006.00654.x.

Hepworth, P.J., Ashelford, K.E., Hinds, J., Gould, K.A., Witney, A.A., Williams, N.J., Leatherbarrow, H., French, N.P., Birtles, R.J., Mendonca, C., Dorrell, N., Wren, B.W., Wigley, P., Hall, N., Winstanley, C., 2011. Genomic variations define divergence of water/wildlife-associated Campylobacter jejuni niche specialists from common clonal complexes. Environ Microbiol 13 (6), 1549–1560. https://doi.org/10.1111/j.1462-2920.2011.02461.x.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9 (5), 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x.

Ito, K., Kubokura, Y., Kaneko, K., Totake, Y., Ogawa, M., 1988. Occurrence of Campylobacter jejuni in free-living wild birds from Japan. J Wildl Dis 24 (3), 467–470. https://doi.org/10.7589/0090-3558-24.3.467.

Jones, K., 2001. Campylobacter in water, sewage and the environment. Applied Microbiology 90 (S6), 68S–79S. https://doi.org/10.1046/j.1365-2672.2001.01355.x.

Katoh, K., Misawa, K., Kuma, K.i, Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30 (14), 3059–3066. https://doi.org/10.1093/nar/gkf436.

KNMI 2018. Zomer 2018 (juni, juli, augustus). https://www.knmi.nl/nederland-nu/ klimatologie/maand-en-seizoensoverzichten/2018/zomer.

KNMI 2020. Archief maand/seizoen/jaaroverzichten. https://www.knmi.nl/ nederland-nu/klimatologie/maand-en-seizoensoverzichten/.

Lange, H.J.d., Lammertsma, D.R. and Keizer-Vlek, H.E. 2013. De invloed van watervogels op de bacteriologische zwemwaterkwaliteit. https://edepot.wur.nl/257045.

Mughini-Gras, L., Penny, C., Ragimbeau, C., Schets, F.M., Blaak, H., Duim, B., Wage- naar, J.A., de Boer, A., Cauchie, H.M., Mossong, J., van Pelt, W., 2016. Quantifying potential sources of surface water contamination with Campylobacter jejuni and Campylobacter coli. Water Res 101, 36–45. https://doi.org/10.1016/j.watres.2016. 05.069.

Mughini Gras, L., Smid, J.H., Wagenaar, J.A., de Boer, A.G., Havelaar, A.H., Friesema, I.H., French, N.P., Busani, L., van Pelt, W., 2012. Risk factors for campylobacteriosis of chicken, ruminant, and environmental origin: a combined case-control and source attribution analysis. PLoS One 7 (8), e42599. https://doi.org/ 10.1371/journal.pone.0042599.

Murphy, C., Carroll, C., Jordan, K.N., 2006. Environmental survival mechanisms of the foodborne pathogen Campylobacter jejuni. Applied Microbiology 100 (4), 623– 632. https://doi.org/10.1111/j.1365-2672.2006.02903.x.

Nachamkin, I., Allos, B.M., Ho, T., 1998. Campylobacter Species and Guillan-Barré Syndrome. Clin Microbiol 11 (3), 555–567. https://doi.org/10.1128/CMR.11.3.555.

Opsteegh, M., van Roon, A., Wit, B., Hagen-Lenselink, R., van Duijkeren, E., Dierikx, C., Hengeveld, P., Franz, E., Bouw, E. and van der Meij, A. 2018. Surveillance zoönosen in de melkgeiten-en melkschapenhouderij in 2016.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research 25 (7), 1043–1055. https://doi.org/10.1101/gr.186072.114.

Pijnacker, R., Friesema, I., Mughini Gras, L., Lagerweij, G., van Pelt, W. and Franz, E. 2019. Disease burden of food-related pathogens in the Netherlands, 2018. https://doi.org/10.21945/RIVM-2019-0086.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A., Lareu, M.V., 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Front Genet 4, 98. doi:10.3389/fgene.2013.00098.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155 (2), 945–959.

Pritchard, J.K., Wen, X., Falush, D., 2009. Documentation for structure software: Version 2.3. Thermo Fisher Scientific, Waltham.

RCT, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org.

Rechenburg, A., Kistemann, T., 2009. Sewage effluent as a source of campylobacter sp. in a surface water catchment.. Int J Environ Health Res 19 (4), 239–249. https://doi.org/10.1080/09603120802460376.

Rosef, O., Rettedal, G., Lågeide, L., 2001. Thermophilic campylobacters in surface water: a potential risk of campylobacteriosis. International journal of environmental health research 11 (4), 321–327. https://doi.org/10.1080/09603120120081791.

Sears, A., Baker, M.G., Wilson, N., Marshall, J., Muellner, P., Cambell, D.M., Lake, R.J., French, N.P., 2011. Marked Campylobacteriosis decline after interventions aimed at poultry. New Zealand. Emerg Infect Dis 17 (6), 1007–1015. https://doi.org/10. 3201/eid1706.101272.

Shrestha, R.D., Midwinter, A.C., Marshall, J.C., Collins-Emerson, J.M., Pleydell, E.J., French, N.P. 2019. Campylobacter jejuni Associated with Wild Birds and those Causing Human Disease in Six High-Use Recreational Waterways in New Zealand. https://doi.org/10.1128/AEM.01228-19.

Smit, L.A., Heederik, D., 2017. Impacts of intensive livestock production on human health in densely populated regions. GeoHealth 1 (7), 272–277. https://doi.org/ 10.1002/2017GH000103.

Sopwith, W., Birtles, A., Matthews, M., Fox, A., Gee, S., Painter, M., Regan, M., Syed, Q., Bolton, E., 2008. Identification of potential environmentally adapted Campylobacter jejuni strain, United Kingdom. Emerg Infect Dis 14 (11), 1769– 1773. https://doi.org/10.3201/eid1411.071678.

Stern, N.J., Hiett, K.L., Alfredsson, G.A., Kristinsson, K.G., Reiersen, J., Hardardottir, H., Briem, H., Gunnarsson, E., Georgsson, F., Lowman, R., Berndtson, E., Lammerding, A.M., Paoli, G.M., Musgrove, M.T., 2003. Campylobacter spp. in Icelandic poultry operations and human disease. Epidemiol Infect 130 (1), 23–32. https://doi.org/10.1017/s0950268802007914.

Ternhag, A., Torner, A., Svensson, A., Ekdahl, K., Giesecke, J., 2008. Short- and longterm effects of bacterial gastrointestinal infections. Emerg Infect Dis 14 (1), 143– 148. https://doi.org/10.3201/eid1401.070524.

Van Den Pol-Van Dasselaar, A., Blokland, P., Gies, T., Holshof, G., De Haan, M., Naeff, H. and Philipsen, A. 2015. Beweidbare oppervlakte en weidegang op melkveebedrijven in Nederland. Wageningen UR (University & Research centre) Livestock Research, page Livestock Research Rapport 917.

Veilinga, A., van loock, F., 2002. The dioxin crisis as experiment to determine poultry-related campylobacter enteritis. Emerg Infect Dis 8 (1), 19–22. https://doi.org/10.3201/eid0801.010129.

Wagenaar, J.A., French, N., Havelaar, A.H., 2013. Preventing Campylobacter at the source: Why is it so difficult? Clin Infect Dis 57 (11), 1600–1602. https://doi. org/10.1093/cid/cit555.

Wagenaar, J.A., Newell, D.G., Kalupahana, R.S. and Mughini-Gras, L. 2015. Campylobacter: animal reservoirs, human infections, and options for control. Zoonoses: Infections affecting humans and animals – A focus on public health aspects., 159-177, https://doi.org/10.1007/978-94-017-9457- 2_6.

Waldenström, J., Broman, T., Carlsson, I., Hasselquist, D., Achterberg, R.P., Wagenaar, J.A., Olsen, B., 2002. Prevalence of Campylobacter jejuni, Campylobacter lari, and Campylobacter coli in different ecological guilds and taxa of migrating birds. Appl Environ Microbiol 68 (12), 5911–5917. https://doi.org/10.1128/aem.68.12.5911-5917.2002.

Whelan, C.D., Monaghan, P., Girdwood, R.W., Fricker, C.R., 1988. The significance of wild birds (Larus sp.) in the epidemiology of Campylobacter infections in humans. Epidemiol Infect 101 (2), 259–267. https://doi.org/10.1017/ s0950268800054170.

Wilson, D.J., Gabriel, E., Leatherbarrow, A.J., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Fearnhead, P., Hart, C.A., Diggle, P.J., 2008. Tracing the source of campylobacteriosis. PLoS Genet 4 (9), e1000203. https://doi.org/10.1371/journal.pgen. 1000203.

Supplementary material

Table S1. Total number of swabs taken per method and number of *Campylobacter jejuni* and *C. coli* isolates isolated per wild bird species.

	Swabs		**Species**		
Wild bird species	Method	Total	*C. jejuni*	*C. coli*	Total
Great cormorant (*Phalacrocorax carbo*)	Dropping	59	8	0	8
Greylag goose (*Anser anser*)	Dropping	50	8	0	8
Common wood pigeon (*Columba palumbus*)	Dropping	43	10	0	10
Lesser black-backed gull (*Larus fuscus*)	Cloaca (37)/Dropping (10)	47	3	1	4
Mute swan (*Cygnus olor*)	Cloaca (47)/Dropping (30)	77	2	9	11
European herring gull (*Larus argentatus*)	Cloaca	9	0	0	0
Eurasian wigeon (*Mareca penelope*)	Cloaca	67	13	5	18
Mallard (*Anas platyrhynchos*)	Dropping (16)/Intestinal contents (2)	18	2	0	2
Total		370	46	15	61

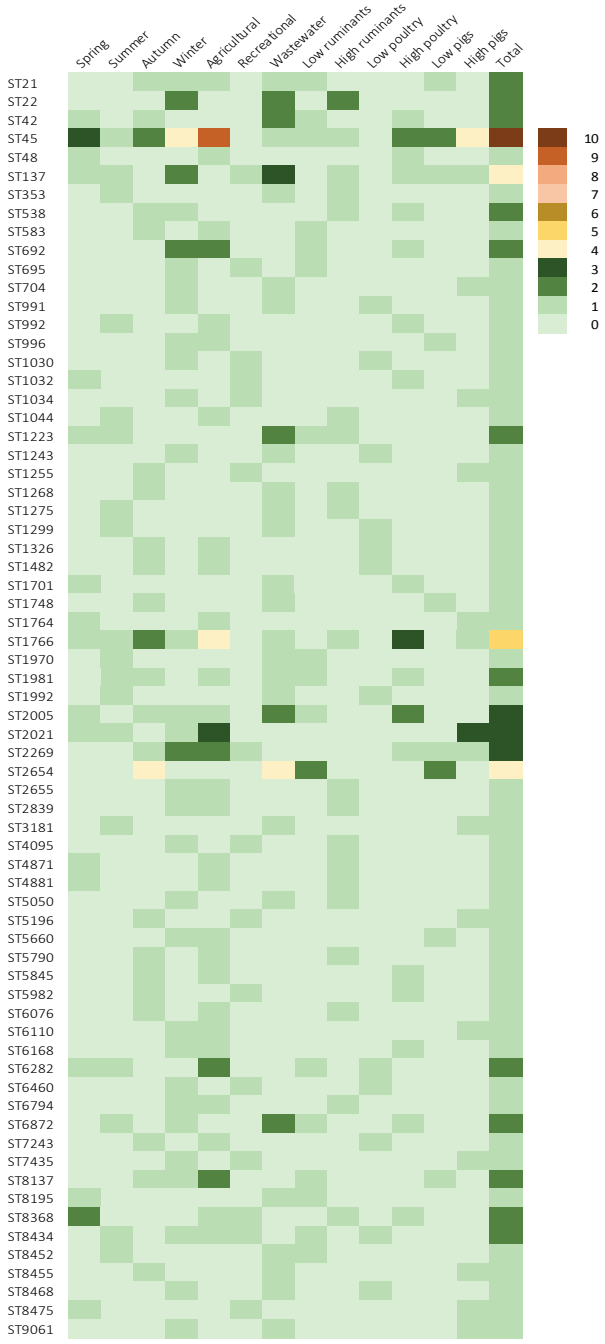


Figure S1. Heat density map showing the variety of sequence types found in the water isolates per category.

Table S2. Genetic heterogeneity of *Campylobacter* isolates between source populations. For each pair of sources, Φ values are displayed above the diagonal and the associated p-values below the diagonal.

p/fst	Broiler	Layer	Turkey	Calf	Meat cattle	Dairy cattle	Swine	Small Ruminants	Birds
Broiler	-	1.17	0.66	1.95	2.06	1.28	13.48	1.58	3.82
Layer	0.034	-	1.70	5.09	3.13	3.44	15.91	2.13	5.42
Turkey	0.381	0.021	-	4.00	1.74	1.96	15.06	1.41	6.11
Calf	0.002	0.001	0.001	-	2.72	1.86	11.50	1.53	7.18
Meat cattle	0.003	0.002	0.022	0.002	-	0.87	17.68	1.21	6.49
Dairy cattle	0.017	0.001	0.006	0.004	0.132	-	16.96	1.45	7.19
Swine	0.001	0.001	0.001	0.001	0.001	0.001	-	11.81	13.99
Small ruminants	0.012	0.006	0.079	0.014	0.068	0.025	0.001	-	4.89
Birds	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	-

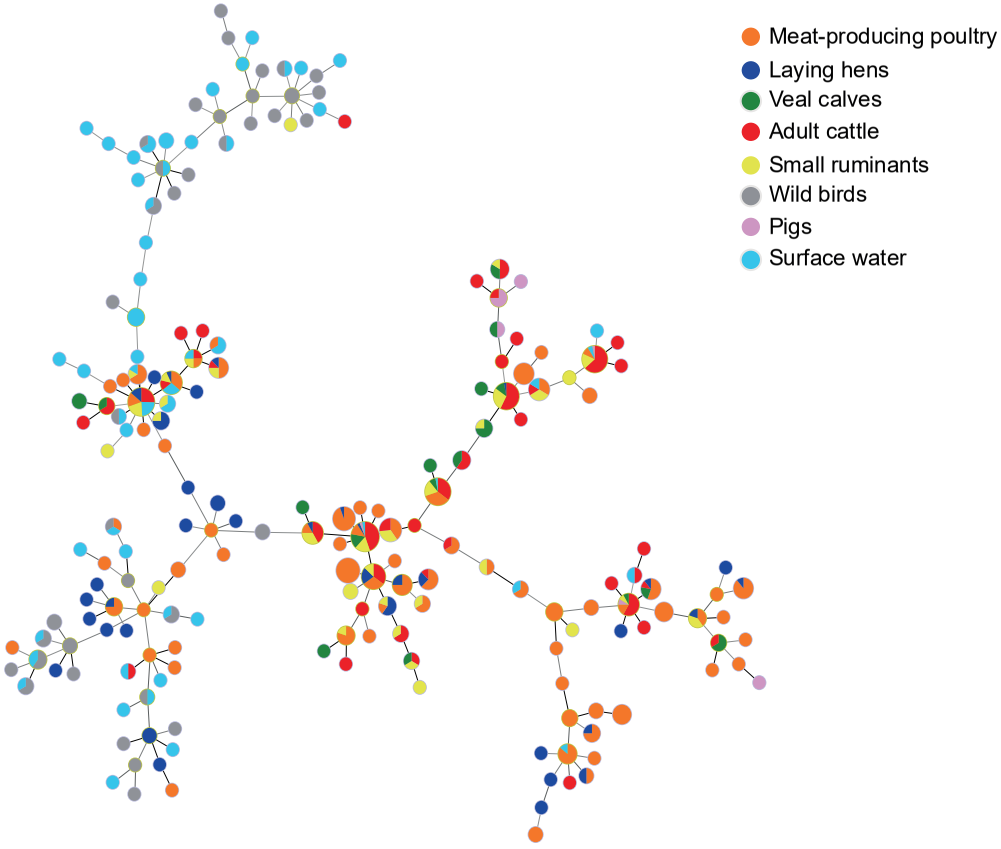


Figure S2. Conventional MLST-based minimum spanning tree showing the population structure of the *C. jejuni* isolates from surface water and from the different animal sources.

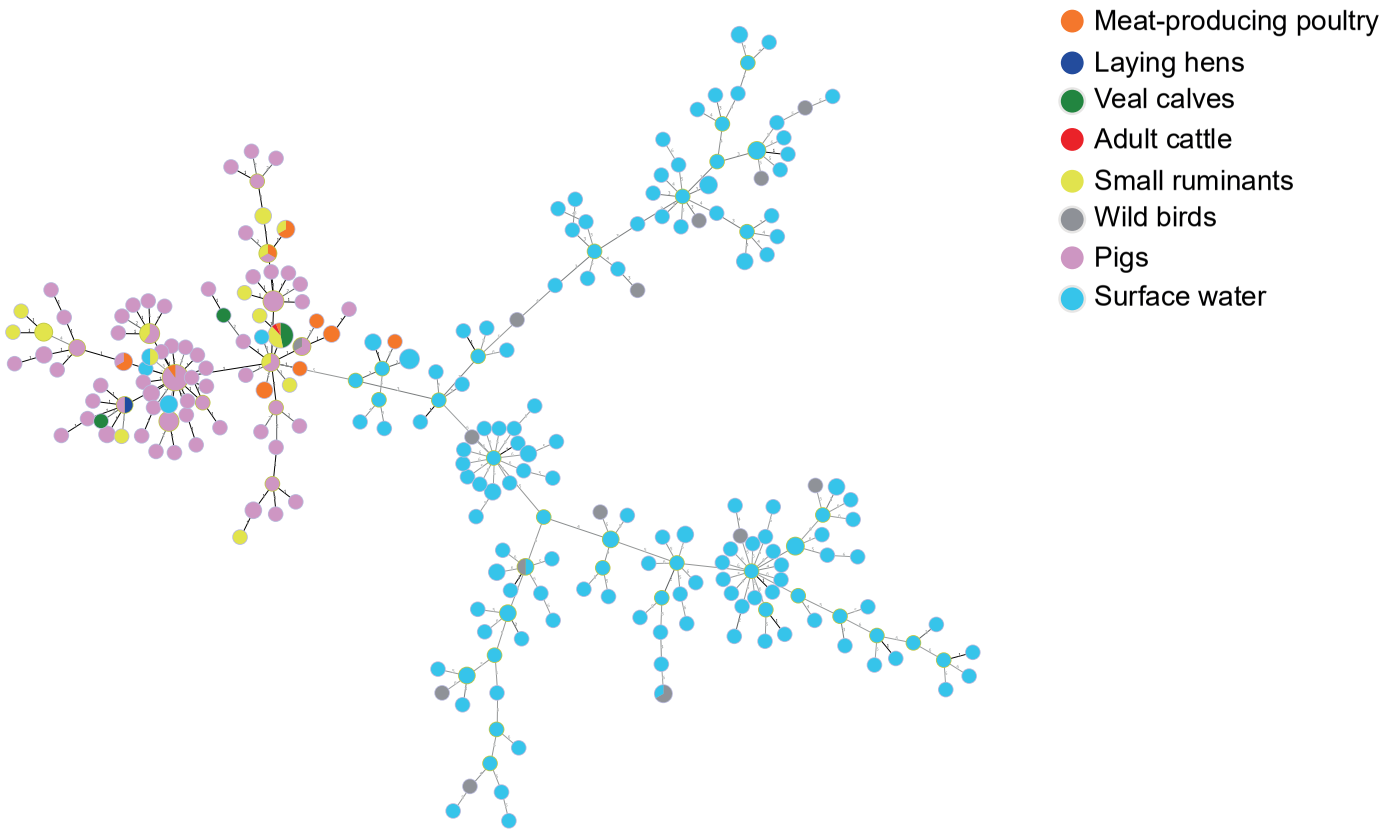


Figure S3. Conventional MLST-based minimum spanning tree showing the population structure of the *C. coli* isolates from surface water and from the different animal sources.

CHAPTER 4

Spatial effects of livestock farming on human infections with Shiga toxin-producing *Escherichia coli* O157 in small but densely populated regions: the case of the Netherlands

Annemieke C. Mulder
Jan van de Kassteele
Dick Heederik
Roan Pijnacker
Lapo Mughini-Gras
Eelco Franz

Published in Geohealth, 2020

Abstract

The role of environmental transmission of typically foodborne pathogens like Shiga toxin-producing *Escherichia coli* (STEC) O157 is increasingly recognized. To gain more insights into spatially restricted risk factors that play a role in this transmission, we assessed the spatial association between sporadic STEC O157 human infections and the exposure to livestock (i.e. small ruminants, cattle, poultry, and pigs) in a densely populated country: the Netherlands. This was done for the years 2007–2016, using a state-of-the-art spatial analysis method in which hexagonal areas with different sizes (90, 50, 25 and 10 km^2) were used in combination with a novel probability of exposure metric: the population-weighted number of animals per hexagon. To identify risk factors for STEC O157 infections and their population attributable fraction (PAF), a spatial regression model was fitted using integrated nested Laplace approximation (INLA). Living in hexagonal areas of 25, 50 and 90 km^2 with twice as much population-weighted small ruminants was associated with an increase of the incidence rate of human STEC O157 infections in summer (RR of 1.09 [95%CI;1.01–1.17], RR of 1.17 [95%CI;1.07–1.28] and RR of 1.13 [95%CI;1.01–1.26]), with a PAF of 49% (95%CI;8–72%). Results suggest exposure to small ruminants to be a risk factor, although no evidence on the mode of transmission is provided. Therefore, the underlying mechanisms warrant further investigation and could offer new targets for control. The newly proposed exposure metric has potential to improve existing spatial modeling studies on infectious diseases related to livestock exposure, especially in densely populated countries like the Netherlands.

Introduction

Food is generally considered to be the most important route of transmission for Shiga toxin-producing *Escherichia coli* (STEC) O157 (Strachan et al., 2001). However, a growing body of evidence suggests that non-foodborne transmission pathways, such as those mediated by the environment, may be important as well (Berry et al., 2015; Elson et al., 2018; Franz et al., 2018; Friesema et al., 2011; ÓHaiseadha et al., 2017; Strachan et al., 2006). A recent source attribution modeling study based on STEC serotyping data revealed that domestic ruminants (cattle, sheep and goats) are important sources of human STEC O157 infections, accounting for approximately three-quarters of reported human STEC infections in the Netherlands (Mughini-Gras et al., 2018). This emphasizes the need for both direct and indirect exposure to different types of livestock to be considered as possible transmission routes for STEC O157.

STEC is a bacterial zoonotic agent associated with human disease with varying clinical manifestations, including diarrhea, haemorrhagic colitis and (occasionally fatal) haemolytic uremic syndrome (HUS), a leading cause of acute renal failure among children (Elson et al., 2018; Franz et al., 2018; Mughini-Gras et al., 2018). Human STEC infections is the third most commonly reported zoonosis in the European Union (EU), with an annual number of laboratory-confirmed STEC infections increasing from 5,901 in 2014 to 8,161 in 2018 (European Food Safety Authority & European Centre for Disease Prevention and Control, 2019). This, in combination with its high virulence and outbreak potential, makes STEC of significant public health concern. Although there are more than a hundred STEC serotypes and their importance is increasingly recognized, STEC O157 is the most important serotype in terms of incidence and clinical significance (Mughini-Gras et al., 2018). In the Netherlands, STEC is a notifiable disease, with an annual incidence between 2 and 7 cases per 100,000 inhabitants (European Centre for Disease Prevention and Control, 2019). The vast majority of cases in the Netherlands are considered sporadic, as outbreaks rarely occur (Franz et al., 2018).

Potential sources of human STEC infection are mainly animals capable of maintaining STEC colonization in absence of continuous exposure to STEC from other sources (i.e. the so-called reservoirs or amplifying hosts, mainly cattle and sheep). But also those that are frequently exposed to STEC from the environment, like birds and other wild animals (Mughini-Gras et al., 2018; Strachan et al., 2006). According to a recent source attribution study, cattle is the primary source of human STEC O157 infection in the Netherlands, followed by small ruminants (sheep and goats) (Mughini-Gras et al., 2018). These animals can shed high quantities (>105/g) of STEC O157, that subsequently are able to survive for extended periods of time (Chase-Topping et al., 2008; Franz et al., 2014; Strachan et al., 2001). This implies that

there is a significant risk of STEC O157 infection linked to environment-mediated transmission to humans (Elson et al., 2018; Strachan et al., 2001).

The Netherlands is one of the world's most densely populated countries, with over 500 inhabitants per km^2 and a remarkably high concentration of intensive livestock farms as well. The presence of livestock in close proximity to residential areas has arisen questions about the associated public health implications (Smit & Heederik, 2017). Since STEC O157 can potentially be contracted from the soil and water environment, and may be spread through the air after periods of drought in the vicinity of its animal reservoirs, it is conceivable that human STEC O157 incidence in the Netherlands might be higher in areas with increased livestock density as well, such as in rural vs. urban areas as shown elsewhere (Berry et al., 2015; Strachan et al., 2006, 2001). This could be tested with different methods, of which examples are: (i) spatial regression analysis to determine the probability of exposure (Elson et al., 2018; Friesema et al., 2011; ÓHaiseadha et al., 2017) or (ii) classical case–control studies including relevant spatial variables to determine the importance of particular types of exposure (e.g. number of animals/km^2) (de Rooij et al., 2019).

As a spatial regression analyses requires less resources, in terms of data needs and financial support, it can be a preferred way of exploring new ideas. However, only a few studies exist that focus on the spatial association between human STEC O157 infections and the probability of exposure to livestock by means of spatial regression analysis (Elson et al., 2018; Friesema et al., 2011; ÓHaiseadha et al., 2017). Most of those studies only include one domestic ruminant species (cattle or sheep or goat) in the analysis (Friesema et al., 2011; ÓHaiseadha et al., 2017), while ignoring other reservoirs that may affect the outcome of those studies. This is especially important in countries like the Netherlands where high numbers of different types of livestock are present on relatively small geographical scales (Smit & Heederik, 2017). Moreover, the probability of exposure in those studies is strictly defined by the number of animals in a given area, while the probability of exposure on a population level is not only determined by the number of animals in a certain area, but also by the number of residents living in that area (Elson et al., 2018; Friesema et al., 2011; Hallisey et al., 2017; Mulder et al., 2016; ÓHaiseadha et al., 2017).

Therefore, the aim of this study was to assess the spatial association between sporadic human STEC O157 infections and the combined exposures to livestock (cattle, goat, sheep, poultry and pigs) in the Netherlands, using different state-of-the-art methods that include population-weighted numbers of animals in the calculation of the probability of exposure to livestock.

Materials and Methods

This study consisted of several parts. First, national surveillance data on notified STEC O157 cases in the Netherlands' general population was gathered together with livestock data (exact locations of registered farms and number of animals therein, per species). Subsequently, the data were transposed into a study-defined spatial division of the Netherlands and we developed a metric for the probability of exposure of the human population to each livestock species that not only includes the number of animals in a certain area, but also the corresponding population number. The last steps involved the spatial regression analysis and calculation of the population attributable fraction (PAF). We used the statistical software environment R (version 3.6.0) (RCT, 2015) and several R packages and functions for data processing and analysis (Arya et al., 2015; Bates et al., 2019; Bivand et al., 2019; De Jonge & Houweling, 2019; Grolemund & Wickham, 2011; Keitt, 2010; Neuwirth, 2015; Pebesma, 2019; Pebesma, Bivand, Racine, et al., 2019; Pebesma, Bivand, Rowlingson, et al., 2019; R-Core, 2017; Rue, 2019; Wickham, 2019; Wickham, Averick, et al., 2019; Wickham, Bryan, et al., 2019; Wickham, Francois, et al., 2019; Wickham, Henry, et al., 2019). An overview is provided in supplementary material, Table S1. The used R scripts can be found at: https:// github.com/mulderac91/R-STECO157-spatialanalysis

Hexagonal Grid and Population-Weighted Interpolation

Hexagons are more suitable than rectangular grids in particular applications of ecological modeling, e.g. connectivity and movement paths (Birch et al., 2007). They have the advantage that the nearest neighborhood in a hexagonal grid is simpler and less ambiguous, because each hexagon has exactly six adjacent hexagons which are in a symmetrically equivalent position. Therefore, there is no need for a setting for the relative weighting of diagonal interactions in a nearest neighborhood analysis, as is the case for rectangular grids (Birch et al., 2000; Birch et al., 2007). Furthermore, the grid is fixed over time (Birch et al., 2007). The latter is a solution for the problem of change of, in this case postal code boundaries over time (supplementary material, Figure S1). Therefore, the Netherlands was divided in a fixed hexagonal grid (Figure 1a). To assess consistency of results and reduce the risk of ecological fallacy, we performed the analyses for hexagonal areas with four different sizes: 10 km^2 (approximately the average area of a four-digit postal code region in the Netherlands), 25 km^2, 50 km^2 and 90 km^2 (approximately the average area of a municipality in the Netherlands) (Shafran-Nathan et al., 2017).

In order to perform the spatial regression analyses on the hexagonal grid, the spatial data needed to be transformed from one regional division to the other (Arsenault et al., 2013). For this purpose, we used population-weighted interpolation. This

approach has the advantage over areal weighted interpolation that it can more accurately estimate the population demographics in transforming small counts by four-digit postal code regions to aggregated counts for large, non-standard study zones (hexagons) (Hallisey et al., 2017). A detailed explanation of this approach can be found in the supplementary material, Text S1.

Population-Weighted Number of Animals

Existing studies have used animals/km^2 to derive the probability of exposure to be able to study the association between STEC O157 infections and livestock densities (Figure 1b) (Elson et al., 2018; Friesema et al., 2011; ÓHaiseadha et al., 2017). Yet, the probability of exposure is not only determined by the number of animals in a certain area, but also by the number and residential addresses of people living in that area and the number of animals in the neighboring areas. For this purpose, we created a new probability of exposure metric: the population-weighted number of animals (Hallisey et al., 2017) (Figure 1c, 1d and 1e).

The metric is constructed as follows. When zooming into one hexagon within the hexagonal grid, the locations of several six-digit postal code points are shown (Figure 1c). Those six-digit postal code points include information about the population numbers at that specific location (Figure 1c). Around these point locations, buffers with a radius of 1 km are constructed (Figure 1d). Farms located within these buffers, also outside the specified hexagon, are included (Figure 1e). The point locations of the farms contain information about the number of animals (Figure 1e). See Figures 1c, 1d and 1e as an example. Within the hexagon, we have five six-digit postal code point locations, each with its own population numbers: 100, 1,000, 10, 5 and 1. We have three farms, each with its own number of animals: A, B and C. The 100 and the 10 individuals on the first and second six-digit postal code point locations are exposed to A animals. The 1,000 individuals in the third six-digit postal code point location are not exposed. The 5 individuals in the fourth six-digit postal code point location are exposed to B animals. The only individual in the fifth six-digit postal code point location is exposed to C animals, but from a farm outside the hexagon. The total exposure in this hexagon is then the population-weighted sum of the number of animals, which can be calculated as follows:

$$Population\ weighted\ animal\ number = \frac{(100 \times A + 10 \times A + 1{,}000 \times 0 + 5 \times B + 1 \times C)}{(100 + 10 + 1{,}000 + 5 + 1)}$$

This was done for each hexagon and for each year, taking into account the number of animals and the changing population numbers. In the end, the data were aggregated over the years, resulting in one hexagon-specific exposure metric.

Spatial Risk Factor Analysis

A Poisson regression model with log-link function was used to assess the associations between human STEC O157 infections and the population-weighted number of animals for cattle, pigs, poultry, and small ruminants (goats and sheep). As the dependent variable in the model was the case count, i.e. the number of human STEC O157 cases (redistributed with the population weighted interpolation technique) within a hexagon, the assumption was that those case counts followed a Poisson distribution. Person-years were used as the offset of the model (the population denominator for each hexagon), and the confounders included were: age category (0–4, 5–9, 10–49 and ≥50 years old), gender (male or female), and period of infection (spring/summer: May–October, autumn/winter: November–April). The different population-weighted number of animals were included as covariates in the model (Friesema et al., 2011).

Because the population-weighted number of animals x could be zero, we applied a $log_2(x+1)$ transformation. Furthermore, several studies have shown a higher risk for human STEC O157 infection in summer (Friesema et al., 2011). Therefore, we performed a stratified analysis based on the period of infection. These variables entered the model as the fixed effect terms. To be able to perform those analyses, it was assumed that residents acquired the infection at or in close proximity to their homes.

It is possible that there is additional variation due to unknown spatially varying risk factors. To account for this, two random-effect terms were added to the model. The first random-effect accounted for the spatially structured variation. This variation represented the possible effect of a common unobserved risk factor that led to neighboring hexagons being more alike. This term was modeled by the intrinsic Conditional Autoregressive Model (CAR) (Besag et al., 1991). The second random-effect term represented the unstructured variation, which was used to correct for possible overdispersion of the data. This variation consisted of possible unobserved variation within hexagons, which was modeled by independent and identically distributed (IID) Gaussian noise (Lawson, 2013).

The spatial regression model was fitted using the integrated nested Laplace approximation technique (INLA) (Rue et al., 2019). For further details we refer to Friesema et al. (2011). Rate ratios (RRs) were calculated from the coefficients of the fixed effects. As the population-weighted animal numbers were transformed, the interpretation of those RRs is as follows: if x increases with a factor two, then the incidence rate increases with a factor $RR = e^{\beta_1}$, provided that x is large enough, approximately >100. When x is smaller, this factor is less than two for the same RR, but the significance stays the same. Supplementary material Text S2 and Figure S2 show a more detailed explanation of this interpretation.

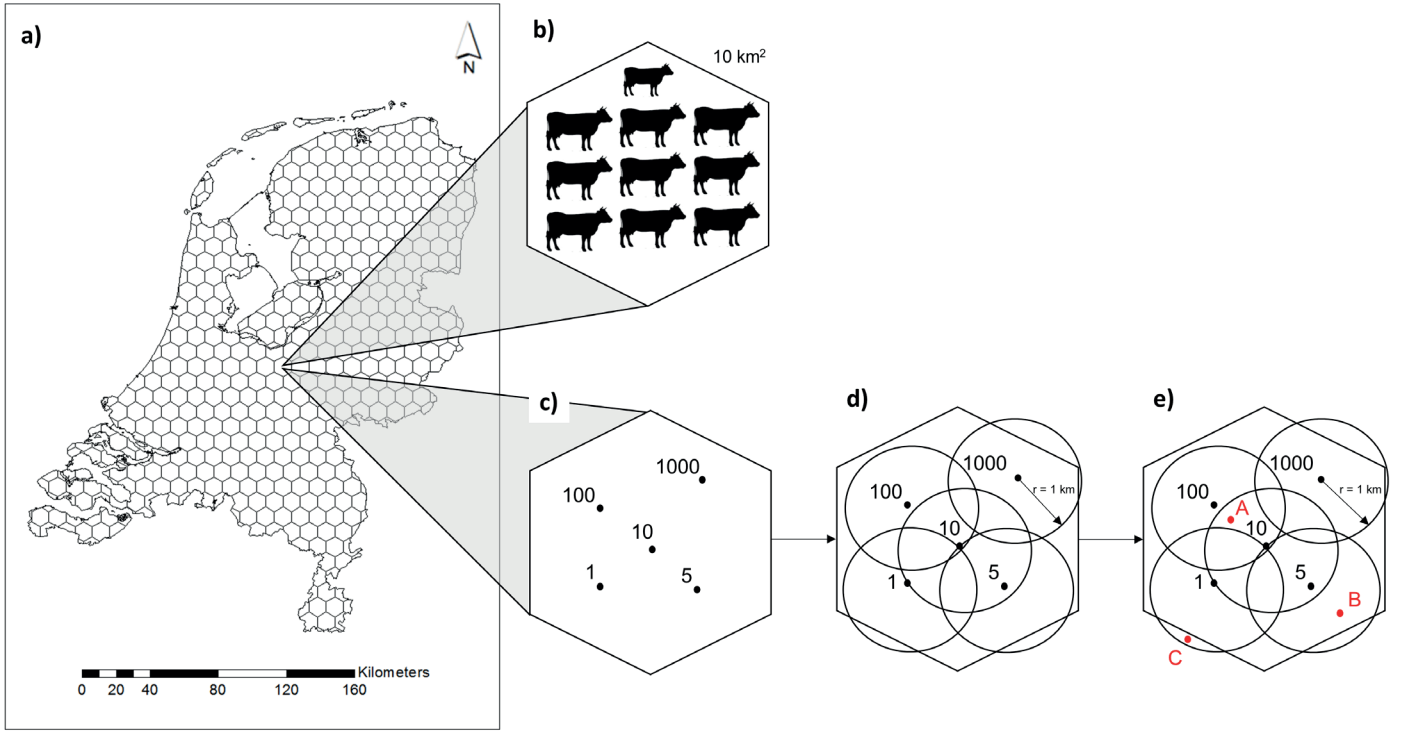


Figure 1. Explanation of the calculation of the old and the new probability of exposure measures. **a)** Hexagonal grid of the Netherlands. **b)** the old probability of exposure measure in a hexagonal grid cell: The number of animals per km². In this figure: 10 cows per 10 km², thus 1 cow/km². Pictures **c**, **d** and **e** visualize the calculation of the new probability of exposure: the population-weighted animal number. **c)** a hexagonal grid cell, including the six-digit postal code point locations within this cell and their corresponding population numbers. **d)** the buffers with a radius *r* of 1 km surrounding the six-digit postal code point locations within the hexagonal grid cell. **e)** the hexagonal grid cell including all the information of Figure 1d. Here, the point locations of the farms of a certain type of animal are added, which also include information about the specific number of animals. This gives the information that is needed to know which six-digit postal code points (and thus which population numbers) are influenced by which farm(s) and the corresponding animal numbers. With this information and the formula given in Section 2.2, the new probability of exposure can be calculated and aggregated per hexagon.

In addition, the population attributable fraction (PAF) and its 95% confidence interval were calculated for the risk factors found (Supplementary material, Text S3). Confidence intervals were obtained by Monte Carlo simulation, using the INLA posterior sampling function with 10,000 samples.

Data

Case Data

Since 1999, it is obligatory for diagnostic laboratories in the Netherlands to notify confirmed human STEC infections to the Municipal Health Services (MHSs) (Friesema et al., 2011). The MHSs reports each laboratory-confirmed case to the national surveillance database at the Dutch National Institute for Public Health and the Environment (RIVM) (Friesema et al., 2011). Furthermore, laboratories are asked (but not obliged) to send STEC isolates to the RIVM for confirmation and further typing for national surveillance purposes (Friesema et al., 2011, 2017).

In this study, a case was defined as an individual with confirmed STEC O157 infection (by the RIVM) during the period 2007–2016. Cases were excluded when they were part of an (inter)national foodborne outbreak, traveled abroad in the week before onset of illness, or when the residential address (postal code) was unknown. A detailed explanation of the different spatial scales (province, municipality and postal code) of the Netherlands and a comparison with the European NUTS classification system is given in the supplementary material Text S4 and Figure S3 (European Commission - Eurostat, 2019). Those data are protected by Dutch privacy regulations and the Dutch Data Protection Authority (Dutch Data Protection Authority, 2020a, 2020b).

Livestock Data

Livestock data of food-producing animals for 2012 was obtained from the Department of Service Arrangements of the Dutch Ministry of Agriculture, Nature and Food Quality. These data are collected yearly, requesting all food-producing farmers to report the number of animals reared (CBS, 2019b; RVO, 2019). In our study, we used the total number of goats, sheep, cattle, poultry and pigs per farm (Table 1). To derive the total number of small ruminants, the total number of goats and sheep per farm were summed together.

Table 1. Total number of food-producing animals, total number of farms and the mean number of animals per farm per type of food-producing animal (goat, sheep, cattle, poultry, pigs) in the Netherlands

Type of animal	Total number of animals	Total number of farms	Mean number of animals per farm
	N	N	N
Goats	398,508	3,954	101
Sheep	1,049,517	13,962	75
Cattle	3,895,657	33,908	115
Poultry	96,802,429	2,889	33,507
Pigs	12,138,896	6,961	1,744

Population Data

The population data per four-digit postal code region per year is available through Statistics Netherlands (www.statline.nl) and consists of the number of inhabitants in five-year age categories and gender. The data were downloaded from this website for the years 2007–2016 (CBS, 2019a). Due to privacy regulations (Dutch Data Protection Authority, 2020b), this information was not available per six-digit postal code point location.

Spatial Data

The four-digit postal code region shapefiles of the Netherlands were obtained for each year (2007–2016) from the geodata portal of the RIVM. For the period 2007–2008, there were no postal code region shapefiles available. Therefore, the shapefile of 2009 was used for those years. The six-digit postal code point location shapefile of the Netherlands from 2016 was also obtained from the geodata portal of the RIVM. This file included population numbers per six-digit postal code point location.

Results

Descriptive Statistics

Between 2007 and 2016, 599 cases of STEC O157 infection were reported. In this period, two national outbreaks of STEC O157 were registered in the Netherlands, one in 2007 involving 41 cases probably caused by lettuce consumption and linked to an outbreak in Iceland (Friesema et al., 2008) and one in 2009 involving 20 cases caused by contaminated raw meat spread (Greenland et al., 2009). Furthermore, there was a regional outbreak in 2007 involving 7 cases, which reported consumption of raw meat spread and all had bought it at the same regional supermarket chain (Friesema et al., 2011). The cases that were involved in those outbreaks were excluded from the

dataset for analysis. Besides, 54 more cases were excluded because information on travel history prior to symptom onset was missing, and 38 cases because there was no data available on geographical location. The remaining 439 cases were included in the analysis, with a median number of 46 cases per year (range 25–63 cases/year, annual incidence 1.5–3.8/100,000 inhabitants).

Table 2. Descriptive Statistics of the STEC O157 Cases.

	STEC O157 cases	
	N	%
Total	439	100
Gender		
Males	167	38
Females	272	62
Age category (years)		
0-4	70	16
5-9	44	10
10-49	200	46
≥ 50	125	28
Period of infection		
Summer	340	77
Winter	99	23

Of all the cases included, 62% (*n* = 272) were female, 38% (*n* = 167) were male (Table 2). The highest number of cases (46%) were between 10 and 49 years of age and most were reported in summer (77%). Figure 2 shows that the incidence varies between hexagons and appears to be highest in the northern and eastern regions of the Netherlands. The west and south of the Netherlands show particularly low incidence of STEC O157.

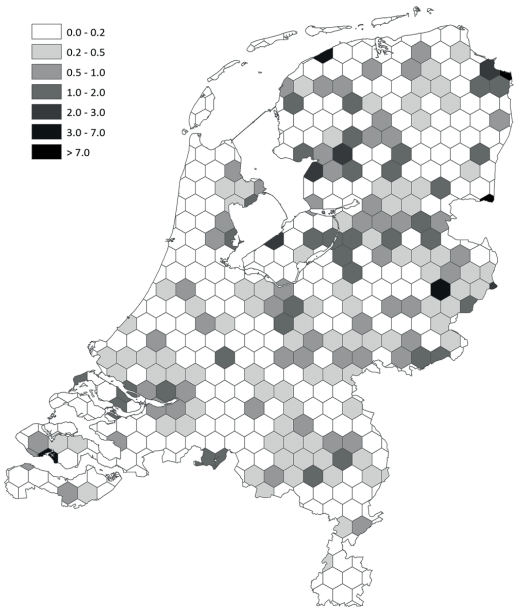


Figure 2. Cumulative incidence rate (x 100,000 person-years) (2007–2016) of STEC O157 infections in the Netherlands.

Figure 3 shows the population-weighted number of small ruminants, cattle, poultry and pigs in the Netherlands. The population-weighted number of small ruminants appeared to be highest in the central north of the country, central south of the country and the island of Texel. For cattle, it was highest in the center, central north and central south of the country and for poultry it was highest in the center, east and south-east (except the region of South-Limburg). Furthermore, the population-weighted number of pigs was highest in the east and south-east (except the region south-Limburg). Visually, the map for small ruminants in Figure 3a seemed to be most comparable with the one for human STEC O157 infections in Figure 2.

Spatial Risk Factor Analysis

Results from the multivariable models for the spatial association between STEC O157 and population-weighted number of animals are presented in Table 3 and Table 4, respectively. For the results of the univariable models, see supplementary material, Table S2.

Living in an hexagonal area of 90 km^2 with twice as much population-weighted small ruminants increased the incidence rate of reporting STEC O157 infection in summer, with a RR of 1.13 (95% CI 1.01–1.26) (Table 3). Other hexagonal areas

have comparable results, except the one of 10 km^2. Here, small ruminants were not significantly associated with STEC O157 infections. To further explore this, the analyses at this spatial scale was repeated with goats and sheep separately. The results showed that goats are still significant in summer, with a RR of 1.07 (95% CI 1.01–1.3), while sheep no longer pose a risk. In both analyses, pigs are marginally associated with STEC O157 infections, with similar RRs. As other studies showed a clear association with cattle density per municipality in summer, the analyses were repeated with only cattle for hexagonal areas of 90 km^2. Here, the population-weighted number of cattle only had a marginal significant association with human STEC O157, with a RR of 1.08 (95% CI 1.00–1.17). In winter, none of the animal types were associated with STEC O157 infections (Table 4). Poultry was never associated with STEC O157 infection. As the population-weighted number of small ruminants in an area was the only consistent significant risk factor for different spatial scales within this study, the PAF was calculated for this factor only. The population-weighted number of small ruminants had a PAF of 49% (95% CI of 8%–72%).

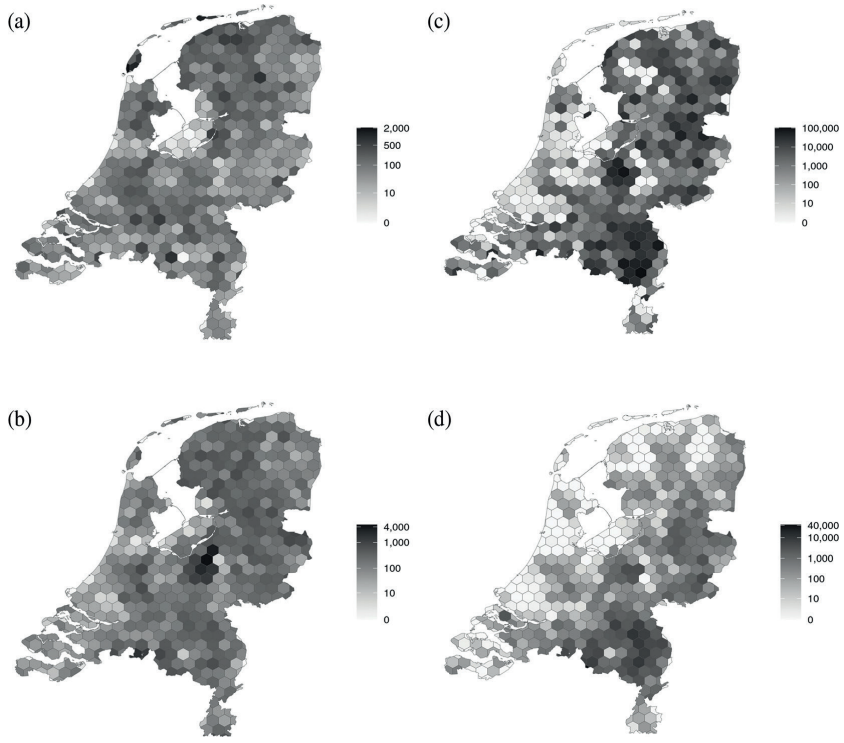


Figure 3. Maps of the population-weighted number of animals in the Netherlands per hexagon (90 km^2) for small ruminants (a), cattle (b), poultry (c) and pigs (d) in 2012.

Table 3. Results of the Multivariable Spatial Analyses for Summer with Different Hexagonal Areas (90, 50, 25 and 10 km²).

Period of infection	Variable	Hexagon 90 km²			Hexagon 50 km²			Hexagon 25 km²			Hexagon 10 km²		
		P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI
Summer	**Gender**												
	Males	*Reference category*											
	Females	**<0.001**	**1.74**	**1.40-2.17**	**<0.001**	**1.74**	**1.40-2.18**	**<0.001**	**1.74**	**1.40-2.17**	**<0.001**	**1.74**	**1.40-2.17**
	Age category (years)												
	0-4	**<0.001**	**4.05**	**2.91-5.59**	**<0.001**	**4.06**	**2.91-5.61**	**<0.001**	**4.06**	**2.91-5.60**	**<0.001**	**4.06**	**2.91-5.60**
	5-9	**<0.001**	**2.01**	**1.32-2.97**	**<0.001**	**2.01**	**1.32-2.98**	**<0.001**	**2.01**	**1.32-2.98**	**<0.001**	**2.01**	**1.32-2.97**
	10-49	0.27	1.15	0.90-1.49	0.27	1.16	0.90-1.49	0.27	1.15	0.90-1.50	0.27	1.16	0.90-1.49
	≥ 50 (ref)	*Reference category*											
	Type of animal [a]												
	Small ruminants	**0.03**	**1.13**	**1.01-1.26**	**<0.001**	**1.17**	**1.07-1.28**	**0.02**	**1.09**	**1.01-1.17**	0.14	1.05	0.99-1.11
	Cattle	0.69	0.97	0.86-1.11	0.20	0.94	0.85-1.03	0.38	0.97	0.89-1.04	0.60	0.98	0.92-1.05
	Poultry	0.50	1.01	0.97-1.06	0.76	0.99	0.96-1.03	0.91	1.00	0.97-1.03	0.96	1.00	0.98-1.03
	Pigs	0.83	1.01	0.94-1.07	0.21	1.04	0.98-1.10	0.28	1.03	0.98-1.08	**0.03**	**1.05**	**1.01-1.09**

[a] *Population-weighted number of animals*

Table 4. Results of the Multivariable Spatial Analyses for Winter with Different Hexagonal Areas (90, 50, 25 and 10 km²).

Period of infection	Variable	Hexagon 90 km²			Hexagon 50 km²			Hexagon 25 km²			Hexagon 10 km²		
		P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI
Winter	**Gender**												
	Males	*Reference category*											
	Females	0.20	1.30	0.87-1.94	0.20	1.30	0.87-1.94	0.20	1.30	0.87-1.94	0.20	1.30	0.87-1.93
	Age category (years)												
	0-4	**<0.01**	**2.80**	**1.39-5.35**	**<0.01**	**2.80**	**1.39-5.35**	**<0.01**	**2.81**	**1.39-5.36**	**<0.01**	**2.80**	**1.39-5.36**
	5-9	**<0.01**	**2.82**	**1.43-5.30**	**<0.01**	**2.82**	**1.43-5.31**	**<0.01**	**2.82**	**1.43-5.30**	**<0.01**	**2.82**	**1.43-5.31**
	10-49	0.67	1.11	0.70-1.78	0.67	1.11	0.70-1.78	0.66	1.11	0.70-1.79	0.66	1.11	0.70-1.78
	≥ 50	*Reference category*											
	Type of animal [a]												
	Small ruminants	0.11	1.15	0.97-1.37	0.14	1.12	0.96-1.30	0.07	1.12	0.99-1.27	0.58	0.97	0.88-1.08
	Cattle	0.39	0.92	0.75-1.12	0.34	0.93	0.79-1.08	0.66	0.97	0.86-1.10	0.19	1.07	0.97-1.19
	Poultry	0.91	1.00	0.94-1.07	0.93	1.00	0.94-1.06	0.56	1.01	0.97-1.07	0.89	1.00	0.95-1.04
	Pigs	0.51	1.03	0.94-1.13	0.27	1.05	0.96-1.14	0.79	0.99	0.92-1.07	0.65	0.99	0.92-1.05

[a] *Population-weighted number of animals*

The variation in the spatially structured residual risks of the main model showed some dependence on region and period of infection (Figure 4), with a slightly increased residual risk for STEC O157 infection in the northern, mid-eastern and south-western regions of the Netherlands in winter and in the mid-eastern region in summer. A lower residual risk was found in the mid-west to north-west and the south-east region for both periods of infection.

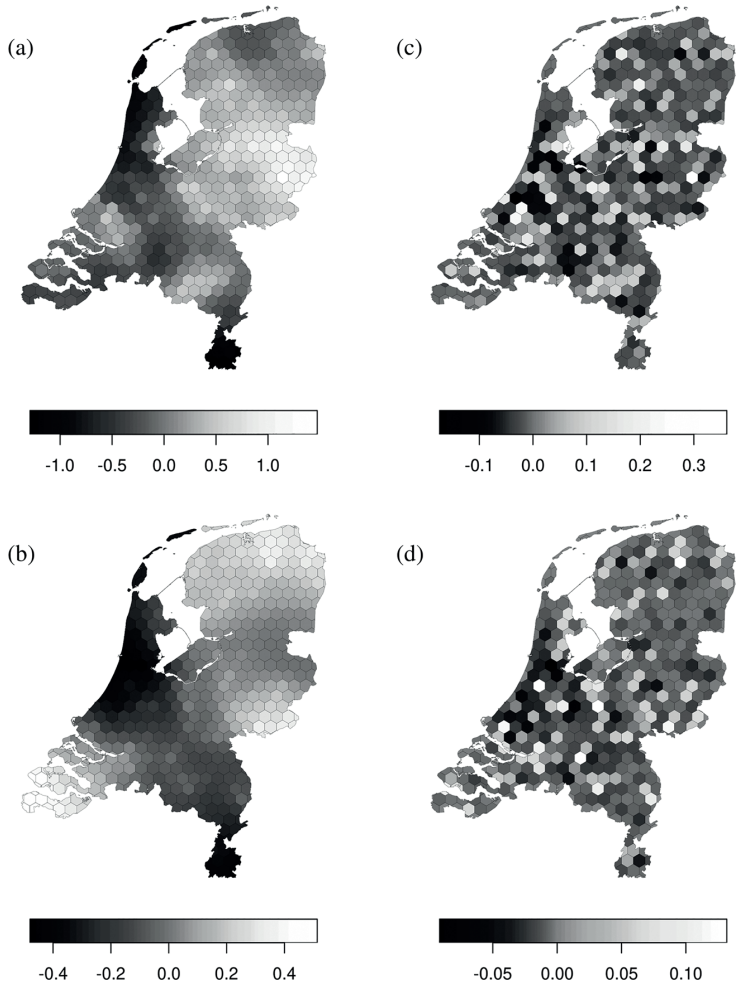


Figure 4. Maps of the spatially structured variation modeled by the conditional autoregressive model (CAR) in summer (a) and winter (b) and maps of the spatially unstructured variation modeled by independent and identically distributed (IID) Gaussian noise in summer (c) and winter (d) for hexagonal areas of 90 km^2.

Discussion

The aim of this study was to explore the spatial association between sporadic human STEC O157 infections and the exposure to livestock (small ruminants, cattle, poultry, and pigs) in the Netherlands, a country with high densities of humans and livestock animals, for the years 2007–2016. This was done using a state-of-the-art spatial analysis method, in which hexagonal areas were used in combination with a novel metric that was applied to define the probability of exposure: the population-weighted number of animals per hexagon.

Results showed that there is a consistent significant association between the population-weighted number of small ruminants and the incidence of reported human STEC O157 infections in summer with a PAF of 49%. This means that in the absence of exposure to small ruminants, the number of sporadic STEC O157 infections should be expected to decrease by 49%, although the uncertainty ranges between 8% and 72%. Since we only accounted for livestock density as a risk factor for infection with STEC O157, we were not able to quantify the relative importance of this spatially restricted risk factor within the broader context of all possible risk factors. Another limitation of only using one risk factor in the analyses is that the risk of ecological biases could not be quantified. Although we corrected for age and gender at an individual level, areas can still differ regarding confounders that are not included in our analyses, as is also suggested by the maps of the spatially structured variation (CAR).

The finding that small ruminants are important contributors to human STEC O157 infections is supported by a Dutch source attribution study (Mughini-Gras et al., 2018), which shows that while domestic ruminants (cattle, sheep, and goats) are responsible for approximately three-quarters of reported human STEC (all serotypes) infections, small ruminants in particular accounted for 25% of all STEC infections. In the Netherlands, STEC O157 has been isolated from sheep and goats (Heuvelink et al., 1998; Heuvelink et al., 2002). Additionally, STEC was detected at almost all dairy goat and sheep farms in the Netherlands that were included in the Dutch surveillance of zoonoses in 2016, although STEC O157 was only detected at one farm (Opsteegh et al., 2018). This reflects a common paradox regarding the results of animal sampling, in which small ruminants are generally considered as a primary reservoir for STEC O157, but their occurrence is infrequently demonstrated. This may reflect the sporadic and/or intermittent nature of STEC O157 carriage and low numbers of bacteria residing in colonized animals, or insufficiently sensitive sampling and culturing approaches (Ferens & Hovde, 2011).

Given the presence of STEC in small ruminants' feces and farms, it is plausible that human infections occur via environmental transmission. In the Netherlands, small

ruminants are usually kept in deep litter houses, with partially open walls or roofs (Schimmer et al., 2011). An initial layer of litter (usually straw or sawdust material) is spread for the animals to use for bedding material and to defecate on. As soon as this layer is soiled, new layers are added, which can build up to a depth of 1–2 meters. This process generates a lot of dust, which is easily spread into the environment through the often (partially) open housing system. As a result, the transport of STEC O157 in dust through the air can be one of the possible environmental transmission routes if infected animals are present on the farm (Chase-Topping et al., 2008; Schimmer et al., 2011). The plausibility of air-borne transmission is supported by a study focusing on microbial air pollution from livestock farms in the Netherlands, where a higher concentration of commensals, among which *Escherichia coli*, in dust particles was found in rural areas with higher farm density (de Rooij et al., 2019). Although no significant associations with the number of goats and sheep were found, the presence of livestock-related microbial markers, such as *Escherichia coli*, indicates that microbial air pollution with *Escherichia coli* is reasonable. The same phenomenon was observed for *Campylobacter*, which coincides with a higher *Campylobacter* incidence in poultry-dense areas, the main reservoir of *Campylobacter* (de Rooij et al., 2019; Poulsen et al., 2018). Furthermore, transmission of STEC O157 to humans may occur via soil or water, since dust precipitates and the stable litter that is stored outside the stable comes into contact with soil and possibly fresh water systems through washout after heavy rainfall (Elson et al., 2018).

Despite the above reasoning, the results did not provide evidence for a particular mode of transmission (e.g. through food, the environment or direct contact with (small) ruminants through petting or feeding animals at 'children farms'). There were no data available on individually reported exposures of the cases. Furthermore, the focus of this study was on food-producing animal farms. However, it has been shown before that visiting a petting farm can be a potential source of STEC O157 infection (Heuvelink et al., 2002; Valkenburgh & Heuvelink, 2006). As these petting farms often host small ruminants as well, it is recommended to dive deeper into the combined effects of petting farms and food-producing animal farms on STEC O157 infection risk in future studies. This could be done, for instance, using a case-control study design with individually reported exposures, which includes risk factors related to direct contact with the animals (e.g. visiting 'children farms'), as well as spatial risk factors (e.g. distance to farms, number of animals in the neighborhood or a combination of those two) to investigate the effect of potentially spatially restricted risk factors. This way, it is also possible to include the consumption of particular food items in the analysis, as transmission through food is considered as the most important risk factor of acquiring a STEC O157 infection (Strachan et al., 2006).

Whilst several studies, including a Dutch one, showed a significant spatial association between cattle and STEC O157 infections (Friesema et al., 2011; ÓHaiseadha et al., 2017; Strachan et al., 2006; Widgren et al., 2018), we did not. This could have several possible explanations.

First, a major difference is the inclusion of small ruminants in this study, next to cattle. Cattle farms are widely distributed in the Netherlands, while small ruminants have a more profound environmental spread. To study whether this could lead to different results, the analyses were performed for a model with only cattle. The results showed that the population-weighted number of cattle had a marginal significant association with human STEC O157, while this effect is not significant anymore after the inclusion of pigs, poultry and small ruminants. This might indicate that the spatial association observed for cattle could be due to its spatial relatedness with small ruminants, the latter which may play a more important role in environmental STEC O157 transmission. This proves that it is meaningful to look at the combined effects of all possible reservoirs for STEC O157. Such a combined analysis is especially important in a country like the Netherlands, which has a peculiar situation in terms of livestock and population density as compared to other countries (Smit & Heederik, 2017). Indeed, it is one of the most densely populated countries in the world in combination with a high density of intensive livestock farms (Smit & Heederik, 2017). An example of such a situation is the Q-fever epidemic in the Netherlands (Schimmer et al., 2011), which became an epidemic because most goat farms were located very close to locations with a high population density (Schimmer et al., 2011). As all the different types of livestock farms in the Netherlands are intertwined and mixed throughout the landscape, spatial inter-relatedness with other animal species does play a role (de Rooij et al., 2019). This makes it complicated to disentangle the effects and to look at each type of livestock separately, emphasizing that a more complete model in terms of possible reservoirs of STEC O157 is necessary for a proper analysis (de Rooij et al., 2019).

Second, livestock farming in the Netherlands underwent several changes in the past few years that could explain the different findings as well (Bos et al., 2013). There was a reduction in the number of farms over the years, which was paralleled by an increase in the number of animals per farm, with cows being increasingly kept inside throughout the year (Bos et al., 2013; Groot & van't Hooft, 2016; Smit & Heederik, 2017). As cattle is more often kept inside and their housing is closed, it is possible that aerial spread of STEC from cattle is reduced over the years and that small ruminants play a more important role nowadays.

Third, this study used a different spatial metric as response variable in order to do the spatial regression analyses on the hexagonal grid. Here, the population-weighted number of animals was used instead of animal density as exposure measure to transform the spatial data from one regional division to the other (Elson et al., 2018; Friesema et al., 2011; Hallisey et al., 2017; ÓHaiseadha et al., 2017). However, our approach has the advantage over areal weighted interpolation that it can more accurately estimate the population demographics in transforming small counts by four-digit postal code regions to aggregated counts for large, non-standard study zones (hexagons) (Hallisey et al., 2017). Moreover, because the probability of exposure on a population level is not only determined by the number of animals in a certain area, but also by the number of residents in a certain area and where they live inside an area, this study is more likely to have captured true environmental exposure, as exposure is less likely to occur when nobody lives in the vicinity of these animals (Mulder et al., 2016). Furthermore, in contrast to other studies, we took into account potential exposure to animals in neighboring hexagons, because pathogen spread is not hold back by "invisible" hexagonal boundaries.

In this study, no associations were found between poultry, pigs and STEC O157 infections in the multivariable model. This supports the finding that STEC has been isolated only sporadically from animals other than ruminants and these animals can merely be seen as spill-over hosts (Caprioli et al., 2005; Mughini-Gras et al., 2018). Also, a low estimated contribution to human STEC infections has previously been found for poultry and pigs in the Netherlands (Mughini-Gras et al., 2018). However, pigs did show a positive association with human STEC infections at a hexagonal size of 10 km^2 in the multivariable model. This could be due to several reasons, such as limitations of power and more limited exposure metric contrasts at this smaller spatial scale (de Rooij et al., 2019).

The association between small ruminants and human STEC O157 infections was only present in the summer. This is in agreement with the incidence of human STEC O157 infections being highest in summer, as well as the seasonality of fecal excretion of STEC in farm animals (Friesema et al., 2011; Heuvelink et al., 1998). Furthermore, humans are more likely to have direct or indirect contact with animals in summer as they probably spend more time outside (Friesema et al., 2011). Similar to what is described globally, women had a higher risk than men to acquire a STEC O157 infection in summer and the incidence of STEC O157 was highest in children <10 years and strongest in children <5 years (Elson et al., 2018; Friesema et al., 2011).

A buffer radius of 1 km was used in the analyses. However, the question remained whether different buffer sizes would influence our results. Therefore, the analyses were repeated for other buffer sizes. Buffers with a radius of 0.10 km, 0.25 km, 0.50 km, 0.75 km, 1.25 km, 1.50 km, 1.75 km and 2 km were used, but they did not show significant changes in the RRs and the 95% CIs were comparable to the results of the analyses with a buffer radius of 1 km. This means that the results of our analyses were not sensitive to the buffer radius size and that the analytical approach used was not suitable for assessing possible dose–response relationships.

Compared to a previous Dutch study (Friesema et al., 2011), underreporting of the human STEC O157 infections and the geographical laboratory bias did not change. Human STEC O157 cases included in this study likely represent the more severe cases, as mild cases often go unnoticed, because they may not always seek medical attention or do not get laboratory tested and hence, do not end up in the surveillance records (Friesema et al., 2011; van den Brandhof et al., 2006). The laboratory surveillance is based on a voluntary system, but despite the fact that the notification is mandatory, it is not guaranteed that all laboratories send in their isolates on a regular basis. Furthermore, the assumption was made that STEC O157 infections were acquired at or in close proximity to the home. However, people travel and it is possible that residents of urban areas went and acquired the infection in the countryside, or vice versa. This could lead to an underestimation of the spatial association between small ruminants and human STEC O157 cases and warrants further research in the future.

Conclusions

Results of this study indicate that living in proximity of small ruminants, is a spatially restricted risk factor for acquiring STEC O157 infection. As this study did not have individually reported exposures available, it could not provide evidence on the specific mode of transmission. Therefore, the exact underlying mechanisms warrant further investigation, and could offer new targets for control. The finding that small ruminants, and not cattle, are significantly associated with human STEC O157 infection is in contradiction with earlier studies. It could be explained by the inclusion of small ruminants in the analysis, a changing farming landscape over the years, and the newly developed exposure metric, the population-weighted number of animals per hexagon, which showed potential to improve existing spatial modeling studies on infectious diseases related to livestock exposure, especially in densely populated regions.

Abbreviations

CAR	Conditional Autoregressive Model
CI	Confidence interval
EU	European Union
IID	Independent and identically distributed INLA Integrated nested Laplace approximation
MHS	Municipal Health Service
PAF	Population attributable fraction
PCR	Polymerase chain reaction
RIVM	Dutch National Institute for Public Health and the Environment
RR	Rate ratio
STEC O157	Shiga Toxin-producing *Escherichia coli*
Stx_1	Shiga Toxin 1
Stx_2	Shiga Toxin 2

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The used R scripts for data analyses can be found at: https://github.com/mulderac91/R-STECO157-spatialanalysis. Livestock data of food-producing animals for 2012 was obtained from the Department of Service Arrangements of the Dutch Ministry of Agriculture, Nature and Food Quality (CBS, 2019b; RVO, 2019). The population data per four-digit postal code region per year is available through Statistics Netherlands (www.statline.nl). The data were downloaded for the years 2007–2016 (CBS, 2019a).

The four-digit postal code region shapefiles and the six-digit postal code point locations of the Netherlands that were used within this study were obtained by the RIVM from the company: Iris International. Those shapefiles can only be given to those for whom permission has been granted by this company. They can be reached at this address: Gr.v. Prinstererlaan 20, 2,271 EN, Voorburg, the Netherlands. Tel: +31(0)70– 3863891, fax: +31(0)70–3873625, e-mail: info@iris-int.nl.

The STEC O157 case data are available within OSIRIS, the Dutch surveillance system and only researchers within the RIVM with access to this database can use those data as it contains privacy sensitive information of cases and therefore are not accessible to the public or research community following the legislation of the Dutch law and the Dutch Data Protection Authority (Dutch Data Protection Authority, 2020a, 2020b).

Author Contributions

Conceptualization: A. C. Mulder, J. van de Kassteele, R. Pijnacker, L. Mughini-Gras, E. Franz
Data curation: A. C. Mulder, J. van de Kassteele, L. Mughini-Gras
Formal analysis: A. C. Mulder
Funding acquisition: L. Mughini- Gras, E. Franz
Investigation: A. C. Mulder, J. van de Kassteele, D. Heederik, R. Pijnacker, L. Mughini-Gras, E. Franz
Methodology: A. C. Mulder, J. van de Kassteele
Project administration: A. C. Mulder, E. Franz
Resources: J. van de Kassteele, L. Mughini-Gras, E. Franz
Software: A. C. Mulder, J. van de Kassteele
Supervision: D. Heederik, L. Mughini-Gras, E. Franz
Validation: J. van de Kassteele
Visualization: A. C. Mulder
Writing - original draft: A. C. Mulder
Writing – review & editing: A. C. Mulder, J. van de Kassteele, D. Heederik, R. Pijnacker, L. Mughini- Gras, E. Franz

Acknowledgments

The authors would like to thank Ingrid Friesema for her valuable input about the STEC O157 surveillance in the Netherlands and Gerrit Koop for sharing his knowledge about the common housing of small ruminants as used by farmers in the Netherlands.

References

Arsenault, J., Michel, P., Berke, O., Ravel, A., & Gosselin, P. (2013). How to choose geographical units in ecological studies: proposal and application to campylobacteriosis. *Spatial and spatio-temporal epidemiology, 7*, 11-24. doi:https://doi.org/10.1016/j.sste.2013.04.004

Arya, S., Mount, D., Kemp, S. E., & Jefferis, G. (2015). RANN: Fast Nearest Neighbour Search (Wraps Arya and Mount's ANN Library). Retrieved from http://cran.rapporter.net/bin/macosx/contrib/3.2/PACKAGES

Bates, D., Maechler, M., Davis, T. A., Oehlschlägel, J., & Riedy, J. (2019). Matrix: Sparse and Dense Matrix Classes and Methods. Retrieved from https://cran.r-project.org/web/packages/Matrix/index.html

Berry, E. D., Wells, J. E., Bono, J. L., Woodbury, B. L., Kalchayanand, N., Norman, K. N., et al. (2015). Effect of Proximity to a Cattle Feedlot on O157:H7 Contamination of Leafy Greens and Evaluation of the Potential for Airborne Transmission. *Applied and Environmental Microbiology, 81*(3), 1101-1110. doi:https://doi.org/10.1128/AEM.02998-14

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics, 43*(1), 1-20. doi:https://doi.org/10.1007/BF00116466

Birch, C. P. D., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling, 206*(3), 347-359. doi:https://doi.org/10.1016/j.ecolmodel.2007.03.041

Birch, C. P. D., Vuichard, N., & Werkman, B. R. (2000). Modelling the effects of patch size on vegetation dynamics: Bracken [Pteridium aquilinum (L.) Kuhn] under grazing. *Annals of Botany, 85*, 63-76. doi:https://doi.org/10.1006/anbo.1999.1081

Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., et al. (2019). spdep: Spatial Dependence: Weighting Schemes, Statistics. Retrieved from https://cran.r-project.org/web/packages/spdep/index.html

Bos, J. F., Smit, A. B. L., & Schröder, J. J. (2013). Is agricultural intensification in The Netherlands running up to its limits? *NJAS-Wageningen Journal of Life Sciences, 66*, 65-73. doi:https://doi.org/10.1016/j.njas.2013.06.001

Caprioli, A., Morabito, S., Brugère, H., & Oswald, E. (2005). Enterohaemorrhagic Escherichia coli: emerging issues on virulence and modes of transmission. *Veterinary research, 36*(3), 289-311. doi:https://doi.org/10.1051/vetres:2005002

CBS. (2019a). Bevolking; geslacht, leeftijd en viercijferige postcode, 1 januari. Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83502NED/table?ts=1575556895525

CBS. (2019b). Landbouwtelling. Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/landbouwtelling

Chase-Topping, M., Gally, D., Low, C., Matthews, L., & Woolhouse, M. (2008). Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. *Nature Reviews Microbiology, 6*(12), 904. doi:https://doi.org/10.1038/nrmicro2029

De Jonge, E., & Houweling, S. (2019). cbsodataR: Statistics Netherlands (CBS) Open Data API Client. Retrieved from https://rdrr.io/cran/cbsodataR/

de Rooij, M. M. T., Hoek, G., Schmitt, H., Janse, I., Swart, A., Maassen, C. B. M., et al. (2019). Insights into Livestock-Related Microbial Concentrations in Air at Residential Level in a Livestock Dense Area. *Environmental Science & Technology, 53*(13), 7746-7758. doi:https://doi.org/10.1021/acs.est.8b07029

Dutch Data Protection Authority. (2020a). Retrieved from https://www.autoriteitpersoonsgegevens.nl/en

Dutch Data Protection Authority. (2020b). Retrieved from https://www.autoriteitpersoonsgegevens.nl/en/about-dutch-dpa/tasks-and-powers-dutch-dpa

Elson, R., Grace, K., Vivancos, R., Jenkins, C., Adak, G. K., O'Brien, S. J., & Lake, I. R. (2018). A spatial and temporal analysis of risk factors associated with sporadic Shiga toxin-producing Escherichia coli O157 infection in England between 2009 and 2015. *Epidemiol Infect, 146*(15), 1928-1939. doi:https://doi.org/10.1017/S095026881800256X

European Centre for Disease Prevention and Control. (2019). *Annual Epidemiological report for 2017 - Shiga toxin/verocytotoxin-producing Escherichia coli (STEC/VTEC) infection - Annual Epidemiological Report for 2017*. Retrieved from Stockholm: https://www.ecdc.europa.eu/en/publications-data/shiga-toxinverocytotoxin-producing-escherichia-coli-stecvtec-infection-annual

European Commission - Eurostat. (2019). Nuts - Nnuts - Nomenclature of territorial units for statistics. Retrieved from https://ec.europa.eu/eurostat/web/nuts/background

European Food Safety Authority, & European Centre for Disease Prevention and Control. (2019). The European Union One Health 2018 Zoonoses Report. *EFSA Journal, 17*(12), e05926. doi:https://doi.org/10.2903/j.efsa.2019.5926

Ferens, W. A., & Hovde, C. J. (2011). Escherichia coli O157: H7: animal reservoir and sources of human infection. *Foodborne pathogens and disease, 8*(4), 465-487. doi:https://doi.org/10.1089/fpd.2010.0673

Franz, E., Rotariu, O., Lopes, B. S., MacRae, M., Bono, J. L., Laing, C., et al. (2018). Phylogeographic analysis reveals multiple international transmission events have driven the global emergence of Escherichia coli O157: H7. *Clinical Infectious Diseases, 69*(3), 428-437. doi:https://doi.org/10.1093/cid/ciy919

Franz, E., Schijven, J., de Roda Husman, A. M., & Blaak, H. (2014). Meta-Regression Analysis of Commensal and Pathogenic Escherichia coli Survival in Soil and Water. *Environmental Science & Technology, 48*(12), 6763-6771. doi:https://doi.org/10.1021/es501677c

Friesema, I., Kuiling, S., van der Voort, M., in 't Veld, P. H., Heck, M. E. O. C., & Franz, E. (2017). Surveillance van shigatoxineproducerende Escherichia coli (STEC) in Nederland, 2016. Retrieved from https://www.rivm.nl/surveillance-van-shigatoxineproducerende-escherichia-coli-stec-in-nederland-2016

Friesema, I., Sigmundsdottir, G., Van Der Zwaluw, K., Heuvelink, A., Schimmer, B., De Jager, C., et al. (2008). An international outbreak of Shiga toxin-producing Escherichia coli O157 infection due to lettuce, September–October 2007. *Euro Surveill., 13*(50), 19065. doi:https://doi.org/10.2807/ese.13.50.19065-en

Friesema, I., Van De Kassteele, J., De Jager, C. M., Heuvelink, A. E., & Van Pelt, W. (2011). Geographical association between livestock density and human Shiga toxin-producing Escherichia coli O157 infections. *Epidemiol. Infect., 139*(7), 1081-1087. doi:https://doi.org/10.1017/S0950268810002050

Greenland, K., De Jager, C., Heuvelink, A., Van Der Zwaluw, K., Heck, M., Notermans, D., et al. (2009). Nationwide outbreak of STEC O157 infection in the Netherlands, December 2008-January 2009: continuous risk of consuming raw beef products. *Euro Surveill., 14*(8). Retrieved from http://researchonline.lshtm.ac.uk/id/eprint/748775/

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software, 40*(3), 1-25.

Groot, M. J., & van't Hooft, K. E. (2016). The Hidden Effects of Dairy Farming on Public and Environmental Health in the Netherlands, India, Ethiopia, and Uganda, Considering the Use of Antibiotics and Other Agro-chemicals. *Frontiers in Public Health, 4*(12). doi:https://doi.org/10.3389/fpubh.2016.00012

Hallisey, E., Tai, E., Berens, A., Wilt, G., Peipins, L., Lewis, B., et al. (2017). Transforming geographic scale: a comparison of combined population and areal weighting to other interpolation methods. *International Journal of Health Geographics, 16*(1), 29. doi:https://doi.org/10.1186/s12942-017-0102-z

Heuvelink, A., Van Den Biggelaar, F., De Boer, E., Herbes, R., Melchers, W., Huis, J., & Monnens, L. (1998). Isolation and characterization of verocytotoxin-producing Escherichia coli O157 strains

from Dutch cattle and sheep. *Journal of Clinical Microbiology, 36*(4), 878-882. doi:https://doi.org/10.1128/JCM.36.4.878-882

Heuvelink, A., Van Heerwaarden, C., Zwartkruis-Nahuis, J., Van Oosterom, R., Edink, K., Van Duynhoven, Y., & De Boer, E. (2002). Escherichia coli O157 infection associated with a petting zoo. *Epidemiology & Infection, 129*(2), 295-302. doi:https://doi.org/10.1017/S095026880200732X

Keitt, T. H. (2010). rgdal: Bindings for the Geospatial Data Abstraction Library, R package version 0.6-28. *http://cran. r-project. org/package= rgdal*.

Lawson, A. B. (2013). *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*: Chapman and Hall/CRC.

Mughini-Gras, L., van Pelt, W., van der Voort, M., Heck, M., Friesema, I., & Franz, E. (2018). Attribution of human infections with Shiga toxin-producing Escherichia coli (STEC) to livestock sources and identification of source-specific risk factors, The Netherlands (2010–2014). *Zoonoses and Public Health, 65*(1), e8-e22. doi:https://doi.org/10.1111/zph.12403

Mulder, A. C., Snabilie, M., & Braks, M. A. (2016). From guessing to GIS-ing: empowering land managers. In *Ecology and prevention of Lyme borreliosis* (Vol. 4, pp. 373-387): Wageningen Academic Publishers.

Neuwirth, E. (2015). Package 'RColorBrewer'. Retrieved from https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf

ÓHaiseadha, C., Hynds, P. D., Fallon, U. B., & O'Dwyer, J. (2017). A geostatistical investigation of agricultural and infrastructural risk factors associated with primary verotoxigenic E. coli (VTEC) infection in the Republic of Ireland, 2008–2013. *Epidemiology and Infection, 145*(1), 95-105. doi:https://doi.org/10.1017/S095026881600193X

Pebesma, E. (2019). st_make_grid. Retrieved from https://www.rdocumentation.org/packages/sf/versions/0.8-0/topics/st_make_grid

Pebesma, E., Bivand, R., Racine, E., Sumner, M., Cook, I., Keitt, T., et al. (2019). sf: Simple Features for R. Retrieved from https://cran.r-project.org/web/packages/sf/index.html

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., Hijmans, R., Sumner, M., et al. (2019). sp: Classes and Methods for Spatial Data. Retrieved from https://cran.r-project.org/web/packages/sp/index.html

Poulsen, M. N., Pollak, J., Sills, D. L., Casey, J. A., Rasmussen, S. G., Nachman, K. E., et al. (2018). Residential proximity to high-density poultry operations associated with campylobacteriosis and infectious diarrhea. *International Journal of Hygiene and Environmental Health, 221*(2), 323-333. doi:https://doi.org/10.1016/j.ijheh.2017.12.005

R-Core. (2017). parallel v3.6.0. Retrieved from https://www.rdocumentation.org/packages/parallel/versions/3.6.0

RCT. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org

Rue, H. (2019). The R-INLA project. Retrieved from http://www.r-inla.org/home

Rue, H., Lindgren, F., Simpson, D., Martino, S., Teixeira Krainski, E., Bakka, H., et al. (2019). INLA: Full Bayesian Analysis of Latent Gaussian Models using Integrated Nested Laplace Approximations. *R package version, 19*(03).

RVO. (2019). Gecombineerde opgave. Retrieved from https://mijn.rvo.nl/gecombineerde-opgave

Schimmer, B., Luttikholt, S., Hautvast, J. L., Graat, E. A., Vellema, P., & van Duynhoven, Y. T. (2011). Seroprevalence and risk factors of Q fever in goats on commercial dairy goat farms in the Netherlands, 2009-2010. *BMC Veterinary Research, 7*(1), 81. doi:https://doi.org/10.1186/1746-6148-7-81

Shafran-Nathan, R., Levy, I., Levin, N., & Broday, D. M. (2017). Ecological bias in environmental health studies: the problem of aggregation of multiple data sources. *Air Quality, Atmosphere & Health, 10*(4), 411-420. doi:https://doi.org/10.1007/s11869-016-0436-x

Smit, L. A. M., & Heederik, D. (2017). Impacts of Intensive Livestock Production on Human Health in Densely Populated Regions. *GeoHealth, 1*(7), 272-277. doi:https://doi.org/10.1002/2017GH000103

Strachan, N. J. C., Dunn, G. M., Locking, M. E., Reid, T. M. S., & Ogden, I. D. (2006). Escherichia coli O157: Burger bug or environmental pathogen? *International Journal of Food Microbiology, 112*(2), 129-137. doi:https://doi.org/10.1016/j.ijfoodmicro.2006.06.021

Strachan, N. J. C., Fenlon, D. R., & Ogden, I. D. (2001). Modelling the vector pathway and infection of humans in an environmental outbreak of Escherichia coli O157. *FEMS Microbiology Letters, 203*(1), 69-73. doi:https://doi.org/10.1111/j.1574-6968.2001.tb10822.x

Valkenburgh, S., & Heuvelink, A. (2006). Children's farms in The Netherlands: hygiene and zoonotic pathogens. *Tijdschrift voor diergeneeskunde, 131*(7), 224.

van den Brandhof, W. E., Bartelds, A. I., Koopmans, M. P., & van Duynhoven, Y. T. (2006). General practitioner practices in requesting laboratory tests for patients with gastroenteritis in the Netherlands, 2001–2002. *BMC Family Practice, 7*(1), 56. doi:https://doi.org/10.1186/1471-2296-7-56

Wickham, H. (2019). stringr v1.4.0. Retrieved from https://www.rdocumentation.org/packages/stringr/versions/1.4.0

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Wickham, H., Bryan, J., Rstudio, Kalicinski, M., Valery, K., Leitienne, C., et al. (2019). Package 'readxl'. Retrieved from https://cran.r-project.org/web/packages/readxl/readxl.pdf

Wickham, H., Francois, R., Henry, L., & Müller, K. (2019). Package 'dplyr'. Retrieved from https://cran.r-project.org/web/packages/dplyr/dplyr.pdf

Wickham, H., Henry, L., & Rstudio. (2019). tidyr: Tidy Messy Data. Retrieved from https://cran.r-project.org/web/packages/tidyr/index.html

Widgren, S., Engblom, S., Emanuelson, U., & Lindberg, A. (2018). Spatio-temporal modelling of verotoxigenic Escherichia coli O157 in cattle in Sweden: exploring options for control. *Veterinary research, 49*(1), 78. doi:https://doi.org/10.1186/s13567-018-0574-2

Supplementary material

This file mainly contains supporting information supporting the materials and methods section of the article, including the following items:

- An explanation of the population-weighted interpolation (**Text S1**)
- An explanation of the interpretation of the rate ratios used (RR) (**Text S2**)
- An explanation of how the population attributable fraction was calculated (PAF) (**Text S3**)
- An explanation of the different spatial scales of the Netherlands compared to the NUTS classification system (**Text S4**)
- A figure showing the changing four-digit postal code regions of the Netherlands over the years (**Figure S1**)
- A figure visualizing this interpretation (**Figure S2**)
- A figure showing those different spatial scales (**Figure S3**)
- An overview of the R packages used (**Table S1**)
- An overview of the univariable spatial regression results (**Table S2**)

Text S1. Population weighted interpolation

The population weighted interpolation was carried out as follows: first, we made an intersection between the four-digit postal code regions and the six-digit postal code points. Next, the four-digit postal code region data (both the STEC O157 cases and the population numbers by age category and gender) were redistributed over the six-digit postal code points, proportional to the number of inhabitants for these six-digit postal code point locations. Then, an intersection was made between the six-digit postal code points and the hexagonal grid. Finally, the redistributed data over the six-digit postal code points were allocated to each hexagon.

Some four-digit postal code regions could not be redistributed, because no six-digit postal code points could be assigned to it. In that case, the nearest six-digit postal code point location was used. Similarly, when a six-digit postal code point could not be assigned to a hexagon, the nearest hexagon was used. The redistribution from the four-digit postal code regions to six-digit postal code points to the hexagonal grid could be done very efficiently by sparse matrix multiplications. For each age category and gender stratum, the same redistribution matrix was used.

Text S2. Rate ratio (RR)

In this study, the exposure measure x can get the value zero. Therefore, the explanatory variable was transformed using the $log_2(x+1)$ transformation. Resulting in the following Poisson regression with log-link function formula:

$$\log(\mu) = \beta_0 + \beta_1 log_2(x+1)$$

By taking the inverse link-function of this, using the exponential function e^x, we obtained:

$$\begin{aligned} \mu &= e^{\beta_0 + \beta_1 (x+1)} \\ &= e^{\beta_0} e^{\beta_1 (x+1)} \end{aligned}$$

The RR_{21} for an exposure at $log_2(x_2+1)$ relative to $log_2(x_1+1)$ then is:

$$RR_{21} = e^{\beta_1 log_2(\frac{x_2+1}{x_1+1})}$$

If $x+1$ grows with a factor two, the rate increases with a factor $RR = e^{\beta_1}$. Fortunately, not much changes when x is large relative to one, as the following applies:

$$\frac{x_2+1}{x_1+1} \approx \frac{x_2}{x_1}$$

This leads to the same "easier" interpretation of the rate ratio as when using a $\log(x)$ transformation: if x increases with a factor two, the incidence rate increases with a factor $RR = e^{\beta_1}$. But what is "large" ? Do we make a big mistake with this approximation? We visualized this in Figure S2. In this figure, x_1 increases from one towards 1,000 and the factor two was chosen as ratio between x_2 and x_1, thus $x_2 = 2x_1$. The x-axis was transformed into a log_{10} scale to make the effect of large values of x_1 on the factor more clear. The constant value of two is what we would have at $\frac{x_2}{x_1} = 2$. The red line is this factor when we add one to x. As Figure S2 shows, this approximation is pretty good when values of x_1 are approximately above 100. This indicates that the "easier" interpretation of the rate ratio can be used.

In summary, if the $log_2(x_2+1)$ is used as explanatory variable in Poisson regression with log-link function, then the interpretation of the rate ratio (RR) is as follows: if x increases with a factor two, then the incidence rate increases with a factor $RR = e^{\beta_1}$, provided that x is large enough, approximately >100. When x is smaller, this factor is less than two for the same RR, but the significance stays the same.

Text S3. Population attributable fraction (PAF)

The PAF is calculated as follows:

$$PAF = \left(\frac{i(E) - i(0)}{i(E)}\right) * 100$$

Here $i(E)$ is the predicted incidence in the exposed population (using the regression model and its estimated coefficients as is) and $i(0)$ is the predicted incidence in the unexposed population (using the same regression model and estimated coefficients, but where the exposure of the risk factor is set to zero). Both predictions can be done simultaneously by augmenting the original dataset, where in the augmented records the exposure of the risk factor is set to zero and the outcome is set to missing. For each group (exposed and non-exposed), the total incidences are calculated as the sum of the individual records.

Text S4. Spatial scales of the Netherlands

To divide the economic territory of the EU, a hierarchical system was developed. This system is called the NUTS classification (Nomenclature of territorial units for statistics) (European Commission - Eurostat, 2019). It contains three levels:

- NUTS 1: major socio-economic regions
- NUTS 2: basic regions for the application of regional policies
- NUTS 3: small regions for specific diagnosis.

The current NUTS 2016 classification is valid from 1 January 2018 and lists 104 regions at NUTS 1, 281 regions at NUTS 2 and 1348 regions at NUTS 3 level (European Commission - Eurostat, 2019). In the Netherlands, the NUTS 1 regions consist of four areas: North of the Netherlands, East of the Netherlands, West of the Netherlands and South of the Netherlands. The NUTS 2 regions are the Dutch provinces (**Figure S3** - a) and the NUTS 3 regions are 40 COROP regions, which consist of a combination of several municipalities of a province. Thus, the municipalities in the Netherlands (~ 90 km2, **Figure S3** - b) are smaller than those NUTS 3 regions and the four-digit postal code regions of the Netherlands (~ 10 km2, **Figure S3** - c) are even smaller than those municipalities. The six-digit postal code point locations of the Netherlands give information about specific locations at street level.

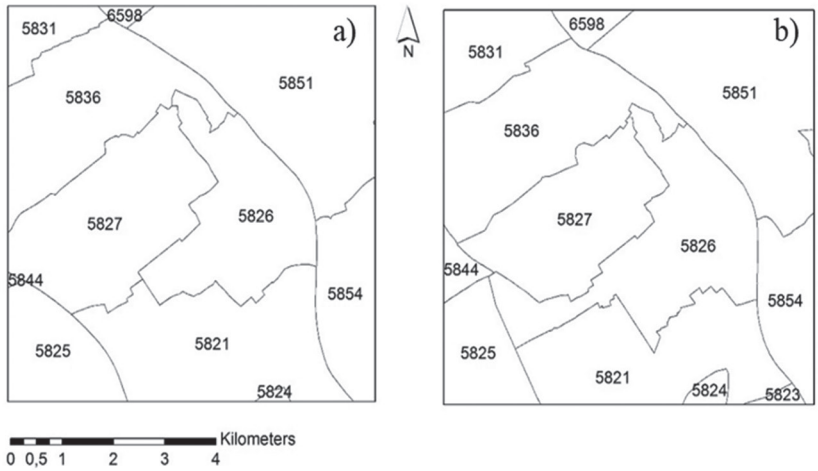


Figure S1. Example of changing four-digit postal code regions of the Netherlands over the years; a) 2009 compared to b) 2016.

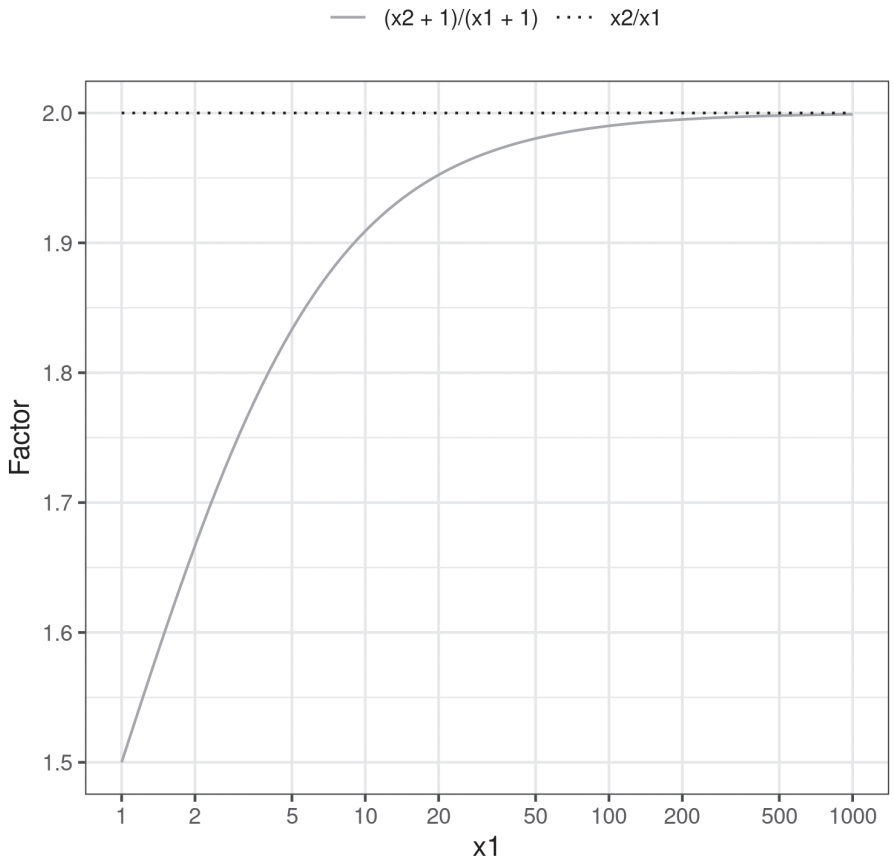


Figure S2. Visualization of interpretation rate ratio (RR) for an exposure measure x_1.

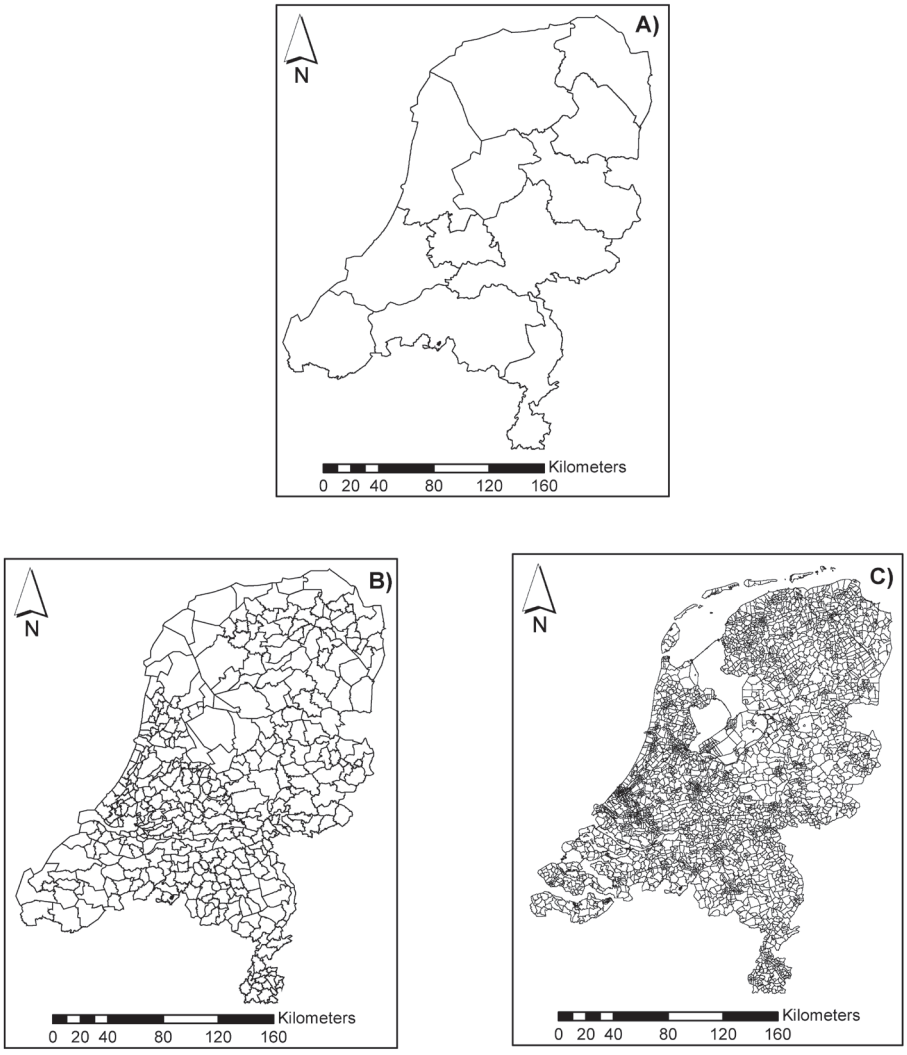


Figure S3. The different administrative boundaries and spatial scales of the Netherlands. a) Provinces (NUTS 2 regions), b) Municipalities, c) Four-digit postal code regions.

Table S1. An overview of the R packages and functions used, including version numbers and references.

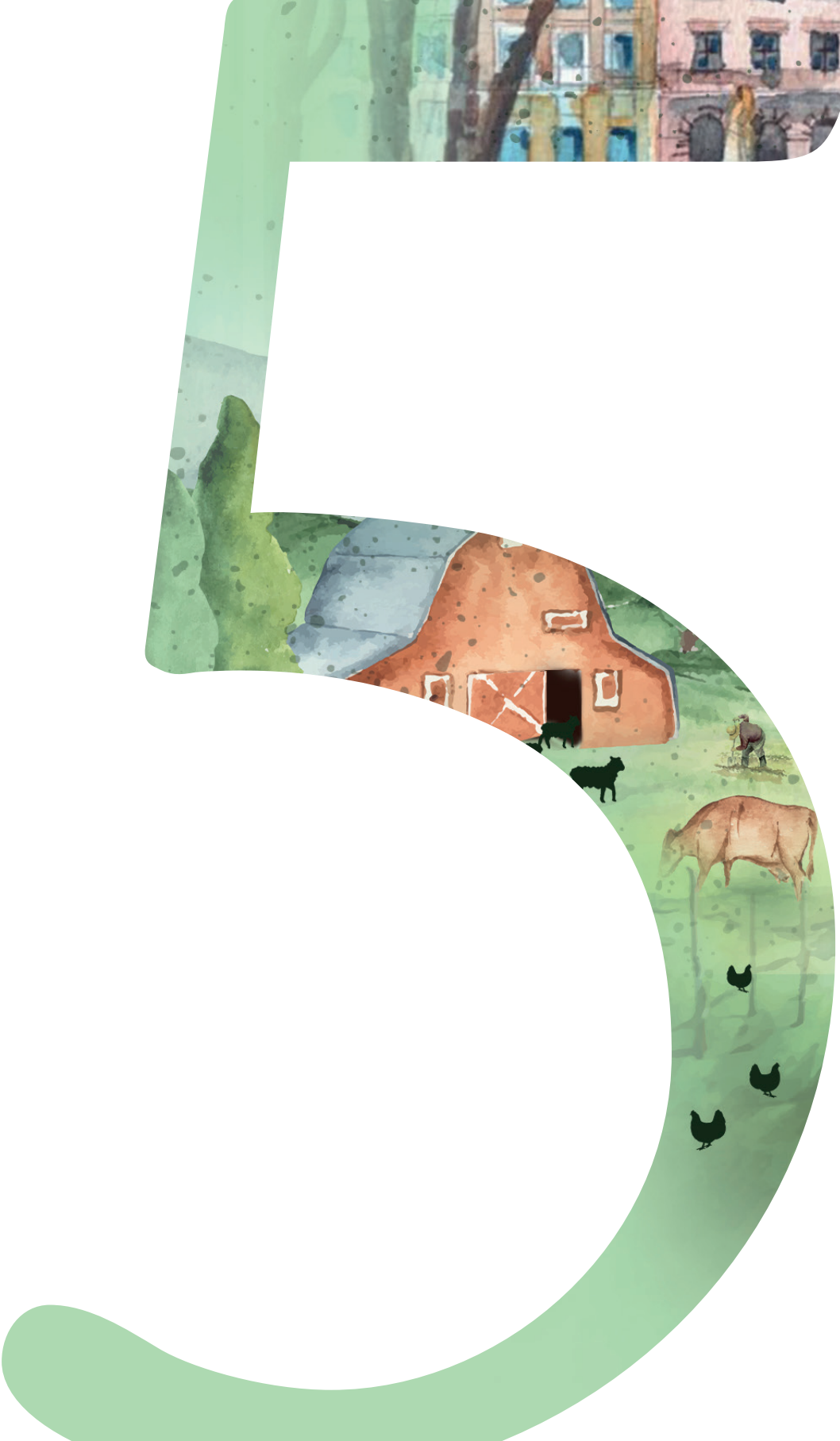
Package/function	Version	Reference*
cbsodataR	0.3.5	(De Jonge & Houweling, 2019)
dplyr	0.8.3	(Wickham, Francois, et al., 2019)
INLA	18.07.12	(Rue, 2019)
lubridate	1.7.4	(Grolemund & Wickham, 2011)
Matrix	1.2-17	(Bates et al., 2019)
Parallel	3.6.0	(R-Core, 2017)
RANN	1.2.6	(Arya et al., 2015)
RColorBrewer	1.1-2	(Neuwirth, 2015)
readxl	1.3.1	(Wickham, Bryan, et al., 2019)
rgdal	1.4-4	(Keitt, 2010)
sf	0.7-7	(Pebesma, Bivand, Racine, et al., 2019)
sp	1.3-1	(Pebesma, Bivand, Rowlingson, et al., 2019)
spdep	1.1-2	(Bivand et al., 2019)
st_make_grid		(Pebesma, 2019)
stringr	1.4.0	(Wickham, 2019)
tidyr	1.0.0	(Wickham, Henry, et al., 2019)
tidyverse	1.3.0	(Wickham, Averick, et al., 2019)

*All literature was included in the reference list belonging to the main text of this article.

Table S2. Univariable spatial analyses results for different hexagonal areas (90, 50, 25 and 10 km^2).

	Hexagon 90 km^2			Hexagon 50 km^2			Hexagon 25 km^2			Hexagon 10 km^2		
Variable	P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI	P-value	RR	95% CI
Period of infection												
Winter	Reference category											
Summer	**<0.001**	**3.43**	**2.76-4.31**	**<0.001**	**3.43**	**2.76-4.31**	**<0.001**	**3.43**	**2.76-4.31**	**<0.001**	**3.43**	**2.76-4.31**
Gender												
Males	Reference category											
Females	**<0.001**	**1.60**	**1.32-1.95**	**<0.001**	**1.60**	**1.32-1.95**	**<0.001**	**1.60**	**1.32-1.95**	**<0.001**	**1.60**	**1.32-1.95**
Age category (years)												
0-4	**<0.001**	**3.70**	**2.75-4.95**	**<0.001**	**3.71**	**2.76-4.96**	**<0.001**	**3.70**	**2.75-4.95**	**<0.001**	**3.70**	**2.75-4.95**
5-9	**<0.001**	**2.16**	**1.52-3.03**	**<0.001**	**2.17**	**1.53-3.04**	**<0.001**	**2.17**	**1.52-3.04**	**<0.001**	**2.17**	**1.52-3.03**
10-49	0.30	1.13	0.90-1.41	0.30	1.13	0.90-1.41	0.31	1.12	0.90-1.41	0.31	1.12	0.90-1.41
≥ 50	Reference category											
Type of animal [a]												
Small ruminants	**<0.01**	**1.12**	**1.04-1.20**	**<0.001**	**1.12**	**1.05-1.19**	**<0.01**	**1.08**	**1.03-1.14**	0.05	1.04	1.00-1.09
Cattle	0.07	1.07	0.99-1.14	0.13	1.05	0.99-1.11	0.10	1.04	0.99-1.09	**0.04**	**1.04**	**1.00-1.08**
Poultry	0.17	1.02	0.99-1.05	0.36	1.01	0.99-1.04	0.23	1.01	0.99-1.04	0.26	1.01	0.99-1.03
Pigs	0.11	1.03	0.99-1.08	**0.03**	**1.04**	**1.00-1.08**	0.10	1.03	0.99-1.06	**0.01**	**1.04**	**1.01-1.07**

[a] *Population weighted number of animals*



CHAPTER 5

Livestock-associated spatial risk factors for human salmonellosis and campylobacteriosis

Annemieke C. Mulder
Lapo Mughini-Gras
Jan van de Kassteele
Sara Lynn Blanken
Roan Pijnacker
Eelco Franz

Revised manuscript submitted to Zoonoses and Public Health, 2022

Abstract

Most human infections with non-typhoid *Salmonella* (NTS) or *Campylobacter* are zoonotic in nature and acquired though consumption of contaminated food of mainly animal origin. However, individuals may also acquire salmonellosis or campylobacteriosis through non-foodborne transmission pathways, such as those mediated by the environment. This emphasizes the need to consider both direct and indirect exposure to livestock sources as a possible transmission route for NTS and *Campylobacter*. Therefore, this study aimed at assessing whether salmonellosis and campylobacteriosis incidence is spatially associated with exposure to livestock (i.e., small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) in the Netherlands for the years 2007-2019 and 2014-2019 respectively. Risk factors (population-weighted number of animals) and their population attributable fractions were determined using a Poisson regression model with log-link function fitted using integrated nested Laplace approximation. Additionally, the analyses accounted for geographical coverage of the diagnostic laboratory catchment areas. Moreover, serological data was used to look into possible effects of acquired immunity due to repeated exposure to the pathogen through the environment that would potentially hinder the analyses based on the incidence of reported cases. A linear mixed-effects model was then built where the postal code areas were included as a random effect. Results showed that living in livestock-rich areas in the Netherlands is not a consistently significant, spatially restricted risk factor for acquiring salmonellosis or campylobacteriosis, thereby supporting current knowledge that human infections with *Salmonella* and *Campylobacter* are mainly foodborne.

Introduction

Most human infections with non-typhoid *Salmonella* (NTS) or *Campylobacter* are zoonotic in nature and acquired though consumption of contaminated food of mainly animal origin (Friesema et al., 2022). However, individuals may also acquire salmonellosis or campylobacteriosis through non-foodborne transmission pathways, such as those mediated by the environment (Guillier et al., 2021). In the Netherlands, pigs and laying hens are the most important livestock sources of *Salmonella* (Mughini-Gras, Enserink, et al., 2014), as broiler chickens and cattle are for *Campylobacter* (Mughini-Gras et al., 2021). Once *Salmonella* and *Campylobacter* are shed into the environment with animal feces, they are able to survive for varying periods of time depending on various environmental parameters (e.g. ≤3 months in manure; ≤1 month in soil) (Guillier et al., 2021; Lee et al., 2019; Mughini-Gras et al., 2021; Nicholson et al., 2005; Schets et al., 2017). This emphasizes the need to consider both direct and indirect exposure to livestock as a possible transmission route for NTS and *Campylobacter*, besides food.

As the Netherlands is a livestock-dense country where residential areas are often in close proximity to intensive livestock farms, environmental transmission via for example air or surface water is plausible (Smit & Heederik, 2017). The association between the concentration of *Campylobacter jejuni* in airborne dust and the presence of poultry farms in the area further supports this (de Rooij et al., 2019). Moreover, there was an observed drop in human campylobacteriosis incidence in the Netherlands after the implementation of massive culling operations in poultry farms in response to the H7N7 avian influenza epidemic that hit the country in 2003 (Friesema et al., 2012). Indeed, it has been hypothesized that a reduced environmental contamination with *Campylobacter* from the culled and therefore temporarily emptied poultry farms, as well as inactive slaughterhouses, could have occurred (Friesema et al., 2012). This was further confirmed in other studies (Mughini-Gras et al., 2021; Mulder, Franz, et al., 2020) showing a positive association between the magnitude of the poultry industry in an area (i.e. poultry farm density) and the probability for *Campylobacter* strains contaminating local surface water to originate from poultry. However, studies looking at the spatial association of human campylobacteriosis and salmonellosis incidence and exposure to livestock are scarce.

Therefore, this study aimed at assessing whether human salmonellosis and campylobacteriosis incidence is spatially associated with local density of small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs in the Netherlands. Moreover, serological data was used to look into possible effects of acquired immunity due to repeated exposure to the pathogen through the environment that would potentially hinder the analyses based on the incidence of reported cases. Analyses were based

on a recently developed spatial regression technique using the population-weighted number of animals per hexagonal area (Mulder, van de Kassteele, et al., 2020), and accounted for geographical coverage of the diagnostic laboratory catchment areas.

Materials and Methods

Data collection

Case data

National surveillance of *Salmonella* consists of a network of Regional Public Health Laboratories (RPHLs), which send *Salmonella* isolates on a voluntary basis to the National Institute for Public Health and the Environment (RIVM) for further typing. This network has an estimated national population coverage of 64% (Friesema et al., 2021). For *Campylobacter*, the RIVM has access to the voluntary reported campylobacteriosis cases based on laboratory surveillance data from the RPHLs which gave consent for using their data via the Infectious Disease Surveillance and Information System for Antibiotic Resistance (ISIS-AR). (Altorf-van der Kuil et al., 2017) The laboratory surveillance has an estimated national population coverage of 52% (Vlaanderen et al., 2019). The analyses were performed on fully deidentified surveillance data, so no ethics approval was required for those two datasets.

In this study, a salmonellosis case was defined as an individual with laboratory-confirmed, NTS infection during the period 2007-2019. A campylobacteriosis case was defined as an individual with laboratory-confirmed *Campylobacter jejuni* or *C. coli* infection during the period 2014-2019. For both *Salmonella* and *Campylobacter*, cases were excluded if age, sex, or residence location (i.e. postal code) was unknown. Additionally, confirmed outbreak-related salmonellosis cases with a proven link to food sources, as well as travel-related salmonellosis cases, were excluded. This information was not available for *Campylobacter*. Duplicates were removed if multiple isolates belonging to the same serovar (*Salmonella*) and species (*Campylobacter*) were obtained from the same patient within six months for *Salmonella* (Mughini-Gras et al., 2020) and two months for *Campylobacter* (Brachman & Evans, 1998; Pasternack, 2002). This resulted in a cleaned database for both *Salmonella* and *Campylobacter*. (Supplementary material, Text S1 and Text S2)

Infection pressure data

The data used for the analysis of infection pressure contained serological data for *Campylobacter* from participants of a population-based, cross-sectional serosurvey in 40 out of 458 municipalities in the Netherlands in 2006-2007 (the 'PIENTER-II' study). Details were described previously (Mollema et al., 2010). The serosurvey was

approved by the Medical Ethics Testing Committee of the Foundation of Therapeutic Evaluation of Medicines (METC-STEG) in Almere (ISRCTN 20164309).

In short, 7,904 participants provided a blood sample and completed an epidemiological questionnaire. Of those participants, 1,304 were tested for anti-*Campylobacter* IgA, IgM and IgG levels using a mixed ELISA (Ang et al., 2007). This dataset has already been used in a previous study (Monge et al. 2018). The serum levels were used to estimate the so-called sero-incidence rates, i.e. the estimated average number of immune response-eliciting exposures (or infections) that a given person experiences in a year (Teunis et al., 2013). For this purpose, the European Centre for Disease Prevention and Control's (ECDC) seroincidence calculator tool (https://ecdc.europa.eu/en/publications-data/seroincidence-calculator-tool) was used. Briefly, this tool uses the combination of IgG, IgM and IgA values at a given point in time to estimate the time since seroconversion, thereby providing an estimate of the annual 'force of infection' for each individual using a Bayesian back-calculation model. (Janneke W Duijster et al., 2019; Teunis et al., 2013) The sero-incidence rate has been used in several studies as a proxy for human exposure to *Campylobacter*, regardless of symptoms (Janneke W Duijster et al., 2019; Monge et al., 2018; Teunis et al., 2013; Teunis et al., 2012) and was used as input for our analyses regarding infection pressure. Participants were excluded when information on age, sex or the residential address (postal code) was unknown.

Animal, human and spatial data

Data on the total number of farm animals per species (per farm location - Supplementary material, Table S1), human population numbers (per postal code) and the postal code region shapefiles were extracted from the same sources as described in Mulder, van de Kassteele, et al. (2020). Animals included in the analyses were small ruminants (sheep and goats), dairy cows, veal calves, laying hens, broiler chickens and pigs (Supplementary material, Text S3).

Methods

The statistical software environment R (version 4.1.3) (R-Core-Team, 2022) was used for data processing and analysis (Supplementary material, Table S2), unless described otherwise. R scripts can be found at: https://github.com/mulderac91/Salmonella_Campylobacter_spatial

Spatial risk factor analysis of human salmonellosis and campylobacteriosis cases

The spatial risk factor analysis for human salmonellosis cases (2007-2019) was performed for two different subsets of the cleaned salmonellosis database. The first subset contained the most frequent serovars *Salmonella* Enteritidis (SE) and *S.* Typhimurium (ST), including its monophasic variants (ST_{mv}). They were analyzed separately because they have their own primary animal source of infection (SE: laying hens; ST: pigs) (Mughini-Gras, Enserink, et al., 2014).

The second subset contained all human salmonellosis cases, thus including all *Salmonella* serovars. These serovars were grouped according to their primary source of infection, including laying hens, broiler chickens, cattle, and pigs (Mughini-Gras, Enserink, et al., 2014). The primary source of infection was based on a *Salmonella* source attribution analyses of the serotyping data. This analyses was performed within this study using a well-documented source attribution model: the modified Dutch model (Mughini-Gras, Barrucci, et al., 2014; Mughini-Gras, Enserink, et al., 2014; Mughini-Gras et al., 2016; Mughini-Gras et al., 2019) using the software @ Risk (Palisade Corp., USA). Supplementary material, Table S3 provides an overview of the primary sources per serovar and the attributions per serovar per source. Additional cases were removed if the human cases were caused by serovars not found in any of the included sources. Serovars with reptiles as primary source were either grouped based on their secondary source, or removed if there was no secondary source. (Supplementary material, Text S4). The spatial analyses then were performed per animal source.

The spatial risk factor analysis for human campylobacteriosis cases (2014-2019) was performed separately for *C. jejuni* and *C. coli*, as those species have their own main livestock sources of infection in the Netherlands (*C. jejuni*: broiler chickens and dairy cows; *C. coli*: broiler chickens and small ruminants) (Mughini-Gras et al., 2021).

All analyses were performed according to the method described in Mulder, van de Kassteele, et al. (2020). In short, risk factors were identified using a Poisson regression model with a log-link function that was fitted using integrated nested Laplace approximation (INLA). The dependent variable in the model was the case count, i.e. the redistributed number of salmonellosis or campylobacteriosis cases within a hexagon based on human population weighted interpolation. The hexagonal grid was used, because their nearest neighborhood is simpler and less ambiguous to identify, as hexagonal areas have six adjacent hexagons which are in a symmetrically equivalent position. In a nearest neighborhood analysis, there is therefore no need for the relative weighting of diagonal interactions, as is the case for rectangular grids. Furthermore, they are fixed over time (Birch et al., 2007; Birch et al., 2000;

Mulder, van de Kassteele, et al., 2020) Log-transformed person-years were used as the offset of the model. Confounders included were: age category (0–4, 5–9, 10–49 and ≥50 years old), sex (male, female) and period of infection (spring/summer: May–October (referred to as summer), autumn/winter: November–April (referred to as winter)). Hexagonal areas with different sizes (10, 25, 50 and 90 km^2) were used in combination with the population-weighted number of farm animals (i.e. dairy cows, veal calves, pigs, laying hens, broiler chickens and small ruminants) per hexagon. With this population weighted number of animals, human exposure was quantified based on the presence of farms in the neighborhood of residents, their corresponding animal numbers per animal type and the number of residents living in that area. (Mulder, van de Kassteele, et al., 2020). They were included as potential risk factors separately in univariate models, after which all animal types were included in the multivariate model as different farm animals are mixed throughout the landscape of the Netherlands and their effects should therefore not be assessed individually. Multicollinearity was assessed by calculating the Pearson's correlation coefficient per pair of animal types together with the variance inflation factor (VIF). Only animal pairs with a VIF below 5 were included in the final multivariate model. Maps showing their distribution are given in the Supplementary material, Figures S1-S12. Scatterplots were created for those transformed population-weighted number of animals and the log-transformed case-counts to check the linearity assumption for continuous variables. Standardized residual plots were used to assess model fit. Spatial auto-correlation was assessed by the addition of two random effect terms that addressed both the spatially structured variation (using the intrinsic Conditional Autoregressive model (CAR)) (Besag et al., 1991) and the unstructured variation (using the independent and identically distributed (IID) Gausian noise) (Lawson, 2013). Population attributable fractions (PAFs) were calculated only for risk factors showing consistent significant results over at least three of the four hexagonal area sizes. Briefly, we used the predicted incidence of *Salmonella/Campylobacter* in the exposed population using the corresponding regression model and its estimated coefficients as is, and the predicted incidence in an unexposed population. For this, the same regression model and coefficients were used, but the exposure of the risk factor was set to zero. (Mulder, van de Kassteele, et al., 2020)

The analyses were performed with and without correction for geographical coverage of the diagnostic laboratories (including only covered hexagonal areas by laboratories vs. all hexagonal areas, i.e. all of the Netherlands), as the surveillance system for both *Salmonella* and *Campylobacter* are sentinel, meaning that they have no nationwide coverage. See Supplementary material, Text S5 for more details. Results of the analysis with correction are the main focus of this study.

Risk factor analysis of human exposure to Campylobacter

To assess whether living in an area with high livestock densities influences the exposure to *Campylobacter*, data on sero-incidence was used together with the different population-weighted animal numbers (dairy cows, veal calves, pigs, laying hens, broiler chickens and small ruminants (goats and sheep). As the sero-incidence data only covers one year, it was not necessary to transpose the data into a hexagonal grid. Therefore, the population-weighted animal numbers were determined per postal code region (Mulder, van de Kassteele, et al., 2020). Similar to hexagonal areas, postal code regions can differ from one another based on an unknown or unobserved factor. This is why a linear mixed model was constructed in which the postal code regions were included as a random effect. We did not correct for spatially structured variation (CAR) in this analysis, because of the incomplete coverage of the Netherlands (only 9% of the Dutch municipalities were included in dataset). The sero-incidence was log-transformed (Monge et al., 2018) and the population-weighted animal numbers were subjected to the same $\log_2(x+1)$ transformation as they were in the previous analyses (Mulder, van de Kassteele, et al., 2020).

Results

Descriptive statistics

Salmonellosis case data

S. Enteritidis (SE) and *S.* Typhimurium (ST/ST_{mv})

In total, 3,734 SE and 5,368 ST/ST_{mv} cases were selected, which had an equal sex distribution (SE: 52%, ST/ST_{mv}: 51%), and a median age of 25 years for SE and 23 years for ST/ST_{mv}. The dominant period of infection was summer (SE: 72%, ST/ST_{mv}: 62%). Details are shown in Table 1. Figure 1 shows that the geographical distribution of the *Salmonella* incidence rate over the Netherlands varies per serovar. The incidence of SE (Figure 1A) is more homogeneously distributed over the Netherlands than the distribution of the incidence of ST/ST_{mv} (Figure 1B). The highest incidence of ST/ST_{mv} is found in the North-Eastern part of the country.

Grouped by source

In total, 4,153 cases were mainly attributable to laying hens, 1,377 to broiler chickens, 6,329 to pigs and 302 to cattle (Supplementary material: Text S2; Table 1). Figure 1 shows that the geographical distribution of the *Salmonella* incidence rate over the Netherlands varies per *Salmonella* source. The incidences of cases attributed to either cattle (Figure 1C) or broiler chickens (Figure 1E) were generally low (< 1.5 cases/100,000).

Table 1. Socio-demographic characteristics of the reported *Salmonella* Enteritidis (SE) and *Salmonella* Typhimurium (ST) cases, including cases from its monophasic variants (ST_{mv}), the reported salmonellosis cases grouped by source (laying hens, broiler chickens, pigs, and cattle), the reported *Campylobacter* (*C.*) *jejuni* and *C. coli* cases and an overview of the infection pressure dataset of *Campylobacter*.

	Salmonella												***Campylobacter***					
	per serovar				**per source**								**per species**				**Infection pressure data**	
	SE‡		**ST/ST_{mv}** ‡		**Laying hens**		**Broiler chickens**		**Pigs**		**Cattle**		***Jejuni***		***Coli***		**Overall**	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total	3,734	100	5,368	100	4,153	100	1,377	100	6,329	100	302	100	14,431	100	1,243	100	1,507	100
Sex																		
Male	1,807	48	2,632	49	1,974	48	623	45	3,085	49	162	54	7,729	54	592	48	661	44
Female	1,927	52	2,736	51	2,179	52	754	55	3,244	51	140	46	6,702	46	651	52	846	56
Age groups																		
0-4 years of age	545	15	930	17	570	14	153	11	982	16	41	14	650	5	20	2	175	12
5-9 years of age	344	9	670	12	347	8	48	3	680	11	24	8	406	3	26	2	117	8
10-49 years of age	1,819	49	2,040	38	2,019	49	680	49	2,372	37	90	30	7,114	49	540	43	966	64
50+ years of age	1,026	27	1,728	32	1,217	29	496	36	2,295	36	147	49	6,261	43	657	53	249	16
Period of infection†																		
Summer (may-oct)	2,686	72	3,345	62	2,933	71	937	68	3,949	62	141	47	9,179	64	842	68	616	41
Winter (nov-april)	1,048	28	2,023	38	1,220	29	440	32	2,380	38	161	53	5,252	36	401	32	891	59

†https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer

‡SE = *Salmonella* Enteritidis, ST/ST_{mv} = *Salmonella* Typhimurium, including its monophasic variants

Campylobacteriosis case data

In total, 14,431 campylobacteriosis cases were caused by *C. jejuni* and 1,243 cases were caused by *C. coli* (Supplementary material: Text S3; Table 1). Figure 2A shows that *C. jejuni* incidence is highest in the North-Eastern part of the country, whereas it is lowest in the North- and South-Western parts of the country. Overall, *C. coli* incidence (Figure 2B) is lower than *C. jejuni*'s, and highest in the Eastern part of the country.

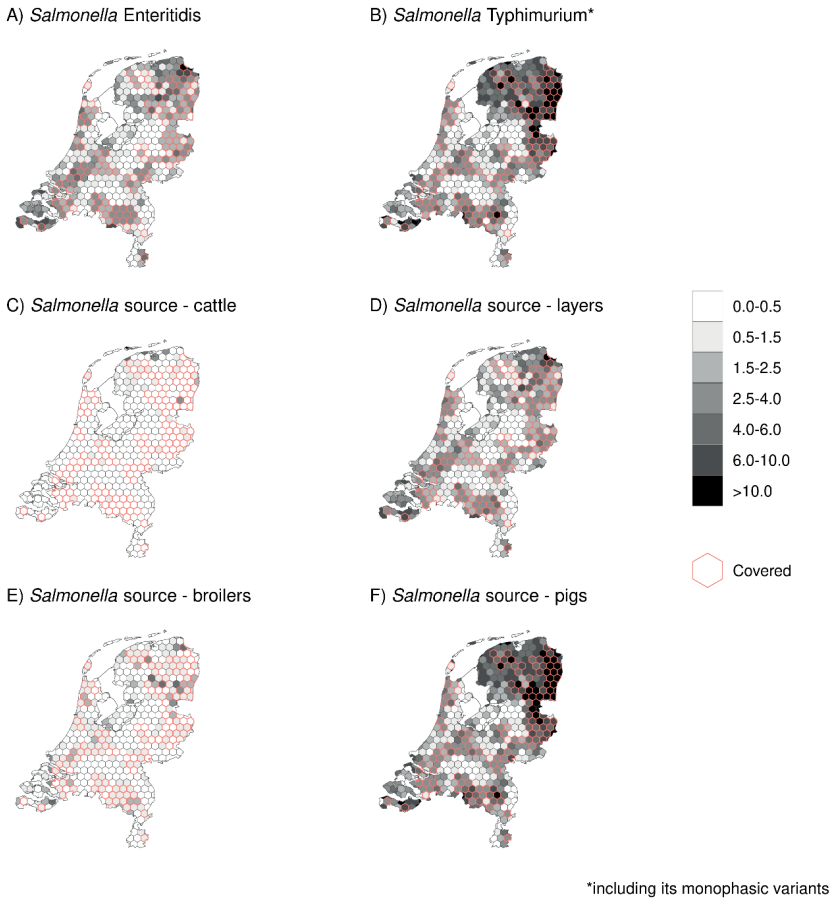


Figure 1. Spatial distribution of the cumulative incidence rate (x 100,000 person-years) (2007-2019) of salmonellosis cases in the Netherlands per serovar: A) *Salmonella* Enteritidis (SE), B) *Salmonella* Typhimurium (ST), including its monophasic variants (ST_{mv}), and grouped serovars per source: C) cattle, D) laying hens, E) broiler chickens, F) pigs, per (covered) hexagon.

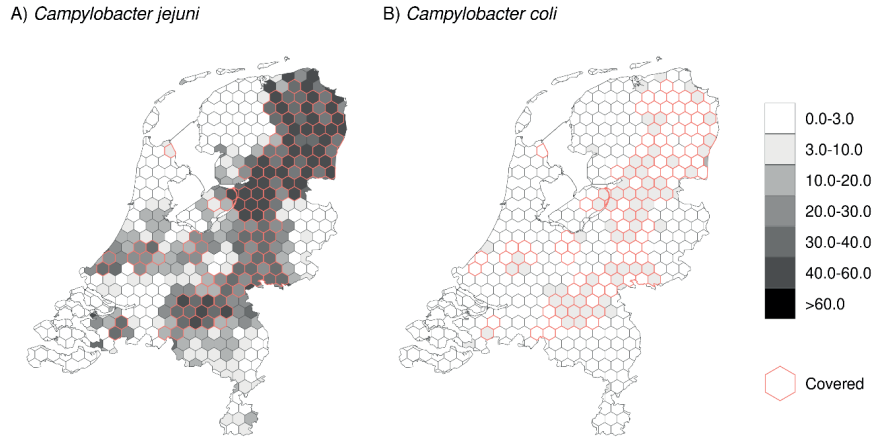


Figure 2. Spatial distribution of the cumulative incidence rate (x 100,000 person-years) (2014-2019) of campylobacteriosis cases in the Netherlands per species: A) *Campylobacter* (*C.*) *jejuni* and B) *C. coli*, per (covered) hexagon.

Campylobacter infection pressure data

Sero-incidence data for *Campylobacter* were available for 1,507 PIENTER-II study participants. Most participants were women (56%) with a median age of 29 years at the time of sampling. A map of the mean sero-incidence per postal code included in the PIENTER-II study is given in the Supplementary material, Figure S13.

Spatial risk factor analysis of case data

Significant results from the adjusted multivariable models for the spatial association between the population-weighted number of animals and SE or ST/STmv, all grouped *Salmonella* serovars (according to their primary sources), as well as *C. jejuni* and *C. coli* are presented in Tables 2, 3 and 4 per hexagonal area (10, 25, 50 and 90 km^2). All other results can be found in Supplementary material, Tables S4-S19.

Salmonella

S. Enteritidis and *S.* Typhimurium

The SE analyses (Table 2) showed a positive significant spatial association between laying hens and SE incidence in humans for a hexagonal area of 50 km^2 in summer and for areas of 90 km^2 in winter. A significant negative spatial association between dairy cows and human SE incidence was found for hexagonal areas of 10 km^2 in summer and winter (Table 2).

Table 2. Significant results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Salmonella* (*S.*) Enteritidis (SE), *S.* Typhimurium (ST) including its monophasic variants (ST_{mv}) and the population-weighted number of animals (small ruminants, dairy cows and laying hens) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), including the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
SE‡	**Age category (years)**						
	0-4	3.04***	2.61-3.53	2.84***	2.45-3.27	2.94***	2.57-3.35
	5-9	1.82***	1.51-2.17	1.56***	1.29-1.86	1.69***	1.44-1.99
	10-49				*Reference category*		
	50+	0.71***	0.63-0.81	0.69***	0.61-0.78	0.73***	0.66-0.81
	Type of animal†						
	Dairy cows	1.04	0.96-1.13	0.97	0.92-1.02	0.98	0.95-1.02
	Laying hens	1.02	1.00-1.04	1.02**	1.01-1.04	1.01	0.99-1.02
ST/ST_{mv} ‡	**Age category (years)**						
	0-4	3.95***	3.45-4.50	3.81***	3.36-4.33	3.78***	3.35-4.24
	5-9	2.93***	2.53-3.37	2.93***	2.56-3.36	2.78***	2.45-3.16
	10-49				*Reference category*		
	50+	0.94	0.85-1.05	0.96	0.87-1.07	0.92	0.83-1.01
	Type of animal†						
	Small ruminants	0.97	0.91-1.02	1.03	0.98-1.08	1.01	0.98-1.04
	Dairy cows	0.99	0.92-1.07	0.94*	0.90-0.99	1.01	0.97-1.04
	Laying hens	1.02	1.00-1.04	1.02**	1.01-1.04	1.01	1.00-1.02

† population-weighted number of animals

‡ IRR = incident rate ratio, 95% CI = 95% confidence interval, SE = *Salmonella* Enteritidis, ST/ST_{mv} = *Salmonella* Typhimurium, including its monophasic variants

* p-value <0.05

** p-value <0.01

*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
2.71***	2.38-3.09	2.92***	2.26-3.74	2.86***	2.24-3.63	3.04***	2.42-3.78	3.01***	2.42-3.72
1.63***	1.39-1.91	1.55**	1.12-2.10	1.59**	1.17-2.12	1.70***	1.28-2.22	1.75***	1.34-2.25
					Reference category				
0.69***	0.62-0.77	1.02	0.85-1.22	0.92	0.77-1.10	0.96	0.81-1.14	0.96	0.82-1.13
0.97*	0.94-1.00	0.98	0.89-1.08	0.96	0.89-1.03	0.97	0.92-1.02	0.95**	0.91-0.99
1.01	1.00-1.03	1.04**	1.01-1.07	1.01	0.99-1.04	1.00	0.98-1.02	1.00	0.97-1.02
3.72***	3.30-4.17	5.09***	4.29-6.03	4.73***	4.01-5.55	5.01***	4.31-5.81	5.09***	4.37-5.91
2.69***	2.37-3.06	3.37***	2.78-4.07	3.19***	2.66-3.82	3.06***	2.57-3.63	3.20***	2.69-3.79
					Reference category				
0.96	0.87-1.05	1.51***	1.32-1.73	1.39***	1.22-1.59	1.46***	1.29-1.65	1.48***	1.31-1.67
1.01	0.99-1.04	0.92**	0.85-0.98	0.95	0.90-1.00	0.98	0.94-1.01	0.99	0.96-1.02
0.98	0.96-1.01	0.98	0.89-1.07	1.01	0.95-1.06	1.00	0.96-1.04	0.96*	0.93-0.99
1.01*	1.00-1.03	1.01	0.98-1.03	1.02*	1.00-1.03	1.02*	1.00-1.03	1.01	0.99-1.02

The ST/ST_{mv} analyses (Table 2) showed a significantly protective spatial association between dairy cows and human ST/ST_{mv} incidence for 50 km^2 in summer and for 10 km^2 in winter. Furthermore, laying hens were a risk factor for acquiring a ST/ST_{mv} infection for 50 km^2 and 10 km^2 in summer and for 50 km^2 and 25 km^2 in winter (Table 2).

Salmonella serovars grouped by source
Small ruminants showed a significant positive association with cattle-associated *Salmonella* cases in summer in hexagonal areas of 90 km^2 and a significant protective association between pigs and cattle-associated *Salmonella* cases in summer (25 km^2) and winter (90 and 25 km^2) (Table 3).

Table 3 shows that laying hens were associated with salmonellosis caused by a pig-related serovar for 50 km^2 in summer and for 50 km^2 and 25 km^2 in winter. Furthermore, both dairy cows (summer: 50 km^2, winter: 90 km^2) and small ruminants (winter: 90 km^2) showed a significant protective effect of pig-related salmonellosis cases. Small ruminants in winter (10 km^2) and laying hens in summer (50 km^2) and winter (90 km^2) showed a significant positive association with layer-related salmonellosis cases (Table 3). A protective effect was observed in summer for dairy cows in hexagonal areas of 10 km^2 and pigs (90 km^2).

The analyses of salmonellosis cases caused by broiler-related serovars had most significant associations with the population-weighted animal numbers. Table 3 shows that broiler chickens had a positive association in winter in hexagonal areas of 90 km^2 and 50 km^2 and a negative association for 25 km^2. Furthermore, a negative spatial association between pigs in summer in hexagonal areas of 90, 50 and 25 km^2 and in winter (90 km^2), and broiler-related salmonellosis cases was observed. Other negative associations for the broiler-related serovars concerned dairy cows (summer: 10 km^2), veal calves (winter: 10 km^2) and laying hens (winter: 25 km^2) (Table 3).

Campylobacter

For both *C. jejuni* and *C. coli,* no consistently significant results were found for the spatial association between the population-weighted number of animals for dairy cows, veal calves, pigs, laying hens, broiler chickens and small ruminants (goats and sheep) and human campylobacteriosis cases over all four hexagonal area sizes after correcting for laboratory coverage (Table 4). However, dairy cows were a significant risk factor for campylobacteriosis caused by both *C. jejuni* and *C. coli* in summer in hexagonal areas of 50 km^2. Furthermore, small ruminants were positively associated with campylobacteriosis caused by *C. coli* in winter (10 km^2) (Table 4).

Associations with infection pressure

Similar to the case-based analysis, no consistently significant associations between *Campylobacter* sero-incidence and the different population-weighted number of animals were observed (Table 5). The estimates of the different animals were close to zero and varied only marginally according to the postal code random effect (Table 5). Results of the univariable analyses are given in Supplementary material, Table S20.

Table 3. Significant results (IRR and 95% CI) of the multivariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (cattle, pigs, laying hens and broiler chickens) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km^2), including the correction for geographical laboratory coverage.

		Summer					
		90 km^2		50 km^2		25 km^2	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Cattle-related	**Age category (years)**						
	0-4	3.60***	1.69-7.20	4.53***	2.29-8.56	3.22***	1.63-6.01
	5-9	1.01	0.27-2.95	0.32	0.03-1.78	1.27	0.46-2.99
	10-49	*Reference category*					
	50+	1.60	0.96-2.69	2.05**	1.28-3.33	1.72*	1.11-2.68
	Type of animal†						
	Small ruminants	0.77**	0.64-0.93	0.92	0.80-1.07	1.02	0.91-1.14
	Pigs	0.92	0.82-1.03	0.96	0.88-1.06	0.92*	0.85-0.99
Pig-related	**Age category (years)**						
	0-4	3.51***	3.09-3.99	3.37***	2.97-3.81	3.35***	2.99-3.75
	5-9	2.55***	2.21-2.93	2.60***	2.28-2.96	2.45***	2.16-2.77
	10-49	*Reference category*					
	50+	1.09	0.98-1.20	1.10*	1.00-1.21	1.04	0.96-1.14
	Type of animal†						
	Small ruminants	0.96	0.91-1.01	1.00	0.96-1.05	1.00	0.97-1.03
	Dairy cows	0.99	0.92-1.07	0.95*	0.91-1.00	1.00	0.97-1.04
	Laying hens	1.01	0.99-1.03	1.02*	1.00-1.03	1.01	1.00-1.02
Laying hen-related	**Sex**						
	Male	*Reference category*					
	Female	1.05	0.95-1.15	1.06	0.97-1.16	1.05	0.96-1.14
	Age category (years)						
	0-4	2.82***	2.42-3.26	2.62***	2.27-3.02	2.77***	2.43-3.15
	5-9	1.63***	1.35-1.94	1.42***	1.18-1.69	1.53***	1.30-1.79
	10-49	*Reference category*					
	50+	0.76***	0.67-0.85	0.72***	0.64-0.80	1.53***	1.30-1.79
	Type of animal†						
	Small ruminants	0.94	0.89-1.00	0.99	0.94-1.04	1.01	0.98-1.05
	Dairy cows	1.05	0.97-1.14	0.96	0.91-1.01	0.98	0.95-1.02
	Laying hens	1.02	1.00-1.04	1.02*	1.00-1.04	1.00	0.99-1.02
	Pigs	0.95*	0.91-1.00	0.98	0.94-1.01	0.98	0.96-1.01

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
3.07***	1.56-5.71	4.64***	2.20-9.28	7.17***	3.55-14.17	4.82***	2.64-8.49	5.34***	2.85-9.70
1.20	0.44-2.82	3.18**	1.36-6.81	3.82**	1.61-8.41	2.40*	1.10-4.80	3.78***	1.88-7.22
					Reference category				
1.67*	1.09-2.58	2.09**	1.25-3.58	2.83***	1.66-4.98	2.23***	1.46-3.44	2.64***	1.70-4.18
0.94	0.85-1.03	0.88	0.72-1.07	0.93	0.80-1.09	0.95	0.85-1.07	1.03	0.93-1.13
0.95	0.88-1.01	0.88*	0.79-0.98	0.96	0.87-1.05	0.93*	0.86-1.00	0.95	0.88-1.02
3.28***	2.93-3.66	4.74***	4.02-5.57	4.48***	3.83-5.23	4.72***	4.08-5.44	4.75***	4.10-5.48
2.35***	2.07-2.66	3.01***	2.49-3.62	2.81***	2.35-3.35	2.78***	2.34-3.28	2.81***	2.37-3.32
					Reference category				
1.07	0.98-1.17	1.77***	1.56-2.00	1.63***	1.45-1.83	1.71***	1.53-1.91	1.76***	1.57-1.96
1.00	0.98-1.02	0.94*	0.88-1.00	0.95	0.91-1.00	0.98	0.94-1.01	0.99	0.97-1.02
0.98	0.96-1.01	0.98	0.89-1.07	1.00	0.95-1.06	1.00	0.96-1.04	0.96**	0.93-0.99
1.01	1.00-1.02	1.01	0.99-1.03	1.02*	1.00-1.03	1.02*	1.00-1.03	1.01	1.00-1.03
					Reference category				
1.06	0.97-1.15	1.15	0.99-1.34	1.16*	1.00-1.34	1.16*	1.01-1.33	1.09	0.96-1.24
2.52***	2.22-2.87	2.80***	2.19-3.55	2.88***	2.28-3.61	2.98***	2.41-3.66	2.90***	2.35-3.55
1.43***	1.22-1.67	1.63**	1.20-2.17	1.64***	1.23-2.16	1.73***	1.33-2.22	1.75***	1.36-2.23
					Reference category				
0.72***	0.65-0.79	1.15	0.97-1.36	1.03	0.87-1.22	1.08	0.92-1.25	1.07	0.92-1.24
1.02	0.99-1.04	0.98	0.91-1.05	1.01	0.95-1.07	1.04	0.99-1.09	1.05***	1.01-1.09
0.97*	0.94-1.00	0.98	0.89-1.08	0.96	0.90-1.02	0.96	0.92-1.00	0.95	0.92-0.99
1.01	0.99-1.02	1.04**	1.01-1.06	1.01	0.99-1.03	1.00	0.98-1.02	0.99	0.97-1.01
0.99	0.97-1.01	0.97	0.92-1.02	0.96	0.92-1.00	1.00	0.97-1.03	0.99	0.97-1.02

		Summer					
		90 km²		50 km²		25 km²	
Salmonella-group	Variable	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Broiler-related	**Sex**						
	Male				*Reference category*		
	Female	1.09	0.92-1.30	1.14	0.97-1.35	1.18*	1.02-1.38
	Age category (years)						
	0-4	2.32***	1.75-3.04	2.57***	1.98-3.32	2.22***	1.74-2.82
	5-9	0.77	0.49-1.17	0.70	0.44-1.07	0.76	0.51-1.09
	10-49				*Reference category*		
	50+	0.89	0.73-1.08	0.94	0.78-1.13	0.86	0.73-1.02
	Type of animal†						
	Dairy cows	0.98	0.88-1.09	0.95	0.88-1.02	0.99	0.94-1.04
	Veal calves	1.02	0.92-1.13	1.03	0.95-1.12	1.01	0.95-1.07
	Laying hens	1.01	0.98-1.04	1.01	0.99-1.04	0.99	0.97-1.01
	Broiler chickens	1.01	0.99-1.04	1.01	0.99-1.04	1.01	0.98-1.03
	Pigs	0.92**	0.87-0.97	0.93**	0.89-0.98	0.95**	0.92-0.99

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
1.18*	1.02-1.37	1.12	0.87-1.44	1.18	0.93-1.51	1.14	0.91-1.43	1.09	0.88-1.35
2.39***	1.87-3.04	1.86**	1.16-2.88	1.85**	1.15-2.86	1.80**	1.17-2.67	1.88**	1.25-2.73
0.76	0.51-1.11	0.24*	0.07-0.66	0.39*	0.15-0.88	0.32*	0.12-0.71	0.47*	0.22-0.91
					Reference category				
1.05	0.89-1.23	1.24	0.95-1.04	1.39*	1.08-1.80	1.13	0.89-1.44	1.18	0.92-1.46
0.95*	0.91-1.00	1.03	0.89-1.19	1.01	0.92-1.11	1.03	0.95-1.11	0.98	0.92-1.04
0.98	0.93-1.04	0.98	0.86-1.12	0.93	0.84-1.04	0.99	0.91-1.08	0.93*	0.86-1.00
1.00	0.97-1.02	0.99	0.95-1.04	0.99	0.96-1.03	0.96*	0.93-0.99	0.97	0.94-1.01
1.00	0.97-1.02	1.07**	1.03-1.11	1.03*	1.00-1.07	0.96*	0.93-0.99	1.01	0.98-1.04
0.99	0.96-1.02	0.92*	0.86-0.99	0.97	0.91-1.03	0.97	0.92-1.02	1.00	0.96-1.05

Table 4. Significant results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Campylobacter* (*C.*) *jejuni* and *C. coli* incidence rate and the population-weighted number of animals (small ruminants and dairy cows) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km^2), including the correction for geographical laboratory coverage.

		Summer					
		90 km^2		50 km^2		25 km^2	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. jejuni	**Sex**						
	Male				*Reference category*		
	Female	0.83***	0.79-0.88	0.82***	0.78-0.86	0.83***	0.80-0.88
	Age category (years)						
	0-4	0.80***	0.70-0.91	0.75***	0.65-0.85	0.78***	0.69-0.88
	5-9	0.48***	0.41-0.56	0.47***	0.40-0.55	0.46***	0.39-0.53
	10-49				*Reference category*		
	50+	1.11***	1.05-1.17	1.09***	1.04-1.15	1.11***	1.05-1.16
	Type of animal†						
	Dairy cows	1.01	0.95-1.06	1.04*	1.01-1.08	0.99	0.97-1.02
C. coli	**Age category (years)**						
	0-4	0.20***	0.08-0.45	0.18***	0.07-0.41	0.27***	0.13-0.52
	5-9	0.65	0.38-1.03	0.48**	0.27-0.79	0.51**	0.30-0.82
	10-49				*Reference category*		
	50+	1.48***	1.25-1.77	1.40***	1.19-1.66	1.34***	1.14-1.58
	Type of animal†						
	Small ruminants	0.94	0.85-1.03	1.00	0.93-1.06	1.00	0.95-1.06
	Dairy cows	0.99	0.89-1.10	1.11**	1.03-1.19	1.04	0.99-1.11

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
0.84***	0.80-0.88	0.87***	0.81-0.93	0.89***	0.83-0.95	0.88***	0.82-0.93	0.87***	0.82-0.93
0.73***	0.65-0.83	0.98	0.83-1.14	1.00	0.85-1.16	0.96	0.82-1.12	0.97	0.84-1.13
0.46***	0.40-0.54	0.49***	0.39-0.60	0.50***	0.40-0.61	0.50***	0.41-0.61	0.55***	0.45-0.66
					Reference category				
1.09***	1.04-1.15	1.20***	1.12-1.28	1.21***	1.13-1.30	1.22***	1.14-1.31	1.22***	1.14-1.30
0.98	0.97-1.00	1.01	0.96-1.08	1.01	0.98-1.05	0.99	0.96-1.02	0.99	0.97-1.01
0.27***	0.13-0.51	0.42	0.14-1.04	0.30*	0.09-0.85	0.10*	0.01-0.54	0.09*	0.01-0.47
0.44**	0.25-0.73	0.50	0.19-1.12	0.57	0.23-1.21	0.56	0.23-1.19	0.33*	0.11-0.81
					Reference category				
1.01	0.98-1.03	2.19***	1.70-2.85	2.21***	1.72-2.86	2.15***	1.67-2.76	2.12***	1.68-2.69
0.99	0.95-1.03	1.06	0.93-1.21	1.02	0.93-1.13	1.06	0.98-1.15	1.09**	1.03-1.15
1.01	0.97-1.06	0.94	0.82-1.07	0.98	0.88-1.09	0.98	0.91-1.07	0.98	0.92-1.03

Table 5. Results (fixed effect estimates (β) and 95% CI) of the multivariable linear mixed model of *Campylobacter* seroincidence and the population-weighted animal numbers (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs).

Variable	β	95% CI
Age (years)		
0-4	-0.30***	-0.42;-0.18
5-9	-0.34***	-0.48;-0.19
10-49	*Reference*	
≥50	0.07	-0.04;0.17
Sex		
Male		
Female	0.06	-0.02;0.13
Type of animal†		
Small ruminants	0.00	-0.01;0.02
Dairy cows	0.00	-0.02;0.02
Veal calves	-0.01	-0.03;0.01
Laying hens	0.01	-0.00;0.02
Broiler chickens	-0.01	-0.03;0.01
Pigs	< 0.005	-0.01;0.01

† population-weighted number of animals
‡ β = fixed effect estimate, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Discussion

The main aim of this study was to assess whether there is a significant spatial association between human salmonellosis or campylobacteriosis incidence and the local level of exposure to different livestock (i.e. population-weighted number of animals) in the Netherlands. Additionally, the effect of such exposure on the *Campylobacter* infection pressure, as described by serology, was assessed.

Overall, no consistently significant associations with a given livestock were found across hexagonal areas with different sizes and periods of infection. This suggests that residents living in areas with a high number of farm animals do not have a significantly increased risk of acquiring a *Salmonella* or *Campylobacter* infection. Results regarding age, sex and period of infection were comparable to previously published studies, as young children and elderly have the highest risk of infection, most human cases are reported in the Netherlands in summer (Acheson & Hohmann, 2001; van Pelt et al., 2003), and exposure can differ per gender. (J. W. Duijster et al., 2019)

Our findings indicate that the potential for environmental transmission of *Salmonella* and *Campylobacter* to Dutch residents is likely to be limited. This could be explained by limited exposure of Dutch residents to viable *Salmonella* or *Campylobacter* present in the environment, e.g. in agricultural waters, which had the highest *Campylobacter* prevalence (77%) of all Dutch surface water types (Mulder, Franz, et al., 2020). Alternatively, there could be limited dispersal of the pathogens from farms into the environment. Indeed, it is possible that the generally high biosecurity levels of Dutch farms (especially pig and broiler farms) relative to other European countries would not only prevent pathogens from entering a farm, but also from being released into the environment outside the farm, or at least reduce the load (Caekebeke et al., 2020; Filippitzi et al., 2018). However, these interpretations should be read with caution, as the complexity of environmental transmission is not easy to capture in a densely populated country like the Netherlands.

The complexity of the interactions between pathogens, their reservoirs and human hosts was also visible in our results, as varying significance levels were found for different risk factors of both salmonellosis and campylobacteriosis for hexagonal areas with different sizes. Possibly, there is an interaction between area size, exposure and incidence and it could be that a different relationship was estimated in a smaller vs. larger area. This could lead to varying results over the different spatial scales and indicates that the interpretation of the results changes over different hexagonal area sizes; thus, different confounders and causal relations could play a role when studying those larger areas. This effect is known as the modifiable areal unit problem (MAUP) which is closely related to ecological fallacy and ecological bias (Jensen & Jensen, 2013; Robinson, 2009).

As both salmonellosis and campylobacteriosis are not notifiable diseases in the Netherlands, a geographical laboratory coverage correction was applied to the national surveillance data in order to minimize potential bias due to inclusion of geographical areas that are not covered by the surveillance system. Because we used relatively strict criteria, this could have led to the exclusion of postal code areas that have a significantly lower than expected number of cases (one of the criteria), but are covered in reality. Therefore, analyses were repeated with and without this correction as a sensitivity analysis. Overall, results were comparable, but the analyses corrected for laboratory coverage found less animal groups to be significantly associated with salmonellosis and campylobacteriosis incidence. The only results that differed consistently between analyses with and without correction were the ones for SE. In the analyses without correction, laying hens were significantly associated with human SE incidence in three out of four hexagonal area sizes (90, 25 and 10 km^2) in summer instead of only one (50 km^2) in the analysis with correction. The model without

correction for an hexagonal area of 90 km^2 showed a significant IRR of 1.02 for laying hens with a corresponding PAF of 7% (Supplementary material, Table S16). If all possible risk factors were included, of which food is the most important one, this PAF will decrease even further. This indicates that if laying hens do play a role as a spatially restricted risk factor for SE infections, this is most likely to be minor.

Interestingly, only one hexagonal area size showed significant results for laying hens in summer in the analysis including the correction, as laying hens are known to be the main source of SE. The population-weighted number of animal maps in Figures S1-S12 of the Supplementary material show that areas with for example a high number of laying hens (Figure S7 and S8) are excluded after the correction. This means that the analysis including the correction is missing one of the extremes of the population-weighted number of laying hens potentially leading to an underestimation of the IRR. Considering that laying hens are not only kept indoors like broiler chickens, but can also be housed in an organic or free-range farm (Bestman et al., 2019), it is plausible that there might be some degree of environment-mediated transmission of SE via the direct surroundings of the farms through fecal contamination, surface water run-off, contaminated soil or airborne (dust) transmission towards residents living nearby those farms. However, this can currently only be hypothesized and would require data on human SE incidence in those areas to be evidenced and further studied. It is also possible that livestock, wildlife and surface water are important from an interactive perspective on human *Campylobacter* and *Salmonella* incidence. (Hazeleger et al., 2018; Mughini-Gras et al., 2021; Mulder, Franz, et al., 2020) It is therefore recommended to include this possible interaction in future studies.

Similar to the case-based analysis for *Campylobacter*, no consistent significant associations were found when analyzing *Campylobacter* exposure using serology data. It is possible that residents were frequently exposed to *Campylobacter* through other routes, such as food consumption. This could lead to limited discriminatory capacity to identify risk factors due to a large proportion of the population having high antibody levels (Monge et al., 2018). Furthermore, it was not possible to discriminate between *C. jejuni* and *C. coli*, while their main sources differ. Despite of those limitations, the results of the *Campylobacter* infection pressure analyses suggest the same as the case-based analyses, i.e. that there is no strong association between *Campylobacter* sero-incidence and farm animals as spatially restricted risk factors. This leaves foodborne transmission to be the main pathway of human exposure to the pathogens considered here.

Acknowledgements

The authors would like to thank Ben Bom for his valuable input about the PIENTER-II study and for providing data on the total number of farm animals per species reared

in each farm in the Netherlands and Maarten Schipper for his help with the creation of the figures. We also want to thank ISIS-AR, the ISIS-AR study group and participating medical microbiological laboratories, as well as Wieke Altorf-van der Kuil for her role in data extraction. This study was supported by the Netherlands' Organization for Health Research and Development (ZonMw) with grant number 50-52200-98-316 (project name: "DEPiCT – Discerning Environmental Pathways of Campylobacter Transmission") .

Conflict of Interest Statement

The authors declare no conflicts of interest relevant to this study.

Data availability statement

The used R scripts for data analyses can be found at: https://github.com/mulderac91/Salmonella_Campylobacter_spatial. Livestock data of food-producing animals for 2012, 2015 and 2018 was obtained from the Department of Service Arrangements of the Dutch Ministry of Agriculture, Nature and Food Quality (CBS, 2021b; RVO, 2021). The population data per four-digit postal code region per year is available through Statistics Netherlands (www.statline.nl). The data were downloaded for the years 2006–2019 (CBS, 2021a).

The four-digit postal code region shapefiles and the six-digit postal code point locations of the Netherlands that were used within this study were obtained by the RIVM from the company: Iris International. Those shapefiles can only be given to those for whom permission has been granted by this company. They can be reached at this address: Gr.v. Prinstererlaan 20, 2,271 EN, Voorburg, the Netherlands. Tel: +31(0)70-3863891, fax: +31(0)70–3873625, e-mail: info@iris-int.nl.

The national surveillance data of *Salmonella* is available for people within the network of Regional Public Health Laboratories (RPHLs) and researchers within the RIVM with access to this network. Case reports of *Campylobacter* from the RPHLs can be accessed by researchers within the RIVM with access to the Infectious Disease Surveillance and Information System for Antibiotic Resistance (ISIS-AR). The PIENTER-II data is available to researchers working within the PIENTER-II project from the PIENTER-II database. Those three datasets contain privacy sensitive information of cases and are therefore not accessible to the public or research community following the legislation of the Dutch law and the Dutch Data Protection Authority (Dutch Data Protection Authority, 2021). However, anonymized data are available from the authors upon reasonable request.

References

Acheson, D., & Hohmann, E. L. (2001). Nontyphoidal Salmonellosis. Clinical Infectious Diseases, 32(2), 263-269. doi: 10.1086/318457

Altorf-van der Kuil, W., Schoffelen, A. F., de Greeff, S. C., Thijsen, S. F., Alblas, H. J., Notermans, D. W., Vlek, A. L., van der Sande, M. A., Leenstra, T., & Group, t. N. A. S. S. (2017). National laboratory-based surveillance system for antimicrobial resistance: a successful tool to support the control of antimicrobial resistance in the Netherlands. Eurosurveillance, 22(46), 17-00062. doi: doi:https://doi.org/10.2807/1560-7917.ES.2017.22.46.17-00062

Andri et mult al 2022. DescTools: Tools for Descriptive Statistics. https://cran.r-project.org/web/packages/DescTools/index.html.

Ang, C. W., Krogfelt, K., Herbrink, P., Keijser, J., van Pelt, W., Dalby, T., Kuijf, M., Jacobs, B. C., Bergman, M. P., Schiellerup, P., & Visser, C. E. (2007). Validation of an ELISA for the diagnosis of recent Campylobacter infections in Guillain-Barré and reactive arthritis patients. Clin Microbiol Infect, 13(9), 915-922. doi: 10.1111/j.1469-0691.2007.01765.x

Arya, S., Mount, D., Kemp, S.E. and Jefferis, G. 2019. RANN: Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric. https://CRAN.R-project.org/package=RANN.

Bakka, H., Rue, H., Fuglstad, G., Riebler, A., Bolin, D., Krainski, E., Simpson, D. and Lindgren, F. 2018. Spatial modelling with R-INLA: A review. WIREs computational statistics 10(06), 1-24, 10.1002/wics.1443.

Bates, D., Maechler, M., Davis, T.A., Oehlschlägel, J. and Riedy, J. 2021. Matrix: Sparse and Dense Matrix Classes and Methods. https://CRAN.R-project.org/package=Matrix.

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. Annals of the institute of statistical mathematics, 43(1), 1-20. doi: https://doi.org/10.1007/BF00116466

Bestman, M., Verwer, C., van Niekerk, T., Leenstra, F., Reuvekamp, B., Amsler-Kepalaite, Z., & Maurer, V. (2019). Factors related to free-range use in commercial laying hens. Applied Animal Behaviour Science, 214, 57-63. doi: 10.1016/j.applanim.2019.02.015

Birch, C. P. D., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecological Modelling, 206(3), 347-359. doi: https://doi.org/10.1016/j.ecolmodel.2007.03.041

Birch, C. P. D., Vuichard, N., & Werkman, B. R. (2000). Modelling the effects of patch size on vegetation dynamics: Bracken [Pteridium aquilinum (L.) Kuhn] under grazing. Annals of Botany, 85, 63-76. doi: https://doi.org/10.1006/anbo.1999.1081

Bivand, R.S., Pebesma, E. and Gomez-Rubio, V. 2013. Applied spatial data analysis with R, Second edition. https://asdar-book.org/.

Brachman, P. S., & Evans, A. S. (1998). Bacterial infections of humans (3th ed.). Springer.

Caekebeke, N., Jonquiere, F. J., Ringenier, M., Tobias, T. J., Postma, M., van den Hoogen, A., Houben, M. A. M., Velkers, F. C., Sleeckx, N., Stegeman, J. A., & Dewulf, J. (2020). Comparing Farm Biosecurity and Antimicrobial Use in High-Antimicrobial-Consuming Broiler and Pig Farms in the Belgian–Dutch Border Region. Frontiers in Veterinary Science, 7, 1-11. doi: 10.3389/fvets.2020.558455

Canty, A. and Ripley, B.D. 2021. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28. https://www.rdocumentation.org/packages/boot/versions/1.3-28/topics/boot.

CBS. (2021a, December 05). Bevolking; geslacht, leeftijd en viercijferige postcode, 1 januari. Retrieved from https://opendata.cbs.nl/#/CBS/nl/dataset/83502NED/table.

CBS. (2021b, December 04). Landbouwtelling. Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/landbouwtelling.

De Jonge, E. and Houweling, S. 2021. cbsodataR: Statistics Netherlands (CBS) Open Data API Client.

https://edwindj.github.io/cbsodataR/.

de Rooij, M. M. T., Hoek, G., Schmitt, H., Janse, I., Swart, A., Maassen, C. B. M., Schalk, M., Heederik, D. J. J., & Wouters, I. M. (2019). Insights into Livestock-Related Microbial Concentrations in Air at Residential Level in a Livestock Dense Area. Environmental Science & Technology, 53(13), 7746-7758. doi: 10.1021/acs.est.8b07029

Dragulescu, A. and Arendt, C. 2020. xlsx-package: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 Files. R package version 0.6.5.

Duijster, J. W., Franz, E., Neefjes, J. J. C., & Mughini-Gras, L. (2019). Occupational risk of salmonellosis and campylobacteriosis: a nationwide population-based registry study. Occup Environ Med, 76(9), 617-624. doi: 10.1136/oemed-2019-105868

Duijster, J. W., Franz, E., Neefjes, J. J. C., & Mughini-Gras, L. (2019). Occupational risk of salmonellosis and campylobacteriosis: a nationwide population-based registry study. Occupational and Environmental Medicine, 76(9), 617-624. doi: 10.1136/oemed-2019-105868

Dutch Data Protection Authority. (2021, October 23). Retrieved from https://www.autoriteitpersoonsgegevens.nl/en.

Filippitzi, M. E., Brinch Kruse, A., Postma, M., Sarrazin, S., Maes, D., Alban, L., Nielsen, L. R., & Dewulf, J. (2018). Review of transmission routes of 24 infectious diseases preventable by biosecurity measures and comparison of the implementation of these measures in pig herds in six European countries. Transboundary and Emerging Diseases, 65(2), 381-398. doi: https://doi.org/10.1111/tbed.12758

Friesema, I., Pijnacker, R., Tulen, R., van den Beld, M., Mughini Gras, L., Bosch, T., & Franz, E. (2021). Jaarrapportage surveillance gastro-intestinale infecties en zoönosen (Annual report on gastrointestinal infection and zoonoses surveillance). doi: 10.21945/rivm-2021-0188

Friesema, I. H. M., Havelaar, A. H., Westra, P. P., Wagenaar, J. A., & van Pelt, W. (2012). Poultry culling and Campylobacteriosis reduction among humans, the Netherlands. Emerging infectious diseases, 18(3), 466-468. doi: 10.3201/eid1803.111024

Friesema, I. H. M., Slegers-Fitz-James, I. A., Wit, B., & Franz, E. (2022). Surveillance and characteristics of food-borne outbreaks in the Netherlands, 2006 to 2019. Eurosurveillance, 27(3). doi: 10.2807/1560-7917.ES.2022.27.3.2100071

Grolemund, G. and Wickham, H. 2011. Dates and times made easy with lubridate. Journal of Statistical Software 40(3), 1-25, 10.18637/jss.v040.i03.

Guillier, L., Thébault, A., Fravalo, P., Mughini-Gras, L., Jourdan-da Silva, N., David, J., Kooh, P., Cadavez, V., & Gonzales-Barron, U. (2021). Risk factors for sporadic salmonellosis: a systematic review and meta-analysis. Microbial Risk Analysis, 17, 1-11. doi: 10.1016/j.mran.2020.100138

Hazeleger, W. C., Jacobs-Reitsma, W. F., Lina, P. H. C., de Boer, A. G., Bosch, T., van Hoek, A. H. A. M., & Beumer, R. R. (2018). Wild, insectivorous bats might be carriers of Campylobacter spp. PLOS ONE, 13(1), e0190647. doi: 10.1371/journal.pone.0190647

Jensen, J. R., & Jensen, R. R. (2013). Introductory geographic information systems. Pearson Education, Inc.

Keitt, T.H. 2021. rgdal: Bindings for the Geospatial Data Abstraction Library, R package version 1.5-27. http://cran.r-project.org/package=rgdal.

Lawson, A. B. (2013). Bayesian disease mapping: hierarchical modeling in spatial epidemiology. Chapman and Hall/CRC.

Lee, D., Tertuliano, M., Harris, C., Vellidis, G., Levy, K., & Coolong, T. (2019). Salmonella Survival in Soil and Transfer onto Produce via Splash Events. J Food Prot, 82(12), 2023-2037. doi: 10.4315/0362-028x.Jfp-19-066

Mollema, L., De Melker, H., Hahné, S., Van Weert, J., Berbers, G., & Van Der Klis, F. (2010). PIENTER 2-project: second research project on the protection against infectious diseases offered by the national immunization programme in the Netherlands (RIVM report 230421001). doi:

Monge, S., Teunis, P., Friesema, I., Franz, E., Ang, W., van Pelt, W., & Mughini-Gras, L. (2018). Immune response-eliciting exposure to Campylobacter vastly exceeds the incidence of clinically overt campylobacteriosis but is associated with similar risk factors: A nationwide serosurvey in the Netherlands. Journal of Infection, 77(3), 171-177. doi: https://doi.org/10.1016/j.jinf.2018.04.016

Mughini-Gras, L., Barrucci, F., Smid, J. H., Graziani, C., Luzzi, I., Ricci, A., Barco, L., Rosmini, R., Havelaar, A. H., Van Pelt, W., & Busani, L. (2014). Attribution of human Salmonella infections to animal and food sources in Italy (2002–2010): adaptations of the Dutch and modified Hald source attribution models. Epidemiology and Infection, 142(5), 1070-1082. doi: 10.1017/S0950268813001829

Mughini-Gras, L., Enserink, R., Friesema, I., Heck, M., van Duynhoven, Y., & van Pelt, W. (2014). Risk Factors for Human Salmonellosis Originating from Pigs, Cattle, Broiler Chickens and Egg Laying Hens: A Combined Case-Control and Source Attribution Analysis. PLOS ONE, 9(2), e87933. doi: 10.1371/journal.pone.0087933

Mughini-Gras, L., Heck, M., & van Pelt, W. (2016). Increase in reptile-associated human salmonellosis and shift toward adulthood in the age groups at risk, the Netherlands, 1985 to 2014. Euro Surveill, 21(34), 1-5. doi: 10.2807/1560-7917.Es.2016.21.34.30324

Mughini-Gras, L., Kooh, P., Fravalo, P., Augustin, J.-C., Guillier, L., David, J., Thébault, A., Carlin, F., Leclercq, A., Jourdan-Da-Silva, N., Pavio, N., Villena, I., Sanaa, M., & Watier, L. (2019). Critical Orientation in the Jungle of Currently Available Methods and Types of Data for Source Attribution of Foodborne Diseases [Review]. Frontiers in Microbiology, 10, 1-22. doi: 10.3389/fmicb.2019.02578

Mughini-Gras, L., Pijnacker, R., Coipan, C., Mulder, A. C., Fernandes Veludo, A., de Rijk, S., van Hoek, A., Buij, R., Muskens, G., Koene, M., Veldman, K., Duim, B., van der Graaf-van Bloois, L., van der Weijden, C., Kuiling, S., Verbruggen, A., van der Giessen, J., Opsteegh, M., van der Voort, M., . . . Franz, E. (2021). Sources and transmission routes of campylobacteriosis: A combined analysis of genome and exposure data. J Infect, 82(2), 216-226. doi: 10.1016/j.jinf.2020.09.039

Mughini-Gras, L., Pijnacker, R., Duijster, J., Heck, M., Wit, B., Veldman, K., & Franz, E. (2020). Changing epidemiology of invasive non-typhoid Salmonella infection: a nationwide population-based registry study. Clinical Microbiology and Infection, 26(7), 941.e949-941.e914. doi: https://doi.org/10.1016/j.cmi.2019.11.015

Mulder, A. C., Franz, E., de Rijk, S., Versluis, M. A. J., Coipan, C., Buij, R., Müskens, G., Koene, M., Pijnacker, R., Duim, B., Bloois, L. v. d. G.-v., Veldman, K., Wagenaar, J. A., Zomer, A. L., Schets, F. M., Blaak, H., & Mughini-Gras, L. (2020). Tracing the animal sources of surface water contamination with Campylobacter jejuni and Campylobacter coli. Water Research, 187, 1-12. doi: https://doi.org/10.1016/j.watres.2020.116421

Mulder, A. C., van de Kassteele, J., Heederik, D., Pijnacker, R., Mughini-Gras, L., & Franz, E. (2020). Spatial Effects of Livestock Farming on Human Infections With Shiga Toxin-Producing Escherichia coli O157 in Small but Densely Populated Regions: The Case of the Netherlands. GeoHealth, 4(11), 1-14. doi: https://doi.org/10.1029/2020GH000276

Neuwirth, E. 2015. Package 'RColorBrewer'. https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf.

Nicholson, F. A., Groves, S. J., & Chambers, B. J. (2005). Pathogen survival during livestock manure storage and following land application. Bioresource Technology, 96(2), 135-143. doi: https://doi.org/10.1016/j.biortech.2004.02.030

Ooms, J. and McNamara, J. 2021. Package 'writexl'. https://cran.r-project.org/web//packages/writexl/writexl.pdf.

Pasternack, M. S. (2002). Impact and management of Campylobacter in human medicine-US perspective. International journal of infectious diseases, 6, S37-S43. doi: 10.1016/S1201-9712(02)90182-7

Pebesma, E. 2019. st_make_grid. https://www.rdocumentation.org/packages/sf/versions/0.8-0/topics/st_make_grid.

Pebesma, E., Bivand, R., Racine, E., Sumner, M., Cook, I., Keitt, T., Lovelace, R., Wickham, H., Ooms, J., Müller, K., Pedersen, T.L. and Baston, D. 2019a. sf: Simple Features for R. https://cran.r-project.org/web/packages/sf/index.html.

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., Hijmans, R., Sumner, M., MacQueen, D., Lemon, J., O'Brian, J. and O'Rourke, J. 2019b. sp: Classes and Methods for Spatial Data. https://cran.r-project.org/web/packages/sp/index.html

R-Core-Team 2021. parallel v3.6.2. https://www.rdocumentation.org/packages/parallel/versions/3.6.0.

R-Core-Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. R-Core-Team. Retrieved from https://www.R-project.org.

Robinson, W. (2009). Ecological Correlations and the Behavior of Individuals*. International Journal of Epidemiology, 38(2), 337-341. doi: 10.1093/ije/dyn357

RVO. (2021). Gecombineerde opgave. Retrieved 04-12-2021, Retrieved from Retrieved from

Schets, F. M., Jacobs-Reitsma, W. F., van der Plaats, R. Q. J., Heer, L. K.-D., van Hoek, A. H. A. M., Hamidjaja, R. A., de Roda Husman, A. M., & Blaak, H. (2017). Prevalence and types of Campylobacter on poultry farms and in their direct environment. Journal of Water and Health, 15(6), 849-862. doi: 10.2166/wh.2017.119

Smit, L. A. M., & Heederik, D. (2017). Impacts of Intensive Livestock Production on Human Health in Densely Populated Regions. GeoHealth, 1(7), 272-277. doi: https://doi.org/10.1002/2017GH000103

Teunis, P., Falkenhorst, G., Ang, C., Strid, M., De Valk, H., Sadkowska-Todys, M., Zota, L., Kuusi, M., Rota, M., & Simonsen, J. (2013). Campylobacter seroconversion rates in selected countries in the European Union. Epidemiology & Infection, 141(10), 2051-2057. doi: 10.1017/S0950268812002774

Teunis, P., van Eijkeren, J., Ang, C., van Duynhoven, Y., Simonsen, J., Strid, M., & van Pelt, W. (2012). Biomarker dynamics: estimating infection rates from serological data. Statistics in Medicine, 31(20), 2240-2248. doi: https://doi.org/10.1002/sim.5322

van Pelt, W., de Wit, M. A., Wannet, W. J., Ligtvoet, E. J., Widdowson, M. A., & van Duynhoven, Y. T. (2003). Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991-2001. Epidemiol Infect, 130(3), 431-441. doi:

Vlaanderen, F., Uiterwijk, M., Cuperus, T., Keur, I., de Rosa, M., Rozendaal, H., Koene, M., Schreurs, H., Nijsse, R., Nielen, M., Friesema, I., van Pelt, W., Franz, E., Hogerwerf, L., Opsteegh, M., & Maassen, K. (2019). Staat van Zoönosen 2018. In State of Zoonotic Diseases 2018: Rijksinstituut voor Volksgezondheid en Milieu RIVM.

Wickham, H. 2019. stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr.

Wickham, H. 2021. tidyr: Tidy Messy Data. R package version 1.1.4. https://CRAN.R-project.org/package=tidyr.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L. and Hester, J. 2019. Welcome to the Tidyverse. Journal of Open Source Software 4(43), 1686, 10.21105/joss.01686.

Wickham, H. and Bryan, J. 2019. readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl.

Wickham, H., Francois, R., Henry, L. and Müller, K. 2022. dplyr: A Grammar of Data Manipulation. R package version 1.0.9. https://CRAN.R-project.org/package=dplyr.

Andri et mult al 2022. DescTools: Tools for Descriptive Statistics. https://cran.r-project.org/web/packages/DescTools/index.html.

Arya, S., Mount, D., Kemp, S.E. and Jefferis, G. 2019. RANN: Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric. https://CRAN.R-project.org/package=RANN.

Bakka, H., Rue, H., Fuglstad, G., Riebler, A., Bolin, D., Krainski, E., Simpson, D. and Lindgren, F. 2018. Spatial modelling with R-INLA: A review. WIREs computational statistics 10(06), 1-24, 10.1002/wics.1443.

Bates, D., Maechler, M., Davis, T.A., Oehlschlägel, J. and Riedy, J. 2021. Matrix: Sparse and Dense Matrix Classes and Methods. https://CRAN.R-project.org/package=Matrix.

Bivand, R.S., Pebesma, E. and Gomez-Rubio, V. 2013. Applied spatial data analysis with R, Second edition. https://asdar-book.org/.

Canty, A. and Ripley, B.D. 2021. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28. https://www.rdocumentation.org/packages/boot/versions/1.3-28/topics/boot.

De Jonge, E. and Houweling, S. 2021. cbsodataR: Statistics Netherlands (CBS) Open Data API Client. https://edwindj.github.io/cbsodataR/.

Dragulescu, A. and Arendt, C. 2020. xlsx-package: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 Files. R package version 0.6.5. https://www.rdocumentation.org/packages/xlsx/versions/0.6.5.

Grolemund, G. and Wickham, H. 2011. Dates and times made easy with lubridate. Journal of Statistical Software 40(3), 1-25, 10.18637/jss.v040.i03.

Keitt, T.H. 2021. rgdal: Bindings for the Geospatial Data Abstraction Library, R package version 1.5-27. http://cran.r-project.org/package=rgdal.

Neuwirth, E. 2015. Package 'RColorBrewer'. https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf.

Ooms, J. and McNamara, J. 2021. Package 'writexl'. https://cran.r-project.org/web//packages/writexl/writexl.pdf.

Pebesma, E. 2019. st_make_grid. https://www.rdocumentation.org/packages/sf/versions/0.8-0/topics/st_make_grid.

Pebesma, E., Bivand, R., Racine, E., Sumner, M., Cook, I., Keitt, T., Lovelace, R., Wickham, H., Ooms, J., Müller, K., Pedersen, T.L. and Baston, D. 2019a. sf: Simple Features for R. https://cran.r-project.org/web/packages/sf/index.html.

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., Hijmans, R., Sumner, M., MacQueen, D., Lemon, J., O'Brian, J. and O'Rourke, J. 2019b. sp: Classes and Methods for Spatial Data. https://cran.r-project.org/web/packages/sp/index.html

R-Core-Team 2021. parallel v3.6.2. https://www.rdocumentation.org/packages/parallel/versions/3.6.0.

Wickham, H. 2019. stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr.

Wickham, H. 2021. tidyr: Tidy Messy Data. R package version 1.1.4. https://CRAN.R-project.org/package=tidyr.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L. and Hester, J. 2019. Welcome to the Tidyverse. Journal of Open Source Software 4(43), 1686, 10.21105/joss.01686.

Wickham, H. and Bryan, J. 2019. readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl.

Wickham, H., Francois, R., Henry, L. and Müller, K. 2022. dplyr: A Grammar of Data Manipulation. R package version 1.0.9. https://CRAN.R-project.org/package=dplyr.

Zeileis, A., Fisher, J.C., Hornik, K., Ihaka, R., McWhite, C.D., Murrell, P., Stauffer, R. and Wilke, C.O. 2020. colorspace: A Toolbox for Manipulating and Assessing Colors and Palettes. Journal of Statistical Software 96(1), 1-49, 10.18637/jss.v096.i01.

Supplementary material

Figures

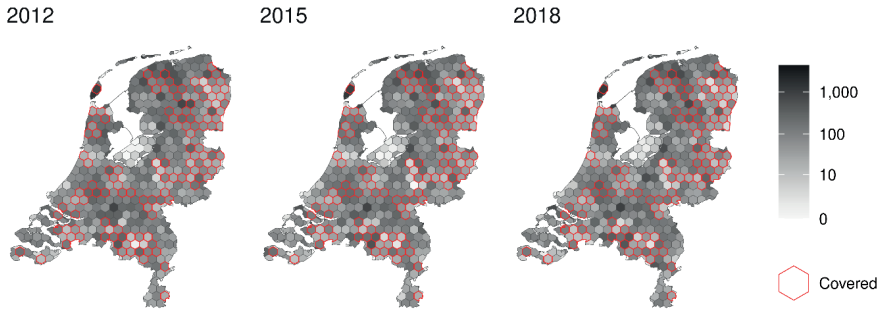


Figure S1. Map of the population-weighted number of small ruminants in the Netherlands per hexagon (90 km^2) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

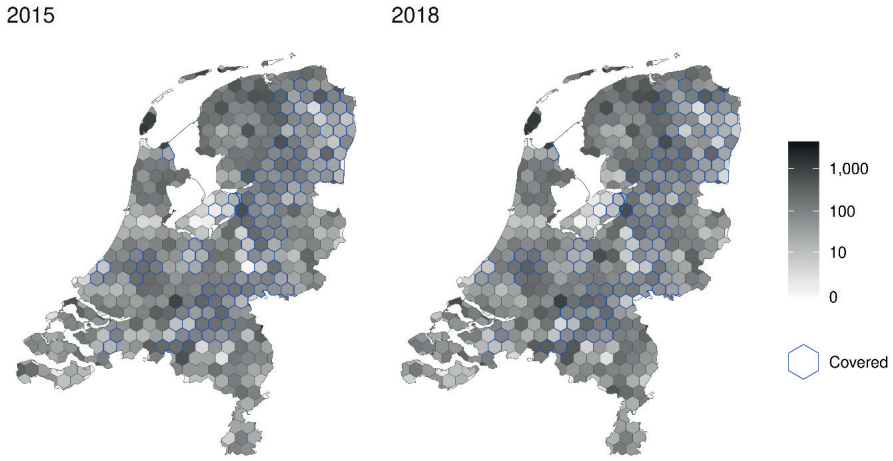


Figure S2. Map of the population-weighted number of small ruminants in the Netherlands per hexagon (90 km^2) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

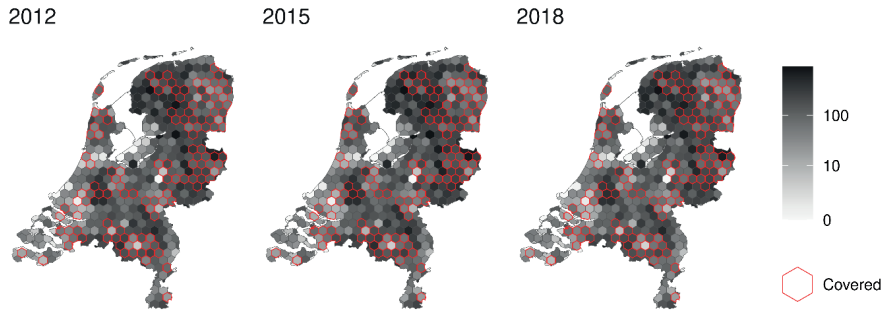


Figure S3. Map of the population-weighted number of dairy cows in the Netherlands per hexagon (90 km^2) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

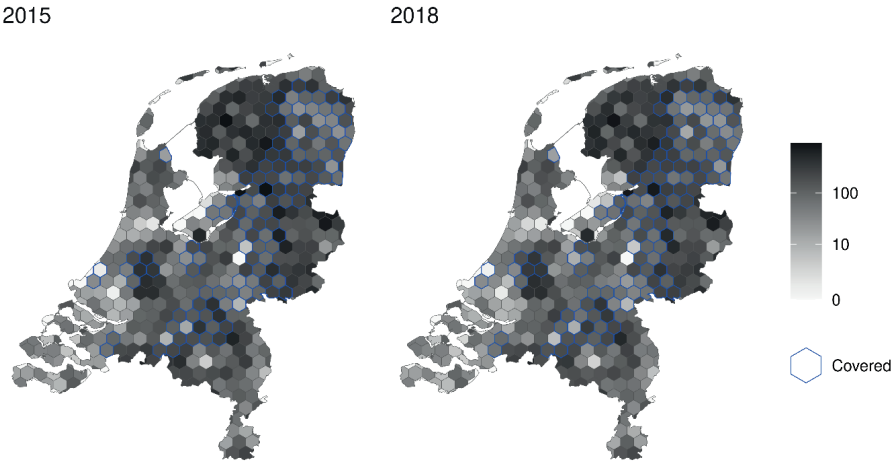


Figure S4. Map of the population-weighted number of dairy cows in the Netherlands per hexagon (90 km^2) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

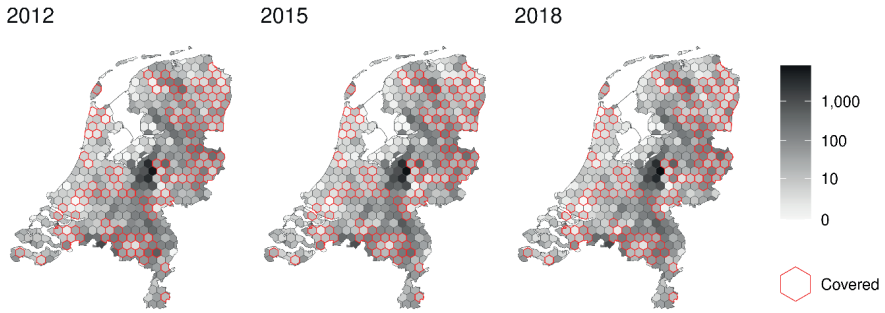


Figure S5. Map of the population-weighted number of veal calves in the Netherlands per hexagon (90 km^2) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

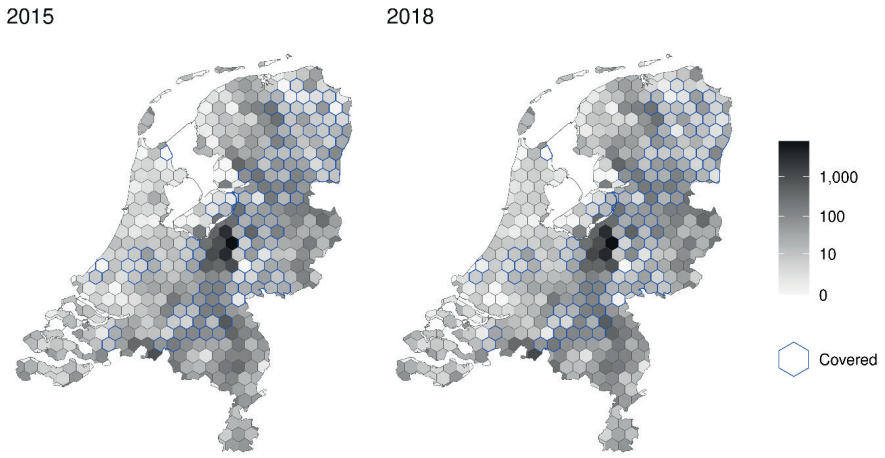


Figure S6. Map of the population-weighted number of veal calves in the Netherlands per hexagon (90 km^2) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

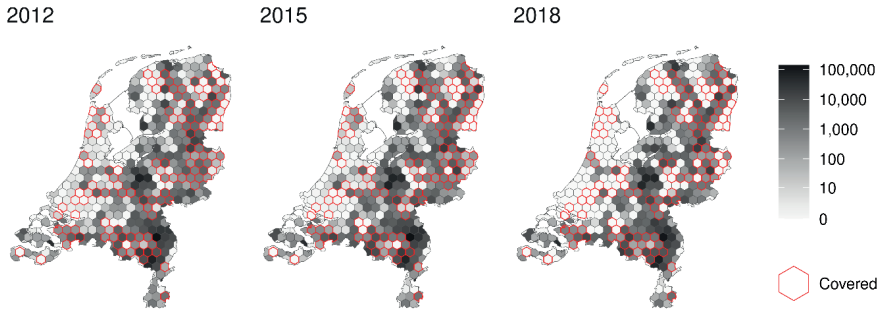


Figure S7. Map of the population-weighted number of laying hens in the Netherlands per hexagon (90 km^2) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

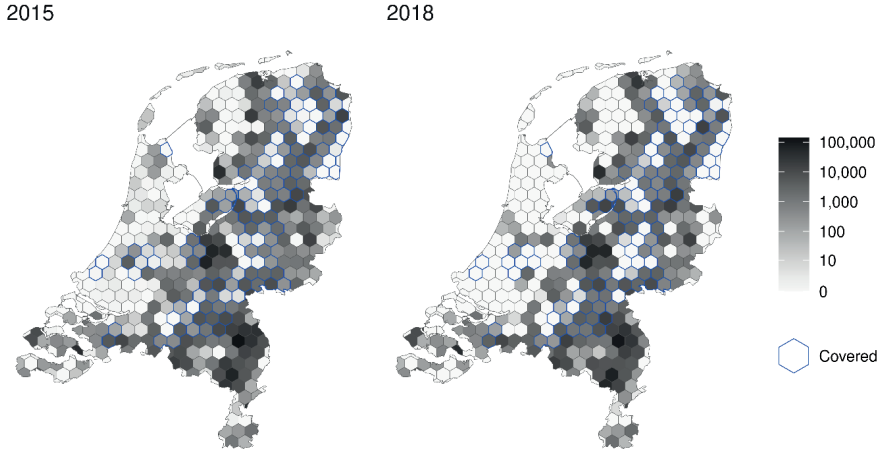


Figure S8. Map of the population-weighted number of laying hens in the Netherlands per hexagon (90 km^2) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

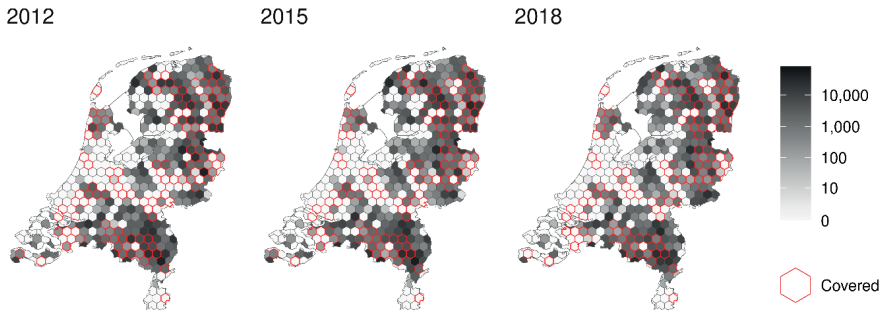


Figure S9. Map of the population-weighted number of broilers in the Netherlands per hexagon (90 km^2) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

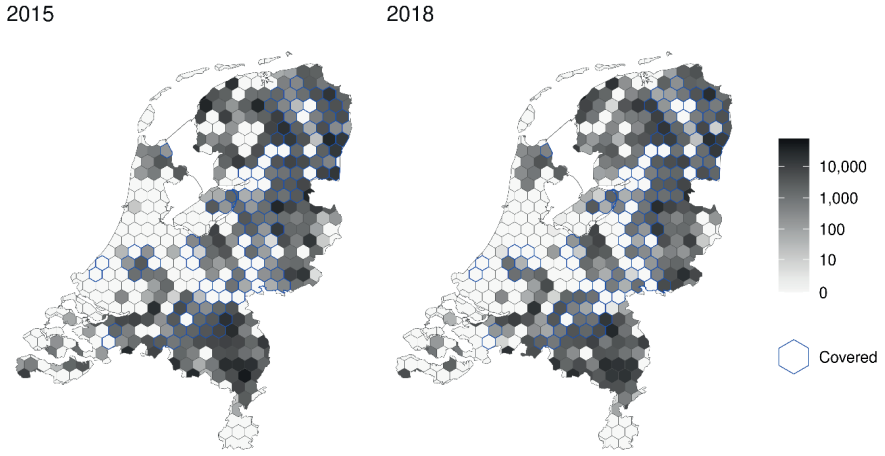


Figure S10. Map of the population-weighted number of broilers in the Netherlands per hexagon (90 km^2) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

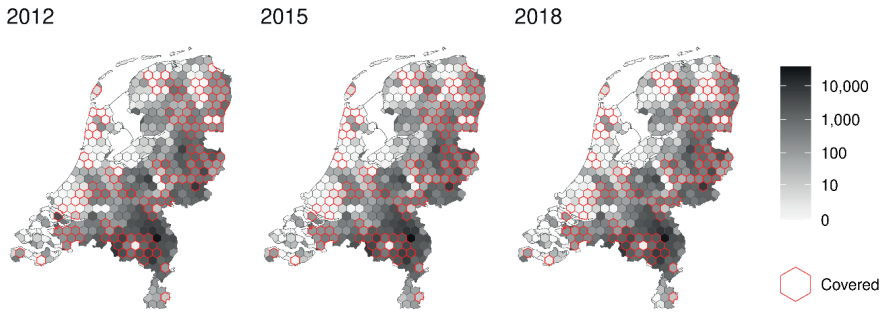


Figure S11. Map of the population-weighted number of pigs in the Netherlands per hexagon (90 km²) for *Salmonella* for the years 2012, 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

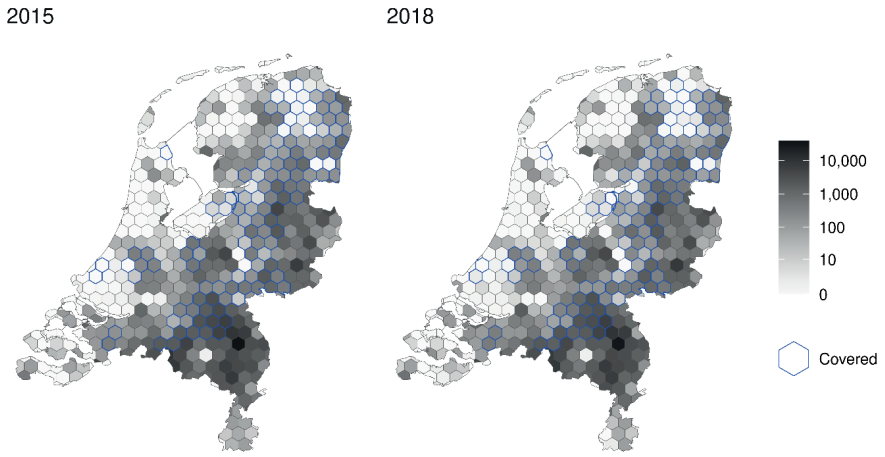


Figure S12. Map of the population-weighted number of pigs in the Netherlands per hexagon (90 km²) for *Campylobacter* for the years 2015 and 2018. The colored hexagonal areas are considered covered by the national surveillance system according to the executed analyses.

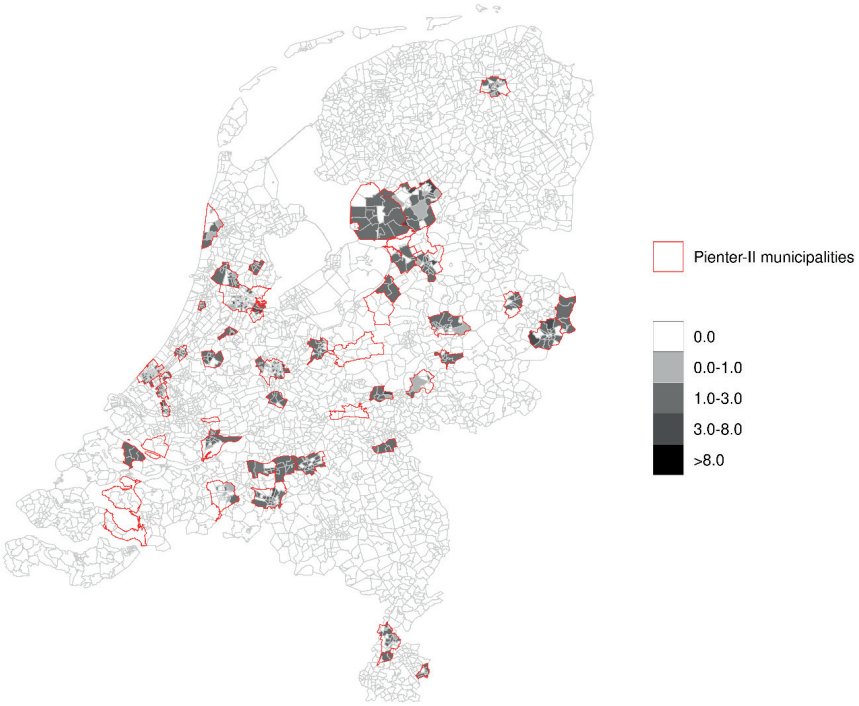


Figure S13. Mean sero-incidence of PIENTER-II participants per sampling region, split by postal code region.

Tables

Table S1. Total and mean number of food-producing animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) in the Netherlands per year (2012, 2015 and 2018).

	Number of animals					
	2012		2015		2018	
Animal group	Total	Mean	Total	Mean	Total	Mean
Small ruminants	1,448,025	90	1,401,549	100	1,543,009	119
Dairy cows	2,680,106	115	2,955,459	131	2,654,711	123
Veal calves	1,100,150	96	1,077,849	117	1,167,139	129
Laying hens	42,288,269	23,798	46,808,385	29,910	45,418,538	41,215
Broiler chickens	44,094,946	71,351	48,872,208	69,619	40,929,799	54,212
Pigs	12,138,896	1,744	12,496,985	2,122	12,329,522	2,297

Table S2. An overview of all packages and functions used, including version numbers and references.

Package/function	Version	Reference*
boot	1.3-28	(Canty and Ripley, 2021)
cbsodataR	0.5.1	(De Jonge and Houweling, 2021)
colorspace	2.0-2	(Zeileis et al., 2020)
Desctools	0.99.45	(Andri et mult al, 2022)
dplyr	1.0.9	(Wickham et al., 2022)
INLA	21.02.23	(Bakka et al., 2018)
lubridate	1.8.0	(Grolemund and Wickham, 2011)
Matrix	1.3-4	(Bates et al., 2021)
Parallel	4.1.2	(R-Core-Team, 2021)
RANN	2.6.1	(Arya et al., 2019)
RColorBrewer	1.1-2	(Neuwirth, 2015)
readxl	1.3.1	(Wickham and Bryan, 2019)
rgdal	1.5-27	(Keitt, 2021)
sf	1.0-3	(Pebesma et al., 2019a)
sp	1.4-5	(Pebesma et al., 2019b)
spdep	1.1-12	(Bivand et al., 2013)
st_make_grid	-	(Pebesma, 2019)
stringr	1.4.0	(Wickham, 2019)
tidyr	1.1.4	(Wickham, 2021)
tidyverse	1.3.1	(Wickham et al., 2019)
writexl	1.4.0	(Ooms and McNamara, 2021)
xlsx	0.6.5	(Dragulescu and Arendt, 2020)

*All literature was included in the reference list belonging to the main text of this article.

Table S3. An overview of the attributions per source per serovar and the primary sources per serovar.

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
1001	Abaetetuba	0	0	0	0	1	reptiles
1002	Aberdeen	0	0	0	0	1	reptiles
1003	Abony	0	0	0.076586	0	0.923414	reptiles
1008	Adelaide	0	0	0	0	1	reptiles
1010	Agona	0.385471	0.026526	0.130762	0.421848	0.035393	laying hens
1014	Alachua	0	0	0	0	1	reptiles
1015	Albany	0	0	0.473358	0	0.526642	reptiles
1023	Anatum	0.605659	0.009911	0.066904	0.212513	0.105013	pigs
1028	Apapa	0	0	0	0	1	reptiles
1038	Baildon	0	0	0	0	1	reptiles
1042	Bardo	0	0	1	0	0	broiler chickens
1043	Bareilly	0	0	0.00569	0.526798	0.467512	laying hens
1055	Bispebjerg	0	0	0	0	1	reptiles
1058	Blockley	0.277079	0	0.167204	0	0.555716	reptiles
1062	Bonariensis	0	0	0	0	1	reptiles
1066	Bovismorbificans	0.661853	0.156496	0.035443	0.075569	0.070639	pigs
1068	Braenderup	0.008496	0.005189	0.015476	0.9363	0.034539	laying hens
1070	Brandenburg	0.878083	0.027479	0.011801	0.080132	0.002505	pigs
1071	Bredeney	0.181617	0.02322	0.209859	0.372347	0.212957	laying hens
1085	Cerro	0.017219	0.007506	0.272779	0.322721	0.379775	reptiles
1091	Chester	0	0	1	0	0	broiler chickens
1092	Chicago	0	0	0	0	1	reptiles
1097	Coeln	0.352155	0	0.020317	0.270815	0.356714	reptiles
1101	Concord	0	0	0	0	1	reptiles
1102	Corvallis	0	0	0.736294	0.263706	0	broiler chickens
1103	Cubana	0.981808	0	0	0	0.018193	pigs
1106	Derby	0.700533	0.187181	0.05125	0.058976	0.00206	pigs
1110	Dublin	0.252024	0.729227	0.001345	0.017404	0	cattle
1112	Duisburg	0	0	0.518994	0.481006	0	broiler chickens
1114	Durban	0	0	0	0	1	reptiles
1117	Eastbourne	0	0	0	0	1	reptiles
1121	Elisabethville	0	0	0	0	1	reptiles
1136	Fresno	0	0	1	0	0	broiler chickens
1139	Gaminara	0.332714	0	0	0	0.667286	reptiles

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
1144	Give	0.685848	0	0.171811	0	0.142341	pigs
1148	Goettingen	0.544929	0	0.042725	0.412347	0	pigs
1157	Hadar	0.044796	0	0.279587	0	0.675617	reptiles
1162	Halle	0	0	0	0	1	reptiles
1167	Havana	0.044576	0	0.164672	0.339441	0.451311	reptiles
1169	Heidelberg	0.023731	0	0.693759	0.273755	0.008755	broiler chickens
1181	Hull	0	0	0	0	1	reptiles
1183	Hvittingfoss	0	0	0	0	1	reptiles
1186	Indiana	0.019705	0.050102	0.57547	0.336516	0.018206	broiler chickens
1187	Infantis	0.345658	0.009064	0.390396	0.249739	0.005143	broiler chickens
1198	Jangwani	0	0	0	0	1	reptiles
1200	Javiana	0	0	0.099227	0	0.900773	reptiles
1204	Johannesburg	0	0	0	0	1	reptiles
1206	Kapemba	1	0	0	0	0	pigs
1208	Kentucky	0.043704	0	0.860506	0.09579	0	broiler chickens
1214	Kingabwa	0	0	0	0	1	reptiles
1215	Kingston	0	0	1	0	0	broiler chickens
1219	Kottbus	0.279888	0.056659	0.116992	0.450126	0.096335	laying hens
1230	Lexington	0	0.307164	0	0.32464	0.368196	reptiles
1235	Litchfield	0.138446	0	0	0	0.861554	reptiles
1237	Livingstone	0.445844	0.096951	0.087379	0.360957	0.008869	pigs
1240	Lomalinda	0	0	0	0	1	reptiles
1242	London	0.880875	0.022993	0.068	0	0.028132	pigs
1252	Manhattan	0.71063	0	0.0015	0	0.287869	pigs
1257	Mbandaka	0.16773	0.036147	0.381628	0.414495	0	laying hens
1258	Meleagridis	0.5	0	0	0.5	0	pigs
1262	Miami	0	0	0	0	1	reptiles
1267	Minnesota	0	0	0.8623	0	0.1377	broiler chickens
1272	Molade	0	0	0.02663	0.97337	0	laying hens
1274	Monschaui	0	0	0	0	1	reptiles
1275	Montevideo	0.06596	0.473201	0.112683	0.078058	0.270098	cattle
1278	Muenchen	0.247868	0.037066	0.02196	0	0.693107	reptiles
1279	Muenster	0.118811	0.112498	0.067433	0	0.701258	reptiles
1282	Napoli	0	0	1	0	0	broiler chickens
1290	Newport	0.251413	0.213912	0.072484	0	0.46219	reptiles

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
1296	Nima	0	0	0	0	1	reptiles
1301	Ohio	0.452821	0.088344	0.398816	0.060019	0	pigs
1308	Oranienburg	0.008082	0	0.024119	0.240117	0.727682	reptiles
1311	Orion	0.356259	0	0.454875	0	0.188866	broiler chickens
1313	Oslo	0	0	0	0	1	reptiles
1327	Plymouth	0	0	0	0	1	reptiles
1328	Pomona	0	0	0	0	1	reptiles
1329	Poona	0	0	0.083059	0.304554	0.612387	reptiles
1335	Putten	0	0	1	0	0	broiler chickens
1338	Reading	0.280547	0	0	0	0.719453	reptiles
1341	Richmond	0	0	0	0	1	reptiles
1342	Rissen	0.373825	0.285629	0.008542	0.314941	0.017064	pigs
1346	Rubislaw	0	0	0	0	1	reptiles
1348	Saintpaul	0.04577	0.005302	0.427322	0.047626	0.47398	reptiles
1350	Sandiego	0	0	0	0	1	reptiles
1355	Schwarzengrund	0.085381	0.027275	0.290498	0	0.596846	reptiles
1359	Senftenberg	0.221998	0.014844	0.243288	0.476098	0.043771	laying hens
1364	Singapore	0	0	0	0	1	reptiles
1373	Stanley	0.215704	0	0.160222	0	0.624074	reptiles
1374	Stanleyville	0.184618	0	0.028885	0	0.786497	reptiles
1382	Takoradi	0	0	0	0	1	reptiles
1387	Telelkebir	0	0	0	0	1	reptiles
1389	Tennessee	0.02296	0.08718	0.095832	0.715633	0.078395	laying hens
1393	Thompson	0.004091	0	0.129342	0.443415	0.423152	laying hens
1401	Uganda	0.545163	0	0.454837	0	0	pigs
1403	Umbilo	0.011434	0	0	0.988566	0	laying hens
1405	Urbana	0	0	0	0	1	reptiles
1415	Virchow	0.006432	0	0.439194	0.436231	0.118142	broiler chickens
1427	Waycross	0	0	0	0	1	reptiles
1429	Weltevreden	0	0	1	0	0	broiler chickens
1436	Wien	0.5	0	0	0.5	0	pigs
1442	Worthington	1	0	0	0	0	pigs
1460	Gloucester	0	0	1	0	0	broiler chickens
1484	Kedougou	0.542928	0	0.266601	0.190471	0	pigs
1501	Agoueve	0	0	0	0	1	reptiles

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
1507	Goldcoast	0.938167	0.04519	0.009101	0.007542	0	pigs
1526	Offa	0	0	0	0	1	reptiles
1531	Altona	1	0	0	0	0	pigs
1539	Belem	0	0	0	0	1	reptiles
1571	Ealing	0	0	0	0	1	reptiles
1584	Fluntern	0	0	0	0	1	reptiles
1587	Banana	0	0	1	0	0	broiler chickens
1630	Bochum	0	0	1	0	0	broiler chickens
1655	Stourbridge	0.333333	0	0.666667	0	0	broiler chickens
1667	Teddington	0	0	0	0	1	reptiles
1674	Ndolo	0	0	0	0	1	reptiles
2050	SII 4,12,27:b:(e,n,x)	0	0	0	0	1	reptiles
2055	SII 42:b:e,n,x,z15	0	0	0	0	1	reptiles
2087	58:l,z13,z28:z6	0	0	0	0	1	reptiles
3002	47:r:z53	0	0	0	0	1	reptiles
3005	SIIIb 48:k:z35	0	0	0	0	1	reptiles
3007	SIIIb 50:z52:z35	0	0	0	0	1	reptiles
3014	SIIIb 48:i:z	0	0	0	0	1	reptiles
3020	SIIIb 61:k:z35	0	0	0	0	1	reptiles
3021	SIIIb 61:l,v:1,5,7	0	0	0	0	1	reptiles
3023	42:(k):z35	0	0	0	0	1	reptiles
3025	61:i:z53	0	0	0	0	1	reptiles
3035	SIIIa 41:z4,z23:-	0	0	0	0	1	reptiles
3039	65:(k):z53	0	0	0	0	1	reptiles
3044	SIIIb 61:c:z35	0	0	0	0	1	reptiles
3059	SIIIb 48:l,v:1,5,(7)	0	0	0	0	1	reptiles
3065	SIIIb 48:z52:z	0	0	0	0	1	reptiles
3068	SIIIb 38:z10:z53	0	0	0	0	1	reptiles
3069	SIIIb 50:(k):z	0	0	0	0	1	reptiles
3072	60:r:z	0	0	0	0	1	reptiles
3073	42:l,v:1,5,7	0	0	0	0	1	reptiles
3086	SIIIb 65:z10:e,n,x,z15	0	0	0	0	1	reptiles
3088	SIIIb 53:z10:z35	0	0	0	0	1	reptiles
3112	SIIIb 61:i:z	0	0	0	0	1	reptiles

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
3115	SIIIa 13,23:z4,z23,[z32]	0	0	0	0	1	reptiles
3117	48:g,z51:-	0	0	0	0	1	reptiles
3129	SIIIb 48:k:1,5,(7)	0	0	0	0	1	reptiles
3142	48:z4,z24:-	0	0	0	0	1	reptiles
4000	Typhimurium : g.f.b.	0.845842	0.093835	0.010594	0.049697	3.16E-05	pigs
4001	Typhimurium : 1	0.823899	0.176101	0	0	0	pigs
4002	Typhimurium : 2	0.313336	0.343817	0.113182	0	0.229665	cattle
4003	Typhimurium : 3	0.214775	0.222939	0.084656	0	0.47763	reptiles
4010	Typhimurium : 10	0.712517	0	0	0	0.287483	pigs
4011	Typhimurium : 11	0	0	0	0.985871	0.014129	laying hens
4012	Typhimurium : 12	0	0	0	1	0	laying hens
4020	Typhimurium : 20	1	0	0	0	0	pigs
4061	Typhimurium : 61	1	0	0	0	0	pigs
4080	Typhimurium : 80	0.31247	0.68753	0	0	0	cattle
4090	Typhimurium : 90	0	0	1	0	0	broiler chickens
4100	Typhimurium : 100	1	0	0	0	0	pigs
4111	Typhimurium : 111	0.633156	0	0	0.366844	0	pigs
4130	Typhimurium : 130	0	0	1	0	0	broiler chickens
4200	Typhimurium : 200	0	0	0	0	1	reptiles
4280	Typhimurium : 280	1	0	0	0	0	pigs
4281	Typhimurium : 281	1	0	0	0	0	pigs
4290	Typhimurium : 290	0.113045	0.886955	0	0	0	cattle
4292	Typhimurium : 292	1	0	0	0	0	pigs
4296	Typhimurium : 296	0.808896	0.111778	0.079327	0	0	pigs
4300	Typhimurium : 300	0	0.933346	0.066654	0	0	cattle
4301	Typhimurium : 301	0.918009	0.081991	0	0	0	pigs
4345	Typhimurium : 345	1	0	0	0	0	pigs
4350	Typhimurium : 350	0.672998	0	0	0	0.327002	pigs
4351	Typhimurium : 351	0.604654	0.395346	0	0	0	pigs
4353	Typhimurium : 353	0.2275	0.7725	0	0	0	cattle
4401	Typhimurium : 401	0.169957	0	0	0.830043	0	laying hens
4501	Typhimurium : 501	1	0	0	0	0	pigs
4504	Typhimurium : 504	0.877013	0.122987	0	0	0	pigs
4506	Typhimurium : 506	0.422553	0.169364	0.012278	0.374407	0.021399	pigs

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
4507	Typhimurium : 507	0.57668	0.353448	0.042196	0	0.027676	pigs
4508	Typhimurium : 508	0.487663	0.512337	0	0	0	cattle
4510	Typhimurium : 510	0.60989	0.084652	0.122104	0	0.183354	pigs
4561	Typhimurium : 561	0.669769	0.330231	0	0	0	pigs
4651	Typhimurium : 651	0.023074	0.540231	0.00361	0.433084	0	cattle
4652	Typhimurium : 652	0	1	0	0	0	cattle
4655	Typhimurium : 655	0.211478	0	0	0.788522	0	laying hens
4658	Typhimurium : 658	0.881452	0	0	0	0.118548	pigs
4690	Typhimurium : 690	1	0	0	0	0	pigs
4995	Typhimurium : Veront..	0	0.466673	0.533327	0	0	broiler chickens
4998	Typhimurium : OS	0.762035	0.206059	0.007059	0	0.024847	pigs
4999	Typhimurium : ARS	0.535353	0.257868	0.008947	0.147263	0.050569	pigs
5000	Panama	0.668237	0.059269	0.003225	0.166241	0.103029	pigs
5100	Enteritidis : g.f.b.	0.039981	0.015237	0.124904	0.819878	0	laying hens
5121	Enteritidis Pt 1b	0.266582	0	0	0.733418	0	laying hens
5122	Enteritidis Pt 1	0.022907	0	0.052343	0.924751	0	laying hens
5125	Enteritidis Pt 21	0.003449	0.022482	0.045837	0.928231	0	laying hens
5133	Enteritidis Pt 3	0.297045	0	0.120291	0.582665	0	laying hens
5134	Enteritidis Pt 4	0.011585	0.002208	0.048707	0.934578	0.002922	laying hens
5135	Enteritidis Pt 4a	0	0	0	1	0	laying hens
5142	Enteritidis Pt 12	0	0	0.059368	0.940632	0	laying hens
5144	Enteritidis Pt 6	0.123221	0	0.105338	0.771442	0	laying hens
5145	Enteritidis Pt 6a	0.139592	0	0.349112	0.462002	0.049295	laying hens
5146	Enteritidis Pt 7a	0	0	0.124661	0.875339	0	laying hens
5147	Enteritidis Pt 7	0	0	0.009348	0.990652	0	laying hens
5150	Enteritidis Pt 8	0.061443	0.02279	0.0385	0.866768	0.010499	laying hens
5153	Enteritidis Pt 9a	0	0	1	0	0	broiler chickens
5156	Enteritidis Pt 28	0	0	0	0	1	reptiles
5157	Enteritidis Pt 13a	0.080794	0	0.287671	0.631535	0	laying hens
5158	Enteritidis Pt 13	0.832433	0	0.167567	0	0	pigs
5160	Enteritidis Pt 23	0	0	0	1	0	laying hens
5161	Enteritidis Pt 11	0.014896	0.548594	0	0.436509	0	cattle
5167	Enteritidis Pt 14b	0.015918	0	0.005021	0.804546	0.174515	laying hens
5181	Enteritidis Pt 6b	1	0	0	0	0	pigs

Serovar	Serovar name	Pigs	Cattle	Broiler chickens	Laying hens	Reptiles	Source
5190	Enteritidis Pt 4b	0	0	0	0	1	reptiles
5191	Enteritidis Pt6C	0.368951	0	0.631049	0	0	broiler chickens
5196	Enteritidis + verontr.	0	0	1	0	0	broiler chickens
5198	Enteritidis : OS	0.24685	0	0.181432	0.421883	0.149836	laying hens
5199	Enteritidis : ARS	0.006167	0.100627	0.099245	0.793961	0	laying hens
5209	Enteritidis Pt 53	0	0	0.111601	0.888399	0	laying hens
6000	SIV rough	0	0	0	0	1	reptiles
6011	SIV 48:g,z51:-	0	0	0	0	1	reptiles
6018	SIV 50:g,z51:-	0	0	0	0	1	reptiles
6025	SIV 45:g,z51:-	0	0	0	0	1	reptiles
6029	SIV 50:z4,z23:-	0	0	0	0	1	reptiles
6033	SIV 44:z4,z23:-	0	0	0	0	1	reptiles
7017	SI 18:NM	1	0	0	0	0	pigs
7032	SI 4,5,12:d:2ef nat	0.172042	0	0.763837	0.064121	0	broiler chickens
7033	SI 4,5,12:e,h:2ef nat	0	0	0	0	1	reptiles
7037	SI 4,5,12:1,2:2ef nat	0.167736	0.440613	0.391651	0	0	cattle
7039	SI 4,5,12:b:2ef nat	0.06631	0	0.299374	0	0.634317	reptiles
7045	SI 6,7:y:2ef nat	0	0	0	0.98365	0.01635	laying hens
7048	SI 6,7:1,5:2ef nat	0.190869	0	0.014146	0.282825	0.51216	reptiles
7054	SI 6,7:k:2ef nat	0	0	0	0	1	reptiles
7062	SI (6),8:e,h:2ef nat	0	0	0	0	1	reptiles
7064	SI 9,12:l,v:2ef nat	0.732027	0.027373	0.005087	0.235514	0	pigs
7065	SI 9,12:1,5:2ef nat	0	0	0	0	1	reptiles
7073	SI 3,10:e,h:2ef nat	1	0	0	0	0	pigs
7114	SI 9,12:z35:2ef nat	0	0	0	0	1	reptiles
7121	SI 6,7:z4,z23:2ef nat	0	0	0	0	1	reptiles
7202	SII 4,(5),12:b:2ef nat	0	0	0	0	1	reptiles
7221	SI 1,4,5,12:i:2ef nat	0.76335	0.130355	0.031623	0.074671	0	pigs
7236	SI 1,4,5,12:i:2ef nat	0.855088	0.095239	0.013484	0.036076	0.000112	pigs
7237	SI 1,4,5,12:i:2ef nat	0.823807	0.10217	0.028098	0.045924	0	pigs
9986	Biochemisch Salm. Rough	0.487981	0	0.127629	0.384389	0	pigs
9987	ORough	0.968488	0	0.031512	0	0	pigs

Table S4. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association between *S.* Enteritidis (SE), *S.*Typhimurium (ST) including its monophasic variants (ST_{mv}) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		90 km²		50 km²		25 km²		10 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
SE‡	**Sex**								
	Male				*Reference category*				
	Female	1.05	0.96-1.14	1.03	0.95-1.12	1.02	0.94-1.10	1.00	0.93-1.08
	Age category (years)								
	0-4	3.01***	2.64-3.42	2.84***	2.50-3.21	2.96***	2.64-3.32	2.79***	2.49-3.11
	5-9	1.74***	1.49-2.03	1.56***	1.33-1.82	1.69***	1.47-1.94	1.66***	1.45-1.90
	10-49				*Reference category*				
	50+	0.80***	0.72-0.88	0.75***	0.68-0.83	0.79***	0.72-0.86	0.76***	0.70-0.83
	Period of infection§								
	Summer (may-oct)	2.44***	2.22-2.69	2.52***	2.30-2.77	2.61	2.40-2.85	2.55***	2.35-2.77
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	0.94**	0.90-0.99	0.98	0.94-1.01	1.01	0.98-1.04	1.02	0.99-1.04
	Dairy cows	0.96	0.91-1.01	0.95**	0.92-0.99	0.97*	0.95-1.00	0.97**	0.95-0.99
	Veal calves	0.95*	0.91-0.99	0.96*	0.93-1.00	0.96**	0.93-0.99	0.98*	0.95-1.00
	Laying hens	1.02*	1.00-1.04	1.00	0.99-1.02	1.00	0.99-1.02	1.01	0.99-1.02
	Broiler chickens	0.99	0.98-1.01	1.00	0.99-1.01	1.00	0.99-1.01	1.00	0.99-1.01
	Pigs	0.97*	0.94-0.99	0.97*	0.95-1.00	0.99	0.97-1.00	0.98	0.97-1.00

		90 km²		50 km²		25 km²		10 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
ST/ST$_{mv}$ ‡	**Sex**								
	Male				*Reference category*				
	Female	1.05	0.97-1.12	1.06	1.00-1.14	1.04	0.97-1.10	1.04	0.98-1.11
	Age category (years)								
	0-4	4.34***	3.91-4.81	4.13***	3.74-4.56	4.19***	3.82-4.60	4.16***	3.80-4.56
	5-9	3.08***	2.74-3.45	3.03***	2.71-3.37	2.88***	2.60-3.19	2.85***	2.57-3.16
	10-49				*Reference category*				
	50+	1.14**	1.05-1.24	1.11**	1.03-1.21	1.10*	1.02-1.19	1.13**	1.05-1.21
	Period of infection§								
	Summer (may-oct)	1.57***	1.46-1.68	1.60***	1.49-1.71	1.58***	1.49-1.69	1.65***	1.54-1.76
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	0.96	0.92-1.01	1.02	0.99-1.06	1.01	0.99-1.04	1.00	0.99-1.02
	Dairy cows	1.02	0.97-1.08	1.01	0.98-1.05	1.01	0.99-1.04	0.98	0.96-1.00
	Veal calves	1.04	0.99-1.08	1.04*	1.01-1.07	1.01	0.99-1.04	1.00	0.98-1.03
	Laying hens	1.03**	1.01-1.04	1.03***	1.01-1.04	1.02**	1.01-1.03	1.01	1.00-1.02
	Broiler chickens	1.02**	1.01-1.04	1.01*	1.00-1.02	1.01	1.00-1.02	1.00	0.99-1.01
	Pigs	1.04**	1.01-1.07	1.04***	1.01-1.06	1.01	0.99-1.03	1.00	0.99-1.01

† population-weighted number of animals

‡ IRR = incident rate ratio, 95% CI = 95% confidence interval, SE = *Salmonella* Enteritidis, ST/ST$_{mv}$ = *Salmonella* Typhimurium, including its monophasic variants

§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer

* p-value <0.05

** p-value <0.01

*** p-value <0.001

Table S5. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (cattle and pigs) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km^2), with the correction for geographical laboratory coverage.

		90 km^2		50 km^2		25 km^2		10 km^2	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Cattle-related	**Sex**								
	Male	*Reference category*							
	Female	0.85	0.61-1.17	0.84	0.62-1.14	0.84	0.64-1.09	0.99	0.76-1.29
	Age category (years)								
	0-4	4.08***	2.41-6.70	5.60***	3.45-8.92	3.99***	2.56-6.10	4.05***	2.58-6.24
	5-9	2.01*	1.00-3.71	1.73	0.81-3.38	1.82*	0.99-3.13	2.30**	1.32-3.83
	10-49	*Reference category*							
	50+	1.83**	1.27-2.65	2.37***	1.66-3.40	1.97***	1.46-2.69	2.08***	1.53-2.84
	Period of infection§								
	Summer (may-oct)	0.94	0.68-1.29	1.06	0.78-1.44	0.85	0.65-1.12	0.90	0.69-1.18
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.89*	0.79-0.99	0.94	0.86-1.03	0.99	0.92-1.06	0.98	0.93-1.03
	Dairy cows	0.98	0.89-1.09	0.99	0.92-1.07	1.00	0.94-1.06	0.98	0.94-1.03
	Veal calves	0.94	0.85-1.03	0.94	0.86-1.02	0.96	0.89-1.03	0.93*	0.88-0.99
	Laying hens	0.96*	0.92-1.00	0.99	0.95-1.02	0.97	0.94-1.00	0.97	0.93-1.01
	Broiler chickens	0.96	0.93-1.00	0.98	0.95-1.01	0.96*	0.92-0.99	0.97	0.92-1.00
	Pigs	0.93**	0.88-0.97	0.96	0.91-1.01	0.93***	0.89-0.97	0.94**	0.90-0.98

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Pig-related	**Sex**								
	Male	*Reference category*							
	Female	1.03	0.97-1.10	1.07	1.00-1.14	1.04	0.98-1.11	1.05	0.99-1.11
	Age category (years)								
	0-4	3.92***	3.55-4.34	3.75***	3.40-4.12	3.80***	3.48-4.15	3.74***	3.42-4.09
	5-9	2.70***	2.42-3.02	2.67***	2.00-2.97	2.56***	2.31-2.83	2.50***	2.26-2.76
	10-49	*Reference category*							
	50+	1.31***	1.22-1.42	1.28***	1.19-1.38	1.27***	1.18-1.35	1.29***	1.21-1.38
	Period of infection§								
	Summer (may-oct)	1.56***	1.46-1.67	1.59***	1.49-1.70	1.58***	1.49-1.68	1.65***	1.55-1.75
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.97	0.93-1.01	1.01	0.98-1.04	1.00	0.98-1.03	1.00	0.98-1.01
	Dairy cows	1.03	0.97-1.08	1.01	0.97-1.04	1.01	0.99-1.03	0.98*	0.96-1.00
	Veal calves	1.03	0.99-1.07	1.04*	1.01-1.07	1.00	0.98-1.03	1.00	0.98-1.02
	Laying hens	1.02**	1.01-1.04	1.02***	1.01-1.04	1.01*	1.00-1.02	1.01*	1.00-1.02
	Broiler chickens	1.02**	1.01-1.03	1.01	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.00
	Pigs	1.04**	1.02-1.07	1.03**	1.01-1.05	1.01	0.99-1.02	1.00	0.98-1.01

† population-weighted number of animals
‡IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S6. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (laying hens and broiler chickens) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Laying hens-related	**Sex**								
	Male	*Reference category*							
	Female	1.06	0.98-1.15	1.07	0.99-1.16	1.06	0.99-1.14	1.05	0.98-1.13
	Age category (years)								
	0-4	2.81***	2.47-3.18	2.69***	2.38-3.03	2.82***	2.53-3.15	2.62***	2.34-2.92
	5-9	1.63***	1.39-1.89	1.47***	1.26-1.71	1.58***	1.37-1.80	1.51***	1.32-1.72
	10-49	*Reference category*							
	50+	0.86**	0.78-0.95	0.80***	0.73-0.88	0.84***	0.77-0.91	0.81***	0.74-0.88
	Period of infection§								
	Summer (may-oct)	2.33***	2.13-2.55	2.41***	2.21-2.63	2.43***	2.24-2.63	2.42***	2.24-2.61
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.94**	0.90-0.98	0.97	0.94-1.01	1.00	0.97-1.03	1.00	0.98-1.02
	Dairy cows	0.96	0.92-1.01	0.95**	0.92-0.98	0.97**	0.94-0.99	0.96***	0.95-0.98
	Veal calves	0.94**	0.90-0.98	0.97	0.93-1.00	0.97**	0.94-0.99	0.98*	0.95-1.00
	Laying hens	1.01	1.00-1.03	1.01	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.01
	Broiler chickens	0.98	0.98-1.01	1.00	0.99-1.01	1.00	0.99-1.01	1.00	0.99-1.01
	Pigs	0.96**	0.94-0.99	0.97**	0.95-0.99	0.98*	0.96-1.00	0.98*	0.97-1.00

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Broiler chicken-related	**Sex**								
	Male	*Reference category*							
	Female	1.10	0.95-1.27	1.16*	1.01-1.33	1.17*	1.03-1.32	1.15*	1.02-1.30
	Age category (years)								
	0-4	2.18***	1.71-2.74	2.35***	1.87-2.93	2.09***	1.69-2.57	2.21***	1.79-2.71
	5-9	0.61*	0.40-0.89	0.61*	0.40-0.88	0.63**	0.44-0.87	0.66*	0.46-0.92
	10-49	*Reference category*							
	50+	1.00	0.86-1.17	1.08	0.93-1.26	0.95	0.82-1.09	1.09	0.95-1.25
	Period of infection§								
	Summer (may-oct)	2.11***	1.81-2.46	2.11***	1.83-2.45	2.27***	1.99-2.61	2.09***	1.83-2.38
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.94*	0.87-1.00	0.94*	0.89-0.99	0.96*	0.92-0.99	0.95**	0.92-0.98
	Dairy cows	0.93	0.87-1.00	0.93*	0.89-0.98	0.96*	0.93-1.00	0.94***	0.91-0.96
	Veal calves	0.93*	0.88-0.99	0.94*	0.89-0.99	0.95*	0.92-0.99	0.93***	0.90-0.96
	Laying hens	0.99	0.96-1.01	0.99	0.97-1.02	0.97**	0.96-1.00	0.98*	0.96-1.00
	Broiler chickens	1.00	0.98-1.03	1.00	0.98-1.02	1.00	0.98-1.01	0.98	0.97-1.00
	Pigs	0.94***	0.91-0.98	0.90*	0.92-0.98	0.95***	0.93-0.98	0.97**	0.94-0.99

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S7. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of *Campylobacter* (*C.*) *jejuni* and *C. coli* and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		90 km²		50 km²		25 km²		10 km²	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. jejuni	**Sex**								
	Male	*Reference category*							
	Female	0.85***	0.82-0.88	0.85***	0.82-0.88	0.85***	0.82-0.89	0.86***	0.82-0.89
	Age category (years)								
	0-4	0.86**	0.78-0.95	0.83***	0.75-0.92	0.84***	0.76-0.93	0.82***	0.74-0.90
	5-9	0.48***	0.43-0.55	0.48***	0.42-0.54	0.48***	0.42-0.54	0.49***	0.44-0.55
	10-49	*Reference category*							
	50+	1.14***	1.09-1.18	1.13***	1.09-1.18	1.14***	1.10-1.19	1.13***	1.09-1.18
	Period of infection§								
	Summer (may-oct)	1.75***	1.68-1.83	1.78***	1.71-1.85	1.79***	1.72-1.86	1.76***	1.70-1.83
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	1.03	0.99-1.06	1.02	1.00-1.04	1.01	0.99-1.03	1.00	0.98-1.01
	Dairy cows	1.03	0.99-1.07	1.04**	1.01-1.07	1.00	0.98-1.02	0.98*	0.97-1.00
	Veal calves	1.00	0.97-1.03	1.01	0.99-1.03	1.01	0.99-1.02	0.98*	0.97-1.00
	Laying hens	1.00	0.99-1.01	1.00	1.00-1.01	1.00	0.99-1.01	1.00	1.00-1.01
	Broiler chickens	1.01*	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.01	0.99	0.99-1.00
	Pigs	1.00	0.98-1.02	1.00	0.99-1.02	1.00	0.98-1.01	0.99*	0.98-1.00

		90 km²		50 km²		25 km²		10 km²	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. coli	**Sex**								
	Male				*Reference category*				
	Female	1.05	0.91-1.20	1.05	0.91-1.20	1.05	0.92-1.20	1.05	0.93-1.20
	Age category (years)								
	0-4	0.26***	0.13-0.49	0.22***	0.10-0.41	0.23***	0.11-0.42	0.22***	0.11-0.40
	5-9	0.60*	0.38-0.91	0.50**	0.31-0.77	0.52**	0.33-0.79	0.41***	0.25-0.64
	10-49				*Reference category*				
	50+	1.69***	1.46-1.95	1.63***	1.42-1.87	1.56***	1.36-1.78	1.61***	1.42-1.84
	Period of infection§								
	Summer (may-oct)	2.08***	1.79-2.41	2.13***	1.84-2.46	2.24***	1.94-2.59	2.11***	1.84-2.42
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	0.99	0.93-1.05	1.02	0.97-1.07	1.00	0.97-1.05	1.00	0.97-1.03
	Dairy cows	1.01	0.95-1.07	1.05*	1.00-1.10	1.00	0.97-1.04	0.98	0.95-1.03
	Veal calves	1.01	0.96-1.06	1.02	0.98-1.07	0.98	0.94-1.01	0.97	0.94-1.00
	Laying hens	1.00	0.98-1.02	0.99	0.98-1.01	0.99	0.97-1.00	0.99	0.98-1.01
	Broiler chickens	1.01	0.99-1.04	0.99	0.98-1.01	1.00	0.98-1.02	0.98	0.96-1.00
	Pigs	1.03	0.99-1.06	1.00	0.97-1.03	0.99	0.96-1.01	0.98*	0.96-1.00

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S8. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Salmonella* (*S.*) Enteritidis (SE), *S.* Typhimurium (ST), including its monophasic variants (ST_{mv}) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
SE‡	**Sex**						
	Male	*Reference category*					
	Female	1.05	0.95-1.17	1.03	0.93-1.13	1.01	0.92-1.10
	Age category (years)						
	0-4	3.04***	2.61-3.53	2.84***	2.45-3.27	2.94***	2.57-3.35
	5-9	1.82***	1.51-2.17	1.56***	1.29-1.86	1.69***	1.44-1.99
	10-49	*Reference category*					
	50+	0.71***	0.63-0.81	0.69***	0.61-0.78	0.73***	0.66-0.81
	Type of animal†						
	Small ruminants	0.95	0.89-1.01	1.00	0.95-1.05	1.02	0.99-1.06
	Dairy cows	1.04	0.96-1.13	0.97	0.92-1.02	0.98	0.95-1.02
	Veal calves	0.95	0.89-1.02	0.96	0.91-1.02	0.96	0.92-1.00
	Laying hens	1.02	1.00-1.04	1.02**	1.01-1.04	1.01	0.99-1.02
	Broiler chickens	1.01	0.98-1.03	1.00	0.99-1.02	1.01	0.99-1.02
	Pigs	0.96	0.92-1.01	0.98	0.95-1.02	0.99	0.96-1.01
ST/ST_{mv}‡	**Sex**						
	Male	*Reference category*					
	Female	1.05	0.96-1.15	1.07	0.98-1.16	1.03	0.95-1.12
	Age category (years)						
	0-4	3.95***	3.45-4.50	3.81***	3.36-4.33	3.78***	3.35-4.24
	5-9	2.93***	2.53-3.37	2.93***	2.56-3.36	2.78***	2.45-3.16
	10-49	*Reference category*					
	50+	0.94	0.85-1.05	0.96	0.87-1.07	0.92	0.83-1.01
	Type of animal†						
	Small ruminants	0.97	0.91-1.02	1.03	0.98-1.08	1.01	0.98-1.04
	Dairy cows	0.99	0.92-1.07	0.94*	0.90-0.99	1.01	0.97-1.04
	Veal calves	1.00	0.94-1.07	1.02	0.97-1.07	1.01	0.97-1.04
	Laying hens	1.02	1.00-1.04	1.02**	1.01-1.04	1.01	1.00-1.02
	Broiler chickens	1.01	1.00-1.03	1.01	1.00-1.02	1.01	0.99-1.02
	Pigs	1.03	0.99-1.08	1.03	1.00-1.06	1.00	0.98-1.02

† population-weighted number of animals

‡ IRR = incident rate ratio, 95% CI = 95% confidence interval, SE = *Salmonella* Enteritidis, ST/ST_{mv} = *Salmonella* Typhimurium, including its monophasic variants

* p-value <0.05

** p-value <0.01

*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
1.02	0.94-1.11	1.08	0.92-1.27	1.08	0.93-1.27	1.11	0.96-1.28	1.02	0.88-1.17
2.71***	2.38-3.09	2.92***	2.26-3.74	2.86***	2.24-3.63	3.04***	2.42-3.78	3.01***	2.42-3.72
1.63***	1.39-1.91	1.55**	1.12-2.10	1.59**	1.17-2.12	1.70***	1.28-2.22	1.75***	1.34-2.25
					Reference category				
0.69***	0.62-0.77	1.02	0.85-1.22	0.92	0.77-1.10	0.96	0.81-1.14	0.96	0.82-1.13
1.03	1.00-1.06	0.97	0.90-1.05	1.01	0.95-1.09	1.05	1.00-1.10	1.07	1.03-1.11
0.97*	0.94-1.00	0.98	0.89-1.08	0.96	0.89-1.03	0.97	0.92-1.02	0.95**	0.91-0.99
0.98	0.95-1.01	0.97	0.89-1.07	1.01	0.93-1.09	0.96	0.91-1.02	0.98	0.93-1.02
1.01	1.00-1.03	1.04**	1.01-1.07	1.01	0.99-1.04	1.00	0.98-1.02	1.00	0.97-1.02
1.01	0.99-1.02	1.01	0.98-1.04	1.01	0.98-1.03	0.99	0.97-1.01	1.00	0.98-1.02
0.99	0.97-1.01	0.96	0.91-1.02	0.96	0.92-1.01	1.00	0.97-1.04	0.99	0.96-1.02
					Reference category				
1.04	0.96-1.13	1.07	0.96-1.20	1.09	0.98-1.22	1.06	0.96-1.18	1.07	0.97-1.19
3.72***	3.30-4.17	5.09***	4.29-6.03	4.73***	4.01-5.55	5.01***	4.31-5.81	5.09***	4.37-5.91
2.69***	2.37-3.06	3.37***	2.78-4.07	3.19***	2.66-3.82	3.06***	2.57-3.63	3.20***	2.69-3.79
					Reference category				
0.96	0.87-1.05	1.51***	1.32-1.73	1.39***	1.22-1.59	1.46***	1.29-1.65	1.48***	1.31-1.67
1.01	0.99-1.04	0.92*	0.85-0.98	0.95	0.90-1.00	0.98	0.94-1.01	0.99	0.96-1.02
0.98	0.96-1.01	0.98	0.89-1.07	1.01	0.95-1.06	1.00	0.96-1.04	0.96*	0.93-0.99
1.01	0.99-1.04	1.03	0.96-1.11	1.02	0.97-1.08	1.00	0.96-1.04	1.01	0.98-1.04
1.01*	1.00-1.03	1.01	0.98-1.03	1.02*	1.00-1.03	1.02*	1.00-1.03	1.01	0.99-1.02
1.00	0.99-1.01	1.02	1.00-1.04	1.01	0.99-1.03	1.01	0.99-1.02	1.00	0.99-1.02
1.00	0.98-1.01	1.03	0.98-1.08	1.02	0.98-1.05	1.00	0.98-1.03	1.01	0.99-1.03

Table S9. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (cattle and pigs) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Cattle-related	**Sex**						
	Male	*Reference category*					
	Female	1.03	0.65-1.64	1.04	0.68-1.59	0.86	0.58-1.28
	Age category (years)						
	0-4	3.60***	1.69-7.20	4.53***	2.29-8.56	3.22***	1.63-6.01
	5-9	1.01	0.27-2.95	0.32	0.03-1.78	1.27	0.46-2.99
	10-49	*Reference category*					
	50+	1.60	0.96-2.69	2.05**	1.28-3.33	1.72*	1.11-2.68
	Type of animal†						
	Small ruminants	0.77**	0.64-0.93	0.92	0.80-1.07	1.02	0.91-1.14
	Dairy cows	1.18	0.96-1.47	1.10	0.94-1.28	1.02	0.91-1.14
	Veal calves	1.06	0.85-1.33	0.95	0.80-1.13	1.01	0.88-1.15
	Laying hens	0.98	0.92-1.05	0.99	0.94-1.05	1.00	0.95-1.05
	Broiler chickens	1.01	0.95-1.07	1.01	0.96-1.06	1.00	0.94-1.05
	Pigs	0.92	0.82-1.03	0.96	0.88-1.06	0.92*	0.85-0.99
Pig-related	**Sex**						
	Male	*Reference category*					
	Female	1.01	0.92-1.09	1.04	0.96-1.13	1.03	0.95-1.11
	Age category (years)						
	0-4	3.51***	3.09-3.99	3.37***	2.97-3.81	3.35***	2.99-3.75
	5-9	2.55***	2.21-2.93	2.60***	2.28-2.96	2.45***	2.16-2.77
	10-49	*Reference category*					
	50+	1.09	0.98-1.20	1.10*	1.00-1.21	1.04	0.96-1.14
	Type of animal†						
	Small ruminants	0.96	0.91-1.01	1.00	0.96-1.05	1.00	0.97-1.03
	Dairy cows	0.99	0.92-1.07	0.95*	0.91-1.00	1.00	0.97-1.04
	Veal calves	1.00	0.94-1.06	1.03	0.99-1.07	1.01	0.97-1.04
	Laying hens	1.01	0.99-1.03	1.02*	1.00-1.03	1.01	1.00-1.02
	Broiler chickens	1.01	0.99-1.03	1.00	0.99-1.02	1.00	0.99-1.01
	Pigs	1.04	1.00-1.08	1.02	0.99-1.05	1.00	0.98-1.02

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Winter									
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
1.05	0.71-1.55	0.69	0.44-1.08	0.65	0.41-1.00	0.80	0.55-1.15	0.92	0.63-1.32
3.07***	1.56-5.71	4.64***	2.20-9.28	7.17***	3.55-14.17	4.82***	2.64-8.49	5.34***	2.85-9.70
1.20	0.44-2.82	3.18**	1.36-6.81	3.82**	1.61-8.41	2.40*	1.10-4.80	3.78***	1.88-7.22
					Reference category				
1.67*	1.09-2.58	2.09**	1.25-3.58	2.83***	1.66-4.98	2.23***	1.46-3.44	2.64***	1.70-4.18
0.94	0.85-1.03	0.88	0.72-1.07	0.93	0.80-1.09	0.95	0.85-1.07	1.03	0.93-1.13
1.06	0.96-1.16	1.24*	1.01-1.52	1.08	0.93-1.27	1.14*	1.02-1.28	1.04	0.95-1.14
0.99	0.88-1.10	1.09	0.88-1.36	0.97	0.80-1.15	0.99	0.87-1.12	0.94	0.83-1.05
1.01	0.96-1.07	0.95	0.89-1.02	1.01	0.95-1.06	0.98	0.93-1.03	0.96	0.89-1.02
1.00	0.94-1.05	0.95	0.89-1.01	0.97	0.91-1.02	0.94*	0.89-1.00	0.95	0.88-1.02
0.95	0.88-1.01	0.88*	0.79-0.98	0.96	0.87-1.05	0.93*	0.86-1.00	0.95	0.88-1.02
					Reference category				
1.04	0.96-1.11	1.08	0.97-1.20	1.12	1.01-1.24	1.08	0.99-1.19	1.09	0.99-1.20
3.28***	2.93-3.66	4.74***	4.02-5.57	4.48***	3.83-5.23	4.72***	4.08-5.44	4.75***	4.10-5.48
2.35***	2.07-2.66	3.01***	2.49-3.62	2.81***	2.35-3.35	2.78***	2.34-3.28	2.81***	2.37-3.32
					Reference category				
1.07	0.98-1.17	1.77***	1.56-2.00	1.63***	1.45-1.83	1.71***	1.53-1.91	1.76***	1.57-1.96
1.00	0.98-1.02	0.94*	0.88-1.00	0.95	0.91-1.00	0.98	0.94-1.01	0.99	0.97-1.02
0.98	0.96-1.01	0.98	0.89-1.07	1.00	0.95-1.06	1.00	0.96-1.04	0.96**	0.93-0.99
1.01	0.98-1.04	1.02	0.95-1.09	1.03	0.98-1.08	0.99	0.96-1.03	1.01	0.98-1.04
1.01	1.00-1.02	1.01	0.99-1.03	1.02*	1.00-1.03	1.02*	1.00-1.03	1.01	1.00-1.03
1.00	0.99-1.01	1.02	1.00-1.04	1.01	0.99-1.02	1.01	0.99-1.02	1.00	0.98-1.01
1.00	0.98-1.01	1.03	0.99-1.08	1.01	0.98-1.05	1.01	0.99-1.04	1.01	0.99-1.03

Table S10. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (laying hens and broiler chickens) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km^2), with the correction for geographical laboratory coverage.

		Summer					
		90 km^2		50 km^2		25 km^2	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Laying hen-related	**Sex**						
	Male				*Reference category*		
	Female	1.05	0.95-1.15	1.06	0.97-1.16	1.05	0.96-1.14
	Age category (years)						
	0-4	2.82***	2.42-3.26	2.62***	2.27-3.02	2.77***	2.43-3.15
	5-9	1.63***	1.35-1.94	1.42***	1.18-1.69	1.53***	1.30-1.79
	10-49				*Reference category*		
	50+	0.76***	0.67-0.85	0.72***	0.64-0.80	1.53***	1.30-1.79
	Type of animal†						
	Small ruminants	0.94	0.89-1.00	0.99	0.94-1.04	1.01	0.98-1.05
	Dairy cows	1.05	0.97-1.14	0.96	0.91-1.01	0.98	0.95-1.02
	Veal calves	0.95	0.88-1.02	0.99	0.93-1.04	0.97	0.94-1.01
	Laying hens	1.02	1.00-1.04	1.02*	1.00-1.04	1.00	0.99-1.02
	Broiler chickens	1.01	0.99-1.03	1.00	0.99-1.02	1.01	0.99-1.02
	Pigs	0.95*	0.91-1.00	0.98	0.94-1.01	0.98	0.96-1.01
Broiler chicken-related	**Sex**						
	Male				*Reference category*		
	Female	1.09	0.92-1.30	1.14	0.97-1.35	1.18*	1.02-1.38
	Age category (years)						
	0-4	2.32***	1.75-3.04	2.57***	1.98-3.32	2.22***	1.74-2.82
	5-9	0.77	0.49-1.17	0.70	0.44-1.07	0.76	0.51-1.09
	10-49				*Reference category*		
	50+	0.89	0.73-1.08	0.94	0.78-1.13	0.86	0.73-1.02
	Type of animal†						
	Small ruminants	1.00	0.92-1.10	0.98	0.91-1.06	0.99	0.94-1.04
	Dairy cows	0.98	0.88-1.09	0.95	0.88-1.02	0.99	0.94-1.04
	Veal calves	1.02	0.92-1.13	1.03	0.95-1.12	1.01	0.95-1.07
	Laying hens	1.01	0.98-1.04	1.01	0.99-1.04	0.99	0.97-1.01
	Broiler chickens	1.01	0.99-1.04	1.01	0.99-1.04	1.01	0.98-1.03
	Pigs	0.92**	0.87-0.97	0.93**	0.89-0.98	0.95**	0.92-0.99

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
1.06	0.97-1.15	1.15	0.99-1.34	1.16*	1.00-1.34	1.16*	1.01-1.33	1.09	0.96-1.24
2.52***	2.22-2.87	2.80***	2.19-3.55	2.88***	2.28-3.61	2.98***	2.41-3.66	2.90***	2.35-3.55
1.43***	1.22-1.67	1.63**	1.20-2.17	1.64***	1.23-2.16	1.73***	1.33-2.22	1.75***	1.36-2.23
					Reference category				
0.72***	0.65-0.79	1.15	0.97-1.36	1.03	0.87-1.22	1.08	0.92-1.25	1.07	0.92-1.24
1.02	0.99-1.04	0.98	0.91-1.05	1.01	0.95-1.07	1.04	0.99-1.09	1.05***	1.01-1.09
0.97*	0.94-1.00	0.98	0.89-1.08	0.96	0.90-1.02	0.96	0.92-1.00	0.95**	0.92-0.99
0.98	0.95-1.01	0.96	0.88-1.05	1.01	0.94-1.08	0.99	0.94-1.04	0.99	0.95-1.03
1.01	0.99-1.02	1.04**	1.01-1.06	1.01	0.99-1.03	1.00	0.98-1.02	0.99	0.97-1.01
1.01	1.00-1.02	1.01	0.99-1.04	1.01	0.99-1.03	0.99	0.98-1.01	1.01	0.99-1.02
0.99	0.97-1.01	0.97	0.92-1.02	0.96	0.92-1.00	1.00	0.97-1.03	0.99	0.97-1.02
					Reference category				
1.18*	1.02-1.37	1.12	0.87-1.44	1.18	0.93-1.51	1.14	0.91-1.43	1.09	0.88-1.35
2.39***	1.87-3.04	1.86**	1.16-2.88	1.85**	1.15-2.86	1.80**	1.17-2.67	1.88**	1.25-2.73
0.76	0.51-1.11	0.24*	0.07-0.66	0.39*	0.15-0.88	0.32*	0.12-0.71	0.47*	0.22-0.91
					Reference category				
1.05	0.89-1.23	1.24	0.95-1.04	1.39*	1.08-1.80	1.13	0.89-1.44	1.18	0.94-1.49
1.00	0.96-1.04	0.89	0.79-1.01	0.97	0.88-1.06	0.95	0.88-1.02	0.99	0.94-1.05
0.95*	0.91-1.00	1.03	0.89-1.19	1.01	0.92-1.11	1.03	0.95-1.11	0.98	0.92-1.04
0.98	0.93-1.04	0.98	0.86-1.12	0.93	0.84-1.04	0.99	0.91-1.08	0.93*	0.86-1.00
1.00	0.97-1.02	0.99	0.95-1.04	0.99	0.96-1.03	0.96*	0.93-0.99	0.97	0.94-1.01
1.00	0.97-1.02	1.07**	1.03-1.11	1.03*	1.00-1.07	0.96*	0.93-0.99	1.01	0.98-1.04
0.99	0.96-1.02	0.92*	0.86-0.99	0.97	0.91-1.03	0.97	0.92-1.02	1.00	0.96-1.04

Table S11. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Campylobacter* (*C.*) *jejuni* and *C. coli* incidence rate and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), with the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. jejuni	**Sex**						
	Male				*Reference category*		
	Female	0.83***	0.79-0.88	0.82***	0.78-0.86	0.83***	0.80-0.88
	Age category (years)						
	0-4	0.80***	0.70-0.91	0.75***	0.65-0.85	0.78***	0.69-0.88
	5-9	0.48***	0.41-0.56	0.47***	0.40-0.55	0.46***	0.39-0.53
	10-49				*Reference category*		
	50+	1.11***	1.05-1.17	1.09***	1.04-1.15	1.11***	1.05-1.16
	Type of animal†						
	Small ruminants	1.03	0.99-1.08	1.02	0.99-1.05	1.02	0.99-1.04
	Dairy cows	1.01	0.95-1.06	1.04*	1.01-1.08	0.99	0.97-1.02
	Veal calves	0.99	0.95-1.03	0.99	0.97-1.02	1.01	0.99-1.04
	Laying hens	1.00	0.99-1.01	1.00	0.99-1.01	1.00	0.99-1.01
	Broiler chickens	1.01	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.01
	Pigs	1.00	0.98-1.03	0.99	0.98-1.01	1.00	0.98-1.01
C. coli	**Sex**						
	Male				*Reference category*		
	Female	0.98	0.82-1.16	1.00	0.85-1.18	0.98	0.83-1.14
	Age category (years)						
	0-4	0.20***	0.08-0.45	0.18***	0.07-0.41	0.27***	0.13-0.52
	5-9	0.65	0.38-1.03	0.48**	0.27-0.79	0.51**	0.30-0.82
	10-49				*Reference category*		
	50+	1.48***	1.25-1.77	1.40***	1.19-1.66	1.34***	1.14-1.58
	Type of animal†						
	Small ruminants	0.94	0.85-1.03	1.00	0.93-1.06	1.00	0.95-1.06
	Dairy cows	0.99	0.89-1.10	1.11**	1.03-1.19	1.04	0.99-1.11
	Veal calves	1.01	0.93-1.10	1.00	0.93-1.07	0.97	0.92-1.03
	Laying hens	1.00	0.97-1.03	1.00	0.97-1.02	0.99	0.97-1.01
	Broiler chickens	1.02	0.99-1.04	0.99	0.97-1.01	1.01	0.98-1.03
	Pigs	1.03	0.98-1.09	0.98	0.94-1.02	0.98	0.94-1.01

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
				Reference category					
0.84***	0.80-0.88	0.87***	0.81-0.93	0.89***	0.83-0.95	0.88***	0.82-0.93	0.87***	0.82-0.93
0.73***	0.65-0.83	0.98	0.83-1.14	1.00	0.85-1.16	0.96	0.82-1.12	0.97	0.84-1.13
0.46***	0.40-0.54	0.49***	0.39-0.60	0.50***	0.40-0.61	0.50***	0.41-0.61	0.55***	0.45-0.66
				Reference category					
1.09***	1.04-1.15	1.20***	1.12-1.28	1.21***	1.13-1.30	1.22***	1.14-1.31	1.22***	1.14-1.30
1.01	0.99-1.03	0.98	0.93-1.03	1.00	0.97-1.04	1.02	0.99-1.04	1.00	0.98-1.02
0.98	0.97-1.00	1.01	0.96-1.08	1.01	0.98-1.05	0.99	0.96-1.02	0.99	0.97-1.01
1.00	0.98-1.01	1.00	0.96-1.05	1.00	0.97-1.03	1.00	0.97-1.03	0.98	0.96-1.00
1.01	1.00-1.01	1.01	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.01	1.01	1.00-1.02
0.99	0.99-1.00	1.01	1.00-1.02	1.00	0.99-1.01	1.00	0.99-1.01	0.99	0.98-1.00
1.00	0.99-1.01	0.98	0.96-1.01	0.99	0.97-1.01	0.98	0.97-1.00	0.99	0.98-1.00
				Reference category					
1.01	0.86-1.18	1.15	0.90-1.47	1.10	0.86-1.39	1.17	0.92-1.49	1.10	0.87-1.38
0.27***	0.13-0.51	0.42	0.14-1.04	0.30*	0.09-0.85	0.10*	0.01-0.54	0.09*	0.01-0.47
0.44**	0.25-0.73	0.50	0.19-1.12	0.57	0.23-1.21	0.56	0.23-1.19	0.33*	0.11-0.81
				Reference category					
1.01	0.98-1.03	2.19***	1.70-2.85	2.21***	1.72-2.86	2.15***	1.67-2.76	2.12***	1.68-2.69
0.99	0.95-1.03	1.06	0.93-1.21	1.02	0.93-1.13	1.06	0.98-1.15	1.09**	1.03-1.15
1.01	0.97-1.06	0.94	0.82-1.07	0.98	0.88-1.09	0.98	0.91-1.07	0.98	0.92-1.03
0.99	0.95-1.03	0.98	0.87-1.10	1.07	0.97-1.18	0.98	0.90-1.06	0.96	0.90-1.02
1.01	0.98-1.03	0.98	0.95-1.02	0.97	0.94-1.01	0.97	0.94-1.01	0.99	0.96-1.02
0.99	0.96-1.01	1.01	0.98-1.05	0.99	0.96-1.02	1.01	0.98-1.04	0.98	0.95-1.01
0.98	0.95-1.01	1.05	0.97-1.13	1.00	0.94-1.06	1.01	0.96-1.06	1.00	0.96-1.04

Table S12. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association between *S.* Enteritidis (SE), *S.* Typhimurium (ST), including its monophasic variants (ST_{mv}), and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km²), without the correction for geographical laboratory coverage.

		90 km²		50 km²		25 km²		10 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
SE	**Sex**								
	Male	*Reference category*							
	Female	1.06	0.99-1.12	1.05	0.99-1.12	1.06	0.99-1.13	1.05	0.98-1.13
	Age category (years)								
	0-4	2.92***	2.66-3.20	2.90***	2.64-3.18	2.92***	2.66-3.21	2.82***	2.56-3.10
	5-9	1.70***	1.52-1.90	1.69***	1.51-1.89	1.70***	1.51-1.90	1.69***	1.50-1.90
	10-49	*Reference category*							
	50+	0.78***	0.73-0.84	0.78***	0.73-0.84	0.78***	0.72-0.84	0.78***	0.72-0.84
	Period of infection§								
	Summer (may-oct)	2.61***	2.43-2.80	2.61***	2.43-2.79	2.62***	2.44-2.81	2.61***	2.43-2.80
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.97	0.93-1.01	0.98	0.95-1.02	1.02	0.99-1.04	1.01	0.99-1.03
	Dairy cows	0.96	0.91-1.00	0.97	0.93-1.00	0.97*	0.94-0.99	0.96***	0.94-0.98
	Veal calves	0.96	0.92-1.00	0.98	0.94-1.01	0.99	0.96-1.01	0.98	0.96-1.01
	Laying hens	1.02*	1.00-1.03	1.01	1.00-1.02	1.01*	1.00-1.03	1.01	1.00-1.02
	Broiler chickens	0.99	0.98-1.01	1.00	0.99-1.02	1.01	1.00-1.02	1.00	0.99-1.02
	Pigs	0.99	0.96-1.01	0.99	0.97-1.02	1.00	0.98-1.02	0.99	0.97-1.00

		90 km²		50 km²		25 km²		10 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
ST/ST$_{mv}$	**Sex**								
	Male				*Reference category*				
	Female	1.03	0.97-1.08	1.03	0.98-1.08	1.02	0.97-1.08	1.02	0.97-1.08
	Age category (years)								
	0-4	4.36***	4.04-4.70	4.41***	4.08-4.75	4.39***	4.06-4.74	4.36***	4.02-4.71
	5-9	2.90***	2.66-3.15	2.89***	2.65-3.15	2.87***	2.63-3.13	2.85***	2.60-3.11
	10-49				*Reference category*				
	50+	1.12***	1.05-1.19	1.12***	1.05-1.19	1.12***	1.05-1.19	1.12***	1.05-1.20
	Period of infection§								
	Summer (may-oct)	1.66***	1.58-1.76	1.66***	1.58-1.76	1.66***	1.57-1.76	1.66***	1.57-1.76
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	1.04	0.99-1.09	1.03	0.99-1.07	1.04*	1.00-1.07	1.01	0.98-1.03
	Dairy cows	0.98	0.92-1.04	1.00	0.95-1.04	0.98	0.94-1.01	0.93***	0.90-0.96
	Veal calves	1.03	0.97-1.09	1.03	0.98-1.08	1.00	0.96-1.04	0.97	0.93-1.00
	Laying hens	1.02	1.00-1.03	1.01	0.99-1.03	1.01	1.00-1.03	1.01	0.99-1.03
	Broiler chickens	1.00	0.99-1.02	1.01	0.99-1.02	1.01	0.99-1.02	1.00	0.99-1.02
	Pigs	1.03	0.99-1.07	1.02	0.99-1.06	0.99	0.96-1.01	0.97**	0.95-0.99

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval, SE = *Salmonella* Enteritidis, ST/ST$_{mv}$ = *Salmonella* Typhimurium, including its monophasic variants
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S13. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (cattle and pigs) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km^2), without the correction for geographical laboratory coverage.

		90 km^2		50 km^2		25 km^2		10 km^2	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Cattle-related	**Sex**								
	Male	*Reference category*							
	Female	0.85	0.68-1.05	0.84	0.67-1.04	0.83	0.66-1.03	0.86	0.68-1.08
	Age category (years)								
	0-4	4.03***	2.77-5.77	4.06***	2.78-5.84	3.97***	2.71-5.72	3.80***	2.54-5.58
	5-9	2.41***	1.54-3.66	2.58***	1.66-3.91	2.59***	1.66-3.92	2.76***	1.77-4.19
	10-49	*Reference category*							
	50+	2.27***	1.77-2.93	2.27***	1.76-2.95	2.23***	1.73-2.90	2.31***	1.78-3.02
	Period of infection§								
	Summer (may-oct)	0.89	0.72-1.11	0.88	0.70-1.09	0.90	0.72-1.12	0.86	0.69-1.08
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.93	0.85-1.01	0.94	0.87-1.01	0.96	0.92-1.01	0.97	0.93-1.01
	Dairy cows	0.99	0.92-1.07	0.97	0.91-1.04	0.98	0.93-1.03	1.00	0.96-1.04
	Veal calves	0.92*	0.85-0.99	0.93*	0.87-0.99	0.95	0.90-1.00	0.94*	0.89-0.99
	Laying hens	0.97	0.94-1.01	0.97	0.95-1.00	0.98	0.95-1.01	0.97	0.94-1.00
	Broiler chickens	0.94***	0.91-0.97	0.95*	0.91-1.00	0.96*	0.94-1.00	0.97	0.93-1.00
	Pigs	0.94**	0.89-0.98	0.95*	0.91-1.00	0.94**	0.90-0.98	0.95**	0.91-0.98

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Pig-related	**Sex**								
	Male				*Reference category*				
	Female	1.04	0.99-1.09	1.04	0.99-1.09	1.04	0.99-1.10	1.05	1.00-1.10
	Age category (years)								
	0-4	3.99***	3.71-4.29	4.04***	3.76-4.34	4.02***	3.73-4.32	3.97***	3.68-4.28
	5-9	2.54***	2.33-2.76	2.53***	2.32-2.75	2.52***	2.31-2.74	2.49***	2.28-2.72
	10-49				*Reference category*				
	50+	1.29***	1.22-1.36	1.29***	1.22-1.37	1.29***	1.22-1.36	1.29***	1.22-1.37
	Period of infection§								
	Summer (may-oct)	1.66***	1.58-1.75	1.67***	1.59-1.75	1.67***	1.59-1.76	1.67***	1.59-1.76
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	1.00	0.97-1.03	1.00	0.97-1.03	1.00	0.98-1.03	0.99	0.97-1.00
	Dairy cows	0.99	0.94-1.03	0.99	0.96-1.02	1.00	0.97-1.02	0.97***	0.95-0.98
	Veal calves	1.03	1.00-1.07	1.02	1.00-1.05	1.01	0.98-1.03	0.99	0.97-1.01
	Laying hens	1.01*	1.00-1.03	1.01*	1.00-1.02	1.01*	1.00-1.02	1.01	1.00-1.02
	Broiler chickens	1.01	0.99-1.02	1.00	0.99-1.01	1.00	0.99-1.01	0.99	0.98-1.00
	Pigs	1.03*	1.00-1.05	1.02*	1.00-1.04	1.00	0.99-1.02	0.99*	0.97-1.00

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S14. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (laying hens and broiler chickens) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km²), without the correction for geographical laboratory coverage.

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Laying hen-related	**Sex**								
	Male	*Reference category*							
	Female	1.09**	1.03-1.16	1.09**	1.03-1.15	1.10**	1.03-1.17	1.09**	1.03-1.16
	Age category (years)								
	0-4	2.74***	2.50-2.99	2.72***	2.48-2.97	2.75***	2.51-3.01	2.65***	2.41-2.91
	5-9	1.54***	1.37-1.72	1.53***	1.37-1.71	1.53***	1.36-1.71	1.53***	1.36-1.71
	10-49	*Reference category*							
	50+	0.83***	0.78-0.89	0.84***	0.78-0.90	0.83***	0.78-0.89	0.83***	0.77-0.89
	Period of infection§								
	Summer (may-oct)	2.43***	2.27-2.59	2.42***	2.27-2.58	2.43***	2.28-2.60	2.44***	2.28-2.61
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.96*	0.92-1.00	0.98	0.95-1.01	1.01	0.98-1.03	1.00	0.98-1.02
	Dairy cows	0.96*	0.91-1.00	0.96*	0.93-1.00	0.96**	0.93-0.98	0.95***	0.93-0.97
	Veal calves	0.97	0.93-1.00	0.98	0.95-1.01	0.98	0.96-1.01	0.98	0.96-1.00
	Laying hens	1.01	1.00-1.03	1.00	0.99-1.01	1.01	1.00-1.02	1.00	0.99-1.01
	Broiler chickens	1.00	0.98-1.01	1.01	0.99-1.02	1.01	1.00-1.02	1.00	0.99-1.02
	Pigs	0.98	0.96-1.01	0.98	0.96-1.01	1.00	0.98-1.02	0.98*	0.96-1.00

		90 km²		50 km²		25 km²		10 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Broiler chicken-related	**Sex**								
	Male	*Reference category*							
	Female	1.17**	1.06-1.29	1.16**	1.05-1.28	1.17**	1.06-1.30	1.16**	1.04-1.29
	Age category (years)								
	0-4	2.23***	1.89-2.63	2.25***	1.90-2.64	2.21***	1.86-2.60	2.29***	1.93-2.70
	5-9	0.66**	0.49-0.86	0.65**	0.49-0.86	0.67**	0.50-0.88	0.69**	0.51-0.90
	10-49	*Reference category*							
	50+	1.00	0.89-1.12	1.01	0.90-1.13	0.99	0.88-1.11	1.02	0.91-1.14
	Period of infection§								
	Summer (may-oct)	2.17***	1.95-2.42	2.14***	1.92-2.38	2.20	1.97-2.45	2.14***	1.91-2.39
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	0.90***	0.85-0.95	0.90***	0.86-0.94	0.92***	0.89-0.96	0.94***	0.91-0.97
	Dairy cows	0.92**	0.87-0.97	0.93***	0.89-0.97	0.94***	0.91-0.98	0.93***	0.90-0.96
	Veal calves	0.93**	0.88-0.98	0.93**	0.89-0.97	0.94**	0.91-0.98	0.93***	0.90-0.96
	Laying hens	0.98	0.97-1.00	0.98*	0.96-1.00	0.98**	0.96-0.99	0.98*	0.96-1.00
	Broiler chickens	1.00	0.98-1.02	1.00	0.98-1.02	1.00	0.98-1.02	0.98	0.97-1.00
	Pigs	0.94***	0.91-0.98	0.94***	0.92-0.97	0.96**	0.93-0.98	0.97**	0.94-0.99

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S15. All results (IRR and 95% CI) of the univariable spatial analyses of the possible association of *Campylobacter* (*C.*) *jejuni* and *C. coli* and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) with different hexagonal areas (90, 50, 25, 10 km^2), without the correction for geographical laboratory coverage.

		90 km^2		50 km^2		25 km^2		10 km^2	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. jejuni	**Sex**								
	Male	*Reference category*							
	Female	0.85***	0.83-0.88	0.85***	0.83-0.88	0.85***	0.83-0.88	0.86***	0.83-0.89
	Age category (years)								
	0-4	0.91*	0.84-0.98	0.91*	0.84-0.98	0.91*	0.85-0.99	0.90*	0.83-0.98
	5-9	0.51***	0.46-0.56	0.51***	0.46-0.56	0.51***	0.46-0.56	0.50***	0.45-0.55
	10-49	*Reference category*							
	50+	1.14***	1.10-1.17	1.13***	1.10-1.17	1.14***	1.10-1.18	1.14***	1.10-1.18
	Period of infection§								
	Summer (may-oct)	1.73***	1.67-1.78	1.73***	1.67-1.78	1.74***	1.68-1.79	1.73***	1.67-1.78
	Winter (nov-april)	*Reference category*							
	Type of animal†								
	Small ruminants	1.05*	1.01-1.09	1.03	1.00-1.05	1.00	0.98-1.02	1.00	0.98-1.01
	Dairy cows	1.13***	1.07-1.19	1.09***	1.06-1.13	1.02	0.99-1.04	0.99	0.97-1.01
	Veal calves	1.01	0.98-1.05	1.00	0.98-1.03	1.01	0.99-1.03	0.99	0.97-1.01
	Laying hens	1.02***	1.01-1.03	1.01*	1.00-1.02	1.02***	1.01-1.02	1.01**	1.00-1.02
	Broiler chickens	1.02**	1.01-1.03	1.00	0.99-1.01	1.00	0.99-1.01	1.00	0.99-1.00
	Pigs	1.03*	1.01-1.06	1.01	0.99-1.03	0.99	0.98-1.01	0.98**	0.97-1.00

		90 km²		50 km²		25 km²		10 km²	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. coli	**Sex**								
	Male				*Reference category*				
	Female	1.07	0.96-1.19	1.09	0.98-1.22	1.07	0.96-1.19	1.07	0.96-1.19
	Age category (years)								
	0-4	0.38***	0.24-0.56	0.38***	0.24-0.56	0.38***	0.25-0.56	0.38***	0.24-0.56
	5-9	0.45***	0.31-0.65	0.43***	0.29-0.62	0.44***	0.30-0.63	0.41***	0.27-0.59
	10-49				*Reference category*				
	50+	1.64***	1.47-1.83	1.63***	1.46-1.82	1.61***	1.44-1.80	1.59	1.43-1.78
	Period of infection§								
	Summer (may-oct)	2.09***	1.87-2.34	2.05***	1.84-2.30	2.08***	1.86-2.33	2.06***	1.84-2.32
	Winter (nov-april)				*Reference category*				
	Type of animal†								
	Small ruminants	1.01	0.94-1.08	1.03	0.98-1.08	0.99	0.95-1.03	0.99	0.96-1.02
	Dairy cows	1.02	0.95-1.11	1.05	0.99-1.11	1.01	0.97-1.05	0.99	0.95-1.02
	Veal calves	1.01	0.95-1.08	1.02	0.97-1.07	0.98	0.94-1.03	0.97	0.94-1.01
	Laying hens	1.01	0.98-1.03	1.00	0.98-1.02	1.00	0.98-1.01	0.99	0.98-1.01
	Broiler chickens	1.00	0.98-1.03	1.00	0.98-1.02	1.00	0.98-1.01	0.98	0.96-1.00
	Pigs	1.03	0.99-1.08	1.02	0.98-1.06	0.99	0.97-1.02	0.98	0.95-1.00

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
§ https://www.knmi.nl/kennis-en-datacentrum/uitleg/Zomer
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S16. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Salmonella* (*S.*) Enteritidis (SE), *S.* Typhimurium (ST), including its monophasic variants, and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), without the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
Serovar	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
SE‡	**Sex**						
	Male	*Reference category*					
	Female	1.06	0.98-1.14	1.06	0.98-1.14	1.05	0.97-1.13
	Age category (years)						
	0-4	2.91***	2.61-3.24	2.88***	2.59-3.21	2.92***	2.61-3.25
	5-9	1.67***	1.46-1.90	1.66***	1.45-1.89	1.66***	1.45-1.90
	10-49	*Reference category*					
	50+	0.71***	0.65-0.78	0.72***	0.65-0.78	0.72***	0.66-0.79
	Type of animal†						
	Small ruminants	0.98	0.93-1.04	0.98	0.94-1.02	1.02	0.98-1.05
	Dairy cows	0.99	0.93-1.06	0.99	0.94-1.03	0.96*	0.93-1.00
	Veal calves	0.96	0.91-1.02	0.97	0.93-1.02	0.98	0.94-1.02
	Laying hens	1.02*§	1.00-1.04	1.01	1.00-1.03	1.02*	1.00-1.03
	Broiler chickens	1.00	0.98-1.01	1.01	0.99-1.02	1.01	1.00-1.03
	Pigs	0.98	0.95-1.02	0.99	0.96-1.02	1.00	0.97-1.02
ST/ST$_{mv}$‡	**Sex**						
	Male	*Reference category*					
	Female	1.02	0.96-1.09	1.03	0.96-1.10	1.02	0.96-1.10
	Age category (years)						
	0-4	4.02***	3.65-4.42	4.09***	3.72-4.50	4.00***	3.63-4.40
	5-9	2.79***	2.50-3.10	2.79***	2.51-3.10	2.76***	2.47-3.07
	10-49	*Reference category*					
	50+	0.96	0.88-1.04	0.96	0.89-1.04	0.94	0.87-1.02
	Type of animal†						
	Small ruminants	1.05	0.98-1.11	1.01	0.96-1.06	1.04	1.00-1.08
	Dairy cows	0.97	0.98-1.05	0.95	0.90-1.01	0.97	0.92-1.01
	Veal calves	1.01	0.94-1.09	1.06*	1.00-1.13	1.02	0.97-1.08
	Laying hens	1.01	0.99-1.03	1.00	0.98-1.02	1.01	1.00-1.03
	Broiler chickens	1.00	0.98-1.02	1.01	0.99-1.02	1.01	1.00-1.03
	Pigs	1.02	0.97-1.07	1.02	0.98-1.06	0.98	0.95-1.01

† population-weighted number of animals

‡ IRR = incident rate ratio, 95% CI = 95% confidence interval, SE = *Salmonella* Enteritidis, ST/ST$_{mv}$ = *Salmonella* Typhimurium, including its monophasic variants

§ The PAF calculated for this model and specifically laying hens was 7%

* p-value <0.05

** p-value <0.01

*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
1.06	0.98-1.14	1.12	0.99-1.26	1.11	0.99-1.25	1.14*	1.01-1.29	1.08	0.95-1.22
2.79***	2.49-3.11	2.94***	2.44-3.52	2.94***	2.44-3.53	2.94***	2.44-3.53	2.94***	2.42-3.54
1.65***	1.44-1.89	1.79***	1.43-2.22	1.81***	1.45-2.24	1.81***	1.45-2.24	1.83***	1.46-2.28
					Reference category				
0.71***	0.65-0.78	0.98	0.86-1.12	0.97	0.85-1.12	0.95	0.82-1.09	0.98	0.85-1.13
1.02	1.00-1.05	0.99	0.93-1.07	1.02	0.97-1.08	1.05*	1.00-1.10	1.05**	1.01-1.09
0.96*	0.93-0.98	0.94	0.86-1.01	0.95	0.90-1.01	0.94*	0.89-0.99	0.94***	0.90-0.97
0.99	0.96-1.02	0.97	0.89-1.04	0.98	0.93-1.05	0.97	0.92-1.02	0.98	0.94-1.03
1.01*	1.00-1.03	1.01	0.99-1.04	0.99	0.97-1.02	1.00	0.98-1.02	1.00	0.98-1.02
1.01	0.99-1.02	1.00	0.98-1.02	1.01	0.99-1.03	1.00	0.98-1.02	1.01	0.98-1.03
0.99	0.97-1.01	1.00	0.95-1.05	1.00	0.96-1.04	1.03	0.99-1.07	0.99	0.96-1.02
					Reference category				
1.01	0.94-1.08	1.07	0.98-1.16	1.06	0.97-1.16	1.06	0.97-1.15	1.07	0.98-1.17
4.00***	3.62-4.42	5.03***	4.43-5.70	5.03***	4.44-5.70	5.19***	4.57-5.89	5.08***	4.45-5.78
2.76***	2.47-3.08	3.12***	2.70-3.60	5.03***	4.44-5.70	3.10***	2.67-3.59	3.04***	2.61-3.53
					Reference category				
0.95	0.87-1.03	1.44***	1.30-1.60	1.44***	1.30-1.59	1.47***	1.33-1.63	1.47***	1.32-1.64
1.01	0.98-1.05	0.99	0.92-1.07	1.00	0.95-1.07	1.03	0.99-1.08	1.04*	1.00-1.09
0.95**	0.91-0.98	0.93	0.85-1.03	0.99	0.92-1.06	0.96	0.91-1.01	0.91***	0.87-0.96
1.00	0.95-1.05	0.98	0.89-1.08	0.95	0.88-1.03	0.97	0.91-1.04	0.95	0.89-1.01
1.01	0.99-1.04	1.00	0.98-1.03	1.01	0.98-1.03	1.01	0.99-1.04	1.01	0.98-1.04
1.01	0.99-1.03	1.00	0.98-1.03	1.01	0.99-1.03	1.01	0.98-1.03	1.00	0.98-1.03
0.98	0.95-1.01	1.04	0.98-1.10	1.02	0.97-1.06	0.99	0.95-1.03	0.99	0.96-1.03

Table S17. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (cattle and pigs) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), without the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Cattle-related	**Sex**						
	Male				*Reference category*		
	Female	0.83	0.60-1.14	0.81	0.59-1.13	0.81	0.59-1.12
	Age category (years)						
	0-4	2.95***	1.68-4.98	3.19***	1.80-5.42	2.90***	1.62-4.95
	5-9	1.47	0.69-2.85	1.77	0.86-3.36	1.72	0.84-3.24
	10-49				*Reference category*		
	50+	1.80**	1.27-2.57	1.90***	1.32-2.73	1.81**	1.26-2.60
	Type of animal†						
	Small ruminants	0.85*	0.74-0.97	0.92	0.82-1.03	0.95	0.86-1.04
	Dairy cows	1.16*	1.01-1.33	1.05	0.94-1.17	1.04	0.95-1.14
	Veal calves	1.00	0.87-1.15	0.98	0.87-1.10	0.98	0.88-1.09
	Laying hens	0.99	0.95-1.04	0.99	0.95-1.04	1.01	0.96-1.05
	Broiler chickens	0.98	0.94-1.02	1.01	0.97-1.05	1.00	0.96-1.05
	Pigs	0.95	0.87-1.02	0.97	0.90-1.04	0.93*	0.87-0.99
Pig-related	**Sex**						
	Male				*Reference category*		
	Female	1.02	0.96-1.08	1.02	0.96-1.09	1.03	0.97-1.10
	Age category (years)						
	0-4	3.59***	3.27-3.93	3.65***	3.33-3.99	3.57***	3.25-3.91
	5-9	2.43***	2.19-2.69	2.43***	2.19-2.69	2.42***	2.17-2.68
	10-49				*Reference category*		
	50+	1.08*	1.00-1.16	1.08*	1.01-1.16	1.07	0.99-1.15
	Type of animal†						
	Small ruminants	1.01	0.97-1.06	0.99	0.96-1.03	1.00	0.97-1.02
	Dairy cows	0.94*	0.89-1.00	0.96	0.92-1.00	0.98	0.95-1.01
	Veal calves	1.03	0.98-1.08	1.04*	1.00-1.07	1.02	0.99-1.05
	Laying hens	1.01	1.00-1.03	1.01	0.99-1.02	1.01	1.00-1.02
	Broiler chickens	1.00	0.99-1.01	1.00	0.99-1.01	1.00	0.99-1.01
	Pigs	1.02	0.99-1.06	1.02	1.00-1.05	1.00	0.98-1.02

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
					Reference category				
0.83	0.59-1.16	0.84	0.62-1.13	0.84	0.62-1.14	0.82	0.60-1.11	0.86	0.63-1.18
2.73***	1.47-4.82	5.46***	3.27-8.94	5.11***	3.04-8.39	5.38***	3.19-8.88	5.16***	3.00-8.68
1.84	0.90-3.50	3.66***	2.04-6.32	3.57***	2.00-6.15	3.75***	2.09-6.49	3.96***	2.20-6.89
					Reference category				
1.83**	1.26-2.66	2.92***	2.03-4.26	2.77***	1.93-4.03	2.84***	1.97-4.17	2.95***	2.03-4.36
0.93	0.86-1.01	1.02	0.88-1.18	0.98	0.87-1.10	0.94	0.85-1.04	0.97	0.89-1.05
1.06	0.98-1.15	1.14	0.99-1.31	1.08	0.96-1.20	1.05	0.96-1.16	1.04	0.96-1.13
0.99	0.90-1.09	0.88	0.77-1.01	0.94	0.83-1.05	1.04	0.94-1.15	0.95	0.86-1.05
1.01	0.96-1.06	0.99	0.95-1.04	0.99	0.94-1.03	0.99	0.95-1.04	0.98	0.92-1.03
1.01	0.96-1.06	0.93**	0.89-0.97	0.96	0.92-1.00	0.96	0.91-1.00	0.95	0.89-1.01
0.94*	0.88-1.00	0.97	0.90-1.05	0.97	0.90-1.04	0.95	0.89-1.02	0.97	0.91-1.04
					Reference category				
1.02	0.96-1.09	1.09*	1.01-1.18	1.09	1.01-1.18	1.08	1.00-1.17	1.10*	1.02-1.20
3.56***	3.23-3.91	4.83***	4.28-5.44	4.87***	4.32-5.49	5.00***	4.42-5.64	4.86***	4.28-5.50
2.40***	2.16-2.68	2.76***	2.39-3.18	2.73***	2.37-3.15	2.75***	2.37-3.18	2.69***	2.31-3.12
					Reference category				
1.07	0.99-1.15	1.73***	1.58-1.90	1.74***	1.58-1.91	1.77***	1.61-1.94	1.77***	1.61-1.95
0.99	0.97-1.01	0.95*	0.90-1.00	0.96	0.92-1.01	0.99	0.96-1.03	1.00	0.97-1.02
0.98	0.96-1.00	0.97	0.91-1.04	0.99	0.95-1.04	0.99	0.96-1.03	0.95**	0.93-0.98
1.01	0.99-1.04	1.02	0.96-1.08	1.01	0.97-1.05	0.99	0.95-1.02	1.00	0.97-1.03
1.01	1.00-1.02	1.00	0.99-1.02	1.01	0.99-1.02	1.01	1.00-1.03	1.01	1.00-1.03
1.00	0.98-1.01	1.01	0.99-1.03	1.00	0.99-1.02	1.00	0.99-1.02	0.99	0.98-1.01
0.99	0.97-1.00	1.02	0.98-1.06	1.01	0.98-1.04	1.00	0.98-1.03	1.00	0.98-1.02

Table S18. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association of grouped *Salmonella* serovars per source (laying hens and broiler chickens) and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km^2), without the correction for geographical laboratory coverage.

		Summer					
		90 km^2		50 km^2		25 km^2	
***Salmonella*-group**	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
Laying hen-related	**Sex**						
	Male				*Reference category*		
	Female	1.08*	1.01-1.16	1.08*	1.01-1.16	1.08*	1.01-1.16
	Age category (years)						
	0-4	2.69***	2.41-2.98	2.67***	2.40-2.96	2.71***	2.43-3.01
	5-9	1.48***	1.29-1.69	1.47***	1.29-1.68	1.46***	1.28-1.67
	10-49				*Reference category*		
	50+	0.74***	0.68-0.81	0.74***	0.68-0.81	0.75***	0.69-0.81
	Type of animal†						
	Small ruminants	0.97	0.93-1.02	0.98	0.94-1.02	1.01	0.98-1.05
	Dairy cows	0.99	0.93-1.05	0.98	0.94-1.02	0.96*	0.93-0.99
	Veal calves	0.97	0.92-1.02	0.98	0.94-1.03	0.98	0.95-1.02
	Laying hens	1.01	1.00-1.03	1.01	1.00-1.02	1.01	1.00-1.02
	Broiler chickens	1.00	0.98-1.02	1.01	0.99-1.02	1.01	1.00-1.03
	Pigs	0.99	0.95-1.03	0.98	0.95-1.01	0.99	0.97-1.02
Broiler chicken-related	**Sex**						
	Male				*Reference category*		
	Female	1.19**	1.06-1.35	1.19**	1.05-1.34	1.19**	1.05-1.35
	Age category (years)						
	0-4	2.37***	1.95-2.87	2.37***	1.95-2.87	2.35***	1.93-2.85
	5-9	0.79	0.58-1.06	0.77	0.56-1.04	0.80	0.59-1.08
	10-49				*Reference category*		
	50+	0.91	0.79-1.04	0.92	0.80-1.05	0.89	0.78-1.03
	Type of animal†						
	Small ruminants	0.95	0.88-1.02	0.93*	0.88-0.99	0.95	0.91-1.00
	Dairy cows	0.97	0.90-1.05	0.99	0.93-1.05	0.99	0.95-1.04
	Veal calves	0.98	0.91-1.06	0.97	0.91-1.03	0.98	0.93-1.03
	Laying hens	1.00	0.97-1.02	1.00	0.97-1.02	0.99	0.97-1.01
	Broiler chickens	1.01	0.99-1.04	1.01	0.99-1.03	1.01	0.99-1.03
	Pigs	0.96	0.91-1.00	0.97	0.93-1.01	0.97	0.93-1.00

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
				Reference category					
1.09*	1.01-1.17	1.16**	1.04-1.30	1.16***	1.04-1.29	1.18**	1.06-1.32	1.15*	1.03-1.29
2.57***	2.30-2.87	2.89***	2.42-3.42	2.86***	2.40-3.40	2.89***	2.42-3.43	2.87***	2.40-3.43
1.46***	1.27-1.67	1.71***	1.38-2.09	1.71***	1.38-2.10	1.73***	1.40-2.13	1.75***	1.41-2.15
				Reference category					
0.74***	0.68-0.80	1.09	0.96-1.23	1.09	0.96-1.24	1.07	0.94-1.21	1.10	0.97-1.25
1.01	0.99-1.04	0.99	0.93-1.05	1.02	0.97-1.07	1.04	0.99-1.08	1.03	0.99-1.06
0.96**	0.93-0.98	0.94	0.87-1.01	0.96	0.91-1.01	0.94**	0.90-0.98	0.94***	0.91-0.97
0.99	0.96-1.02	0.99	0.92-1.06	0.99	0.94-1.04	0.97	0.93-1.02	1.00	0.96-1.04
1.01	1.00-1.02	1.01	0.99-1.03	0.99	0.97-1.01	1.00	0.98-1.02	0.99	0.97-1.01
1.01	1.00-1.02	1.00	0.98-1.03	1.01	0.99-1.03	1.00	0.98-1.02	1.01	0.99-1.03
0.99	0.97-1.01	0.99	0.95-1.04	1.00	0.96-1.03	1.02	0.99-1.06	0.99	0.96-1.02
				Reference category					
1.19**	1.05-1.35	1.12	0.94-1.35	1.11	0.93-1.33	1.13	0.94-1.36	1.11	0.92-1.34
2.42***	1.98-2.94	1.93***	1.40-2.63	1.97***	1.42-2.68	1.88***	1.34-2.58	2.02***	1.45-2.76
0.81	0.58-1.10	0.36**	0.18-0.67	0.41**	0.21-0.73	0.38**	0.18-0.70	0.43**	0.22-0.77
				Reference category					
0.94	0.82-1.09	1.19	0.98-1.44	1.21*	1.00-1.47	1.20	0.99-1.46	1.18	0.97-1.44
0.99	0.95-1.03	0.91*	0.83-1.00	0.93	0.87-1.01	0.93*	0.87-0.99	0.98	0.93-1.03
0.95*	0.91-0.99	1.00	0.91-1.11	0.97	0.90-1.04	0.96	0.90-1.03	0.95	0.91-1.01
0.97	0.92-1.01	1.01	0.92-1.11	1.01	0.93-1.09	1.01	0.94-1.08	0.97	0.91-1.03
0.99	0.97-1.02	1.00	0.97-1.03	0.99	0.96-1.02	0.97	0.94-1.00	0.98	0.95-1.01
1.00	0.98-1.02	1.02	0.99-1.05	1.02	0.99-1.05	1.02	1.00-1.05	1.00	0.97-1.03
0.99	0.96-1.02	0.95	0.89-1.01	0.97	0.92-1.03	1.00	0.95-1.05	1.01	0.97-1.05

Table S19. All results (IRR and 95% CI) of the multivariable spatial analyses of the possible association between *Campylobacter* (*C.*) *jejuni* and *C. coli* incidence rate and the population-weighted number of animals (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs) for both summer and winter with different hexagonal areas (90, 50, 25, 10 km²), without the correction for geographical laboratory coverage.

		Summer					
		90 km²		50 km²		25 km²	
Species	**Variable**	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
C. jejuni	**Sex**						
	Male				*Reference category*		
	Female	0.84***	0.81-0.87	0.84***	0.81-0.87	0.84***	0.81-0.87
	Age category (years)						
	0-4	0.85**	0.77-0.94	0.85**	0.77-0.94	0.85**	0.77-0.94
	5-9	0.51***	0.45-0.57	1.08***	1.04-1.13	0.50***	0.44-0.56
	10-49				*Reference category*		
	50+	1.09***	1.05-1.14	1.01***	1.00-1.02	1.09***	1.04-1.13
	Type of animal†						
	Small ruminants	1.03	0.98-1.08	1.01	0.98-1.05	1.00	0.98-1.03
	Dairy cows	1.09*	1.02-1.16	1.09***	1.04-1.14	1.01	0.98-1.05
	Veal calves	0.97	0.93-1.02	0.98	0.94-1.01	1.00	0.98-1.03
	Laying hens	1.01	1.00-1.02	1.01	1.00-1.02	1.01**	1.00-1.02
	Broiler chickens	1.01	0.99-1.02	0.99	0.98-1.00	1.00	0.99-1.01
	Pigs	1.01	0.98-1.05	1.00	0.97-1.02	0.99	0.97-1.01
C. coli	**Sex**						
	Male				*Reference category*		
	Female	1.04	0.91-1.18	1.05	0.92-1.19	1.03	0.90-1.17
	Age category (years)						
	0-4	0.41***	0.25-0.64	0.41***	0.25-0.64	0.42***	0.25-0.65
	5-9	0.44***	0.27-0.67	0.41***	0.25-0.63	0.42***	0.26-0.65
	10-49				*Reference category*		
	50+	1.47***	1.28-1.67	1.44***	1.27-1.65	1.45***	1.27-1.65
	Type of animal†						
	Small ruminants	0.96	0.87-1.06	1.00	0.93-1.07	0.97	0.92-1.02
	Dairy cows	1.03	0.93-1.15	1.07	1.00-1.16	1.05	0.99-1.11
	Veal calves	0.98	0.90-1.08	0.98	0.91-1.05	0.99	0.93-1.04
	Laying hens	0.99	0.96-1.02	1.00	0.97-1.02	1.00	0.98-1.02
	Broiler chickens	1.00	0.96-1.03	0.99	0.97-1.01	0.99	0.97-1.01
	Pigs	1.05	0.99-1.13	1.02	0.97-1.07	1.00	0.96-1.04

† population-weighted number of animals
‡ IRR = incident rate ratio, 95% CI = 95% confidence interval
* p-value <0.05
** p-value <0.01
*** p-value <0.001

		Winter							
10 km²		**90 km²**		**50 km²**		**25 km²**		**10 km²**	
IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡	IRR‡	95% CI‡
				Reference category					
0.84***	0.81-0.87	0.87***	0.83-0.91	0.87***	0.82-0.91	0.87***	0.82-0.91	0.88***	0.83-0.92
0.84***	0.76-0.92	1.01	0.90-1.14	1.01	0.90-1.14	1.03	0.91-1.16	1.02	0.90-1.15
0.49***	0.43-0.55	0.52***	0.44-0.61	0.53***	0.45-0.62	0.53***	0.45-0.62	0.52***	0.44-0.61
				Reference category					
1.09***	1.05-1.14	1.23***	1.17-1.30	1.24***	1.17-1.31	1.25***	1.18-1.32	1.24***	1.17-1.31
1.00	0.99-1.02	1.01	0.96-1.08	1.02	0.98-1.05	1.00	0.97-1.03	1.00	0.98-1.02
0.99	0.97-1.01	1.03	0.96-1.11	1.06*	1.01-1.11	1.00	0.97-1.04	1.00	0.97-1.02
1.00	0.98-1.02	0.99	0.93-1.04	0.99	0.95-1.04	1.01	0.98-1.04	0.98	0.96-1.01
1.01*	1.00-1.02	1.01	1.00-1.03	1.00	0.99-1.02	1.01*	1.00-1.03	1.01*	1.00-1.03
1.00	0.99-1.00	1.02*	1.00-1.04	1.00	0.99-1.01	1.00	0.99-1.01	0.99	0.98-1.01
0.99	0.97-1.00	1.01	0.97-1.05	0.99	0.96-1.02	0.98	0.96-1.01	0.98*	0.97-1.00
				Reference category					
1.05	0.91-1.19	1.09	0.90-1.31	1.13	0.94-1.36	1.10	0.91-1.32	1.07	0.88-1.29
0.43***	0.26-0.67	0.29**	0.11-0.65	0.28**	0.11-0.64	0.29**	0.11-0.64	0.23**	0.08-0.57
0.41***	0.25-0.64	0.49*	0.24-0.92	0.49*	0.24-0.91	0.49*	0.24-0.92	0.39*	0.17-0.78
				Reference category					
1.43***	1.25-1.64	2.06***	1.69-2.50	2.05***	1.69-2.48	1.98***	1.63-2.42	1.97***	1.62-2.40
0.98	0.94-1.02	1.07	0.95-1.22	1.04	0.95-1.13	1.05	0.98-1.13	1.05	1.00-1.12
1.01	0.97-1.06	0.96	0.84-1.09	1.00	0.91-1.10	0.99	0.92-1.06	0.99	0.94-1.05
0.99	0.94-1.04	0.94	0.84-1.06	1.00	0.91-1.09	0.94	0.87-1.01	0.95	0.89-1.01
1.00	0.98-1.02	1.03	0.99-1.06	0.99	0.95-1.02	0.99	0.97-1.02	1.00	0.97-1.03
0.98	0.96-1.01	1.00	0.96-1.04	0.99	0.96-1.02	1.00	0.98-1.03	0.99	0.96-1.02
0.99	0.96-1.02	1.04	0.96-1.04	1.03	0.97-1.10	1.01	0.96-1.07	0.99	0.95-1.03

Table S20. Results (fixed effect estimates (β) and 95% CI) of the univariable linear mixed model of *Campylobacter* seroincidence and the population-weighted animal numbers (small ruminants, dairy cows, veal calves, laying hens, broiler chickens and pigs), including the standard deviation and 95% CI of the PC4- and residual random effects.

Variable	Model		PC4 random effect		Residual random effect	
	β‡	95% CI‡	Std. dev‡	95% CI‡	Std. dev‡	95% CI‡
Age (years)						
0-4	-0.31***	-0.43;-0.19	0.06	0.00;0.13	0.75	0.72;0.77
5-9	-0.35***	-0.49;-0.21	0.06	0.00;0.13	0.75	0.72;0.77
10-49			*Reference*			
≥50	0.07	-0.04; 0.17	0.06	0.00;0.13	0.75	0.72;0.77
Sex						
Male						
Female	0.09*	0.01; 0.17	0.07	0.00;0.13	0.76	0.73;0.78
Type of animal†						
Small ruminants	< 0.001	-0.01; 0.01	0.07	0.00;0.13	0.76	0.73;0.79
Dairy cows	<-0.005	-0.02; 0.01	0.07	0.00;0.13	0.76	0.73;0.79
Veal calves	< 0.005	-0.02; 0.01	0.07	0.00;0.13	0.76	0.73;0.79
Laying hens	0.01	-0.01; 0.02	0.07	0.00;0.13	0.76	0.73;0.79
Broiler chickens	-0.01	-0.03; 0.01	0.06	0.00;0.13	0.76	0.73;0.79
Pigs	<-0.001	-0.01; 0.01	0.07	0.00;0.13	0.76	0.73;0.79

† population-weighted number of animals
‡ β = fixed effect estimate, 95% CI = 95% confidence interval, std. dev = standard deviation
* p-value <0.05
** p-value <0.01
*** p-value <0.001

Text

Text S1:
Overall, 19,771 salmonellosis cases were reported in the Netherlands between 2007 and 2019. 1,686 of those cases were outbreak-related and therefore excluded. Other reasons for exclusion were: unknown age (n = 81), sex (n = 563) or postal code (n = 500), being a travel-related case (n = 2,336), having a chronic infection (n = 672) or being a duplicate isolate from the same patient (n = 119). A chronic infection was defined when multiple isolates belonging to the same *Salmonella* serovar were obtained from the same patient within six months. (Mughini-Gras et al., 2020)* After applying correction for geographical laboratory coverage, 12,833 out of 13,814 salmonellosis cases remained with a median of 982 cases per year (range 790-1,372 cases/year, annual incidence 4.7-8.4/100,000 inhabitants).

Text S2:
The 13,814 cases in the cleaned salmonellosis database were grouped based on the primary source of the *Salmonella* serovars causing the infection as derived from the source attribution analysis. During this grouping, 461 additional cases were removed, as their serovars did not match any of the included sources (i.e. laying hens, broiler chickens, cattle, pigs or reptiles). Out of the remaining 13,353 cases, 1,261 cases were mainly attributable to reptiles. Of those, 298 were removed, as they did not have a secondary (livestock) source. The other 963 cases with reptiles as primary source were redistributed over their secondary livestock sources (pigs: 431 cases, broiler chickens: 369 cases, laying hens: 157 cases and cattle: 6 cases). This resulted in the following distribution over the different groups of the remaining 12,161 cases after the correction for laboratory coverage was applied: 4,153 cases were mainly attributable to laying hens, 1,377 to broiler chickens, 6,329 to pigs and 302 to cattle.

Text S3:
For the *Salmonella* analyses (2007-2019), animal data from the 2012, 2015 and 2018 censuses were used for the years 2007-2013, 2014-2016 and 2017-2019, respectively, while the annual human population data and postal code region shapefiles were used. The postal code region shapefile of 2009 was used for 2007-2008 (Mulder et al., 2020).

For *Campylobacter*, the periods of the data differed between case data (2014-2019) and infection pressure data (2006-2007). For the campylobacteriosis cases, animal data from 2015 (for the years 2014-2016) and 2018 (for the years 2017-2019) was included, together with the annual human population data and the postal code region shapefiles from 2014-2019. For the *Campylobacter* infection pressure data, animal

data from 2012 were used together with the annual human population data from 2006-2007. Here, the postal code region shapefile of 2009 was used for 2006-2007 (Mulder et al., 2020).

Text S4

Between 2014 and 2019, 17,549 *Campylobacter* isolates representing individual cases of campylobacteriosis were reported in the Netherlands. Cases with unknown postal code (n = 1) or with a chronic infection (n = 6) were excluded. A chronic infection was defined when multiple isolates belonging to the same *Campylobacter* species were obtained from the same patient within two months. (Brachman & Evans, 1998; Pasternack, 2002)* After correcting for laboratory coverage, a total of 15,674 campylobacteriosis cases remained as input for the spatial analyses (median: 2,633 cases/year; range 2,293-2,851 cases/year, annual incidence 13.3-16.9/100,000 inhabitants). Of these, 14,431 were *C. jejuni* and 1,243 *C. coli*.

Text S5:

The surveillance systems for *Salmonella* and *Campylobacter* are both sentinel, meaning that they do not have nationwide coverage. This introduces a certain type of bias in the case- and exposure dataset due to an incomplete geographical coverage of the diagnostic laboratories. To correct for this incomplete geographical coverage, it is important to know which areas are covered by at least one laboratory in the Netherlands. Therefore, a method was developed for this within this study.

To make sure that we only included postal code areas in our analyses that were covered by the surveillance systems, we had to identify those postal codes. This was done separately for *Campylobacter* and *Salmonella*. As a first step, only postal code areas were selected that included cases that were reported by laboratories which reported cases consistently over all study years (Salmonella: 2007-2019; Campylobacter: 2014-2019). Next, the expected number of cases was calculated for those selected postal code areas by multiplying the incidence rate of *Salmonella* or *Campylobacter* during the study period with the number of inhabitants for each postal code area. The number of expected cases was then compared with the number of observed cases using a Poisson model in STATA version 16.0 (StataCorp, College Station, TX, USA), with the number of inhabitants as offset variable. The latter was obtained from the Central Bureau for Statistics (https://www.cbs.nl). Postal code areas that had no cases or had significantly less observed cases than expected based on the Poisson model ($p\text{-value}<0.05$) were considered not covered by the surveillance. Additionally, postal code areas that were defined as covered, but had adjacent postal code areas that were all not covered, were redefined as not-covered. Conversely, postal code areas that were defined as not covered, were considered covered if all adjacent postal code

areas were classified as covered. This was done in ArcMap (version 10.6.1). Lastly, an overlay was created between the covered postal code areas and the hexagonal areas in R. Only hexagon's of which >80% of its area consisted of covered postal code areas were included in further analyses. Thus, the dataset was corrected for the geographical coverage of the diagnostic laboratories in this way.

*All literature was included in the reference list belonging to the main text of this article.

CHAPTER 6

Spatial association between human gut microbial community and agricultural land coverage: a cross-sectional population-based study in the Netherlands

Annemieke C. Mulder
David Boverhoff
Mari-Lee Odendaal
Susana Fuentes
Jolanda Kool
Sudarshan A. Shetty
Liza Tymchenko
Dick Heederik
Lapo Mughini-Gras
Roan Pijnacker
Eelco Franz

In preparation

Abstract

Although all individual microorganisms within the human gut microbiome strive towards a relatively stable equilibrium, its composition and diversity can be influenced by many factors throughout life, of which the human living environment is an example. These environmental effects, particularly the urban-rural gradient, are poorly studied despite increasingly availability of high-resolution microbiome and spatial data. The main aim of this study was to assess the potential effects of agricultural land coverage on both diversity and composition of microbial communities in the human gut of adults (aged 40 years or older). This was done using a comprehensive data set from a nationwide cross-sectional population-based study (n=1,852) coupled with highly detailed land use data with a spatial resolution of 25x25 meters for 2016-2017. Results showed a significant increase in bacterial diversity and two microbial clusters, one dominated by the genus *Blautia* and one dominated by *Bifidobacterium, Collinsella* and *Akkermansia*, to be significantly positively associated with higher agricultural land coverage. Furthermore, we identified two genera, *Barnesiella* and *Leuconostoc*, with a significant higher relative abundance in participants living in areas with a high agricultural land coverage. Thus, we observed an effect of the urban-rural gradient on diversity and composition of the adult human gut microbiome. As this study was based on an ecological design, causal inference is limited and further research using individual-level data is recommended.

Introduction

Microorganisms are everywhere, both inside and outside the human body. These microorganisms, together with their genomes, constitute the human microbiome, whose diversity and composition differ depending on the anatomical sites where they reside, e.g., the skin, airways, intestines, vagina, mouth. The human gut microbiome plays an important role in health and disease, both physically and mentally.[1-5] Although all microorganisms within the human gut microbiome strive towards a relatively stable equilibrium, its composition and diversity can be influenced by many factors throughout life, such as the living environment.[6-8]

Studies focusing on effects of environmental microbial exposure on human health gained popularity when it was suggested that a change in lifestyle (particularly 'westernization') influenced the occurrence of asthma.[9-13] Several studies identified microorganism-related factors (e.g. endotoxins, β(1→3)glucans and muramic acid) that influence the development of the immune system, which in turn leads to protective effects against asthma and atopy.[13-23] More recently, the reduced risk for asthma and atopy development in farmers' children was associated with diversity of microbial exposure.[1] Protective effects of the farming environment were more specifically related to growing up or living on a farm and having regular contact with farm animals[5,6]. This gives rise to the hypothesis that these associations could be explained by an increased environmental microbial exposure in the farm environment.

The increased environmental microbial exposure in agricultural areas is not restricted to farms alone. Outdoor microorganisms or microorganisms found in farmhouses are also present in surrounding air, soil, dust, water, and even on surfaces and in homes.[24-26] Besides, several studies have observed direct transfer of those microorganisms from farms to residential areas through air.[15,25,26] Those microorganisms mainly deposit in the human respiratory organ after inhalation[9] and are partially ingested and cleared through the gastrointestinal tract, which leads to both oral and respiratory uptake of microorganisms through which they contribute to the transient members of the microbiome.[9]

A limited number of studies investigated and showed the potential health effects of the farming environment on local residents beyond farmers and their families. For example, living in close proximity to livestock farms was associated with protection against asthma and atopy.[27,28] When focusing more specifically on the effects of environmental microbial exposure on the diversity and composition of the human gut microbiome, differences in both airway and gut microbiotas between infants in urban and rural living environments have been observed.[29] Besides, an elevated risk of asthma and atopic traits with urbanization-related changes in the microbiota have

been shown.[29] As environmental microbial exposure affects the human microbiome throughout life[6,30], it would be interesting to study whether human gut microbiome diversity and composition in adults is affected by the urban-rural gradient throughout the Netherlands.

In this study, we assessed the potential effects of agricultural land coverage on both the diversity and composition of microbial communities in the human gut of adults (aged 40 years or older). To this end, extensive data were used from a nation-wide cross-sectional population-based study (n=1,852) coupled with a highly detailed land use database with a spatial resolution of 25x25 meters in the Netherlands in 2016-2017.

Methods

Study population

To assess the effects of agricultural land coverage on the diversity and composition of the human gut microbiome of adults, fecal samples from 1,852 people aged 40 years or older were collected within the framework of a large cross-sectional population-based study in the Netherlands (PIENTER-3). Sample collection was accompanied by self-administered questionnaires with extensive metadata on demographics, health status and dietary habits. More detailed information about the PIENTER-3 study can be found elsewhere[31]. Although the PIENTER-3 study was primarily designed to obtain population-based seroprevalence estimates of vaccine-preventable diseases, it also included the collection of fecal samples amongst others.[31] The study was approved by the Medical Ethics Committee of North-Holland (METC number M015–022). Written informed consent was obtained from all participants. Fecal samples were collected in 2016-2017, kept in the participants' freezer at home until they were delivered in cold packs to the mobile study team, kept on dry ice during transport to the Dutch National Institute for Public Health and the Environment (RIVM) and stored at -80°C until processing.[31]

Fecal DNA extraction and sequencing

Mechanical cell disruption was used for fecal DNA extraction. This method included pre-assembled bead-beating tubes that contained a combination of 0.5g zirconia/silica beads (0.1 mm) and 5 glass beads (2.7 mm) (Biospec products, Bartlesville, OK, USA). Approximately 0.25 g of stool material and 700 ul STAR buffer were added to the pre-assembled bead-beating tubes.

Extraction controls (positive and negative) were added by using prepared samples (Supplementary material, Control samples). Repeated bead-beating was established

by the Fastprep-24 (MP Biomedicals, Irvine, USA) and was performed 3 x 1 minute for 5.5 ms. After this procedure, samples were heated to a temperature of 95˚C for 15 minutes and subsequently centrifuged till lysates could be collected. After addition of 350 ul Stool Transport and Recovery (S.T.A.R.) buffer, the bead-beating step was repeated, followed by a heat step of 95˚C for 7 minutes. Lysates were collected and purified using the Maxwell.

The Maxwell RSC blood DNA kit and the Maxwell RSC instrument (Promega, Madison, USA) were used for purification of DNA extraction, while DNA was eluted in 60 µl elution buffer and additionally purified using the OneStep PCR Inhibitor Removal Kit (ZymoBIOMICS, Zymo Research, Irvine, CA, USA).

Quantification

The total DNA concentration was determined by the Quantus fluorometer (Promega, Madison, USA) after which it was stored at -20˚C until further processing. Bacterial content in purified DNA samples was measured by a quantitative PCR (qPCR) (StepOnePlus Real-Time PCR System, Thermo Fisher Scientific, the Netherlands) using a universal primer set targeting the 16S rRNA gene (forward Eub341F: CCTACGGGAGGCAGCAG, reverse Eub534R: ATTACCGCGGCTGCTGGC).[32] qPCR was performed by addition of a SYBR Green (25µl) reaction mix containing: 12.5 µl Maxima SYBR Green/ROX qPCR Master Mix (ThermoFisher scientific, Waltham, MA, USA), 0.5 µM forward primer Eub341F, 0.5 µM reverse primer Eub534R, 2 µl DNA (500 times diluted in HPLC grade water) and 8 µl HPLC grade water. (PCR program: denaturation (95˚C; 10 min), 40 cycles of denaturation (95˚C; 15 sec), annealing (60˚C; 15 sec), extension (72˚C; 15 sec), holding stage (95˚C; 1 min and 60˚C; 1 min))

Amplification of the V4 16S rRNA fragment

qPCR concentrations were used to calculate bacterial input for the V4 16S rRNA PCR. Primers used included the specific V4 primers: 515F (5'- GTG CCA GCM GCC GCG GTA A-3') and 806R (5'-GGA CTA CHV GGG TWT CTA AT-3'), the Illumina flow cell adapter and a unique 8-nt index key.[33] PCR mixture: 0.5 µl (1 U) Phusion Hot Start II High-Fidelity DNA Polymerase, 5 µl 5x Phusion HF Buffer (Thermo Fisher Scientific), 7 µl HPLC grade water, 2.5 µl of 2mM dNTP mix (ThermoFisher scientific, Waltham, MA, USA), 0.5 µM of forward primer 515F, 0.5 µM of reverse primer 806R and 5 µl template DNA. (PCR program: After denaturation (98˚C; 30 sec), 30 cycles were performed consisting of denaturation (98˚C; 10 sec), annealing (55˚C; 30 sec), extension (72˚C; 30 sec) and a holding stage (72 ˚C; 30 sec))

The size and concentration of the amplified product was measured on the Qiaxcel Advanced System (Qiagen, Hilden, Germany) using QIAxcel DNA High Resolution Kit. Concentrations of the specific V4 fragment were used to pool the amplified product equimolar, 144 samples were pooled in every sequence pool. The pool was purified twice by using 0.9x AMPure XP magnetic beads (Beckman Coulter, the Netherlands) to remove primer dimers. The purified pool was quantified by qPCR using the KAPA library quantification kit (Roche, USA). The pool was sequenced paired-end on an Illumina Miseq instrument (Illumina Inc., San Diego, CA, USA) using a V3 Miseq reagent kit (600 cycles).

Bioinformatics

For the current study, a modified version of the DADA2 pipeline[34] (11 Triumph: Run pipeline | Bioinformatics Pipelines for Bacteriology (IDS, RIVM) (rivm-bioinformatics.github.io) was used to process the paired-end fastq files obtained from the Illumina Miseq instrument (Illumina Inc., San Diego, CA, USA), after demultiplexing and primer removal (v1.16.0; maxEE = 2; truncLen = 220/100). The Silva v138 (version 2; August 2020) reference database and the naïve Bayesian classifier were used for taxonomic assignment by which the reverse-complement sequence was utilized for classification when a better match was found.[35]

Definition of agricultural land coverage

The LGN7 database (grid file) contains land use information of the Netherlands in 2012 with a spatial resolution of 25x25 meters.[36] For this study, the database with general classes was used, distinguishing eight land use classes: agriculture, greenhouse or horticulture, orchards, forest, water, built-up area, infrastructure, and nature. The percentage of land coverage by land use class was calculated on the level of postal codes for each individual participant by creating a mask from the LGN7 grid file using a spatial data set of postal codes in the Netherlands from 2012. The preparation and merge of the spatial data with the individual participants were performed using the sf (version 1.0.8)[37] and raster package (version 3.6.3)[38] in R version 4.2.1 with RStudio Server version 2022.07.2, build 576 (Boston, MA).

Statistical analyses

All statistical analyses were performed in R version 4.2.1 with RStudio Server version 2022.07.2, build 576 (Boston, MA). For handling the microbiota data, three packages were used: the microbiome package (version 1.16.0)[39], the biomeUtils package (version 0.13)[40] and the phyloseq package (version 1.38.0)[41].

Filtering

Phyla with an abundance of less than 0.005% ($n_{samples}$ = 14) were removed from the phyloseq before α-diversity was calculated. Additionally, spurious amplicon sequence variants (ASVs; n_{taxa} = 20,313) were removed, which were not eligible for inclusion in the analyses of microbial community composition (β-diversity). This was done based on a detection threshold for absence/presence count of 5 and a prevalence threshold of 0.1%. Cut-off values were based on both literature[42,43] and expert opinion, depending on our data and biological relevance. Furthermore, relative abundances were calculated and used as input for all analyses involving β-diversity.

Bacterial diversity

Observed richness and Shannon diversity index were used as metrics for α-diversity. Both indices of the samples from the different agricultural land coverage categories (no agriculture, 0-25%, 25-50%, 50-75% and 75-100%) were compared using the non-parametric Wilcoxon rank sum test (wilcox.test) from the R Stats package (version 4.2.1)[44] including Benjamini-Hochberg corrections for multiple testing[45].

Microbial community composition

Microbial community compositions were visualized using principle coordinate analysis (PCoA) based on Bray-Curtis dissimilarity[46] and related differences between groups were tested with permutational multivariate ANOVA using the vegan package (version 2.6.2)[47] with 999 permutations at genus level for both agricultural land coverage categories and other covariates (i.e. age, gender, ethnical background, monthly income, SES-WOA score, country of birth of the father, country of birth of the mother, and country of birth of the participant). The SES-WOA score is an average score for social-economic wellbeing, based on welfare, level of education, and recent work experience which was matched with the participants based on postal code. Data was obtained from: SES per postal code between 2014-2019, excl. studenten (cbs.nl). Homogeneity of variances was checked using the permutational multivariate of dispersion analysis with the "permdisp2" function.[48] To assess collinearity between covariates, correlations and their associations were computed based on the Chi-Square test of independence for the aforementioned categorical covariates. Pearson correlation coefficients were calculated to assess collinearity between categorical and continuous variables. As country of birth of the father, country of birth of the mother, country of birth of the participant, and ethnical background were significantly associated with each other and with agricultural land coverage, only ethnical background was included as covariate. Results of the selected covariates from the permutational multivariate ANOVA are shown in the main text. All other results regarding the observed richness and Shannon diversity, the PCoA

plots and the permutational multivariate ANOVA are provided in the supplementary material (Other results: Figure S1-S17).

Microbial clusters

To study the distribution of genera across agricultural land coverage in more detail, a hierarchical clustering method with complete linkage was applied using the Bray-Curtis dissimilarity matrix. To determine the optimal number of clusters, the Calinski-Harabasz index for K-means clustering evaluation was used.[49] Generalized linear models (GLM) with a binomial family were used to study associations between the aforementioned clusters and the different agricultural land coverage categories and adjusted for multiple testing using the Benjamini-Hochberg correction[50].

Associations at genus level

Associations of the relative abundance of the specific microbial genera with agricultural land coverage were identified using the multivariable analysis by linear models (MaAsLin2) statistical framework which relies on general linear models by implementing the Maaslin2 package (version 1.8.0)[51], which were always adjusted for age (years), gender (woman, man), monthly income (< €971, €971-1,355, €1,356-1,969, €1,970-3,314, €3,315-3,500, > €3,500, missing) and ethnical background (Dutch, non-Dutch). P-values were subjected to multiple hypothesis testing with correction using the Benjamini-Hochberg method[50] with an FDR threshold of 0.25. As monthly income was included in the calculation of the SES-WOA score, an extra MaAsLin2 analysis was performed in which SES-WOA score replaced monthly income in the model (Supplementary material, Other results: Table S1). These analyses yielded similar results and therefore only results of the model with monthly income are shown in the main text.

Role of the funding source

This study was supported by the research project "TRIuMPH – The RIVM mIcrobiome and Metagenome facility for Public Health", funded by the Strategic Program RIVM (SPR). The funding source had no involvement in the design, conduct and analysis of the study and/or preparation of the article.

Results

Descriptive statistics

The 1,852 participants had a median age of 60 years (25-75% percentile P_{25-75} 50 – 68). Of all participants included, 55% (n = 1,019) were women, 45% (n = 833)

were men (Table 1). Most participants were native Dutch (75%, n = 1,386) whereas 25% (n = 466) had another ethnical background. Most participants had a monthly income between €1,970 and €3,314 and a median SES-WOA score of 0.05 (25-75% percentile P_{25-75} = -0.14;0.14). Overall, participants lived in postal code areas with a median agricultural land coverage of 23% (25-75% percentile P_{25-75} = 1 – 53). Figure 1A shows the geographical distribution of the agricultural land coverage categories in percentage per unique postal code area of the PIENTER-3 participants. This map showed that the postal code areas with the highest agricultural land coverage were in the Northern part of the Netherlands, whereas most postal code areas with the lowest agricultural land coverage were in the mid-Western part of the country. The microbiome data set consisted of 4,194 taxa (2 kingdoms, 11 phyla, 18 classes, 44 orders, 80 families, 269 genera and 310 species).

Table 1. Socio-demographic characteristics of the participants included in this study

	Overall	
	N	%
Total	1,852	100
Gender		
Men	833	45
Women	1,019	55
Ethnical background		
Dutch	1,386	75
non-Dutch	466	25
Monthly income		
< €970	83	4
€971 - 1,355	191	10
€1,356 - 1,969	326	18
€1,970 - 3,314	573	31
€3,315 - 3,500	118	6
> €3,501	364	20
Missing	197	11
Agricultural land coverage		
No agriculture	317	17
0-25%	638	34
25-50%	296	16
50-75%	430	23
75-100%	171	9

Gut microbial composition and agricultural land coverage

The differences in microbial composition and structure in relation to the agricultural land coverage categories are visualized in the PCoA plot of Figure 1B. The PCoA defined by the agricultural land coverage categories explained 7.7% of the variation in gut microbiota composition amongst participants with the first principle coordinate axis and 6.2% with the second axis. Overall, the participants shifted gradually to the upper-right segment of the plot with an increase in agricultural land coverage. Figure 1C shows that this increase is statistically significant with an explained variance of 0.7%. Ethnical background was found to be the variable with the highest variance explained (1.4%), followed by monthly income (0.9%), SES-WOA score (0.8%), age (0.8%) and gender (0.6%). Those variables were used as covariates in the multivariate analyses for the associations between the relative abundances of the genera and the agricultural land coverage.

Overall, the participants within this study had a mean Shannon index of 4.02 and a mean observed richness of 212 ASVs per sample. Figure 2 presents those metrics per agricultural land coverage category. Figure 2A shows a significant increase in Shannon index with an increasing percentage of agricultural land coverage in the postal code area of participants after correcting for multiple testing with the Benjamini-Hochberg correction (no agriculture: median Shannon index = 3.95, Q_1-Q_3 = 3.60-4.21; 0-25%: median Shannon index = 4.06, Q_1-Q_3 = 3.79-4.31; 25-50%: median Shannon index = 4.12, Q_1-Q_3 = 3.88-4.32; 50-75%: median Shannon index = 4.08, Q_1-Q_3 = 3.78-4.27; 75-100%: median Shannon index = 4.19, Q_1-Q_3 = 3.93-4.37). The median Shannon index of 50-75% agricultural land coverage is lower when compared to the previous (25-50%) and latter (75-100%) agricultural land coverage categories. Figure 2B shows similar results for richness (no agriculture: median richness = 187, Q_1-Q_3 = 143-230; 0-25%: median richness = 208, Q_1-Q_3 = 165-249; 25-50%: median richness = 213, Q_1-Q_3 = 178-262; 50-75%: median richness = 205, Q_1-Q_3 = 164-259; 75-100%: median richness = 221, Q_1-Q_3 = 179-274).

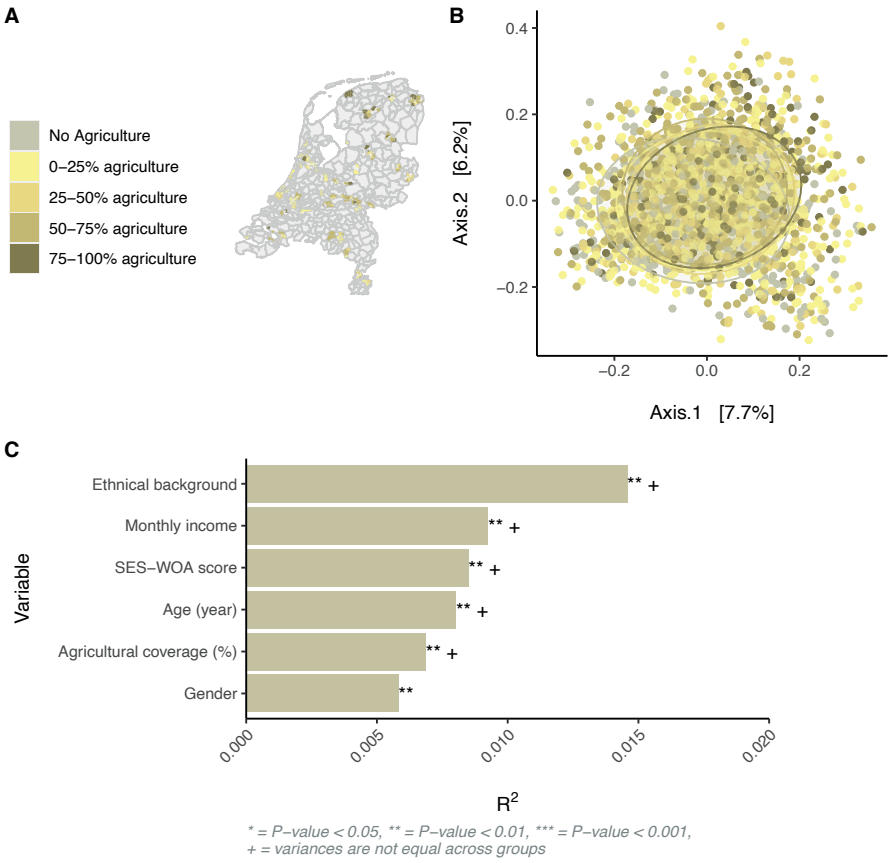


Figure 1. A) Geographical distribution of the agricultural land coverage categories in percentage (no agriculture, 0-25% agricultural coverage, 25-50% agricultural coverage, 50-75% agricultural coverage, 75-100% agricultural coverage) per unique postal code area of the study participants. **B)** PCoA visualizing the β-diversity based on Bray-Curtis dissimilarity of the study participants, colored by agricultural land coverage category. **C)** Microbiome variance explained per variable, including its significance level (* = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001) based on the permutational multivariate ANOVA results and whether the condition of homogeneity of variances was met, which was checked using the PERMDISP2 method (+ = variances are not equal across groups).

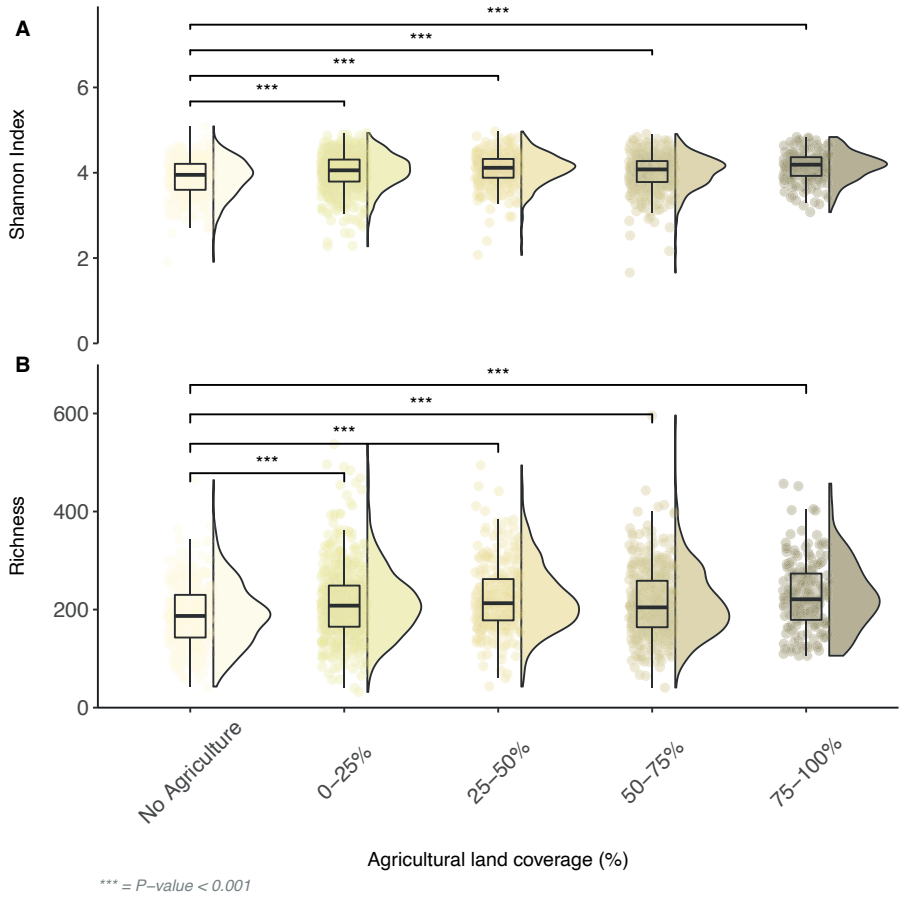


Figure 2. A) Combined bar- and violin plot of the Shannon index distribution over the different agricultural land coverage categories (no agriculture, 0-25%, 25-50%, 50-75%, 75-100%). **B)** Combined bar- and violin plot of the observed richness over the different agricultural land coverage categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

Microbial clusters and agricultural land coverage

Results of the hierarchical clustering analysis at genus level in relation to agricultural land coverage showed that the microorganisms present in the gut of the study participants could be grouped into seven major clusters (Figure 3). Each cluster was defined by and therefore named after one or two dominant genera. The main cluster among participants aged 40 years or older was dominated by *Blautia* with a relative abundance of 12% (Figure 3A), which was also significantly positively associated with three agricultural land coverage categories: 0-25% (adjusted odds ratio [aOR] = 1.42, 95% confidence interval [CI]: 1.07-1.89), 50-75% (aOR = 1.46, 95%CI: 1.07-1.99) and 75-100% (aOR = 1.69, 95%CI: 1.12-2.57) (Figure 3B). The other cluster, which was dominated by *Bifidobacterium*, followed by *Collinsella* and *Akkermansia*, with relative abundances of 11%, 6% and 6%, respectively, was significantly positively associated with two agricultural land coverage categories: 25-50% (aOR = 3.55; 95%CI: 1.48-9.88) and 75-100% (aOR = 3.91; 95%CI: 1.49-11.42). The cluster dominated by *Prevotella (9)* (relative abundance = 30%) was significantly negatively associated with 75-100% agricultural land coverage (aOR = 0.38; 95%CI: 0.17-0.78). The other four clusters did not show significant associations with agricultural land coverage categories.

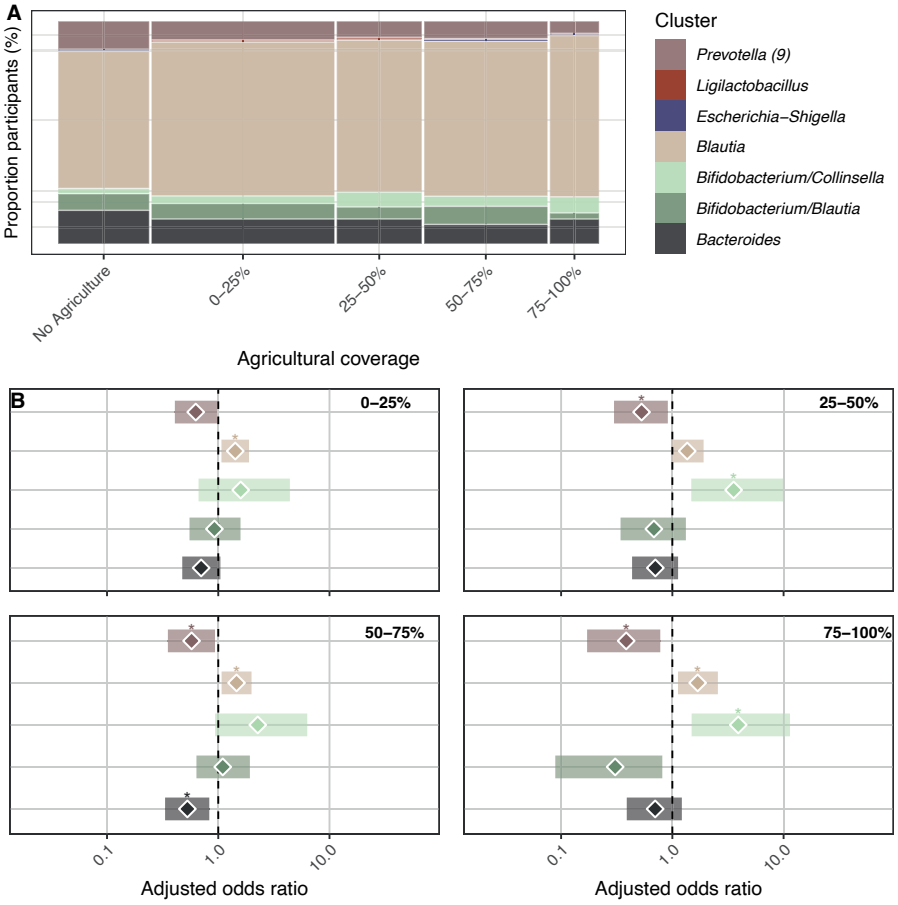


Figure 3. A) Mosaic plot representing the results of the hierarchical clustering analyses with complete linkage (n = 7 clusters) based on the Calinski-Harabasz index for K-means clustering evaluation. Each cluster was named after the most dominant genus within that cluster. This plot visualized the proportion of participants within each cluster per agricultural land coverage category (no agriculture, 0-25%, 25-50%, 50-75%, 75-100%). **B)** Forest plot containing the univariate analyses results of the associations between each agricultural land coverage category and each cluster (n > 10 participants) separately. P-values were corrected for multiple testing using the Benjamini-Hochberg correction. The results include its significance level (* = p-value < 0.05, ** = p-value < 0.01)

Associations between genera and agricultural land coverage

Fourteen genera were significantly associated with the percentage of agricultural land coverage in the postal code areas of the participants, after adjusting the p-values for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) with an FDR threshold of 0.25 and correcting for the confounders age, gender, ethnical

background and monthly income (Table 2). The three genera with a significant p-value of < 0.05 were *Slackia* (β = -0.21, FDR p-value = 0.02), *Barnesiella* (β = 0.23, FDR p-value = 0.03) and *Leuconostoc* (β = 0.14, FDR p-value = 0.03).

Table 2. Results of the multivariable analysis studying the associations between specific genera in the gut microbiome of study participants and agricultural land coverage (%) as continuous variable while adjusting for age, gender, ethnical background and monthly income using the multivariable analysis by linear models (MaAsLin2) statistical framework (coefficient and false discovery rate (FDR) p-value (subjected to multiple hypothesis testing correction using the Benjamini-Hochberg method with an FDR threshold of 0.25)).

Genus	β*	FDR p-value**
Slackia	-0.21	0.02
Barnesiella	0.23	0.03
Leuconostoc	0.14	0.03
Incertae Sedis	0.11	0.04
Lachnospiraceae UCG 010	0.17	0.10
Phascolarctobacterium	-0.23	0.10
Prevotellaceae NK3B31 group	-0.13	0.11
Acidaminococcus	-0.12	0.12
Prevotella (7)	0.16	0.13
Victivallis	0.07	0.14
Romboutsia	0.16	0.16
Anaerofustis	0.07	0.16
Allisonella	0.09	0.17
Ligilactobacillus	-0.11	0.18
Odoribacter	0.13	0.19
X Eubacterium ruminantium group	-0.15	0.23
Lachnospira	0.13	0.23
Alloprevotella	-0.12	0.23
Defluviitaleaceae UCG 011	0.10	0.24

* β = coefficient; confounders included in the multivariate model were: age (years), gender (woman, man), ethnical background (Dutch, non-Dutch) and monthly income (< €971, €971-1,355, €1,356-1,969, €1,970-3,314, €3,315-3,500, > €3,500, missing).
** FDR p-value = false discovery rate p-value, subjected to multiple hypothesis testing correction using the Benjamini-Hochberg method with an FDR threshold of 0.25.

Discussion

To our knowledge, this study is the first comprehensive study combining gut microbiome, highly detailed land use and relevant epidemiological data to explore the potential associations of agricultural land coverage on the diversity and composition of human gut microbial communities in Dutch adults aged 40 years or older. Data were collected within the framework of a large nation-wide cross-sectional population-based study (PIENTER-III, n = 1,852 participants).

We observed a significant association between agricultural land coverage and several aspects of the adult human gut microbiome such as bacterial diversity, richness and composition. First, the bacterial diversity and richness significantly increased with increasing agricultural land coverage. These associations might indicate a positive relation between land use and health status, because, in general, there seems to be consensus in scientific literature that more diverse gut microbiotas are associated with a higher health status.[6,52-54] The underlying reasoning is that with increased diversity there is more support for a diverse array of beneficial functions of the gut microbiota regarding absorption and production of essential nutrients, and regulating our immune, metabolic and nervous systems.

Second, the overall microbial composition significantly changed with an increase in agricultural land coverage. Third, we identified two microbial clusters (one dominated by the genus *Blautia* and one dominated by *Bifidobacterium, Collinsella* and *Akkermansia*) to be significantly positively associated with higher agricultural land coverage. The species within the genus *Blautia* are known to produce short-chain fatty acids, which have anti-inflammatory, antiproliferative, and antineoplastic properties and are implicated in protection against colorectal cancer.[55,56] The microbial cluster dominated by *Prevotella (9)* was negatively associated with the agricultural land coverage. This is in contrast to a recently published study on environmental factors shaping the human gut microbiome in the Netherlands, which showed an increase in *Prevotella* (*P.*) *cropi* in more rural areas.[6] Differences between the two studies could be explained by factors like level of detail of both microbiome and land use data, as well as focus on different age groups and regions of the Netherlands.[6] Despite these differences, the overall message that rurality of an area can shape the human gut microbiome was comparable.

Fourth and last, we identified two genera, *Barnesiella* and *Leuconostoc*, with a significant higher relative abundance in participants living in areas with a high agricultural land coverage. *Barnesiella* is a mucus specialist like *Akkermansia muciniphila*[57] and its abundance is negatively associated with activity levels of colitis in mice[58]. It could therefore be a key protective intestinal bacterium by removing

harmful bacteria from the intestines.[59] *Barnesiella* possibly protects the intestinal tract from pathogen infections and likely plays a key role in immunomodulation. Some species from the second genus, *Leuconostoc*, are producers of dextran (LM742 - *Leuconostoc mesenteroides* SPCL742), which has prebiotic potential.[60] Furthermore, *Leuconostoc pseudomesenteroides* improves microbiota dysbiosis and liver metabolism imbalance.[61] However, short read lengths did not allow us identify specific species. Further investigation with species specific primer will be required to investigate their presence in the population we studied here.

A limited number of other studies into the association between the living environment and the diversity and composition of the human gut microbiome are conducted. Although the comparison of studies is difficult because of differences in geographical area, demographics, and methodological aspects, the above mentioned Dutch study on environmental factors shaping the gut microbiome in the Dutch population[6] as well as a Danish study that focused on infants[29] also concluded that different microbiome signatures exist along the urban-rural gradient. Additionally, the Danish study showed that urbanization-related changes in the infant microbiota may elevate the risk of asthma and atopic traits, probably via cross talk with the developing immune system[29]. The focus on children is a result of previous findings that microbial exposure in early-life is associated with a protective effect regarding human health in later life.[9] However, this protective effect likely only holds with continued microbial exposure to keep the most optimal protection from developing disease.[9] Furthermore, the protective effect in adults can also be a result of high occupational endotoxin exposure levels later in life without early-life exposure.[62] Therefore, major strengths of this current study are the large number of PIENTER-3 participants from whom fecal samples were collected across all ages (n_{all_ages} = 3,746, $n_{40\_years\_and\_older}$ = 1,852) making it possible to focus on adults specifically, in combination with the detailed land use database (resolution = 25x25 meters).[31,36]

Because the PIENTER-3 cohort was not specifically designed for microbiome research, the ecological design of this study makes it prone to ecological fallacy. No data was available on home addresses of PIENTER-3 participants, their history of movement or longitudinal microbiome data. For those reasons, we cannot infer causality, which would require follow-up studies using longitudinal individual-level data combined with high resolution spatial data (e.g. animal numbers per farm). To study whether there are other unobserved spatial risk factors affecting our results leading to neighboring postal code areas being more alike, such as those related to antibiotic usage[63] or the presence of farm animals in the neighborhood, we recommend creating a map of the spatially structured variation (CAR) as a first step towards further unravelling the effects of different spatially restricted variables, after which specific variables can be studied in more detail.[64]

Using highly detailed land use data, this study gave insights into the associations of the gut microbial diversity and composition of adults with the agricultural land coverage in their neighborhood.

Conclusions & Recommendations

This study showed that bacterial diversity, richness and composition of the adult gut microbiome are associated with the urban-rural gradient in the Netherlands. As this study was based on an ecological design, causal inference is limited and further research using individual-level data is recommended. Another step forward would be to include additional information on farm animals to get a more complete overview of the effects rural areas have on the human gut microbiota and to perform sensitivity analyses. Additionally, we would recommend to study this population with metagenomics in the near future to be able to look for function differences or to set up an *in vitro*, *in vivo* investigation to study the impact of specific bacteria on human health and immune function.

Acknowledgements

The authors would like to thank everyone who actively shared their thoughts and questions relevant for this study. We also would like to thank the TRIuMPH project team members and all colleagues involved in PIENTER-3.

Literature

1 Ege, M. J. et al. Exposure to Environmental Microorganisms and Childhood Asthma. New England Journal of Medicine 364, 701-709, doi:10.1056/NEJMoa1007302 (2011).

2 Gautam, A. et al. Altered fecal microbiota composition in all male aggressor-exposed rodent model simulating features of post-traumatic stress disorder. Journal of Neuroscience Research 96, 1311-1323, doi:https://doi.org/10.1002/jnr.24229 (2018).

3 Molina-Torres, G., Rodriguez-Arrastia, M., Roman, P., Sanchez-Labraca, N. & Cardona, D. Stress and the gut microbiota-brain axis. Behav Pharmacol 30, 187-200, doi:10.1097/fbp.0000000000000478 (2019).

4 Oswald, T. K., Rumbold, A. R., Kedzior, S. G. E., Kohler, M. & Moore, V. M. Mental Health of Young Australians during the COVID-19 Pandemic: Exploring the Roles of Employment Precarity, Screen Time, and Contact with Nature. International Journal of Environmental Research and Public Health 18, 5630 (2021).

5 Riedler, J. et al. Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey. The Lancet 358, 1129-1133, doi:10.1016/S0140-6736(01)06252-3 (2001).

6 Gacesa, R. et al. Environmental factors shaping the gut microbiome in a Dutch population. Nature 604, 732-739, doi:10.1038/s41586-022-04567-7 (2022).

7 Spector, T., Wolf, J. & Hadjigeorgiou, G. Rethinking the future of nutrition - Using large-scale biological data, AI and novel digital technologies to improve human health [White paper]. Chapter 5 - The gut microbiome plays an important role in our health. 25-30 (ZOE, 2020).

8 Stege, P. B. et al. Impact of long-term dietary habits on the human gut resistome in the Dutch population. Sci Rep 12, 1892, doi:10.1038/s41598-022-05817-4 (2022).

9 Heederik, D. & von Mutius, E. Does diversity of environmental microbial exposure matter for the occurrence of allergy and asthma? J Allergy Clin Immunol 130, 44-50, doi:10.1016/j.jaci.2012.01.067 (2012).

10 Hirsch et al. Prevalence and determinants of house dust mite allergen in East German homes. Clinical & Experimental Allergy 28, 956-964, doi:10.1046/j.1365-2222.1998.00340.x (1998).

11 Hugg, T., Ruotsalainen, R., Jaakkola, M. S., Pushkarev, V. & Jaakkola, J. J. Comparison of allergic diseases, symptoms and respiratory infections between Finnish and Russian school children. Eur J Epidemiol 23, 123-133, doi:10.1007/s10654-007-9217-z (2008).

12 Hugg, T. T., Jaakkola, M. S., Ruotsalainen, R., Pushkarev, V. & Jaakkola, J. J. Exposure to animals and the risk of allergic asthma: a population-based cross-sectional study in Finnish and Russian children. Environ Health 7, 28, doi:10.1186/1476-069x-7-28 (2008).

13 Pakarinen, J. et al. Predominance of Gram-positive bacteria in house dust in the low-allergy risk Russian Karelia. Environ Microbiol 10, 3317-3325, doi:10.1111/j.1462-2920.2008.01723.x (2008).

14 Braun-Fahrländer, C. et al. Environmental exposure to endotoxin and its relation to asthma in school-age children. N Engl J Med 347, 869-877, doi:10.1056/NEJMoa020057 (2002).

15 de Rooij, M. M. T. et al. Insights into Livestock-Related Microbial Concentrations in Air at Residential Level in a Livestock Dense Area. Environmental Science & Technology 53, 7746 - 7758, doi:10.1021/acs.est.8b07029 (2019).

16 Douwes, J. et al. Does early indoor microbial exposure reduce the risk of asthma? The Prevention and Incidence of Asthma and Mite Allergy birth cohort study. J Allergy Clin Immunol 117, 1067-1073, doi:10.1016/j.jaci.2006.02.002 (2006).

17 Gehring, U. et al. Bacteria and mould components in house dust and children's allergic sensitisation. Eur Respir J 29, 1144-1153, doi:10.1183/09031936.00118806 (2007).

18 Gehring, U. et al. Asthma and allergic symptoms in relation to house dust endotoxin: Phase Two of the International Study on Asthma and Allergies in Childhood (ISAAC II). Clin Exp Allergy 38, 1911-1920, doi:10.1111/j.1365-2222.2008.03087.x (2008).

19 Gereda, J., Leung, D. Y. M. & Liu, A. H. Levels of Environmental Endotoxin and Prevalence of Atopic Disease. JAMA : the journal of the American Medical Association 284, 1652-1653, doi:10.1001/jama.284.13.1652 (2000).

20 Gereda, J. E. et al. Relation between house-dust endotoxin exposure, type 1 T-cell development, and allergen sensitisation in infants at high risk of asthma. The Lancet 355, 1680-1683, doi:10.1016/S0140-6736(00)02239-X (2000).

21 Sordillo, J. E. et al. Multiple microbial exposures in the home may protect against asthma or allergy in childhood. Clin Exp Allergy 40, 902-910, doi:10.1111/j.1365-2222.2010.03509.x (2010).

22 Tischer, C. et al. Respiratory health in children, and indoor exposure to (1,3)-β-D-glucan, EPS mould components and endotoxin. Eur Respir J 37, 1050-1059, doi:10.1183/09031936.00091210 (2011).

23 van Strien, R. T. et al. Microbial exposure of rural school children, as assessed by levels of N-acetyl-muramic acid in mattress dust, and its association with respiratory health. Journal of Allergy and Clinical Immunology 113, 860-867, doi:10.1016/j.jaci.2004.01.783 (2004).

24 Korthals, M., Ege, M. J., Tebbe, C. C., von Mutius, E. & Bauer, J. Application of PCR-SSCP for molecular epidemiological studies on the exposure of farm children to bacteria in environmental dust. J Microbiol Methods 73, 49-56, doi:10.1016/j.mimet.2008.01.010 (2008).

25 Normand, A. C., Vacheyrou, M., Sudre, B., Heederik, D. J. & Piarroux, R. Assessment of dust sampling methods for the study of cultivable-microorganism exposure in stables. Appl Environ Microbiol 75, 7617-7623, doi:10.1128/aem.01414-09 (2009).

26 Roussel, S. et al. Exposure to moulds and actinomycetes in Alpine farms: A nested environmental study of the PASTURE cohort. Environmental Research 111, 744-750, doi:10.1016/j.envres.2011.05.002 (2011).

27 Borlée, F. et al. Residential proximity to livestock farms is associated with a lower prevalence of atopy. Occupational and Environmental Medicine 75, 453, doi:10.1136/oemed-2017-104769 (2018).

28 Müller-Rompa, S. E. K. et al. An approach to the asthma-protective farm effect by geocoding: Good farms and better farms. Pediatr Allergy Immunol 29, 275-282, doi:10.1111/pai.12861 (2018).

29 Lehtimäki, J. et al. Urbanized microbiota in infants, immune constitution, and later risk of atopic diseases. J Allergy Clin Immunol 148, 234-243, doi:10.1016/j.jaci.2020.12.621 (2021).

30 Caporaso, J. G. et al. Moving pictures of the human microbiome. Genome Biol 12, R50, doi:10.1186/gb-2011-12-5-r50 (2011).

31 Verberk, J. D. M. et al. Third national biobank for population-based seroprevalence studies in the Netherlands, including the Caribbean Netherlands. BMC Infectious Diseases 19, 470, doi:10.1186/s12879-019-4019-y (2019).

32 Muyzer, G., Waal, E. C. d. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology 59, 695-700, doi:10.1128/aem.59.3.695-700.1993 (1993).

33 Caporaso, J. G. et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108, 4516-4522, doi:10.1073/pnas.1000080107 (2011).

34 Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13, 581-583, doi:10.1038/nmeth.3869 (2016).

35 Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41, D590-D596, doi:10.1093/nar/gks1219 (2012).

36 LGN7 (2012). Alterra. <https://www.wur.nl/en/Research-Results/Research-Institutes/Environmental-Research/Facilities-Tools/Kaarten-en-GIS-bestanden/Land-use-database-of-the-Netherlands/Different-Versions-of-Databases/LGN7.html>.

37 Pebesma, E. J. Simple features for R: standardized support for spatial vector data. The R journal 10, 439-446 (2018).

38 Hijmans, R. J. raster: Geographic Data Analysis and Modeling. (2022).

39 Lahti, L. & Shetty, S. microbiome R package. (2012-2019).

40 Shetty, S. biomeUtils: Utilities for microbiome analytics. (2021).

41 McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8, e61217, doi:10.1371/journal.pone.0061217 (2013).

42 Cao, Q. et al. Effects of Rare Microbiome Taxa Filtering on Statistical Analysis. Frontiers in Microbiology 11, doi:10.3389/fmicb.2020.607325 (2021).

43 Jervis-Bardy, J. et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. Microbiome 3, 19, doi:10.1186/s40168-015-0083-8 (2015).

44 R Core Team. R: A language and Environment for Statistical Computing. (2022).

45 Chen, S.-Y., Feng, Z. & Yi, X. A general introduction to adjustment for multiple comparisons. Journal of thoracic disease 9, 1725, doi:10.21037/jtd.2017.05.34 (2017).

46 Bray, J. R. & Curtis, J. T. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs 27, 325-349, doi:10.2307/1942268 (1957).

47 Oksanen, J. et al. vegan: community ecology package. (2019).

48 Anderson, M. J. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics 62, 245-253, doi:10.1111/j.1541-0420.2005.00440.x (2006).

49 Desgraupes, B. Clustering indices. University of Paris Ouest-Lab Modal'X 1, 34 (2013).

50 Thissen, D., Steinberg, L. & Kuang, D. Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False Positive Rate in Multiple Comparisons. Journal of Educational and Behavioral Statistics 27, 77-83, doi:10.3102/10769986027001077 (2002).

51 Mallick, H. et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol 17, e1009442, doi:10.1371/journal.pcbi.1009442 (2021).

52 Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. Nature 489, 220-230, doi:10.1038/nature11550 (2012).

53 Manor, O. et al. Health and disease markers correlate with gut microbiome composition across thousands of people. Nat Commun 11, 5206, doi:10.1038/s41467-020-18871-1 (2020).

54 Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in disease. Curr Opin Gastroenterol 31, 69-75, doi:10.1097/MOG.0000000000000139 (2015).

55 Ríos-Covián, D. et al. Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health. Front Microbiol 7, 185, doi:10.3389/fmicb.2016.00185 (2016).

56 Tanaka, S., Yamamoto, K., Yamada, K., Furuya, K. & Uyeno, Y. Relationship of Enhanced Butyrate Production by Colonic Butyrate-Producing Bacteria to Immunomodulatory Effects in Normal Mice Fed an Insoluble Fraction of Brassica rapa L. Appl Environ Microbiol 82, 2693-2699, doi:10.1128/aem.03343-15 (2016).

57 Schroeder, B. O. Fight them or feed them: how the intestinal mucus layer manages the gut microbiota. Gastroenterol Rep (Oxf) 7, 3-12, doi:10.1093/gastro/goy052 (2019).

58 Ye, J. et al. Bacteria and bacterial rRNA genes associated with the development of colitis in IL-10(-/-) mice. Inflamm Bowel Dis 14, 1041-1050, doi:10.1002/ibd.20442 (2008).

59 Ubeda, C. et al. Intestinal microbiota containing Barnesiella species cures vancomycin-resistant Enterococcus faecium colonization. Infect Immun 81, 965-973, doi:10.1128/iai.01197-12 (2013).

60 Kim, G. et al. Prebiotic activities of dextran from Leuconostoc mesenteroides SPCL742 analyzed in the aspect of the human gut microbial ecosystem. Food & Function 13, 1256-1267, doi:10.1039/D1FO03287A (2022).

61 Sun, M. et al. Leuconostoc pseudomesenteroides improves microbiota dysbiosis and liver metabolism imbalance and ameliorates the correlation between dihydroceramide and strains of Firmicutes and Proteobacteria in high fat diet obese mice. Food Funct 11, 6855-6865, doi:10.1039/d0fo01009j (2020).

62 Smit, L. A. M., Heederik, D., Doekes, G., Lammers, J. W. J. & Wouters, I. M. Occupational Endotoxin Exposure Reduces the Risk of Atopic Sensitization but Increases the Risk of Bronchial Hyperresponsiveness. International Archives of Allergy and Immunology 152, 151-158, doi:10.1159/000265536 (2010).

63 Roof, I., van der Hoek, W., Oude Boerrigter, L., Wielders, C. C. H. & Smit, L. A. M. Use of Antibiotics among Residents Living Close to Poultry or Goat Farms: A Nationwide Analysis in The Netherlands. Antibiotics 10, 1346 (2021).

64 Mulder, A. C. et al. Spatial Effects of Livestock Farming on Human Infections With Shiga Toxin-Producing Escherichia coli O157 in Small but Densely Populated Regions: The Case of the Netherlands. GeoHealth 4, e2020GH000276, doi:https://doi.org/10.1029/2020GH000276 (2020).

Supplementary material

Control samples

In total, six types of control samples were used within this study to validate the participant samples: three DNA extraction controls and three PCR controls. The first DNA extraction control was a bacterial mixture of the true samples (MSB). It is a positive control which was constructed by using five randomly picked healthy participant stool samples, equally pooled based on mass. The MSB was stored at -80°C after being uniformly aliquoted to be ready for their usage on several random extraction days (n=23). The first PCR control contained DNA of the true sample mixture. This control was used twice in every sequence library (n = 54). It was provided by the DNA extraction from the aliquots obtained from the MSB.

The second DNA extraction control was a bacterial mixture of the ZYMO Mock (ZMB) and the second PCR control was DNA of the ZYMO mock (ZMD). Those are positive controls in the form of a microbial community standard (ZymoBIOMICS, Zymo Research, Irvine, CA, USA) for both the ZymoBIOMICS Microbial Community Standard in DNA extractions (ZMB) and DNA Standard (ZMD). Those controls were added for library preparation.

Third, a blank (B) was added as control for both DNA extraction (BB) and PCR (BD). The specific buffer linked to the collection method in absence of any stool in all DNA extractions (BB) as a negative control while for the library preparation DNAse free water was used (BD).

Other results

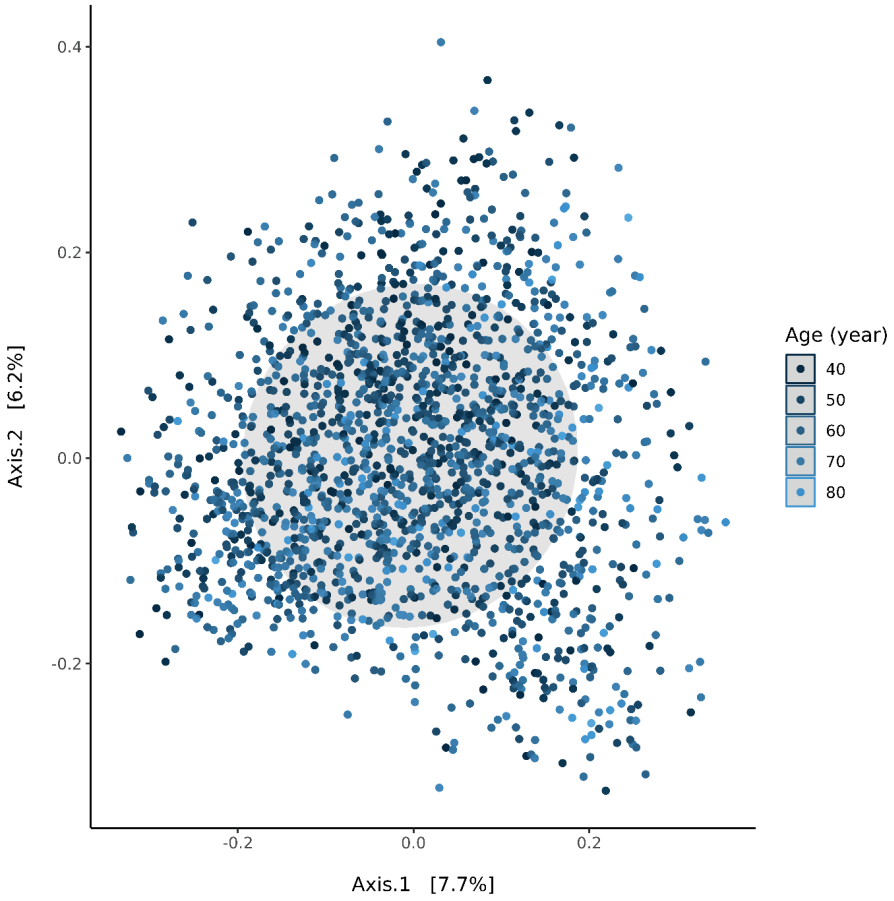


Figure S1. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by age (in years).

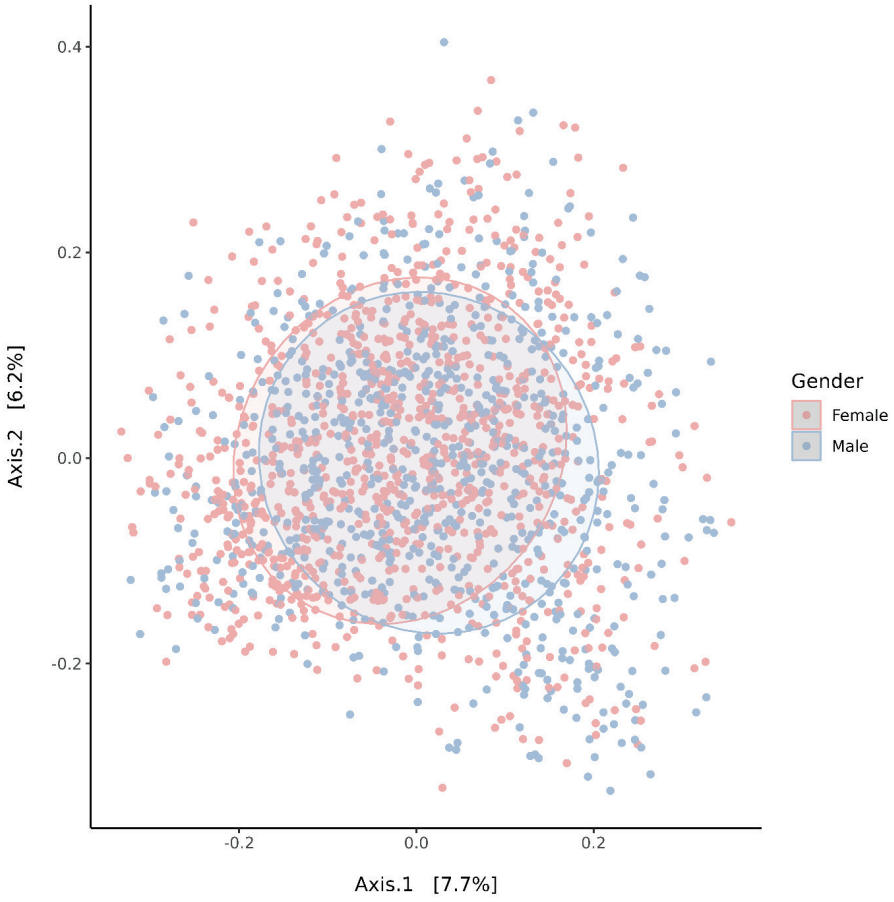


Figure S2. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by gender (female, male).

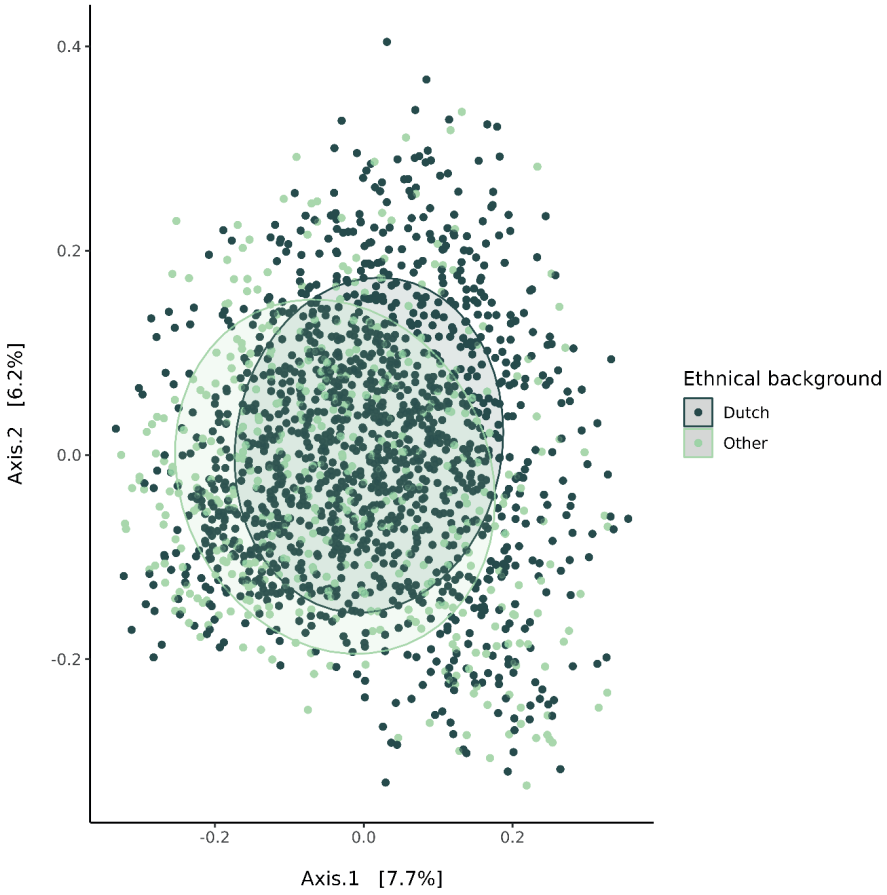


Figure S3. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by ethnical background (Dutch, Other).

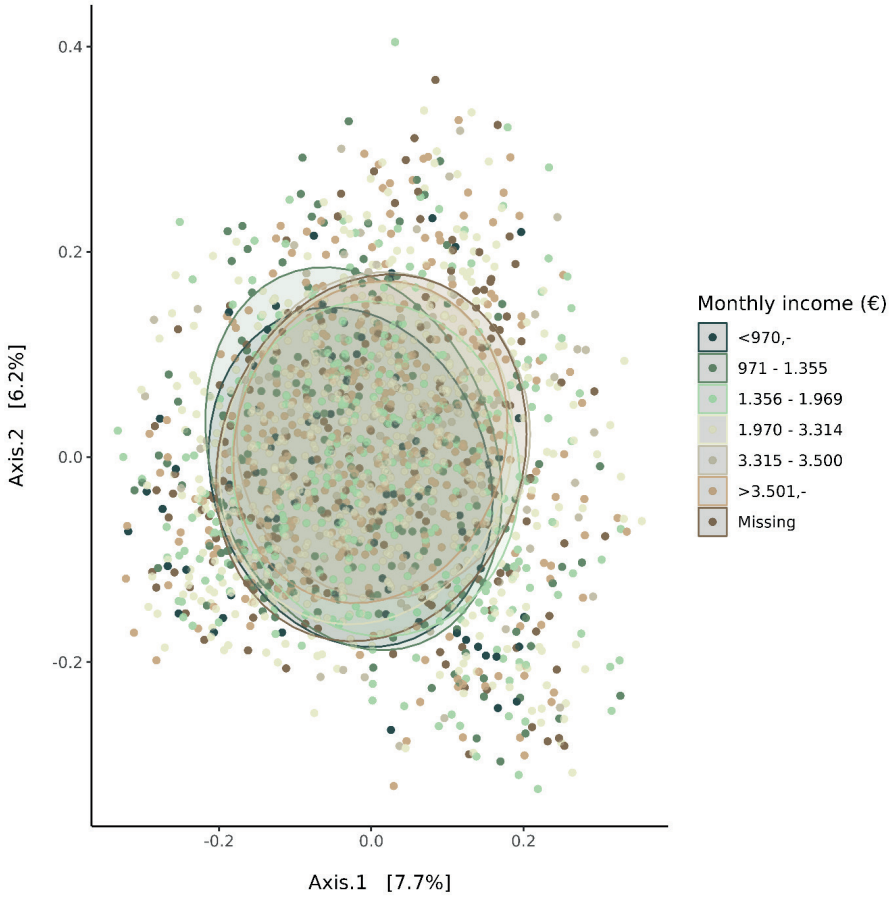


Figure S4. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by monthly income (< €971, €971-1,355, €1,356-1,969, €1,970-3,314, €3,315-3,500, > €3,500, missing).

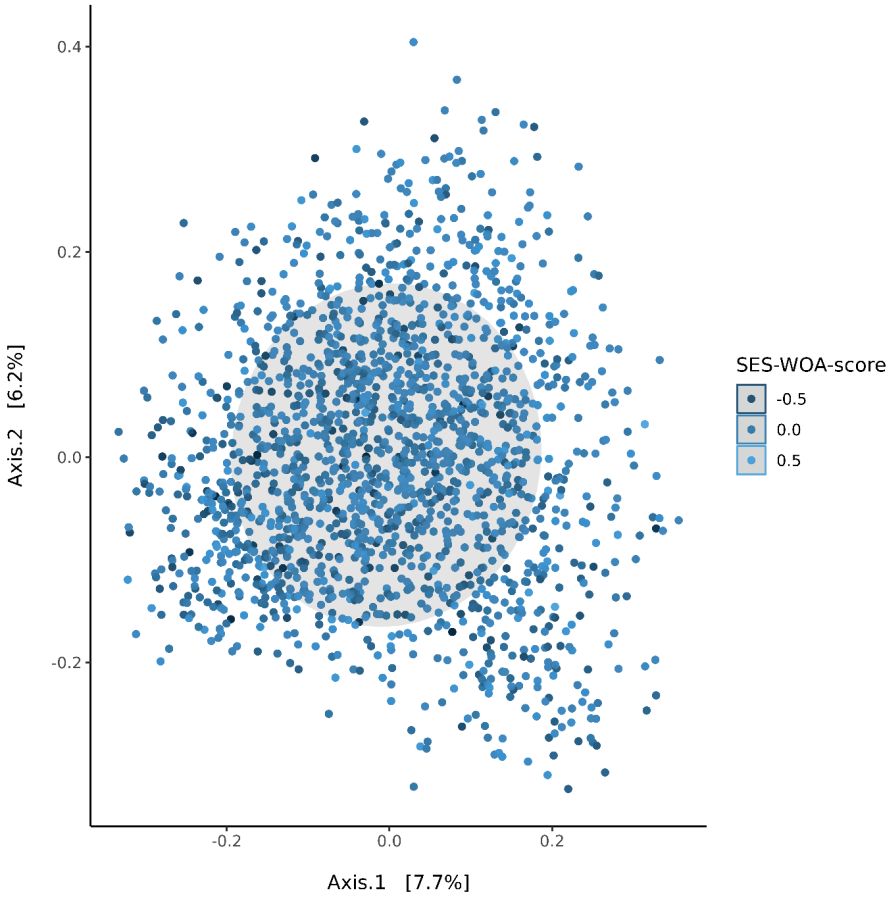


Figure S5. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by SES-WOA score.

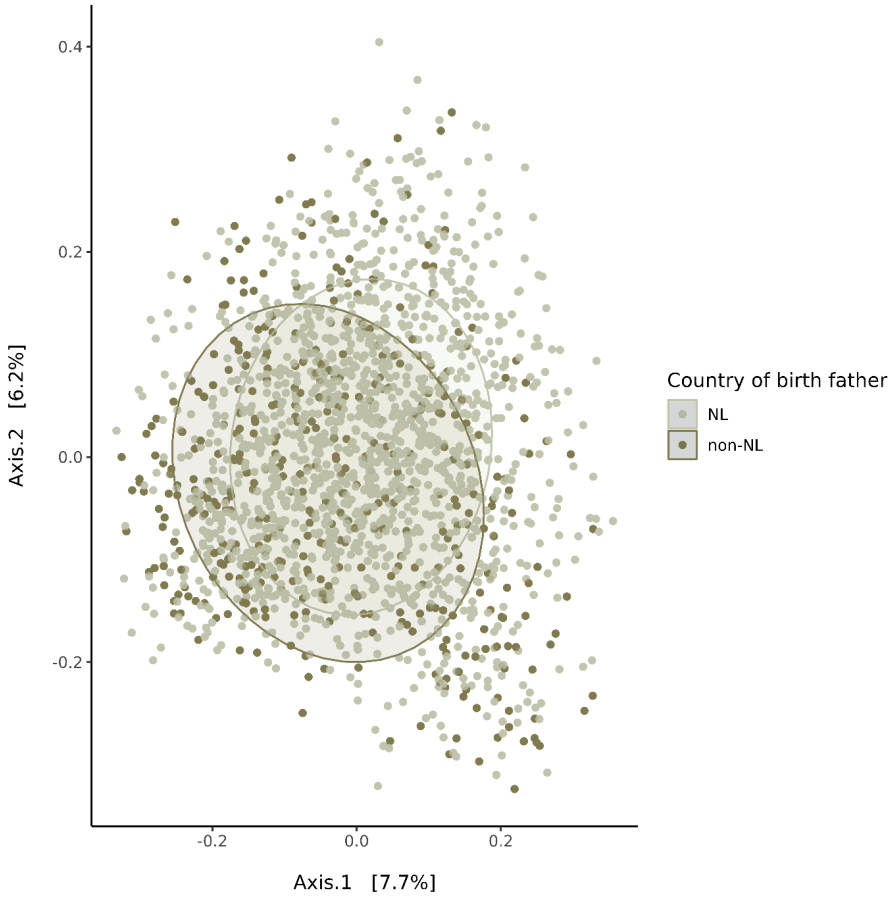


Figure S6. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by country of birth of father (NL, non-NL).

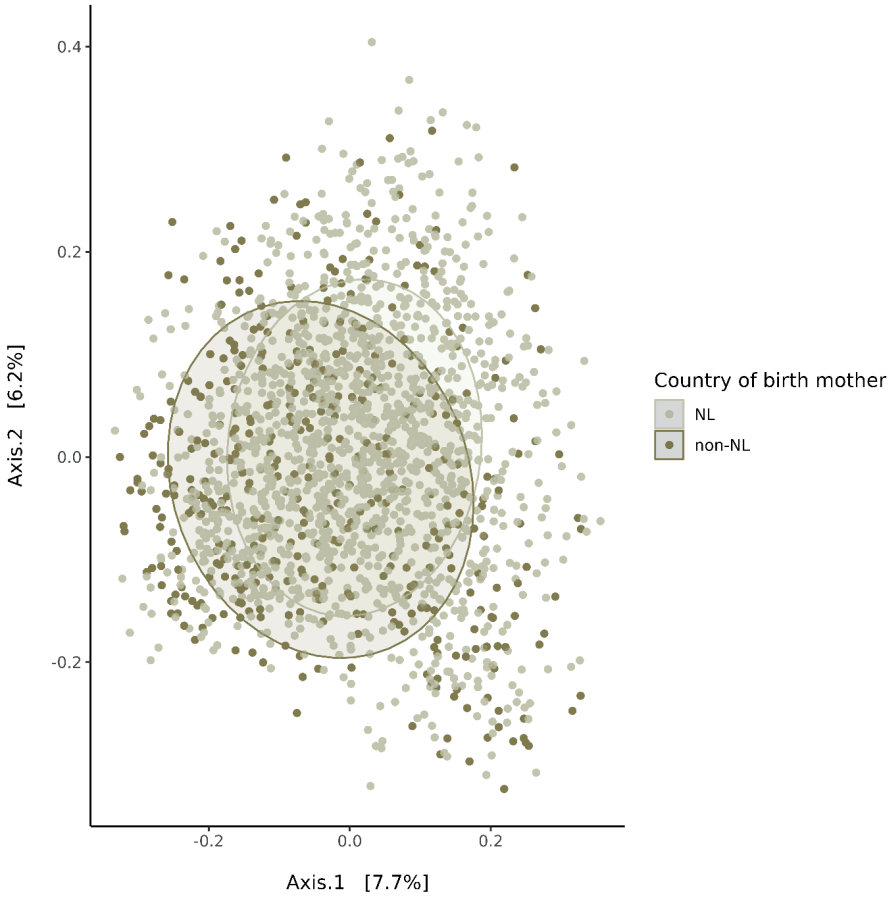


Figure S7. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by country of birth of mother (NL, non-NL).

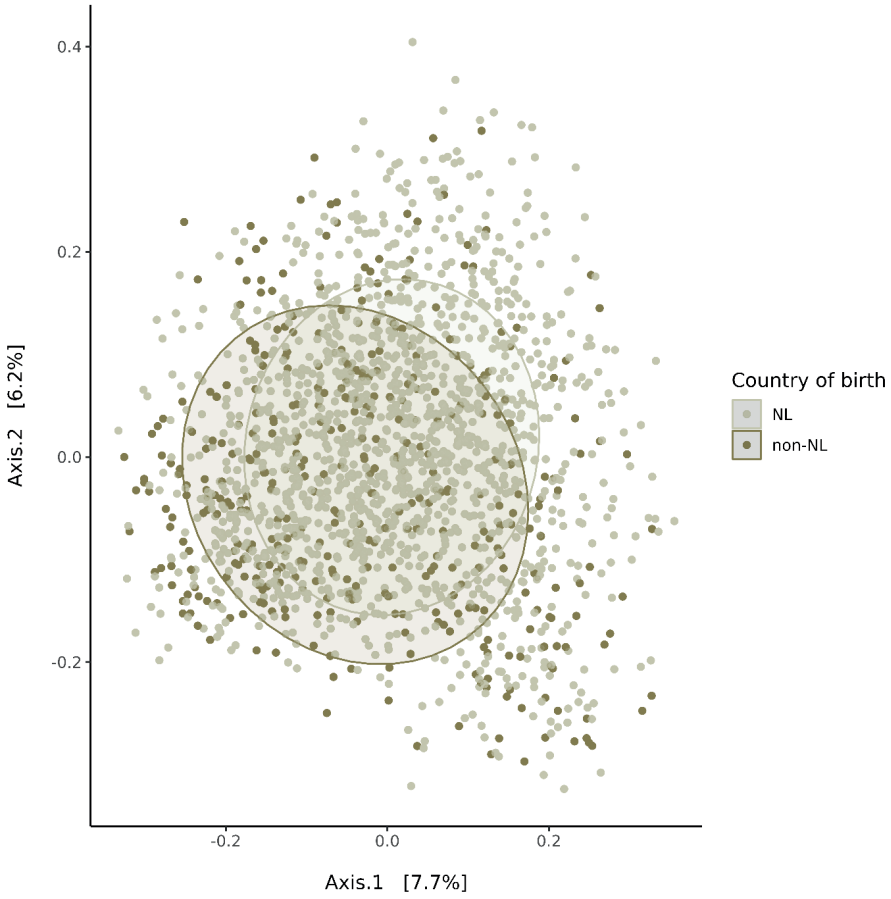


Figure S8. PCoA visualizing the β-diversity based on Bray-Curtis dissimilarities of the study participants, colored by country of birth (NL, non-NL).

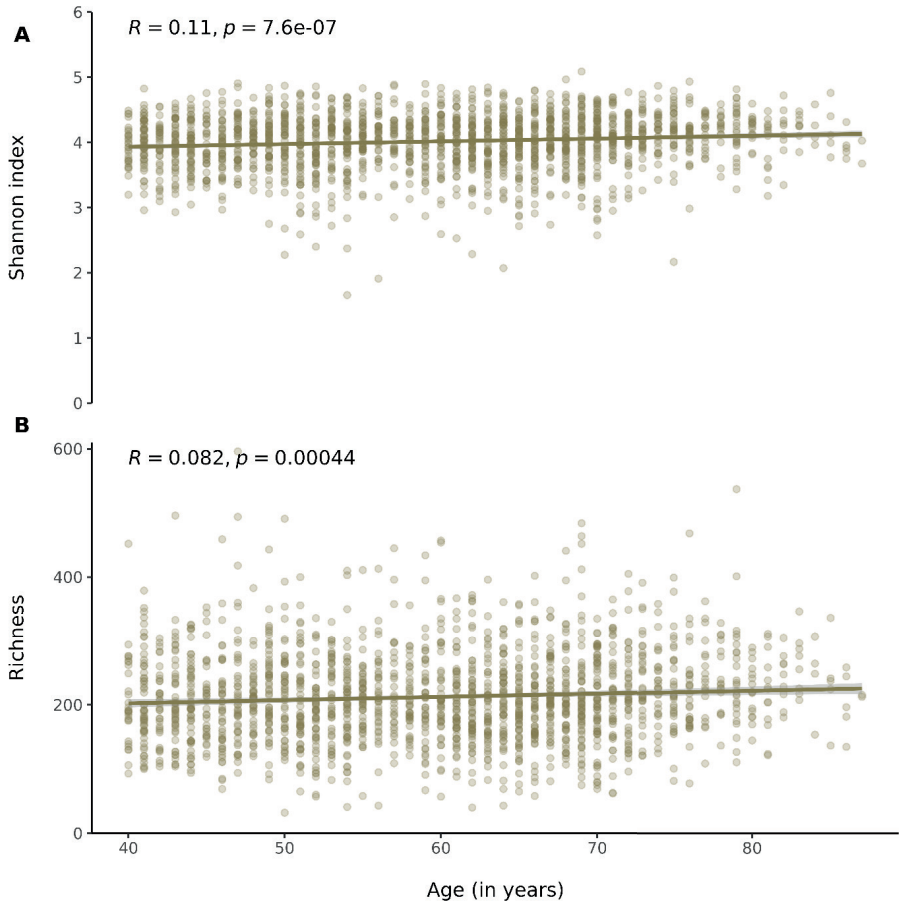


Figure S9. A) Plot of the Shannon index distribution over different ages in years (continuous) **B) P**lot of the observed richness over the different ages in years (continuous). In both plot **A)** and **B)**, the regression line is based on a linear model, points indicate individual participants.

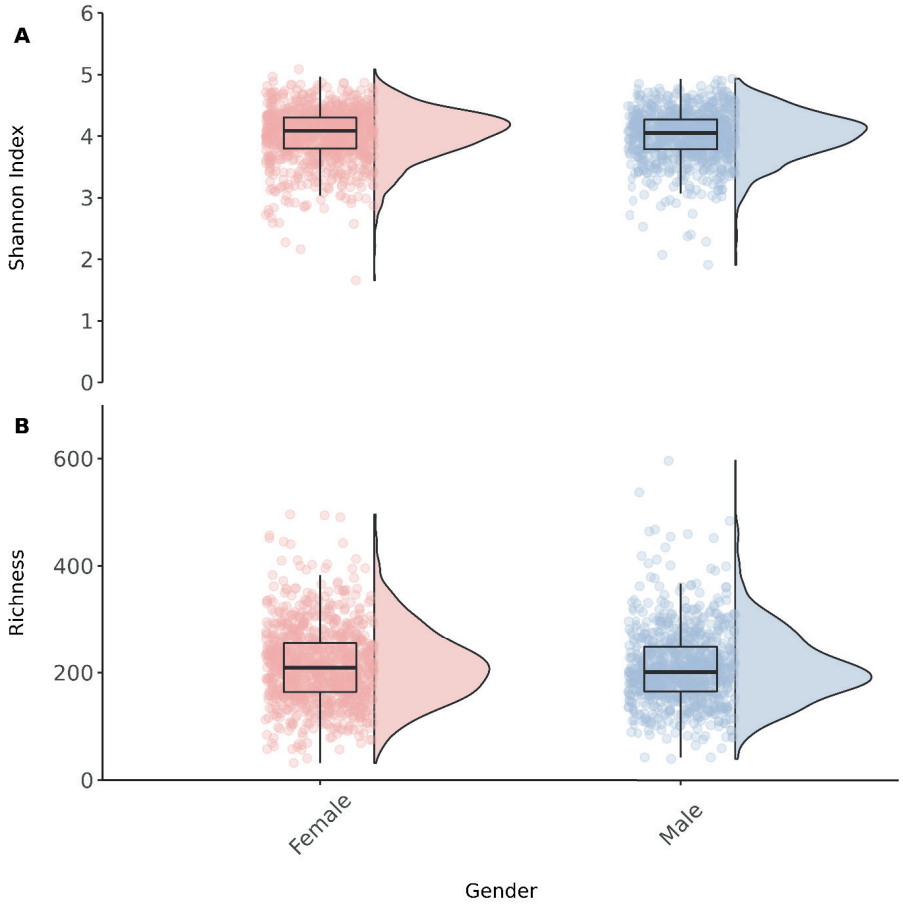


Figure S10. A) Combined bar- and violin plot of the Shannon index distribution over the different gender categories (female, male). **B)** Combined bar- and violin plot of the observed richness over the different gender categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. Differences between groups were not displayed as they were not significant according to the Wilcoxon rank sum test after adjusting the p-value for multiple testing using the Benjamini-Hochberg correction.

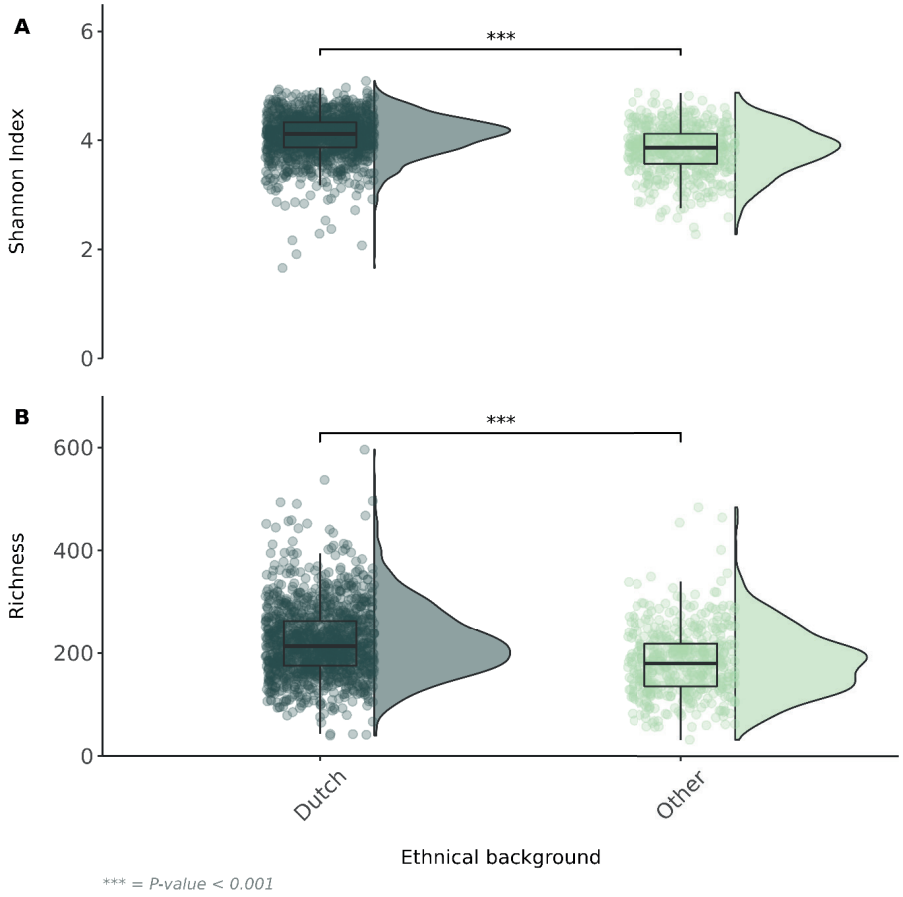


Figure S11. A) Combined bar- and violin plot of the Shannon index distribution over the different ethnical background categories (Dutch, other). **B)** Combined bar- and violin plot of the observed richness over the different ethnical background categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

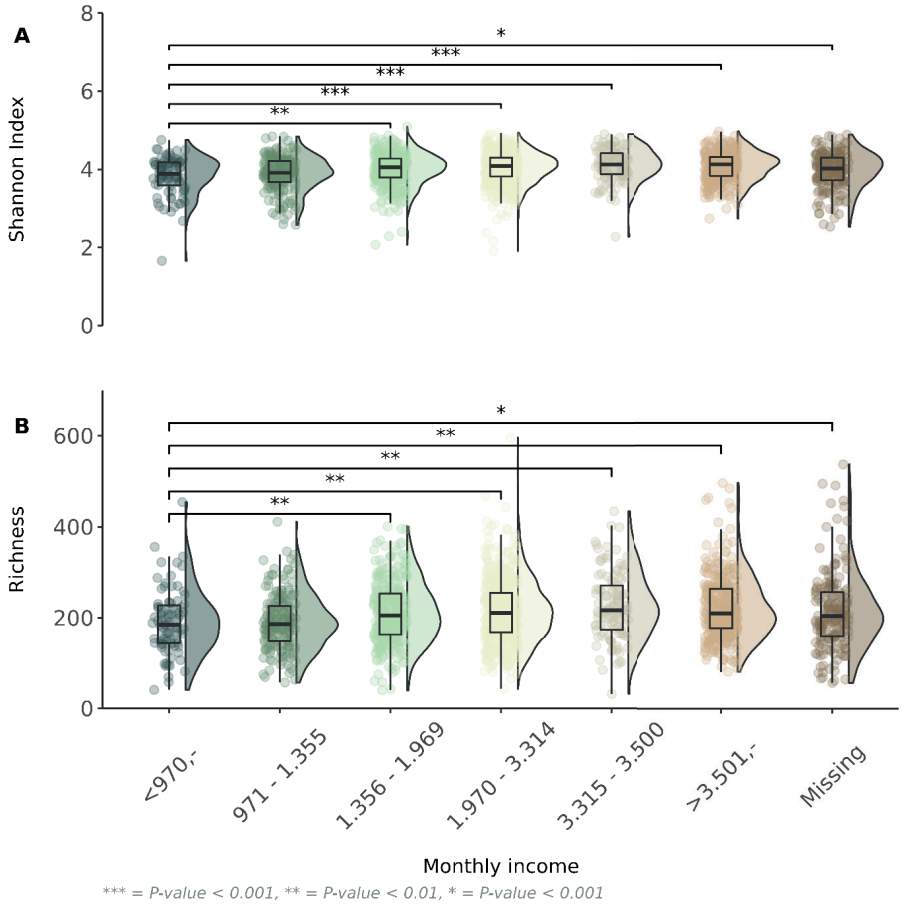


Figure S12. A) Combined bar- and violin plot of the Shannon index distribution over the different monthly income categories (< €971, €971-1,355, €1,356-1,969, €1,970-3,314, €3,315-3,500, > €3,500, missing). **B)** Combined bar- and violin plot of the observed richness over the different monthly income categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 (** = p-value < 0.01, * = p-value < 0.05) when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

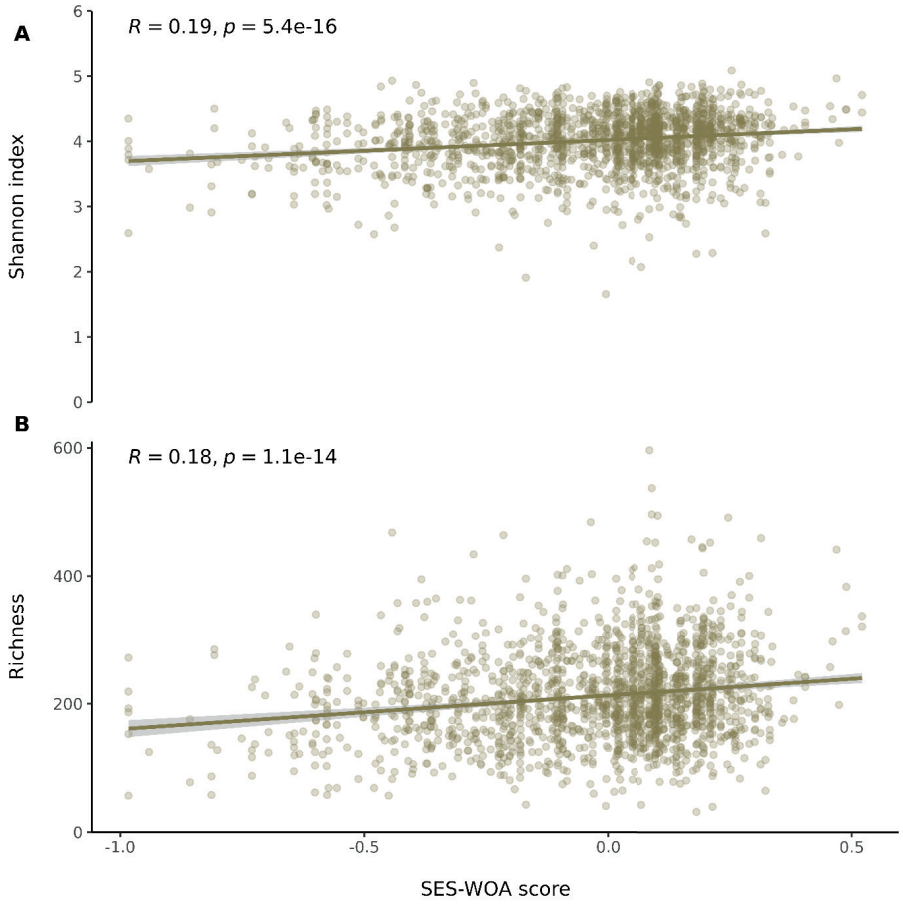


Figure S13. A) Plot of the Shannon index distribution over different SES-WOA scores **B)** Plot of the observed richness over the different SES-WOA scores. In both plot **A)** and **B)**, the regression line is based on a linear model, points indicate individual participants.

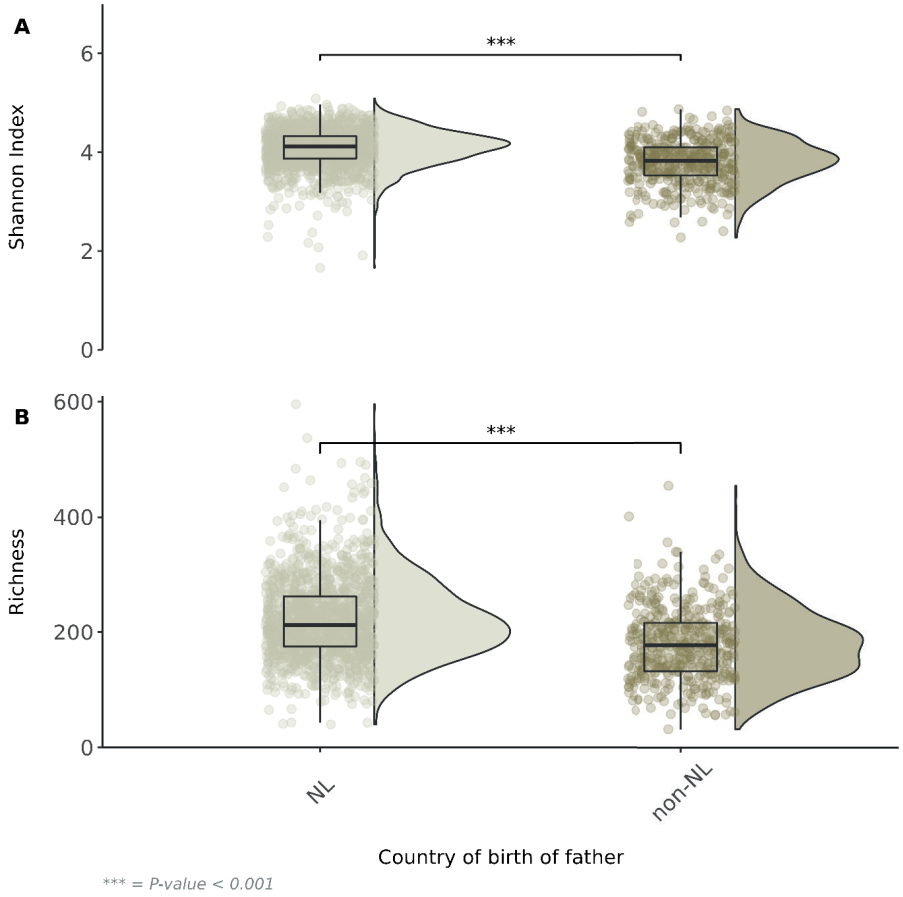


Figure S14. A) Combined bar- and violin plot of the Shannon index distribution over the different country of birth of father categories (NL, non-NL). **B)** Combined bar- and violin plot of the observed richness over the different country of birth of father categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 (** = p-value < 0.01, * = p-value < 0.05) when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

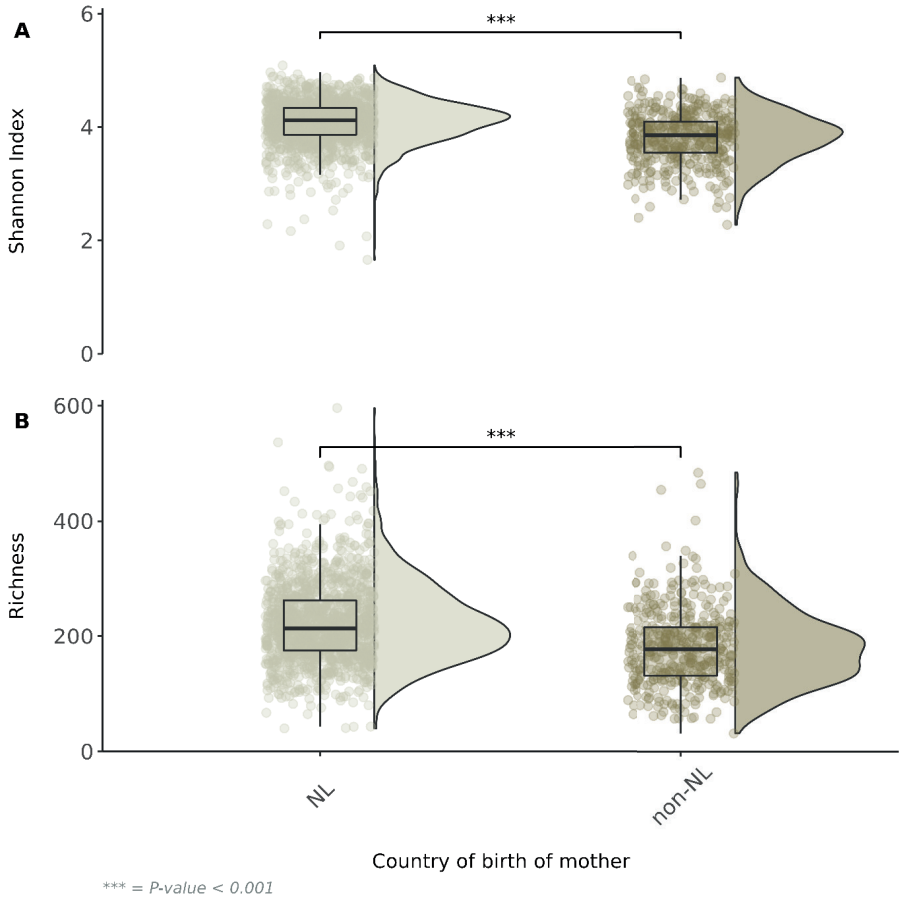


Figure S15. A) Combined bar- and violin plot of the Shannon index distribution over the different country of birth of mother categories (NL, non-NL). **B)** Combined bar- and violin plot of the observed richness over the different country of birth of mother categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 (** = p-value < 0.01, * = p-value < 0.05) when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

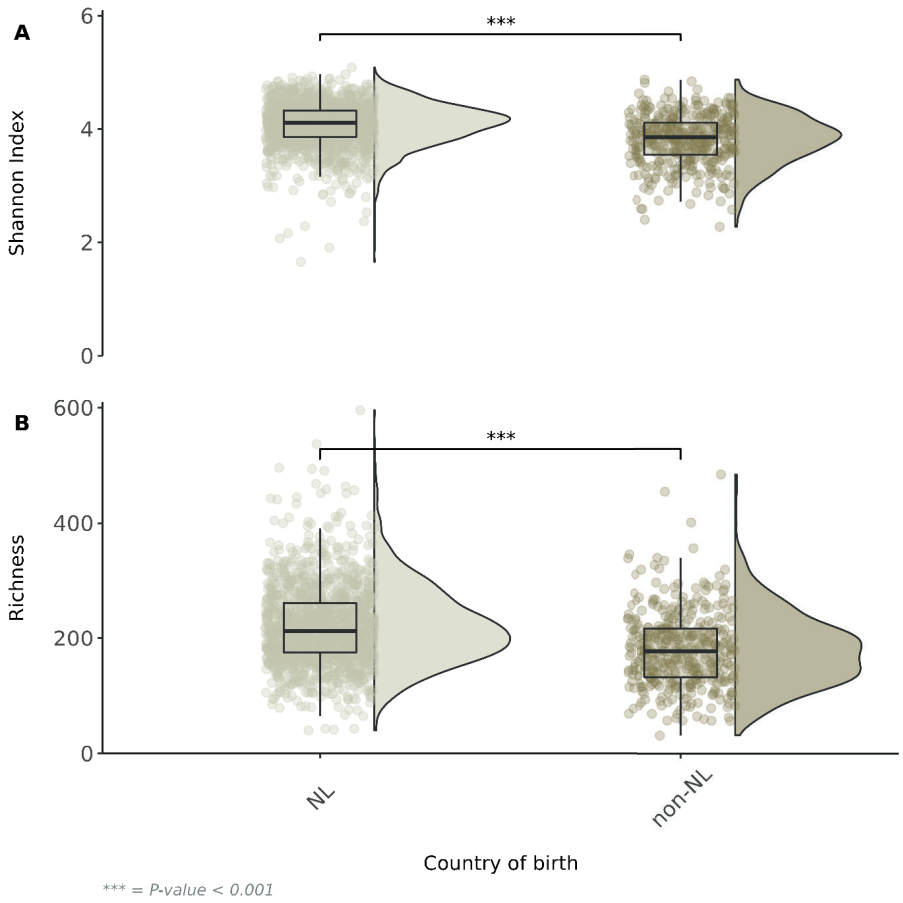


Figure S16. A) Combined bar- and violin plot of the Shannon index distribution over the different country of birth categories (NL, non-NL). **B)** Combined bar- and violin plot of the observed richness over the different country of birth categories. In both plot **A)** and **B)**, the center line indicates the median, box limits are the upper and lower quartiles, whiskers show 1.5x interquartile range, points indicate individual participants and the violin plot outline displays the distribution of the data. *** indicates the adjusted p-value < 0.001 (** = p-value < 0.01, * = p-value < 0.05) when comparing the different groups using the Wilcoxon rank sum test, which are corrected for multiple testing using the Benjamini-Hochberg correction.

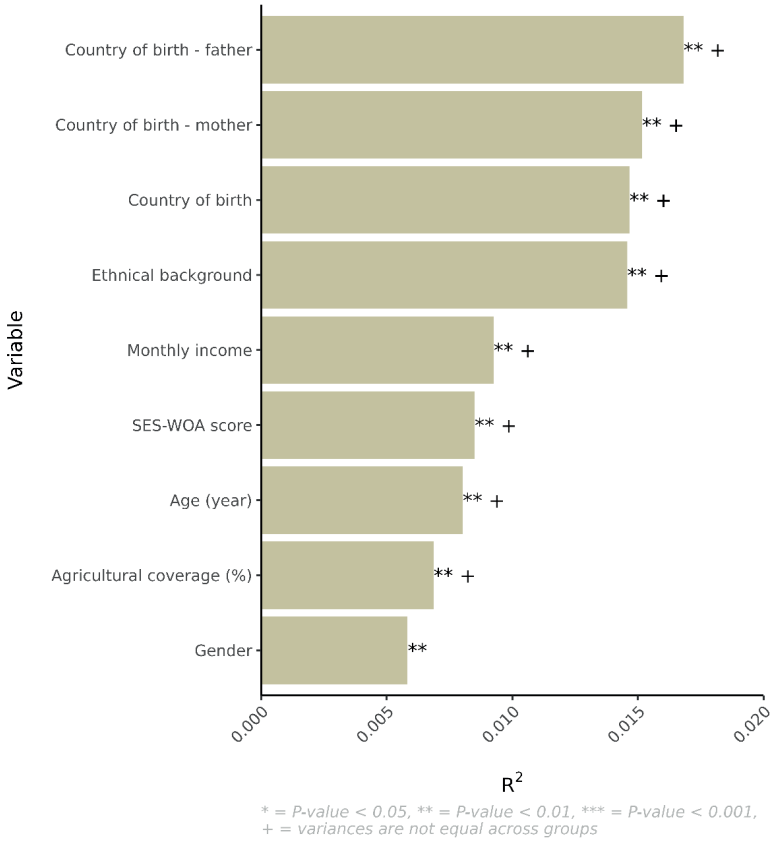
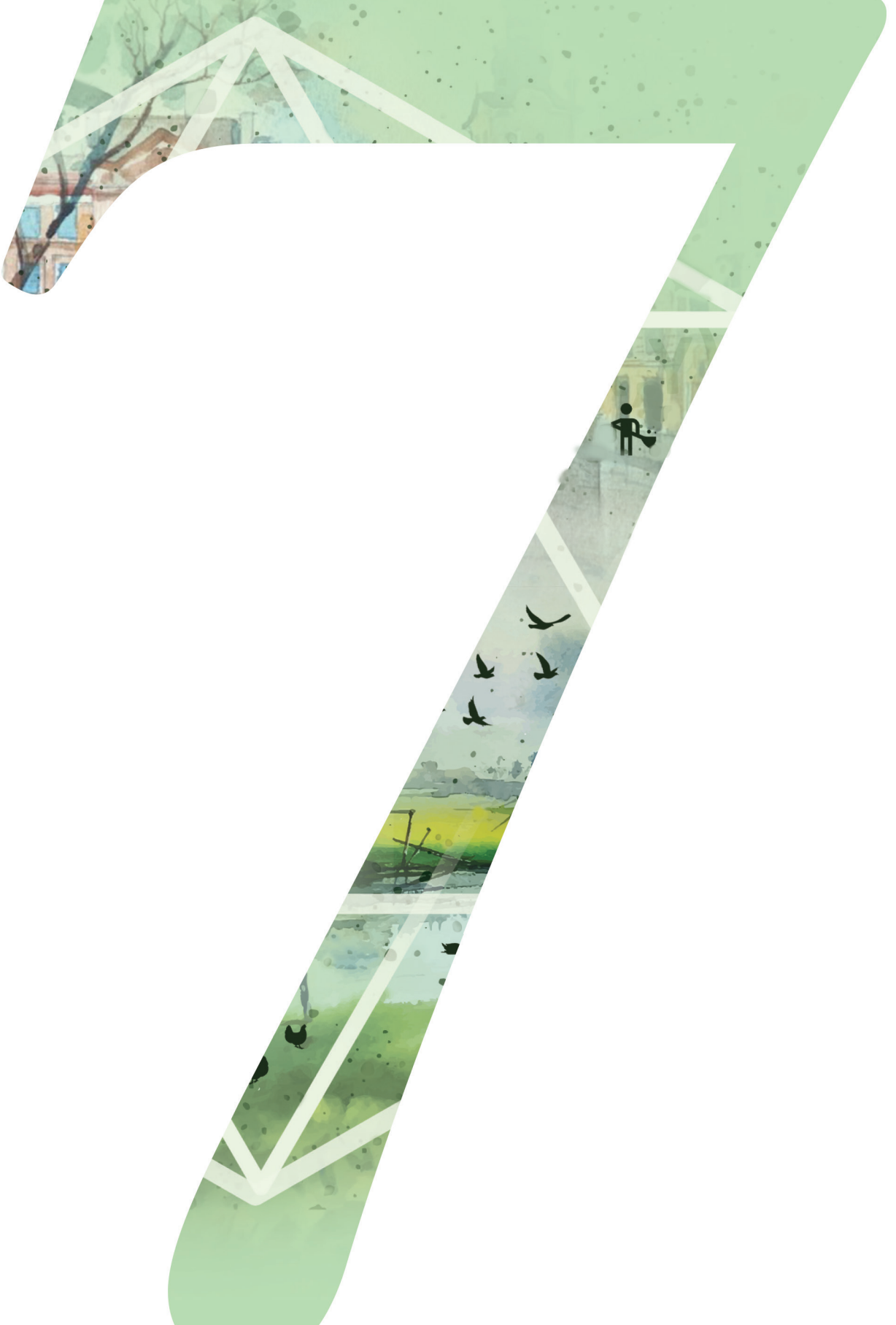


Figure S17. Microbiome variance explained per variable, including its significance level (* = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001) based on the permutational multivariate ANOVA results and whether the condition of homogeneity of variances was met, which was checked using the PERMDISP2 method (+ = variances are not equal across groups).

Table S1. Results of the multivariable analysis studying the associations between specific genera in the gut microbiome of study participants and agricultural land coverage (%) as continuous variable while adjusting for age, gender, ethnical background and SES-WOA score, using the multivariable analysis by linear models (MaAsLin2) statistical framework (coefficient and false discovery rate (FDR) p-value (subjected to multiple hypothesis testing correction using the Benjamini-Hochberg method with an FDR threshold of 0.25)).

Genus	β*	FDR p-value**
Slackia	-0.24	0.01
Leuconostoc	0.14	0.01
Barnesiella	0.22	0.03
Prevotellaceae NK3B31 group	-0.16	0.03
Anaerofustis	0.09	0.07
Lachnospiraceae UCG 010	0.17	0.07
Odoribacter	0.15	0.08
X Eubacterium ruminantium group	-0.19	0.08
Allisonella	0.10	0.08
Phascolarctobacterium	-0.22	0.09
Prevotella (7)	0.16	0.09
Victivallis	0.07	0.09
X Eubacterium siraeum group	-0.18	0.11
Incertae Sedis	0.08	0.12
Papillibacter	-0.06	0.14
Defluviitaleaceae UCG 011	0.10	0.16
Methanobrevibacter	-0.23	0.16
Sutterella	0.14	0.16
Alloprevotella	-0.12	0.17
Dialister	0.18	0.19
Lachnospira	0.12	0.21
Romboutsia	0.12	0.24

* β = coefficient; confounders included in the multivariate model were: age (years), gender (woman, man), ethnical background (Dutch, non-Dutch) and SES-WOA score.
** FDR p-value = false discovery rate p-value, subjected to multiple hypothesis testing correction using the Benjamini-Hochberg method with an FDR threshold of 0.25.



CHAPTER 7

General discussion

The research presented in this thesis provides novel insights into the potential role of environmental transmission of selected exemplary zoonotic pathogens that are typically considered as foodborne pathogens. This was done by studying some of their characteristics in the environment and the spatial associations between the related human disease occurrence and local livestock density. In addition, this thesis explored the spatial association between the urban-rural gradient and gut microbial diversity and composition. As "*Everything comes from somewhere*"[1], but much is yet to know, several knowledge gaps were identified, providing the basis for the aims of this thesis, which used methods and approaches linked to the concept of eco-epidemiology. Broadly speaking, the aims of the different chapters focused on a number of drivers of change fitting the category "globalization and the environment".[2] These drivers were climate change and the human-made environment. The first theme focuses mainly on climate change as driver for infectious diseases and aimes specifically at studying the associations between health risks associated with exposure to urban pluvial flooding. Here, health risks were defined as the possible risk for the development of acute gastroenteritis (AGE) or acute respiratory infection (ARI) after such exposure. The subsequent theme covers four chapters of this thesis related to the human-made environment, which is also relevant for the health risks of floodwater exposure given that pluvial flooding is also influenced by water management in urban planning. Three of the chapters deal specifically with the human-made environment theme and focus on the role of potential environmental transmission of zoonotic pathogens (*Campylobacter*, STEC O157 and *Salmonella*) from livestock farming to humans. The first chapter aimed to study the prevalence and genotype diversity of *Campylobacter* in surface water, as well as the relative contributions of several putative (domestic and wild) animal sources to surface water contamination with two *Campylobacter* species of public health significance (*C. jejuni* and *C. coli)*. Additionally, potential effects of local livestock density, type of surface water, and season were assessed. The next two chapters focused on studying the spatial associations between the incidence of human STEC O157, *Campylobacter* and *Salmonella* infection and livestock exposure using the population-weighted number of farm animals as exposure measure. The last chapter covers a more holistic approach. The focus was not on a specific infectious disease, but on the human gut microbiome as readout. The gut microbiome is generally assumed to play an important role in the overall health status of individuals. Here, potential (indirect) effects of the urban-rural gradient on the human gut microbiome was studied.

Together, these themes created a flow throughout this thesis, from syndrome-level to pathogen-specific, as well as more holistic approaches focused on the potential effects of the environment on human infectious diseases and gut microbiota. The main findings of the chapters are summarized and discussed below. The last part is discussed from a One Health perspective, followed by some concluding remarks.

Climate change as driver for foodborne diseases

"Direct exposure to pluvial floodwater, and activities leading to such exposure, are associated with an increased risk of developing AGE and ARI"

Countries in the Northern hemisphere face an increasing trend in extreme rainfall events causing recurrent pluvial flooding in especially urban areas due to climate change.[3] When the drainage capacity of urban sewage drainage systems (~20 mm of rainfall per hour) is overwhelmed[3,4], this can lead to accumulation of up to 10-50 cm of pluvial floodwater at the surface.[3,4] In principle, pluvial water is clean. However, during flooding events in urban areas, pluvial water mixes with sewage water contaminated with fecal material and its associated pathogens.[5] It is therefore plausible to assume that urban residents can develop health complaints following direct exposure to contaminated floodwater. This assumption was supported by the results of Chapter 2.

In this Chapter, we were able to disentangle the effects of exposure to combined sewage overflows/street flooding on human health, and disease syndromes (AGE and ARI) in particular, in the Netherlands. This was necessary, because most studies into flooding and associated health risks are focused on developing countries and infectious diseases related to poor hygiene conditions, such as *Cholera*[6,7]. Other studies did focus on the association between sanitary sewer overflows (SSOs)[8] or combined sewer overflows (CSOs)[9] and its association with gastrointestinal illness (GI; emergency room (ER) visits with a primary diagnosis of GI). But specific risk factors leading to exposure and eventually GI, like swimming or drinking contaminated drinking water, were not identified.[9] Besides, the design of this study (retrospective with self-reported data) allowed us to collect information on the bottom part of the disease triangle where usually most cases are found. Thus, we were able to gather data about ARI and AGE cases that are generally not reported to the General Practitioner (GP), i.e. cases that do not require medical attention. Only one other study did focus on the association between self-reported health complaints, like diarrhea or coughing, and exposure to pluvial floodwater in the Netherlands.[3] As our study was performed on a larger scale (699 households *versus* 149 households[3]), we were able to provide insights into the associations between exposure and well-defined (standardized) AGE or ARI syndrome definitions instead of separate health complaints with more precise estimates instead.

The associations between direct exposure to urban pluvial floodwater via different risk factors (i.e. skin contact with floodwater, post-flooding cleaning operations and cycling through floodwater) and the occurrence of both AGE and ARI, as presented in Chapter 2, may reflect, to some extent, the primary transmission routes of the pathogens in question, i.e. inhalation for ARI and ingestion for AGE. Indeed, several

pathogens are typically associated with flood-ravaged settings in developed countries. For example, the foodborne pathogen *Campylobacter* was found in all water plaza samples in which people recreate[10], leading to a risk of developing AGE for those people. Moreover, pluvial floodwater is always contaminated with feces and its related pathogens (*E. coli*, intestinal enterococci, and enteropathogens, such as enterovirus, norovirus and *Campylobacter*).[5] Exposure to contaminated pluvial floodwater therefore can potentially lead to environment-mediated (foodborne) pathogen transmission and the development of AGE and ARI.

As extreme rainfall events are increasing in frequency, more residents are expected to experience urban pluvial flooding in the future, leading to an increased public health risk.[11] In the Netherlands, urban flooding will not pass easily unnoticed, because it is a densely populated country (> 500 inhabitants/km^2).[12] Awareness for urban flooding was raised in recent years, leading to Dutch residents demanding municipal health services to improve drainage systems and reduce cementation as to prevent pluvial flooding.[13] Indeed, to facilitate the natural drainage of water in soil, governmental authorities are now promoting removal of pavement and the increase of greenness in gardens and urban areas as a response to the citizen's call for attention for this topic. However, people in developed countries tend not to perceive the associated health risks of those flooding events[3]. Therefore, the health risks have not been part of the debate regarding urban pluvial flooding. In short, this chapter shows that it is necessary to take proper care of water drainage systems/sewage systems in terms of their drainage capacity to mitigate health risks in urban areas. With this, it adds another perspective to the ongoing debate about the consequences of climate change and urban flooding.

The human-made environment as driver for foodborne disease and gut microbiota

Spatial distribution of foodborne pathogens from livestock

Besides climate change, the human-made environment can be a driver for infectious disease occurrence as well. Markers for this driver, such as urbanization, built environment and intensive agriculture, can enable emergence and spread of pathogens, especially in a densely populated country like the Netherlands, where residents live in close proximity to intensive livestock farms and a large livestock population in terms of numbers of cattle, pigs and poultry.[14] Different farm animals can be sources for different zoonotic pathogens causing gastrointestinal illness[15-17], which typically transmit via food derived from these animals. However, it is also plausible that there might be some degree of environment-mediated transmission via fecal shedding.

When livestock is not only kept indoors or housed in a (partially) open housing system[18,19], fecal shedding can lead to environment-mediated transmission via the direct surroundings of the farm through fecal contamination of soil and water, surface water run-off or airborne (dust) transmission towards residents living nearby those farms.[16,18,20-27] The possibility of environment-mediated transmission of typically foodborne pathogens, such as *Campylobacter*, STEC O157 and *Salmonella* in the Netherlands, and more specifically their spatial distribution, was therefore further studied in Chapters 3, 4 and 5 of this thesis.

Livestock-associated *Campylobacter* in surface water

"Surface water is mainly contaminated with Campylobacter originating from meat-producing poultry (i.e. broilers and turkeys) in areas with high poultry densities and from wild birds in recreational waters and WWTP discharge points"

Campylobacter is considered to be a typically foodborne pathogen that is mainly transmitted through the consumption of contaminated chicken meat or beef.[16] Although most campylobacteriosis cases can be attributed to the poultry reservoir, several epidemiological studies have shown that chicken meat consumption only explains about half of the human campylobacteriosis cases attributable to poultry.[28-33] Most interventions to control *Campylobacter* infections have been focused on reducing the pathogen spread through the food production chain, particularly poultry meat. Still, there has been no decrease in human campylobacteriosis cases so far.[34] It is therefore possible that other pathways are important for the transmission of *Campylobacter* to humans as well.

In the human living environment, *Campylobacter* is commonly found in surface water contaminated with animal feces, sewage effluent and agricultural runoff. This surface water has the potential to act as a vehicle for transmission for *Campylobacter* among animals, from animals to humans and *vice versa*.[35,36] A combined source attribution study from the Netherlands and Luxembourg showed that poultry and wild birds are the main contaminators of surface water with *C. jejuni* and *C. coli*, of which the relative contributions seemed to vary with season, water type, and the magnitude of local poultry production.[37] Those results suggest dissemination of *Campylobacter* from poultry farms into the environment in poultry rich regions, but interpretation of these findings is limited due to the extensive use of animal isolates which were not all from the same countries as surface water isolates and also not from the same years as water isolates, retail food data, and course spatial resolution of the analyses. Overall, the aquatic environment seemed to contribute to the transmission of *Campylobacter* to humans[37], but other evidence of environment-mediated transmission of *Campylobacter* is limited. Studies using more detailed data for quantifying the prevalence and

genotype diversity of this pathogen in surface water and the relative contributions of several putative animal sources to surface water contamination with *C. jejuni* and *C. coli* are therefore essential to further investigate the possibility of environment-mediated transmission of *Campylobacter* as a first step. Therefore, this became the subject of study of Chapter 3.

Our study allowed for the quantification of the level of contamination of different surface water types (i.e. recreational water, agricultural waters and wastewater treatment plant (WWTP) discharge points) throughout the Netherlands using recent and local data with a high spatial resolution and typed using whole-genome sequencing (WGS) to derive high-throughput genomic data from the different animal sources when compared to previous studies. We also were able to include geographical variation in the density of different livestock species in the Netherlands in the water sampling scheme to fecal surface water contamination. Research in this chapter revealed a widespread presence of *C. jejuni* and *C. coli* in surface water in the Netherlands. Wild birds were the dominant source of this surface water contamination, followed by meat-producing poultry (i.e. broilers and turkeys). Therefore, surface water can be seen as a collection vessel of strains from multiple hosts as a result of fecal pollution from different animal sources.[16] The role of wild birds in the epidemiology of *Campylobacter* infections in both animals and humans, however, is largely unknown, but should not be underestimated for various reasons. First, drug resistant *Campylobacter* strains (especially those resistant to tetracycline and fluoroquinolones) are increasingly isolated from various wild bird species throughout the world, which is part of the growing global public health concern regarding antibiotic resistance.[38-46] Second, source tracking (molecular-based) studies have linked both human and animal cases of *Campylobacter* infections to wild birds.[47-49] Third, human and animal exposure is possible via equipment or surface contamination with wild birds' fecal material (like in parks or children's playgrounds), but also through contact with contaminated surface water with wild bird-associated *Campylobacter* strains.[16,38,50,51] This is supported by our finding that wild bird-associated strains were mostly found in recreational waters and further emphasized by the finding that open-water swimming is a risk factor for human campylobacteriosis cases attributable to environmental sources (surface water and wild birds).[16] Thus, the role of wild birds in *Campylobacter* epidemiology should be studied more extensively in the near future to address their role in both human and animal health and its related environment-mediated route of transmission. The difficulty is that virtually nothing can be done to control *Campylobacter* in wildlife. In this regard, the finding that > 90% of *Campylobacter* strains from recreational waters are attributable to wild birds, and that the higher contribution of wild birds to recreational water contamination relative to other types of water is significant, implies that the risk of acquiring campylobacteriosis from, e.g., swimming in official recreational water sites in the Netherlands, is largely beyond human control.

Meat-producing poultry, however, is still the secondary most likely source of *Campylobacter* contamination in Dutch surface water and poultry-associated strains are mostly found in agricultural waters and densely populated poultry areas.[37] This may have implications for public health, because even if poultry meat is *Campylobacter*-free when consumed, human exposure can still occur through environmental pathways and specifically the aquatic environment in poultry dense areas. Therefore, interventions aimed at controlling environmental dissemination of *Campylobacter* at primary livestock production and WWTPs are necessary, provided that residents are indeed exposed via this route and cost-benefit analyses show that public health benefits outweigh the costs of such interventions.

Livestock-associated spatial risk factors for human disease

"Human STEC O157 infections can be caused by environmental exposure to small ruminants besides food, while the route of transmission of Salmonella and Campylobacter to humans is mainly foodborne"

In the Netherlands, domestic ruminants (cattle, sheep and goats) are important livestock sources of human STEC O157 infections[17], pigs and laying hens are the most important sources of *Salmonella*[15], as broilers and cattle are for *Campylobacter*[16,29]. Those animals can shed high quantities of those pathogens (e.g. >105/g of STEC O157) via their feces that are able to survive for varying periods of time in their new environment (e.g. ≤3 months in manure; ≤1 month in soil).[16,22-25,27,52,53] Pathogens may also be spread through air. The transition in the laying hen sector from cage housing to alternative housing systems, for example, led to an increased PM_{10} transmission, creating potential opportunities for aerial spread of pathogens.[54] Furthermore, a Dutch study found an association between the concentration of *Campylobacter jejuni* in airborne dust and the presence of poultry farms in the area.[21] The presence of these pathogens in different environmental reservoirs is one of the prerequisites for environment-mediated transmission, together with residential exposure. This is also suggested by the H7N7 avian influenza epidemic in the Netherlands in 2003 and the observed drop in human campylobacteriosis incidence after the implementation of massive culling operations in poultry farms.[28] It has been hypothesized that a reduced environmental contamination with *Campylobacter* from the culled and therefore temporarily emptied poultry farms, as well as inactive slaughterhouses, could have occurred. That a high poultry density in an area is associated with an increased prevalence of poultry-associated *Campylobacter* strains further supports this hypothesis (Chapter 3). However, studies focusing on spatial associations between exposure to different livestock densities in an area and the corresponding human incidence of STEC O157, *Campylobacter* and *Salmonella* infections to further explore the potential role of environment-mediated transmission of foodborne pathogens, are limited.

Chapter 4 and 5 provided information on the spatial association between human STEC O157, *Campylobacter* and *Salmonella* infections and the combined exposures to livestock. Chapter 4 showed that exposure to small ruminants (goats and sheep) is associated with increased incidence of human STEC O157 infections, namely in the summer. Given the presence of STEC in small ruminants' feces and farms[55], it is plausible that human infections occur via environmental transmission, especially when zooming into their housing system.[55-58] The housing system of small ruminants is often partially open and deep litter layers are spread for the animals to use for bedding material and to defecate on, which generates a lot of dust.[19,52] As a result, the transport of STEC O157 in dust through air can be one of the possible environmental transmission routes if infected animals are present on the farm.[19,52] This is further supported by the finding that rural areas with a higher farm density had a higher concentration of commensals, among which STEC in dust particles when compared to other areas.[21] Although no significant associations with the number of goats and sheep were found, the presence of livestock-related microbial markers, such as STEC, indicated that microbial air pollution with STEC seems a plausible route of exposure.

Interestingly, our main finding regarding STEC O157 differed from other studies, as these studies found an association with cattle and not small ruminants.[26,59-61] This could have several explanations. The first major difference with previous studies is the use of a different probability of exposure metric as mentioned in the previous paragraph. Instead of animal density[62,63], we used the population-weighted number of animals, as the level of exposure is not only determined by the number of animals in an area, but also by the number of residents in a certain area and where they live in that area. Therefore, our study is more likely to have captured environmental exposure better, as exposure is less likely to occur when nobody lives in the vicinity of these animals. Second, we included small ruminants in this study, besides cattle, while including several potential reservoirs of the pathogen in question in the models. Last, changes in livestock sectors over the past few years could explain the differences as well.[64] Examples are the increase in number of animals per farm, paralleled by a reduction in the number of farms over the years.[12,64,65] Furthermore, cows are increasingly kept inside in closed housing throughout the year.[12,64,65] Thus, it is possible that, for example, surface runoff and surface water contamination of STEC from cattle reduced over the years and that aerial spread from small ruminants plays a more important role nowadays.

Although Chapter 4 showed that the population-weighted number of small ruminants is a risk factor for acquiring human STEC O157, we were not able to quantify the relative importance of our finding within the broader context of all possible risk factors, such as food consumption, as we assumed that the effect of

food is comparable for all individuals within the population and that this effect has no spatial pattern. A recent source attribution analysis based on a meta-analysis of case-control studies showed that food still is the dominant transmission pathway of human STEC O157 infections (37%), followed by contact with animals (15%) and the environment (10%).[66] This further highlights the role of other sources besides food. STEC O157 transmission from small ruminants to humans should therefore be studied more extensively in the near future to further unravel all possible transmission pathways and the specific role of the environment therein.

Chapter 5 did not show consistent significant associations between the local level of exposure to different livestock species (i.e. the population-weighted number of animals) and human campylobacteriosis or salmonellosis incidence, which is also confirmed by the analyses of *Campylobacter* exposure using serological data. In contrast, Chapter 3 showed that *C. jejuni* and *C. coli* have a widespread presence in surface water in the Netherlands, with wild birds and meat-producing poultry being the dominant sources. It is therefore likely that *Campylobacter* is disseminated from meat-producing poultry farms into the environment, which is also suggested by the association between *C. jejuni* DNA presence and poultry density (high *versus* low poultry density).[21] The findings of Chapter 5, however, indicate that the potential for environmental transmission of *Salmonella* and *Campylobacter* from livestock to Dutch residents is likely to be limited despite an ubiquitous presence of livestock-associated *Campylobacter* strains in surface water and the mixed and densely populated characteristics regarding both humans and livestock in a country like the Netherlands. This supports current knowledge that human infections with *Salmonella* and *Campylobacter* are mainly foodborne.

As we only used population-weighted number of animals as a proxy for human exposure to study *Campylobacter* and *Salmonella* in Chapter 5 (*Campylobacter* and *Salmonella*), future studies need to confirm whether human exposure to livestock-associated *Campylobacter* indeed is limited. Therefore we recommend further research to focus on possible environment-mediated transmission pathways of *Campylobacter* from livestock, and especially meat-producing poultry farms, to neighboring residents. This could be a prospective cohort study, which includes individually reported exposures, including consumption of particular food items, besides risk factors related to direct contact with animals and surface water, as well as spatial risk factors. Ultimately, also human, animal and environmental samples need to be gathered to be able to compare *Campylobacter* isolates based on whole-genome sequencing (WGS) to be able to study the direct link between them. The presence of *Campylobacter* in the environment (Chapter 3), however, remains a concern for public health, because residents' exposure is still possible. To further limit the chance

of such exposure, we do recommend measures to prevent pathogen spread from meat-producing poultry farms into the environment. Keeping animals in fully closed housing systems to prevent pathogen spread as much as possible is not always an acceptable solution because of other considerations such as animal welfare. Therefore, vaccination of poultry flocks could be considered as a solution as well.[67] This has the advantage that it controls the pathogen, for example *C. jejuni*, at the pre-harvest level, which is critical to also reduce foodborne infections with *Campylobacter*.[67]

The urban-rural gradient and gut microbial diversity and composition

"Bacterial diversity, richness and composition of the adult gut microbiome are associated with the urban-rural gradient in the Netherlands"

The human living environment is one of the factors that can influence the human gut microbiome. Especially in an agricultural environment, exposure to micro-organisms is higher than in urbanized areas, for example due to a higher exposure to farm animals in a rural area.[68] This can potentially lead to an association between the level of urbanization and microbial diversity and composition. Previously, an association was found between urbanization and gut microbial diversity and composition in a study performed in the Northern part of the Netherlands.[69] Furthermore, the diversity and composition of especially airway, but also gut microbiotas, has been shown to differ between urban and rural infants in Denmark.[70] Additionally, an association was found between an urbanized structure of the airway and gut microbiotas with an increased risk of asthma coherently during multiple time points.[70] However, whether these relations are causal is not yet clear.[68]

Chapter 6 revealed that the human-made living environment is not only a driver of change for specific pathogens, but it also has the potential to drive changes in diversity and composition of the human gut microbiome. To our knowledge, this study is the first comprehensive study associating the diversity and composition of human gut microbial communities with highly detailed land use information and relevant epidemiological data. In short, we showed consistently that there are significant associations between agricultural land coverage and bacterial diversity, richness and composition of the adult gut microbiome. Both bacterial diversity and richness significantly increased and microbial composition significantly changed with increasing agricultural land coverage. The limited number of studies focusing on the potential effects of the living environment on the diversity and composition of the human gut microbiome showed similar results as Chapter 6 and concluded that different microbiome signatures exist along the urban-rural gradient.[69,70]

It is possible that the effect of agricultural areas is potentially caused by farm animals housed in those areas, we therefore recommend to repeat those analyses with information on different types of farm animals. If possible, it would even be more preferable to combine the aforementioned study with exposure studies, because information on environmental exposure to and transmission of livestock-associated micro-organisms is still limited. Our hypothesis is that this will result in more prominent associations between living in an area with a higher number of farm animals and the human gut diversity and composition. It is also highly plausible that different effects will be observed for different animal types, as microbial spread depends on, for example, the type of animal, its related housing system and amount of dust produced,[19,21,52] which was already discussed for zoonotic pathogens in Chapters 3 and 4 of this thesis. As the analyses were performed on genus level, it was difficult to get insights into potential positive effects of specific genera and related ASVs on human health. To address this, we would recommend to study this population with metagenomics in the near future to be able to look for function differences[69] or to set up an *in vitro*, *in vivo* investigation to study the impact of specific bacteria on human health and immune function.[71]

This thesis in a One Health perspective

"The role of the environment should not be forgotten when using a One Health approach"
The definition of One Health basically describes that the health of humans, animals and the environment is interconnected, in line with the concept that "*everything comes from somewhere*".[1,72] The One Health paradigm is often presented as a disease triangle, where the pathogen is usually central, and three main categories: humans, animals and the environment, each are covering one corner while summarizing their related drivers of change, as they are all interacting with each other and the disease causing pathogen in question.[2,73] To achieve optimal health outcomes, it is therefore necessary to use a collaborative, multisectoral, and transdisciplinary approach by working on different levels (local, national and global levels).[72] As described in Chapter 1, this view on One Health has some history, but has been moved forward as a concept during the "Stone Mountain Meeting" in 2010.

In the past decade, studies have become more inclusive with regard to the One Health approach and foodborne pathogens, as some studies included the environment as driver besides humans, food and animals.[26,60,74] One study set-up a One Health, collaborative, interdisciplinary network and sequence data repository for enhanced hepatitis E virus (HEV) molecular typing, characterization and epidemiological investigations (HEVnet).[74] In Europe, hepatitis E is considered a zoonosis

transmitted via contaminated pork or other pig-derived products. Genotype 3 and, to a lesser degree, genotype 4 are prevalent in the animal reservoir (pigs, wild boar and deer), as well as in humans.[75] Within the HEVnet data repository, HEV sequences from human, animal or environmental sources can still be uploaded and shared with the professional network for further studies. As HEV is considered to be mainly foodborne, sequences from food sources are also gathered. In the first two years (2017-2018), 1,650 sequences were uploaded, of which 89% of the samples had a human origin, 5% an animal origin, 6% a food origin and 0% (5 samples) an environmental origin.[74] Although the idea of a One Health approach was there, it remained difficult to prevent anthropocentric thinking and properly include animals (i.e. wild animals or animals in complex ecosystems) and especially the environment. Even studies related to animal health are conducted from an anthropocentric perspective, as they mostly focus on companion or farm animals, thus, animals with a clear (economic) value for humans. Studies focusing on animal health or health of animals in more complex ecosystems, including the environment, are clearly less frequently conducted and more limited. While the recently emerged zoonoses (Zika virus, SARS-CoV-2)[76,77] made the role of ecosystems explicit, the One Health thinking still has an anthropocentric focus by usually seeking sources of human disease or to prevent transmission to humans. It is time to take a step back and focus on how and why the animal source got infected or the environment got contaminated in the first place, to prevent spread via the environment towards both animals and humans.

This thesis showed that both the environment and wildlife cannot be ignored when studying (transmission of and risk factors for) foodborne pathogens. Animal sources like small ruminants are a spatial risk factor for human STEC O157 infections in the Netherlands (Chapter 4). Given the presence of this pathogen in small ruminants' feces and farms[55-58] together with the open housing system of small ruminants including the deep litter layers which generate a lot of dust, it is plausible that human infections occur via environmental transmission. Although Chapter 5 showed no consistent significant relations between livestock sources and *Campylobacter* or *Salmonella* in the Netherlands, it is too early to discard the role of the environment here. This is especially true for *Campylobacter*, as we found a high prevalence of *Campylobacter* in Dutch surface water (Chapter 3) and wild birds are the main source of this contamination, especially in recreational waters. A follow-up study identified swimming in recreational water as a risk factor for humans acquiring campylobacteriosis associated with environmental sources (i.e. surface water and wild birds).[16] Thus, not only livestock, but also wildlife should be considered as a potential source leading to environment-mediated transmission of foodborne pathogens.

To understand where we stand, every chapter of this thesis was connected to its general drivers of change (human, animal, environment) from the One Health disease triangle as described in Figure 1. This figure shows that basically all chapters focused on two out of three main categories, including some of the related drivers of change. A similar strategy was chosen for the definition of a One Health Surveillance System (OHSS, Matrix – Matrix framework (ejp-matrix.eu)), which is a system were collaborative efforts should exist across at least two sectors (among human health, animal health, food safety and environment).[78] Going back to the definition of One Health, this definition describes that everything and every discipline related to human, animal and environmental health is interconnected. Therefore, the remaining question is whether involving only two main categories and some of its related drivers in One Health-focused studies is enough to disentangle all complex, ongoing interactions between animals, humans and the environment related to specific pathogens.

Although drivers related to each chapter of this thesis are clear, other interactions are present in the background, but not studied simultaneously. Chapter 4 and Chapter 5 for example studied the association between livestock and human infections with foodborne pathogens. However, no evidence was provided for a particular mode of transmission (e.g. air-borne, soil or water), as no epidemiological data concerning possible exposure pathways was available for inclusion in the analyses. Although we identified a spatial risk factor concerning STEC O157 infections related to small ruminants, which is likely to be related to environmental transmission, we could not determine the relative contribution of this pathway compared to transmission via food that is still considered the main source of infections. Furthermore, Chapter 6 identified the association between diversity of the human gut microbiome and the urban-rural gradient. Here, animals, and particularly livestock, are missing in the equation, while it could be hypothesized that these farm animals lead to an increased environmental microbial exposure for humans in agricultural areas. Following these research gaps, it can be stated that future studies into zoonotic pathogens should be as inclusive as possible, as is also stated by the One Health definition. Therefore, a perfect study would include all drivers and determinants that interact with the pathogen in question. To do so, different disciplines have to work together to get the answers they need related to human-, animal- and/or environmental health, which is increasingly recognized by professionals and already resulted in several One Health related projects in the European Union, for example, those that were part of the European Joint Programme (EJP) on One Health (e.g. MATRIX, COHESIVE, BIOPIGEE, ADONIS, BeONE)[79] and United4Surveillance as part of EU4Health[80].

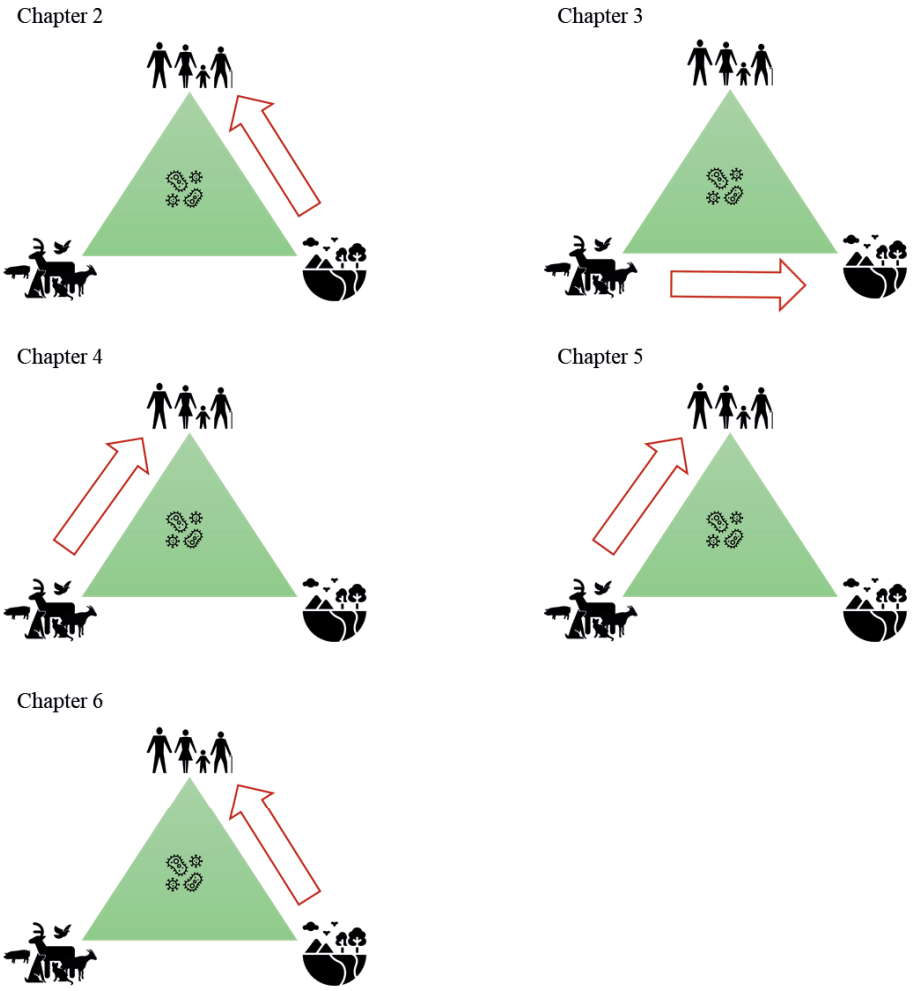


Figure 1. A visual representation of the One Health disease triangle per chapter of this thesis. The red arrow indicates which of the three main categories (humans, animals, environment) of the "drivers of change" were included per chapter and in what direction the association was studied. *Icons from flaticon.com*

Instead of using the disease triangle as presented in Chapter 1 (Figure 1) for visualization and mind-mapping previous to setting up an all-inclusive One Health study, we would like to propose a similar, but adapted approach developed by Paolo Zucca in which the three general drivers of the theoretical framework of One Health (humans, animals and the environment) are merged with the illusory white triangle developed by the Italian perception psychologist Gaetano Kanisza (1955).[81] The idea behind this approach is that we see the white triangle, while it does not exist.[81] This demonstrates the Gestalt idea that the sum of visual perception is more than its

parts, because the observer views all objects together as a single image and not every object separately.[81] The same is true for One Health, as it is more than the sum of its three main categories and related drivers. It is also everything that is in the area of the white triangle and that connects these three categories; i.e. all underlying interactions.[81] By adapting Figure 2 for a specific pathogen, it could result in the development of a mere all-inclusive One Health approach when studying zoonotic pathogens and their role in, for example, human or animal health. Therefore, we want to recommend to implement this newly developed One Health framework of Paolo Zucca (2021) in future One Health related studies.

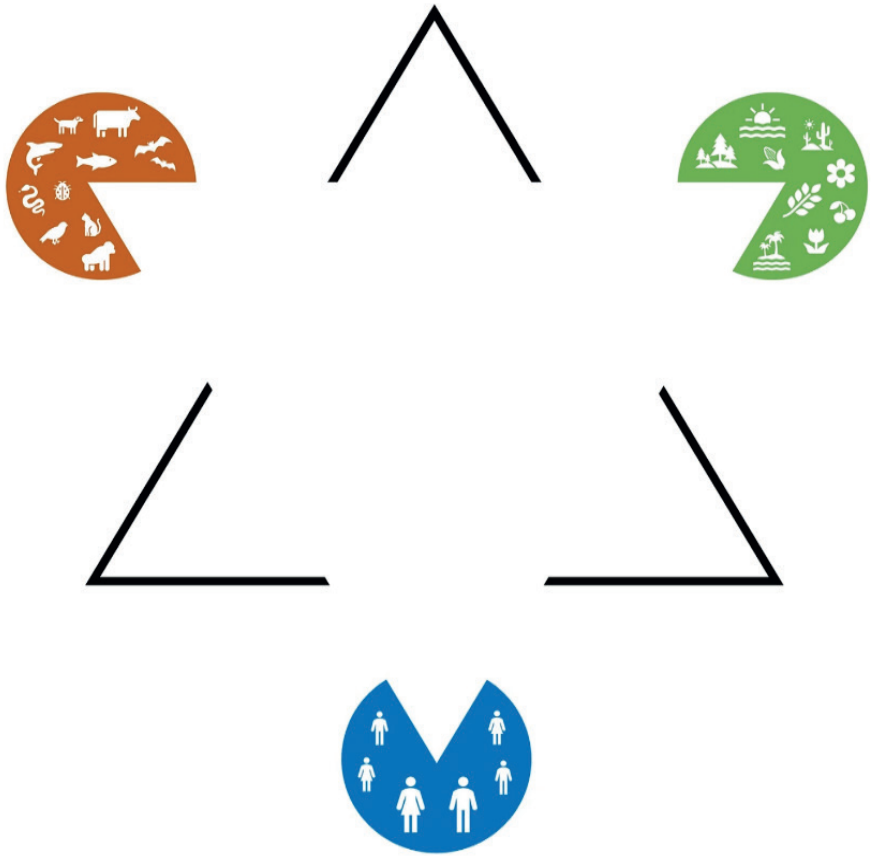


Figure 2. A visualization of the three general drivers of the theoretical framework of One Health (humans, animals and the environment) merged with the illusory white triangle developed by the Italian perception psychologist Gaetano Kanisza (1955). Retrieved from Zucca (2021)[81] *(original image source Kanisza 1955 available also as open source on Wikimedia, modified and merged with the One Health triad framework by Paolo Zucca)*

Concluding remarks

With increasing attention to zoonoses, the need for a multilevel and multidisciplinary (One Health) approach to study the complex interactions between people, animals, and their shared environment to achieve optimal health outcomes for all, increased as well. While the statement "*Everything comes from somewhere*" is self-explanatory, studies including the potential role of the environment in (foodborne) zoonosis transmission, and more holistically shaping the human gut microbiome, were scarce. To date, it has become clear that the environment can play a role in the transmission of zoonotic pathogens from both livestock and wildlife, and in this thesis we investigated in particular surface water and contaminated urban pluvial floodwater. However, evidence about specific routes of transmission (e.g. air-borne, via soil or water) and their relative importance was not yet provided. Furthermore, evidence is generally scarce in the scientific literature. To study the interactions between different drivers of zoonotic pathogens, we recommend future studies to focus on pathogen transmission and their related transmission pathways to be able to look beyond the three drivers of change included in the One Health disease triangle. This can be visualized by including a white triangle based on the Gestalt idea that the sum of visual perception is more than its parts. This triangle illustrates all underlying interactions that have to be included to study zoonotic pathogens and get a complete picture of all interactions influencing one another. Although the realization of such an all-inclusive study is complex, we believe that cutting-edge scientific advances continue to elucidate such complex mechanisms and will help to set-up such a study in the future. Some steps have already been taken, as the foundation with regard to transdisciplinary collaboration between experts already exists. However, we want to emphasize to stop anthropocentric thinking when using the One Health concept, as animal- and environmental health are just as important as human health in preventing the emergence and transmission of zoonotic pathogens in the future.

References

1 Quammen, D. Spillover - Animal infections and the next human pandemic. 1st edn, (W. W. Norton & Company, Incl., 2012).

2 Semenza, J. et al. Determinants and Drivers of Infectious Disease Threat Events in Europe. Emerging Infectious Disease journal 22, 581, doi:10.3201/eid2204.151073 (2016).

3 De Man, H. et al. Gastrointestinal, influenza-like illness and dermatological complaints following exposure to floodwater: a cross-sectional survey in The Netherlands. Epidemiology and Infection 144, 1445-1454, doi:10.1017/S0950268815002654. (2015).

4 Ten Veldhuis, J., Clemens, F., Sterk, G. & Berends, B. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water research 44, 2910-2918, doi:10.1016/j.watres.2010.02.009. (2010).

5 De Man, H. et al. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. Water research 48, 90-99, doi:10.1016/j.watres.2013.1009.1022 (2014).

6 Bett, B., Tumusiime, D., Lindahl, J., Roesel, K. & Delia, G. The Handbook of Environmental Chemistry The Role of Floods on Pathogen Dispersion, (Springer, 2021).

7 Taseen, S., Jawed, S. & Mohammad Ameen, A. Cholera outbreak amidst urban flooding in Karachi: An emergency condition. Annals of Medicine and Surgery 81, 104365, doi:10.1016/j.amsu.2022.104365 (2022).

8 Jagai, J. S., DeFlorio-Barker, S., Lin, C. J., Hilborn, E. D. & Wade, T. J. Sanitary Sewer Overflows and Emergency Room Visits for Gastrointestinal Illness: Analysis of Massachusetts Data, 2006–2007. Environmental Health Perspectives 125, 117007, doi:10.1289/EHP2048 (2017).

9 Miller, A. G., Ebelt, S. & Levy, K. Combined Sewer Overflows and Gastrointestinal Illness in Atlanta, 2002–2013: Evaluating the Impact of Infrastructure Improvements. Environmental Health Perspectives 130, 057009, doi:10.1289/EHP10399 (2022).

10 Sales-Ortells, H. & Medema, G. Microbial health risks associated with exposure to stormwater in a water plaza. Water Research 74, 34-46, doi:10.1016/j.watres.2015.01.044 (2015).

11 Min, S.-K., Zhang, X., Zwiers, F. W. & Hegerl, G. C. Human contribution to more-intense precipitation extremes. Nature 470, 378-381, doi:10.1038/nature09763 (2011).

12 Smit, L. A. M. & Heederik, D. Impacts of Intensive Livestock Production on Human Health in Densely Populated Regions. GeoHealth 1, 272-277, doi:10.1002/2017GH000103 (2017).

13 Meindertsma, B. & Van der Parre, H. in https://nos.nl/artikel/2230728-burgers-met-natte-voeten-kijken-naar-overheid-maar-willen-niet-meer-betalen.html (2018).

14 Heederik, D. J. J., Smit, L. A. M. & Vermeulen, R. C. H. Go slow to go fast: a plea for sustained scientific rigour in air pollution research during the COVID-19 pandemic. Eur Respir J 56, doi:10.1183/13993003.01361-2020 (2020).

15 Mughini-Gras, L. et al. Risk Factors for Human Salmonellosis Originating from Pigs, Cattle, Broiler Chickens and Egg Laying Hens: A Combined Case-Control and Source Attribution Analysis. PLOS ONE 9, e87933, doi:10.1371/journal.pone.0087933 (2014).

16 Mughini-Gras, L. et al. Sources and transmission routes of campylobacteriosis: A combined analysis of genome and exposure data. J Infect 82, 216-226, doi:10.1016/j.jinf.2020.09.039 (2021).

17 Mughini-Gras, L. et al. Attribution of human infections with Shiga toxin-producing Escherichia coli (STEC) to livestock sources and identification of source-specific risk factors, The Netherlands (2010–2014). Zoonoses and Public Health 65, e8-e22, doi:10.1111/zph.12403 (2018).

18 Bestman, M. et al. Factors related to free-range use in commercial laying hens. Applied Animal Behaviour Science 214, 57-63, doi:10.1016/j.applanim.2019.02.015 (2019).

19 Schimmer, B. et al. Seroprevalence and risk factors of Q fever in goats on commercial dairy goat farms in the Netherlands, 2009-2010. BMC Veterinary Research 7, 81, doi:10.1186/1746-6148-7-81 (2011).

20 Berry, E. D. et al. Effect of Proximity to a Cattle Feedlot on O157:H7 Contamination of Leafy Greens and Evaluation of the Potential for Airborne Transmission. Applied and Environmental Microbiology 81, 1101-1110, doi:10.1128/AEM.02998-14 (2015).

21 de Rooij, M. M. T. et al. Insights into Livestock-Related Microbial Concentrations in Air at Residential Level in a Livestock Dense Area. Environmental science & technology 53, 7746-7758, doi:10.1021/acs.est.8b07029 (2019).

22 Guillier, L. et al. Risk factors for sporadic salmonellosis: a systematic review and meta-analysis. Microbial Risk Analysis 17, 1-11, doi:10.1016/j.mran.2020.100138 (2021).

23 Lee, D. et al. Salmonella Survival in Soil and Transfer onto Produce via Splash Events. J Food Prot 82, 2023-2037, doi:10.4315/0362-028x.Jfp-19-066 (2019).

24 Nicholson, F. A., Groves, S. J. & Chambers, B. J. Pathogen survival during livestock manure storage and following land application. Bioresource Technology 96, 135-143, doi:10.1016/j.biortech.2004.02.030 (2005).

25 Schets, F. M. et al. Prevalence and types of Campylobacter on poultry farms and in their direct environment. Journal of Water and Health 15, 849-862, doi:10.2166/wh.2017.119 (2017).

26 Strachan, N. J. C., Dunn, G. M., Locking, M. E., Reid, T. M. S. & Ogden, I. D. Escherichia coli O157: Burger bug or environmental pathogen? International Journal of Food Microbiology 112, 129-137, doi:10.1016/j.ijfoodmicro.2006.06.021 (2006).

27 Strachan, N. J. C., Fenlon, D. R. & Ogden, I. D. Modelling the vector pathway and infection of humans in an environmental outbreak of Escherichia coli O157. FEMS Microbiology Letters 203, 69-73, doi:10.1111/j.1574-6968.2001.tb10822.x (2001).

28 Friesema, I. H. M., Havelaar, A. H., Westra, P. P., Wagenaar, J. A. & van Pelt, W. Poultry Culling and Campylobacteriosis Reduction among Humans, the Netherlands. Emerg Infect Dis. 18, 466-468, doi:10.3201/eid1803.111024 (2012).

29 Mughini Gras, L. et al. Risk factors for campylobacteriosis of chicken, ruminant, and environmental origin: a combined case-control and source attribution analysis. PLoS One 7, e42599, doi:10.1371/journal.pone.0042599 (2012).

30 Veilinga, A. & van loock, F. The dioxin crisis as experiment to determine poultry-related campylobacter enteritis. Emerg Infect Dis 8, 19-22, doi:10.3201/eid0801.010129 (2002).

31 Wagenaar, J. A., French, N. & Havelaar, A. H. Preventing Campylobacter at the source: Why is it so difficult? Clin Infect Dis 57, 1600-1602, doi:10.1093/cid/cit555 (2013).

32 Wagenaar, J. A., Newell, D. G., Kalupahana, R. S. & Mughini-Gras, L. Campylobacter: animal reservoirs, human infections, and options for control. Zoonoses: Infections affecting humans and animals – A focus on public health aspects., 159-177, doi:10.1007/978-94-017-9457-2_6 (2015).

33 Wilson, D. J. et al. Tracing the source of campylobacteriosis. PLoS Genet 4, e1000203, doi:10.1371/journal.pgen.1000203 (2008).

34 European Food Safety and European Centre of Disease Prevention and Control. The European Union summary report on trends and sources of zoonosis, zoonotic agents and food-borne outbreaks in 2017. (2018).

35 Collins, D. M. & Colwell, R. R. Viable but nonculturable stage of Campylobacter jejuni and its role in survival in the natural aquatic environment. American Society for Microbiology Journals 52, 531-538, doi:10.1128/aem.52.3.531-538.1986 (1986).

36 Murphy, C., Carroll, C. & Jordan, K. N. Environmental survival mechanisms of the foodborne pathogen Campylobacter jejuni. Applied Microbiology 100, 623-632, doi:10.1111/j.1365-2672.2006.02903.x (2006).

37 Mughini-Gras, L. et al. Quantifying potential sources of surface water contamination with Campylobacter jejuni and Campylobacter coli. Water Res 101, 36-45, doi:10.1016/j.watres.2016.05.069 (2016).

38 Ahmed, N. A. & Gulhan, T. Campylobacter in Wild Birds: Is It an Animal and Public Health Concern? Frontiers in Microbiology 12, doi:10.3389/fmicb.2021.812591 (2022).

39 Du, J. et al. Emergence of Genetic Diversity and Multi-Drug Resistant Campylobacter jejuni From Wild Birds in Beijing, China. Frontiers in Microbiology 10, doi:10.3389/fmicb.2019.02433 (2019).

40 Indykiewicz, P., Andrzejewska, M., Minias, P., Śpica, D. & Kowalski, J. Prevalence and Antibiotic Resistance of Campylobacter spp. in Urban and Rural Black-Headed Gulls Chroicocephalus ridibundus. EcoHealth 18, 147-156, doi:10.1007/s10393-021-01540-0 (2021).

41 Jurinović, L. et al. Occurrence of Campylobacter jejuni in Gulls Feeding on Zagreb Rubbish Tip, Croatia; Their Diversity and Antimicrobial Susceptibility in Perspective with Human and Broiler Isolates. Pathogens 9, 695, doi:10.3390/pathogens9090695 (2020).

42 Kürekci, C. et al. Characterization of Campylobacter spp. Strains Isolated From Wild Birds in Turkey. Front Microbiol 12, 712106, doi:10.3389/fmicb.2021.712106 (2021).

43 Kwon, Y.-K. et al. Prevalence of Campylobacter species in wild birds of South Korea. Avian Pathology 46, 474-480, doi:10.1080/03079457.2017.1315048 (2017).

44 Mencía-Gutiérrez, A. et al. Prevalence and antimicrobial resistance of Campylobacter from wild birds of prey in Spain. Comparative Immunology, Microbiology and Infectious Diseases 79, 101712, doi:10.1016/j.cimid.2021.101712 (2021).

45 Moré, E. et al. Seabirds (Laridae) as a source of Campylobacter spp., Salmonella spp. and antimicrobial resistance in South Africa. Environmental Microbiology 19, 4164-4176, doi:10.1111/1462-2920.13874 (2017).

46 Russo, T. P. et al. Prevalence and Antimicrobial Resistance of Enteropathogenic Bacteria in Yellow-Legged Gulls (Larus michahellis) in Southern Italy. Animals 11, 275, doi:10.3390/ani11020275 (2021).

47 Cody, A. J. et al. Wild bird-associated Campylobacter jejuni isolates are a consistent source of human disease, in Oxfordshire, United Kingdom. Environmental Microbiology Reports 7, 782-788, doi:10.1111/1758-2229.12314 (2015).

48 Gardner, T. J. et al. Outbreak of Campylobacteriosis Associated With Consumption of Raw Peas. Clinical Infectious Diseases 53, 26-32, doi:10.1093/cid/cir249 (2011).

49 Kwan, P. S. L. et al. Multilocus Sequence Typing Confirms Wild Birds as the Source of a Campylobacter Outbreak Associated with the Consumption of Raw Peas. Applied and Environmental Microbiology 80, 4540-4546, doi:10.1128/AEM.00537-14 (2014).

50 French, N. P. et al. Molecular Epidemiology of Campylobacter jejuni Isolates from Wild-Bird Fecal Material in Children's Playgrounds. Applied and Environmental Microbiology 75, 779-783, doi:10.1128/AEM.01979-08 (2009).

51 Schönberg-Norio, D. et al. Swimming and Campylobacter infections. Emerg Infect Dis 10, 1474-1477, doi:10.3201/eid1008.030924 (2004).

52 Chase-Topping, M., Gally, D., Low, C., Matthews, L. & Woolhouse, M. Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. Nature Reviews Microbiology 6, 904, doi:10.1038/nrmicro2029 (2008).

53 Franz, E., Schijven, J., de Roda Husman, A. M. & Blaak, H. Meta-Regression Analysis of Commensal and Pathogenic Escherichia coli Survival in Soil and Water. Environmental Science & Technology 48, 6763-6771, doi:10.1021/es501677c (2014).

54 Winkel, A., Mosquera, J., Groot Koerkamp, P. W. G., Ogink, N. W. M. & Aarnink, A. J. A. Emissions of particulate matter from animal houses in the Netherlands. Atmospheric Environment 111, 202-212, doi:https://doi.org/10.1016/j.atmosenv.2015.03.047 (2015).

55 Opsteegh, M. et al. Surveillance zoönosen in de melkgeiten-en melkschapenhouderij in 2016. doi:10.21945/RIVM-2018-0059 (2018).

56 Heuvelink, A. et al. Isolation and characterization of verocytotoxin-producing Escherichia coli O157 strains from Dutch cattle and sheep. Journal of Clinical Microbiology 36, 878-882, doi:10.1128/JCM.36.4.878-882 (1998).

57 Heuvelink, A. et al. Escherichia coli O157 infection associated with a petting zoo. Epidemiology & Infection 129, 295-302, doi:10.1017/S095026880200732X (2002).

58 Nichols, M. C. et al. Agritourism and Kidding Season: A Large Outbreak of Human Shiga Toxin-Producing Escherichia coli O157 (STEC O157) Infections Linked to a Goat Dairy Farm—Connecticut, 2016. Frontiers in Veterinary Science 8, doi:10.3389/fvets.2021.744055 (2021).

59 Friesema, I., Van De Kassteele, J., De Jager, C. M., Heuvelink, A. E. & Van Pelt, W. Geographical association between livestock density and human Shiga toxin-producing Escherichia coli O157 infections. Epidemiol. Infect. 139, 1081-1087, doi:10.1017/S0950268810002050 (2011).

60 ÓHaiseadha, C., Hynds, P. D., Fallon, U. B. & O'Dwyer, J. A geostatistical investigation of agricultural and infrastructural risk factors associated with primary verotoxigenic E. coli (VTEC) infection in the Republic of Ireland, 2008–2013. Epidemiology and Infection 145, 95-105, doi:10.1017/S095026881600193X (2017).

61 Widgren, S., Engblom, S., Emanuelson, U. & Lindberg, A. Spatio-temporal modelling of verotoxigenic Escherichia coli O157 in cattle in Sweden: exploring options for control. Veterinary research 49, 78, doi:10.1186/s13567-018-0574-2 (2018).

62 Elson, R. et al. A spatial and temporal analysis of risk factors associated with sporadic Shiga toxin-producing Escherichia coli O157 infection in England between 2009 and 2015. Epidemiol Infect 146, 1928-1939, doi:10.1017/S095026881800256X (2018).

63 Hallisey, E. et al. Transforming geographic scale: a comparison of combined population and areal weighting to other interpolation methods. International Journal of Health Geographics 16, 29, doi:10.1186/s12942-017-0102-z (2017).

64 Bos, J. F., Smit, A. B. L. & Schröder, J. J. Is agricultural intensification in The Netherlands running up to its limits? NJAS-Wageningen Journal of Life Sciences 66, 65-73, doi:10.1016/j.njas.2013.06.001 (2013).

65 Groot, M. J. & van't Hooft, K. E. The Hidden Effects of Dairy Farming on Public and Environmental Health in the Netherlands, India, Ethiopia, and Uganda, Considering the Use of Antibiotics and Other Agro-chemicals. Frontiers in Public Health 4, doi:10.3389/fpubh.2016.00012 (2016).

66 Mughini-Gras, L. et al. A statistical modelling approach for source attribution meta-analysis of sporadic infection with foodborne pathogens. Zoonoses and Public Health 69, 475-486, doi:10.1111/zph.12937 (2022).

67 Jeon, B., Saisom, T., Sasipreeyajan, J. & Luangtongkum, T. Live-Attenuated Oral Vaccines to Reduce Campylobacter Colonization in Poultry. Vaccines (Basel) 10, doi:10.3390/vaccines10050685 (2022).

68 Alexandre-Silva, G. M. et al. The hygiene hypothesis at a glance: Early exposures, immune mechanism and novel therapies. Acta Tropica 188, 16-26, doi:10.1016/j.actatropica.2018.08.032 (2018).

69 Gacesa, R. et al. Environmental factors shaping the gut microbiome in a Dutch population. Nature 604, 732-739, doi:10.1038/s41586-022-04567-7 (2022).

70 Lehtimäki, J. et al. Urbanized microbiota in infants, immune constitution, and later risk of atopic diseases. J Allergy Clin Immunol 148, 234-243, doi:10.1016/j.jaci.2020.12.621 (2021).

71 Ye, J. et al. Bacteria and bacterial rRNA genes associated with the development of colitis in IL-10(-/-) mice. Inflamm Bowel Dis 14, 1041-1050, doi:10.1002/ibd.20442 (2008).

72 Centers for Disease Control and Prevention. One Health Basics - History, <https://www.cdc.gov/onehealth/basics/history/index.html> (2022).

73 Scholthof, K.-B. G. The disease triangle: pathogens, the environment and society. Nature Reviews Microbiology 5, 152-156, doi:10.1038/nrmicro1596 (2007).

74 Mulder, A. C. et al. HEVnet: a One Health, collaborative, interdisciplinary network and sequence data repository for enhanced hepatitis E virus molecular typing, characterisation and epidemiological investigations. Eurosurveillance 24, 1800407, doi:10.2807/1560-7917.ES.2019.24.10.1800407 (2019).

75 Ricci, A. et al. Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen. Efsa j 15, e04886, doi:10.2903/j.efsa.2017.4886 (2017).

76 Wilder-Smith, A. & Osman, S. Public health emergencies of international concern: a historic overview. J Travel Med 27, 1-13, doi:10.1093/jtm/taaa227 (2020).

77 Worobey, M. et al. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. Science, 8715, doi:10.1126/science.abp8715.

78 Filter, M. et al. One Health Surveillance Codex: promoting the adoption of One Health solutions within and across European countries. One Health 12, 100233, doi:10.1016/j.onehlt.2021.100233 (2021).

79 Brown, H. L. et al. The One Health European Joint Programme (OHEJP), 2018-2022: an exemplary One Health initiative. J Med Microbiol 69, 1037-1039, doi:10.1099/jmm.0.001228 (2020).

80 Commission, E. EU4Health and Horizon 2020 projects preparing the Union for future health crises, <https://hadea.ec.europa.eu/news/eu4health-and-horizon-2020-projects-preparing-union-future-health-crises-2022-12-27_en>

81 Zucca, P. The One Health Triangle. doi:10.13140/RG.2.2.15974.14401 (2021).

APPENDIX

Summary

The emergence and spread of infectious diseases, especially zoonoses, is complex and depends on various determinants or drivers, including their interrelations at the human-animal-environment interface (HAEI). This thesis provides novel insights into the potential role of environmental transmission in the epidemiology of selected exemplary zoonotic pathogens (i.e. *Campylobacter*, *Shiga* toxin-producing *E. coli* O157, and non-typhoid *Salmonella*) that are typically considered as foodborne pathogens. This was done mainly by studying some of their characteristics in the environment and the spatial associations between their incidence among human cases and the level of exposure to local livestock density, using novel methods within an eco-epidemiological framework. In addition, this thesis explored the potential human health effects linked to extreme rainfall events and the spatial association between the urban-rural gradient and the diversity and composition of the human gut microbiome.

Climate is one of the drivers of change in the (re-)emergence of infectious diseases and affects exposure pathways of foodborne and waterborne diseases amongst others. Climate change is expected to increase the likelihood of extreme rainfall events in the Northern Hemisphere, thereby increasing the chance of pluvial flooding in urban areas. Urban pluvial flooding is often characterized by roads flooded with rain water and/or overflows of combined sewerage systems, leading to (fecal) contamination of the floodwater with several gastrointestinal and other pathogens. In **Chapter 2**, we determined the risks of two syndromes, acute gastroenteritis (AGE) and acute respiratory infection (ARI), associated with exposure to pluvial floodwater in the Netherlands. Furthermore, specific risk factors for AGE and ARI in pluvial flood-ravaged urban areas were identified. To this end, we performed a retrospective cross-sectional study in the summer of 2015 at 60 locations in the Netherlands with reported pluvial flooding. We used data from questionnaires about information on self-reported AGE and ARI symptoms after floodwater exposure: contact with floodwater was significantly associated with AGE and ARI. Risk factors for AGE were skin contact with floodwater, performing post-flooding cleaning operations and cycling through floodwater. Skin contact with floodwater and performing post-flooding cleaning operations were identified as risk factors for ARI as well. In short, we showed that extreme rainfall events, as a potential consequence of climate change, can be a driver for AGE- and ARI-causing pathogen spread in the Netherlands. As these events are increasing in frequency, more people are expected to be exposed to urban pluvial flooding, which constitutes a significant threat to public health. Therefore, there is a need for flood-proof solutions in urban development and increased awareness among stakeholders and the public about the potential health risks. Yet, future prospective studies are recommended to confirm our results.

In **Chapters 3, 4 and 5,** we focused on the role of potential environmental transmission of three zoonotic - and typically foodborne - pathogens (i.e. *Campylobacter*, *Salmonella* and STEC O157) from livestock farms in the Netherlands. In **Chapter 3** in particular we quantified *Campylobacter* prevalence and genotype diversity in surface water, as well as the relative contributions of several putative (domestic and wild) animal sources to surface water contamination with the two *Campylobacter* species of public health significance, *C. jejuni* and *C. coli*, in the Netherlands. Additionally, potential effects of local livestock (poultry, pig, ruminant) density, type of surface water (i.e. agricultural water, surface water at discharge points of wastewater treatment plants [WWTPs], and official recreational water) and season, were assessed. Surface water of 30 locations spread over six areas with either high or low density of poultry, ruminants, or pigs, were sampled once every season in 2018-2019 for each surface water type. Whole-genome sequencing (WGS) was then used for tracing the animal sources of *Campylobacter* isolates from surface water. This study revealed a widespread presence of *C. jejuni* and *C. coli* in surface water in the Netherlands. Wild birds appeared the dominant source of surface water contamination with *Campylobacter*, followed by meat-producing poultry (i.e. broilers and turkeys). Surface water can be seen as a collection vessel of strains from multiple hosts as a result of fecal pollution from different animal sources. Wild bird contribution was high among isolates from recreational waters and WWTP discharge points, and in areas with low poultry density (for *C. coli*) or high ruminant density (for *C. jejuni*). The contribution of poultry for meat production was high in areas with high poultry density, in springtime, and in agricultural waters and WWTP discharge points. We concluded that while wild birds and poultry were the main contributors to *Campylobacter* contamination in surface water in the Netherlands, their contribution varied significantly by water type, season, and local densities of poultry and ruminants, which are also the main sources of human campylobacteriosis.

Livestock-associated spatial risk factors for human STEC O157 infection (**Chapter 4**), campylobacteriosis (**Chapter 5**) and salmonellosis (**Chapter 5**) were determined using a state-of-the-art spatial analysis method. This method was developed and explained in **Chapter 4** based on reported human STEC O157 infections in the years 2007-2016. For the purpose of the analyses, hexagonal areas with different sizes (90, 50, 25, 10 km^2) were defined and used in combination with population-weighted-numbers of animals in each hexagon in the calculation of the probability of exposure to livestock (i.e. cattle, small ruminants, poultry and pigs). Results showed that exposure to small ruminants was associated with an increased incidence of human STEC O157 infections in summer. **Chapter 5** directly builds on **Chapter 4** as the newly developed method was applied to two other (typically foodborne) zoonotic pathogens: *Campylobacter* and non-typhoid *Salmonella* (NTS). Thus, we assessed whether human

salmonellosis and campylobacteriosis incidence were spatially associated with local density of small ruminants, dairy cows, veal calves, laying hens, broilers and pigs in the Netherlands in the years 2007-2019 and 2014-2019, respectively. Additionally in **Chapter 5**, we accounted for geographical coverage of the diagnostic laboratory catchment areas. Furthermore, we included serological data from the years 2006-2007 to look at possible effects of acquired immunity due to potential repeated exposure to the pathogen through the environment, which may confound the analyses based on the incidence of reported cases. Results showed that living in livestock-rich areas in the Netherlands is not a consistently significant, spatially restricted risk factor for acquiring salmonellosis or campylobacteriosis, thereby supporting current knowledge that human infections with NTS and *Campylobacter* are mainly foodborne. **Chapter 4** and **Chapter 5**, however, did not provide evidence on the exact mode of transmission. The underlying mechanisms, therefore, warrant further investigation and could offer new targets for control. The newly proposed exposure metric has potential to improve existing spatial modeling studies on infectious diseases related to livestock exposure, especially in densely populated countries like the Netherlands.

Although all microorganisms within the human gut microbiome strive towards a relatively stable equilibrium, its composition and diversity can be influenced by many factors throughout life, of which the human living environment is an example. In **Chapter 6**, we explored the potential (indirect) effects of the urban-rural gradient on the diversity and composition of the human gut microbiome in the Netherlands using data on gut microbiome and corresponding metadata from participants' questionnaires of the PIENTER-III cohort (2016-2017). Results showed a significant increase in bacterial diversity and two microbial clusters, one dominated by the genus *Blautia* and one dominated by *Bifidobacterium, Collinsella* and a*kkermansia*, to be significantly positively associated with higher agricultural land coverage. Furthermore, we identified two genera, *Barnesiella* and *Leuconostoc*, with a significant higher relative abundance in participants living in areas with a high agricultural land coverage. Thus, we observed an effect of the urban-rural gradient on diversity and composition of the adult human gut microbiome. As this study was based on an ecological design, causal inferences cannot be made and further research using individual-level data is recommended.

Finally, a general discussion of the findings of the present thesis is given in **Chapter 7**, putting the results of this thesis in a One Health perspective, followed by some concluding remarks. This thesis showed that we cannot ignore the role of the environment in mediating the transmission of (foodborne) zoonotic pathogens from livestock and wildlife, and more holistically in shaping the human gut microbiome. However, evidence about specific routes of transmission (e.g. air-borne, via soil or

water) and their relative importance, was not provided. To study all interactions between different drivers of zoonotic pathogens, we recommend future studies to focus on pathogen transmission and their related transmission pathways to be able to look beyond the three drivers of change included in the One Health disease triangle (humans, animals, environment). This can be visualized by including a white triangle based on the Gestalt idea that the sum of visual perception is more than its parts. This triangle illustrates the interactions to include to study zoonotic pathogens and provide a complete picture of how they may influence one another. This visualization can be considered as a first step towards further unraveling the complexity of zoonotic pathogens by showing "the complete picture" or at least an attempt to do so, to, in the end, achieve optimal health outcomes.

Samenvatting

Het ontstaan en de verspreiding van infectieziekten, en met name zoönosen, is complex. Het hangt af van verschillende determinanten, drijfveren en al hun onderlinge relaties op het vlak van mens, dier en milieu. Dit proefschrift biedt nieuwe inzichten in de mogelijke rol van milieutransmissie in de epidemiologie van geselecteerde voedsel-overdraagbare zoönotische pathogenen (*Campylobacter*, *Shiga*-toxin producing *E. coli* O157, en non-typhoidale *Salmonella*). Dit werd voornamelijk onderzocht door enkele kenmerken van deze pathogenen in de omgeving en de ruimtelijke verbanden tussen humane incidentie en de mate van blootstelling aan landbouwhuisdieren te bestuderen. Hiervoor werden nieuwe methoden binnen een eco-epidemiologisch kader gebruikt. Daarnaast is de ruimtelijke associatie tussen de stedelijk-rurale gradiënt en de diversiteit en samenstelling van het darm microbioom onderzocht.

Het klimaat is een van de drijfveren van verandering bij het (opnieuw) opduiken van infectieziekten. Het beïnvloedt onder meer de blootstellingsroutes van voedsel- en wateroverdraagbare ziekten. De verwachting is dat het noordelijk halfrond vaker te maken krijgt met grote hoeveelheden regen door klimaatverandering, waarbij de kans op overstromingen in stedelijk gebied toeneemt. Wanneer het rioolstelsel dit water niet af kan voeren, kunnen straten (deels) onder water komen te staan. Hierdoor kan verdund rioolwater op straat terechtkomen. Dit water is verontreinigd met verschillende gastro-intestinale en/of respiratoire ziekteverwekkers en kan vooral in stedelijke gebieden tot overlast leiden. In **Hoofdstuk 2** bepaalden we de risico's op het krijgen van twee ziektebeelden: acute gastro-enteritis (AGE) en acute respiratoire infectie (ARI), geassocieerd met contact met overstromingswater. Verder werden specifieke risicofactoren voor AGE en ARI geïdentificeerd in gebieden waar overstroming plaatsvond. Hiervoor hebben we in de zomer van 2015 een retrospectief cross-sectioneel onderzoek uitgevoerd op 60 locaties in Nederland met gemelde wateroverlast. Om dit te kunnen doen, gebruikten we gegevens uit vragenlijsten met informatie over zelf gerapporteerde AGE- en ARI-symptomen na blootstelling aan overstromingswater. Contact met overstromingswater was significant geassocieerd met AGE en ARI. Risicofactoren voor AGE waren huidcontact met overstromingswater, het schoonmaken na overstromingen en fietsen door overstromingswater. Daarnaast werden huidcontact met overstromingswater en schoonmaken na overstromingen ook geïdentificeerd als risicofactoren voor ARI. Kortom, we hebben laten zien dat klimaatverandering en de bijbehorende toename van extreme regenval in Nederland een aanjager kan zijn voor de verspreiding van infectieziekten. Naarmate deze gebeurtenissen vaker voorkomen, worden er elk jaar meer inwoners blootgesteld aan stedelijke wateroverlast. Dit leidt tot een toename van de risico's voor de volksgezondheid. Daarom is er behoefte aan overstromingsbestendige oplossingen in stedelijke ontwikkeling en een groter bewustzijn bij belanghebbenden en het publiek

over de mogelijke gezondheidsrisico's. Toekomstige prospectieve studies worden aanbevolen om onze resultaten te bevestigen.

In **Hoofdstuk 3, 4 en 5** hebben we ons gericht op de mogelijke rol van milieu transmissie van zoönotische pathogenen (*Campylobacter*, *Salmonella*, STEC O157) vanuit de veehouderij. Deze rol wordt steeds meer erkend voor typisch voedseloverdraagbare pathogenen zoals de bovengenoemde drie. In **Hoofdstuk 3** hebben we de *Campylobacter* prevalentie en genotype diversiteit in oppervlaktewater gekwantificeerd, evenals de relatieve bijdragen van verschillende mogelijke (gedomesticeerde en wilde) dierlijke bronnen aan oppervlaktewaterverontreiniging met de twee *Campylobacter* soorten, *C. jejuni* en *C. coli* die van belang zijn voor de volksgezondheid in Nederland. Daarnaast werden mogelijke effecten van lokale landbouwhuisdierdichtheden (pluimvee, varkens, herkauwers), type oppervlaktewater (agrarische sloten, oppervlaktewater bij lozingspunten van afvalwaterzuiveringsinstallaties [AWZI's] en officieel recreatiewater) en seizoen beoordeeld. Hiervoor zijn in 2018-2019 op 30 locaties in Nederland monsters genomen van oppervlaktewater. Deze locaties waren verspreid over zes gebieden met een hoge of lage dichtheid aan pluimvee, varkens en herkauwers. Er werd één monster per seizoen genomen voor elk type oppervlaktewater. Whole-genome sequencing (WGS) werd gebruikt voor het opsporen van de dierlijke bronnen van *Campylobacter* isolaten in oppervlakte water. Dit hoofdstuk liet zien dat *C. jejuni* en *C. coli* wijdverspreid aanwezig zijn in Nederlands oppervlaktewater. Wilde vogels waren de dominante bron van deze oppervlaktewaterverontreiniging, gevolgd door pluimvee bestemd voor de vleesproductie. Oppervlaktewater kan worden gezien als een verzamelbak van stammen afkomstig van meerdere gastheren als gevolg van fecale vervuiling door verschillende dierlijke bronnen. De bijdrage van wilde vogels was hoog onder recreatiewater isolaten, isolaten afkomstig van AWZI lozingspunten en in gebieden met een lage dichtheid aan pluimvee (*C. coli*) of een hoge dichtheid aan herkauwers (*C. jejuni*). De bijdrage van pluimvee bestemd voor de vleesproductie was hoog in gebieden met een hoge pluimvee dichtheid, in het voorjaar, in agrarische sloten en bij AWZI lozingspunten. Hoewel wilde vogels en pluimvee de grootste bijdrage leverden aan *Campylobacter* besmetting van oppervlaktewater, verschilde hun bijdrage aanzienlijk per watertype, seizoen en lokale dichtheden van pluimvee en herkauwers.

In **Hoofdstuk 4** en **Hoofdstuk 5** werden met landbouwhuisdieren-geassocieerde ruimtelijke risicofactoren voor humane STEC O157 (**Hoofdstuk 4**), campylobacteriose (**Hoofdstuk 5**) en salmonellose (**Hoofdstuk 5**) bepaald met behulp van een nieuw ontwikkelde ruimtelijke analyse methode. Deze methode is ontwikkeld en beschreven in **Hoofdstuk 4** op basis van gerapporteerde humane STEC O157 infecties in de jaren 2007-2016. Hiervoor zijn hexagonen met verschillende groottes (90, 50, 25, 10 km^2) gebruikt in combinatie met de populatie-gewogen aantallen die-

ren in de berekening van de kans op blootstelling aan landbouwhuisdieren (rundvee, kleine herkauwers, pluimvee en varkens). Dit hoofdstuk laat zien dat blootstelling aan kleine herkauwers geassocieerd is met een verhoogde STEC O157 incidentie in de zomer. **Hoofdstuk 5** bouwt rechtstreeks voort op **Hoofdstuk 4**. Hier hebben we de nieuw ontwikkelde methode toegepast op twee andere voedseloverdraagbare zoönotische pathogenen: *Campylobacter* en non-typhoidale *Salmonella* (NTS). We hebben hiermee beoordeeld of de incidentie van campylobacteriose en salmonellose bij de mens ruimtelijk geassocieerd was met de lokale dichtheid van kleine herkauwers, melkkoeien, vleeskalveren, leghennen, vleeskuikens en varkens in Nederland in de jaren 2007-2019 en 2014-2019. Daarnaast hebben we in **Hoofdstuk 5** rekening gehouden met de geografische dekking van de diagnostische laboratoria. Ook hebben we serologische gegevens uit de jaren 2006-2007 geanalyseerd om te kijken naar de mogelijke effecten van verworven immuniteit. Dit werd gedaan, omdat herhaaldelijke blootstelling aan het pathogeen via het milieu mogelijk de analyses gebaseerd op de incidentie kan beïnvloeden door middel van confounding. Resultaten toonden aan dat wonen in gebieden met veel landbouwhuisdieren in Nederland geen consequent significante, ruimtelijk beperkte risicofactor is voor het krijgen van campylobacteriose of salmonellose. Dit ondersteunt de huidige kennis dat humane infecties met *Campylobacter* en *Salmonella* voornamelijk via voedsel worden overgedragen. **Hoofdstuk 4** en **Hoofdstuk 5** gaven geen informatie over de wijze van transmissie. Daarom is het noodzakelijk om vervolgonderzoek te doen naar onderliggende mechanismen die mogelijk kunnen leiden tot nieuwe controle maatregelen. De nieuw voorgestelde blootstellingsmaat kan bestaande ruimtelijke modelleringsstudies over infectieziekten verbeteren die verband houden met blootstelling aan landbouwhuisdieren, vooral in dichtbevolkte gebieden zoals Nederland.

Hoewel alle individuele micro-organismen in het menselijke darm microbioom streven naar een relatief stabiel evenwicht, kan de diversiteit en samenstelling gedurende het hele leven door vele factoren worden beïnvloed. Een voorbeeld van zo'n factor is de menselijke leefomgeving. In **Hoofdstuk 6** onderzochten we de mogelijke (indirecte) effecten van de gradiënt tussen stad en platteland op de diversiteit en samenstelling van het menselijke darm microbioom in Nederland met behulp van gegevens over het darm microbioom en bijbehorende metadata van deelnemersvragenlijsten van het PIENTER-III cohort (2016-2017). De resultaten toonden een significante toename van de bacteriële diversiteit en twee microbiële clusters die significant positief geassocieerd zijn met een hoger percentage landbouwgrond in de omgeving van de deelnemer: één gedomineerd door het genus *Blautia* en één gedomineerd door *Bifidobacterium*, *Collinsella* en *Akkermansia*. Verder identificeerden we twee genera: *Barnesiella* en *Leuconostoc* die significant meer aanwezig waren in het darm microbioom van deelnemers die in gebieden woonden met een hoger percentage

landbouwgrond in hun omgeving. Er is dus een effect waargenomen van de gradiënt tussen stad en platteland op de diversiteit en samenstelling van het darm microbioom van volwassenen. Aangezien deze studie gebaseerd was op een ecologisch ontwerp, is causale gevolgtrekking beperkt en wordt verder onderzoek met gegevens op individueel niveau aanbevolen.

Ten slotte worden in **Hoofdstuk 7** de bevindingen van dit proefschrift bediscussieerd, en plaatsten we de resultaten van dit proefschrift in een One Health perspectief. Samenvattend laat dit proefschrift zien dat we de rol van het milieu bij de transmissie van (door voedsel overgedragen) ziekteverwekkers afkomstig van landbouwhuisdieren en wilde dieren niet zomaar kunnen negeren. Daarnaast lijkt de leefomgeving ook een rol te spelen in het vormen van het darm microbioom. Bewijzen over specifieke transmissieroutes (bijvoorbeeld via de lucht, via bodem of water) en hun relatieve belang zijn echter nog niet geleverd. Om alle interacties tussen de verschillende drijfveren van specifieke zoönotische pathogenen te bestuderen, raden we daarom aan om verder te kijken dan de drie drijvende krachten achter verandering die zijn opgenomen in de One Health driehoek (mensen, dieren, milieu). Dit kan bereikt worden door in toekomstige studies te focussen op transmissie van deze pathogenen en de bijbehorende transmissieroutes wat gevisualiseerd kan worden door een witte driehoek te includeren in de bestaande driehoek. Deze witte driehoek is gebaseerd op het Gestalt-idee dat de som van de visuele perceptie meer is dan zijn onderdelen. Deze vernieuwde driehoek illustreert alle interacties die moeten worden meegenomen om zoönotische pathogenen te bestuderen en om een compleet beeld te krijgen van alle interacties die elkaar beïnvloeden. Daarom kan deze visualisatie worden beschouwd als een eerste stap om de complexiteit van zoönotische pathogenen verder te ontrafelen door “het complete plaatje” te laten zien, of op zijn minst een poging daartoe, om uiteindelijk tot optimale gezondheidsresultaten te komen.

Affiliations of co-authors

Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven:
Eelco Franz, Lapo Mughini-Gras, Roan Pijnacker, Jan van de Kassteele, Wilfrid van Pelt, Sharona de Rijk, Claudia Coipan, Franciska M. Schets, Hetty Blaak, Sara Lynn Blanken, David Boverhoff, Marie-Lee Odendaal, Susana Fuentes, Jolanda Kool, Sudarshan A. Shetty, Liza Tymchenko

Department of Infectious Diseases and Immunology (I&I), Utrecht University, Utrecht: Birgitta Duim, Linda van der Graaf-van Bloois, Jaap Wagenaar, Aldert L. Zomer

Department of Medical Microbiology and Infection Prevention, Virology and Immunology Research Group, University Medical Center Groningen, Groningen:
Sudarshan A. Shetty

Department Population Health Sciences, Institute for Risk Assessment Sciences (IRAS), University Utrecht, Utrecht:
Dick J.J. Heederik, Lapo Mughini-Gras, Moyke Versluis

Sanitas Water, Zeist, Utrecht:
Heleen de Man

Wageningen Bioveterinary Research (WBVR), Wageningen University & Research (WUR), Lelystad:
Miriam Koene, Kees Veldman, Jaap A. Wagenaar

Wageningen Environmental Research (WENR), Wageningen University & Research (WUR), Wageningen:
Ralph Buij, Gerard Müskens

WHO Collaborating Centre for *Campylobacter* / OIE Reference Laboratory for Campylobacteriosis, Utrecht:
Birgitta Duim, Linda van der Graaf-van Bloois, Jaap A. Wagenaar, Aldert L. Zomer

Dankwoord

Tja, dan heb je een heel proefschrift geschreven.. en dan nog blijft die eerste zin van dit stuk tekst een uitdaging. Want ook dit wil ik graag goed doen, dat heb je als je perfectionistisch bent. Gelukkig heb ik al het werk van de afgelopen vijf jaar niet alleen hoeven doen en is gebleken dat deze eigenschap, zeker in het doen van onderzoek, helemaal niet zo verkeerd is. Ondanks dat epidemiologie niet het vakgebied is waarin ik ben opgeleid, heeft het een plekje in mijn hart veroverd. Want ook voor infectieziekten en zeker zoönosen blijkt dat het milieu een belangrijke rol speelt. Er zijn dan ook veel mensen geweest die meegedacht en meegeholpen hebben om mijn kennis van het milieu, de ecologie en de ruimtelijke analyses te kunnen verweven in de epidemiologische onderzoeken van dit proefschrift. Hen wil ik dan ook graag bedanken.

Eelco, ik begin toch echt met jou. Als mijn copromotor, maar ook als mijn leidinggevende had je eigenlijk twee petten op. Zo heb ik het echter nooit ervaren. Altijd stond je deur voor me open (en die heb ik behoorlijk plat gelopen), even sparren over wat analyses, uitkomsten, of gewoon even vragen hoe het ging. We zijn rond dezelfde tijd gestart in de GEZ groep en na een tijdje heb jij er samen met mijn andere copromotor Lapo voor gezorgd dat er budget was om een eigen plan te schrijven voor een PhD. Na het afleggen van een paar meter van jouw kantoor naar het mijne kwam de vraag: “of ik daarin wel interesse had?”. Natuurlijk! En wat ben ik blij dat ik die kans met beide handen heb aangegrepen. Samen met Lapo schreven we het plan en vol vertrouwen kon ik beginnen aan deze nieuwe uitdaging. Bedankt voor deze kans en alles wat ik van je heb geleerd. Je dacht altijd een stapje verder en groter en tilde me een niveau hoger.

Lapo, ook jou als copromotor wil ik ontzettend bedanken voor deze kans. Het was een genoegen om met je samen te werken, want wat betreft statistische analyses en met name bron attributie analyses was je een ware encyclopedie waar ik ontzettend veel van heb geleerd. Naast je vakkennis waren je verhalen over thuis altijd heerlijk om aan te horen; met drie dochters heb je heel wat te stellen. Als ware Italiaan nam je me samen met een andere collega-PhD in Italië mee op sleeptouw tijdens een congres. Want die watertoren daar, die moesten we gezien hebben. Ik beschouwde je altijd als een alleskunner, maar tijdens dit tripje leerde ik dat ook jij gewoon mens bent. Die trappen van de watertoren waren toch wat steiler dan wat je je van vroeger herinnerde. Ook werd Eelco daar omgedoopt tot “legmen”. Wat heb ik naast alle serieuze bezigheden met jullie gelachen. Dank daarvoor!

Ook mijn promotor Dick wil ik graag bedanken. Vol enthousiasme werkte en dacht je mee over het PhD plan en elke keer als we elkaar spraken was er wel weer stof tot nadenken (en natuurlijk meer werk). Bedankt voor al het meedenken en je feedback op verschillende delen van dit proefschrift. Je oprechte interesse en belangstelling zorgden ervoor dat ik ontzettend gemotiveerd was om dit project tot een goed einde te brengen. Graag bedank ik ook alle andere coauteurs voor het meedenken en meeschrijven aan de artikelen. En natuurlijk ook de leden van de commissie voor hun bereidheid om mijn proefschrift te beoordelen.

Jan, jou wil ik ontzettend bedanken voor het meedenken over de toegepaste ruimtelijke analyses in dit proefschrift. Zonder jouw kritische blik en hulp had ik wellicht mijn laptop echt een keer uit het raam gegooid, want na de zoveelste INLA error kwam jouw hulp dan als geroepen. Meerdere keren zaten we samen, zowel fysiek als online, naar het scherm te staren en ons weer eens af te vragen hoe we dit nou eens zouden kunnen oplossen. Gelukkig had je altijd goede ideeën en werkte het (meestal) daarna weer. Dank voor je geduld en oprechte interesse in alles wat ik deed. Zo fijn om te weten dat ik altijd aan kon en kan kloppen als mijn kennis van R of de statistiek even tekort schiet.

En dan kom ik aan bij mijn paranimfen en GEZ collega's Lola en Roan. Ik wil jullie beiden ontzettend bedanken voor alles wat jullie voor mij betekend hebben in de afgelopen jaren. Lola, de lunchwandelingen en het fitnessen samen met jou zorgde voor de welkome afwisseling tussen serieuze werk-gerelateerde gesprekken en de luchtige verhalen over onze zoontjes. Gedurende de jaren is er een waardevolle vriendschap ontstaan, waarbij ik zelfs functioneer als spinnenvanger en levend hondenschild. Roan, ook jou wil ik graag bedanken voor alle kopjes koffie en het "even" meekijken en denken over een bepaalde analyse. Vaak werd dit "even" nog wat langer, want "oh ja, hoe zit dat eigenlijk?". Het moest altijd even uitgezocht worden. Ook over de kinderen raken we niet uitgepraat. Bedankt voor je altijd luisterende oor en het doorbreken van de dag met een online kopje koffie tijdens het thuiswerken in de COVID-19 lockdowns.

Natuurlijk wil ik ook graag alle andere collega's van de GEZ groep bedanken. Velen dachten mee bij bepaalde onderdelen van dit proefschrift en de oprechte interesse in mijn onderzoek zal mij altijd bijblijven. Twee collega's wil ik in het bijzonder nog bedanken. Mari-Lee, zeker het afgelopen jaar hebben we elkaar regelmatig gesproken via Teams: "Hey, jij ook nog aan het werk?' – 22:08…. Liep ik ergens in vast, dan kon ik je altijd even een berichtje doen en kreeg ik altijd iets motiverends terug, omgedraaid natuurlijk precies hetzelfde. Het was niet altijd makkelijk met een kleintje die vaak ziek was en het daardoor werken op tijdstippen waarop de meeste

mensen vrij zijn. Maar onze gesprekken waren altijd een opsteker (en zeker de PhD comics). Agnetha, ook al werk je niet meer bij GEZ, ik wil je toch bedanken voor alle waardevolle tips. Als bijna buren op de gang was het altijd makkelijk om even bij je binnen te lopen met een vraag en (bijna) altijd liep ik je deur weer uit met iets waar ik mee verder kon.

Inge ook jou wil ik graag bedanken. Als coördinator van de PhDget2gether groep heb ik samen met jou iets ontzettend moois en waardevols neer kunnen zetten voor alle PhD kandidaten die binnen epi rondlopen. Naast het bijeen brengen van alle PhD's tijdens de COVID-19 periode is het ons gelukt om elk half jaar opnieuw een afwisselend programma te bieden van inhoudelijke presentaties en het delen van PhD ervaringen. Sinds driekwart jaar zijn we inmiddels directe collega's door mijn overstap van GEZ naar RES en ik kijk uit naar verdere samenwerkingen met jou. Graag bedank ik ook iedereen uit het tuberculose-team binnen RES, waar ik sinds april 2022 deel van uit maak. Altijd waren jullie geïnteresseerd in hoe het ging met mijn proefschrift en hoe het met mij als persoon eigenlijk ging. Bedankt voor jullie warme onthaal en steun tijdens de afrondende fase! Rianne, jouw naam kan in dit rijtje zeker niet ontbreken. Als mijn leidinggevende binnen RES was je altijd geïnteresseerd in hoe het ervoor stond. Jouw aanmoedigingen, en de flexibiliteit en vrijheid die ik van je kreeg, hebben er mede aan bijgedragen dat de afronding daadwerkelijk volgens schema is verlopen.

Evelien, wat een topvriendin ben jij! Bedankt voor alle avonden sushi waarbij ik even mijn ei kwijt kon. Druk of niet, als er iets was, was jij er altijd en maar een appje verder weg. Ontzettend fijn ook dat je het Nederlandse deel van mijn proefschrift wilde nakijken, wat moet ik toch zonder je. Ook alle andere vrienden, vriendinnen en familie wil ik graag bedanken voor jullie interesse, gezellige afleiding en stabiele basis.

In het bijzonder wil ik graag nog mijn ouders en zus Marit bedanken voor alle steun en het vertrouwen in mij. Bedankt voor alle lieve berichtjes, kaartjes, appjes, knuffels, het doorlopen van het Nederlandse stuk uit dit proefschrift en nog veel meer.. Marit, bedankt voor het zijn van je heerlijke zelf, jouw blik werkt heerlijk ontnuchterend en zet me vaak weer met beide benen op de grond.

Wesley, wat zou ik moeten zonder jou! Tijdens deze rollercoaster was je mijn rots in de branding. Schoot ik even in de stress, was jij mijn veilige haven. Samen maakten we de afgelopen vijf jaar heel wat mee. We kregen onze hond Pip, die tijdens COVID-19 regelmatig online haar entree maakte. We trouwden, kregen onze lieverd en ondeugd Timo na een spannende zwangerschap en we verhuisden naar ons droomhuis aan de bosrand. Het was niet altijd makkelijk om werk te combineren met het vele ziek

zijn van Timo, maar wat ben ik trots op ons! We vonden altijd samen een oplossing. Afgelopen november kluste jij alleen in ons nieuwe huis, zodat ik alle tijd had voor het afronden van mijn proefschrift en we toch over konden. Als dat geen liefde is.. bedankt dat je me altijd steunt en er altijd voor me bent! Ik ga verder met het uitpakken van de verhuisdozen als dit afgerond is, beloofd! En Timo, jij hebt het nog niet door, maar ook jou wil ik bedanken. Jouw brede lach en harde uitroep "MAMA" als ik je na een dag hard werken weer op kom halen of thuis kom, samen met een ontzettend dikke knuffel, zorgen ervoor dat werk gelijk naar de achtergrond verdwijnt. Wat wil je nou nog meer dan thuiskomen in zo'n warm bad? Ik hou van jullie!

Curriculum Vitae

Annemieke (Christine) Mulder was born on August 31st, 1991 in 's-Hertogenbosch, the Netherlands. After graduating from pre-university education (VWO) at the Pierson College in 's-Hertogenbosch, she studied Forest and Nature Conservation at Wageningen University & Research (WUR) in Wageningen where she also obtained her Bachelor's degree in 2012. While continuing the master Forest and Nature Conservation and spending several months in the field to gather data for her master thesis about both deer and rabbit density as determinants for tick density, she also pursued a master in Geo-information Science at Wageningen University. For this master, she traveled to Kenya for her master thesis about agricultural land use patterns and malaria vector abundance at Rusinga island, where she collaborated with the International Centre for Insect Physiology and Ecology (ICIPE). Both masters where finalized in 2015, after the successful completion of a combined internship regarding both master programs, where she developed a GIS tool to estimate the risks of acquiring a tick bite for visitors of nature areas in the Netherlands at the department for Zoonoses and Environmental Microbiology of the National Institute for Public Health and the Environment (RIVM). Upon completion she started working for the European consultancy firm TAUW as an ecology and GIS trainee. There, she learned that research suited her more and therefore she started to work as junior epidemiologist at the department for Epidemiology and Surveillance - Enteric Infections & Zoonoses in 2016. Here, she got the chance to write her own project plan for a PhD position, which she started in 2018 in collaboration with Utrecht University. During this position, she was nominated for the Jim van Steenbergen price in 2019 for her work on urban flooding and the development of disease syndromes after contact with floodwater. In 2020, she won the price for best poster presentation at the One Health European Joint Programme annual meeting which covered the topic of *Campylobacter* in Dutch surface water. During the COVID-19 pandemic, she worked for two months in the COVID-19 surveillance R-team within the RIVM. Furthermore, she coordinated the PhDget2gether within the center for Epidemiology and Surveillance of Infectious Diseases during her PhD. She continues to work as an epidemiologist at the RIVM at the department for Epidemiology and Surveillance – respiratory infections.

List of publications

Part of this thesis:

Mulder, A.C., Mughini-Gras, L., van de Kassteele, J., Blanken, S. L., Pijnacker, R., Franz, E. Livestock-associated spatial risk factors for human salmonellosis and campylobacteriosis. Revision submitted to: *Zoonoses and Public Health (2022).*

Mulder, A. C., Franz, E., de Rijk, S., Versluis, M. A. J., Coipan, C., Buij, R., Müskens, G., Koene, M., Pijnacker, R., Duim, B., Bloois, L. V. G., Veldman, K., Wagenaar, J. A., Zomer, A. L., Schets, F. M., Blaak, H., & Mughini-Gras, L. (2020). Tracing the animal sources of surface water contamination with *Campylobacter jejuni* and *Campylobacter coli*. *Water research*, 187, 116421.

Mulder, A. C., van de Kassteele, J., Heederik, D., Pijnacker, R., Mughini-Gras, L., & Franz, E. (2020). Spatial Effects of Livestock Farming on Human Infections with Shiga Toxin-Producing Escherichia coli O157 in Small but Densely Populated Regions: The Case of the Netherlands. *GeoHealth*, 4, e2020GH000276.

Mulder, A.C., Pijnacker, R., de Man, H., van de Kassteele, J., van Pelt, W., Mughini-Gras, L., Franz, E. (2019). "Sickenin' in the rain" – increased risk of gastrointestinal and respiratory infections after urban pluvial flooding in a population-based cross-sectional study in the Netherlands. *BMC Infectious Diseases,* 19, 377.

Outside this thesis:

Mughini-Gras, L., Pijnacker, R., Coipan, C., **Mulder, A. C.**, Veludo, A. F., de Rijk, S., van Hoek, A. H. A. M., Buij, R., Muskens, G., Koene, M. Veldman, K., Duim, B., van der Graaf-van Bloois, L., van der Weijden, C., Kuiling, S., Verbruggen, A., van der Giessen, J., Opsteegh, M., van der Voort, M., Castelijn, G. A. A., Schets, F. M., Blaak, H., Wagenaar, J. A., Zomer, A. L., Franz, E. (2021) Sources and transmission routes of campylobacteriosis: A combined analysis of genome and exposure data, *Journal of Infection*, 82 (2), 216-226.

Mulder, A. C., Kroneman, A., Franz, E., Vennema, H., Tulen, A. D., Takkinen, J., ... & Adlhoch, C. (2019). HEVnet: a One Health, collaborative, interdisciplinary network and sequence data repository for enhanced hepatitis E virus molecular typing, characterisation and epidemiological investigations. *Eurosurveillance*, 24(10), 1800407.

Brandwagt, D., van den Wijngaard, C., Tulen, A.D., **Mulder, A. C.**, Hofhuis, A., Jacobs, R., Heck, M., Verbruggen, A., van den Kerkhof, H., Slegers-Fitz-James, I., Mughini-Gras, L., Franz, E. (2018). Outbreak of *Salmonella* Bovismorbificans associated with the consumption of uncooked ham products, the Netherlands, 2016 to 2017. *Eurosurveillance*, 23(1), 1700335.

Montizaan, M. & **Mulder, A. C.** (2017). Zieke hazen in het veld. *Vakblad Natuur bos landschap*, april 2017 (134): 21-23.

Mulder, A. C., Snabilie, M., & Braks, M. A. (2016). 26. From guessing to GIS-ing: empowering land managers. *Ecology and prevention of Lyme borreliosis*, 4, 373.

Braks, M. A. H., **Mulder, A. C**., Swart, A. & Wint, W. (2016). 25. Grasping risk mapping. *Ecology and prevention of Lyme borreliosis*, 4, 351.