

Mathematical modelling approaches to
understanding the spread of
Campylobacter in broiler flocks.



Thomas Rawson
Pembroke College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Hilary 2021

Acknowledgements

Many thanks are due to the multitude of people who have graciously made this thesis possible. Firstly, I must thank my supervisors: Marian Stamp Dawkins, for her unwavering enthusiasm and insight, and Mike Bonsall, for his steadfast support and leadership. Both have kept me confident and focused, and have made me a better scientist through their example. Thanks also to Fran Colles for generously offering her time, expertise, and kindness throughout, and to Caroline Bergmann for her inspiration and guidance.

Many thanks are due to Nina Stamps, Jo Burch, Lorraine Damerell, and Heather Green, for making the journey so painless. Heartfelt thanks to Peter Chamberlain and Calvin Smith for equipping me with the skill set and confidence to make this possible.

The time spent on this work has been a delight, thanks to the friendship and collaboration of my research group. Special thanks to Tom Brewer, Rob Paton, Katie Heath, Beth Raine, Chris Terry, and Sean Hackett, for the laughter, the lessons, the camaraderie, and the walk.

My most sincere thanks to the ever-supportive Tom Snowden, Claire Collins, Jack Alcock, and Josie Dodd. Their friendship and support has provided both joy and refuge throughout the journey.

My gratitude always to Sophie Daniels, for walking every step with me, and a few of them for me. You can't quit.

Thanks to my mum, Julie, my step-dad, Max, and my dad, Rauf. Thanks to Judy Spaniel for being a good dog. My eternal gratitude to Barry and Caroline, whose patience, guidance, shelter, and encouragement enabled everything. Thanks for the first push.

And to Georgia, for giving meaning to all of it.

Format of Chapters

This thesis is submitted as an integrated thesis, with Chapters 2-5 all free-standing manuscripts. Chapters 2 and 3 have been published within *Frontiers in Microbiology*, while Chapters 4 and 5 are currently under review at *Veterinary Microbiology* and *Poultry Science* respectively. Each has been formatted to meet the requirements of their respective journals, and are presented as published herein. The supplementary materials for each chapter are provided in Appendices A-D. All associated code is made available within Code Availability sections at the end of each chapter.

Funding

This thesis was funded through an Engineering and Physical Sciences Research Council (EPSRC) (<https://epsrc.ukri.org/>) Systems Biology studentship award (EP/G03706X/1).

Statement of Authorship

Each paper that constitutes the main chapters of this thesis were co-authored between myself, my supervisors (Michael B. Bonsall and Marian Stamp Dawkins), and others who provided helpful interpretation of results. In each instance I was the first and corresponding author. Below I list the authors and ‘Author Contributions’ statements associated with each chapter.

- **Chapter 2: Thomas Rawson, Marian Stamp Dawkins, and Michael B. Bonsall** - TR designed the model and performed the presented case studies. TR wrote the manuscript. MBB and MSD supervised the project and proofread the manuscript.
- **Chapter 3: Thomas Rawson, Robert Stephen Paton, Frances M. Colles, Martin C.J. Maiden, Marian Stamp Dawkins, and Michael B. Bonsall** - FMC collected the data. TR, RSP, MCJM, and MBB conceived the study. TR and RSP built the models and wrote all associated code. TR wrote the manuscript. MSD, FMC, and MBB supervised the project. All authors reviewed the manuscript.
- **Chapter 4: Thomas Rawson, Frances M. Colles, J. Christopher D. Terry, and Michael B. Bonsall** - FMC collected the data. TR, JCDDT, and MBB conceived the study. TR built the models and wrote all associated code. TR wrote the manuscript. FMC, JCDDT, and MBB supervised the project. All authors aided in interpretation of results and reviewed the manuscript.
- **Chapter 5: Thomas Rawson, Frances M. Colles, Adrian L. Smith, Marian Stamp Dawkins, and Michael B. Bonsall** - FMC performed the microbiological sampling. All authors interpreted the results. TR, MBB, and MSD conceived the study. TR built the models and wrote all associated code. TR wrote the manuscript. MSD, FMC, and MBB supervised the project. All authors reviewed the manuscript.

Chapters 1 and 6 were written by myself with comments from Marian Stamp Dawkins, Michael B. Bonsall, and Barry S. Robbins.

Publications

Other work was performed within the funding period that was not within the scope of this thesis. Below I list all publications and submissions supported by this studentship.

Published

Rawson, T., Dawkins, M. S. & Bonsall, M. B. (2019). “A mathematical model of *Campylobacter* dynamics within a broiler flock.” In: *Frontiers in Microbiology*, 10 p.1940, DOI: 10.3389/fmicb.2019.01940

Rawson, T., Wilkins, K. E., & Bonsall, M. B. (2020). “Optimal control approaches for combining medicines and mosquito control in tackling dengue.” In: *Royal Society Open Science*, 7.4, DOI: 10.1098/rsos.181843

Rawson, T., Paton, R. S., Colles, F. M., Maiden, M. C. J., Dawkins, M. S. & Bonsall, M. B. (2020). “A mathematical modeling approach to uncover factors influencing the spread of *Campylobacter* in a flock of chickens.” In: *Frontiers in Microbiology*, 11 p.2481, DOI: 10.3389/fmicb.2020.576646

Rawson, T., Brewer, T., Veltcheva, D., Huntingford, C. & Bonsall, M. B. (2020). “How and when to end the COVID-19 lockdown: an optimisation approach” In: *Frontiers in Public Health*, 8.262, DOI: 10.3389/fpubh.2020.00262

Rawson, T., Huntingford, C. & Bonsall, M. B. (2020). “Temporary ‘circuit breaker’ lockdowns could effectively delay a COVID-19 second wave infection peak to early spring.” In: *Frontiers in Public Health*, 8.899, DOI: 10.3389/fpubh.2020.00262

Amit, A. M. L., Pepito V. C. F., Gutierrez, B. & **Rawson, T.** (2021). “Tracking changes in reporting of epidemiological data during the COVID-19 pandemic in Southeast Asia: an observational study during the first wave” Accepted in: *Frontiers in Public Health*, Provisional DOI: 10.3389/fpubh.2021.662842

Under Review

Rawson, T., Colles, F. M., Smith, A. L., Dawkins, M. S. & Bonsall, M. B. (2021). “Can good broiler flock welfare prevent colonization by *Campylobacter*?” In review: *Poultry Science*, Pre-print DOI: 10.1101/2021.03.30.437710

Rawson, T., Colles, F. M., Terry, J. C. D. & Bonsall, M. B. (2021). “Mechanisms of biodiversity between *Campylobacter* sequence types in a flock of broiler-breeder chickens.” In review: *Proceedings of the Royal Society B: Biological Sciences*, Pre-print DOI: 10.1101/2021.04.14.439797

Rawson, T. “A hierarchical Bayesian quantitative microbiological risk assessment model for *Salmonella* in the sheep meat food chain.” In review: *Risk Analysis*

Bonsall, M. B., Huntingford C. & **Rawson, T.** (2021). “Optimal time to return to normality: parallel use of COVID-19 vaccines and circuit breakers” In review: *Journal of Mathematical Analysis and Applications*, Pre-print DOI: 10.1101/2021.02.01.21250877

Huntingford C., **Rawson, T.**, & Bonsall, M. B. “Optimal distribution of COVID-19 vaccine sharing under extensive travel between two nations.” In review: *Frontiers in Public Health*

Abstract

Campylobacter spp. are one of the leading causes of human gastroenteritis globally, with the primary route of transmission being poultry meat. Flocks of broilers, chickens grown for their meat, will rapidly become colonised by *Campylobacter* on the farm, and this infection persists throughout the food chain. Recent attempts to prevent the spread of *Campylobacter* in broiler flocks have focused on biosecurity measures on farms to reduce the chance of initial infection. Such measures have proved ineffective, resulting in calls for more fundamental research into the ecology of *Campylobacter* in chicken flocks, and greater investigation of the bacteria-host relationship. This thesis presents a multitude of mathematical frameworks to explore these dynamics, driven by on-farm data from the last several decades. Firstly, I present a mechanistic model system of stochastic differential equations to demonstrate the population dynamics of multiple strains of *Campylobacter* at both in-host and flock level. I find that in-host dynamics are different from those observed in the flock at large, with single strains of *Campylobacter* seen to dominate the gut of a host-bird at any given time, compared to the more stable patterns of co-occurrence seen at the wider flock level. I show that *Campylobacter* outbreaks are most influenced by the demographic capabilities of other native gut microflora within host-birds, raising the hypothesis that further attention to bird gut health may be key to limiting the spread of *Campylobacter*. Building on this, I investigate data detailing the occurrence of multiple *Campylobacter* strains within a flock of chickens across a year. By using Bayesian transition models, competition matrix models, and patch-occupancy models, I demonstrate the existence of variation in the response to bacterial challenge amongst birds in the flock, and that this variation may drive the proliferation of *Campylobacter*. This is further supported by models I present showing that *Campylobacter* strains exist in a hierarchical system of competition, and that the diversity of strains found cannot be supported without the inclusion of host-bird factors. Lastly, I investigated the causal relationships between a variety of management and welfare factors for multiple commercial flocks by building and comparing logistic regression models with Bayesian network models. I show that measures of poor bird health and *Campylobacter* status are conditionally dependent upon one another, revealing that decisions in the rearing of these birds simultaneously jeopardise their health, and increase the risk of *Campylobacter* colonisation.

Contents

1	Introduction	1
	<i>Campylobacter</i> and chickens	3
	Control attempts to date	4
	Thesis outline	9
2	A mathematical model of <i>Campylobacter</i> dynamics within a broiler flock	13
	Abstract	14
	Introduction	15
	Modelling Frameworks	18
	Deterministic Model	18
	Stochastic Model	23
	Model Validation	26
	Simulations	29
	Staggered Strain Infection	29
	Stochastic Model - One strain in one broiler	30
	Stochastic Model - One strain in multiple broilers	33
	Stochastic Model - Five strains in multiple broilers	34
	Sensitivity Analysis	36
	Discussion	41
3	A mathematical modelling approach to uncover factors influencing the spread of <i>Campylobacter</i> in a flock of broiler-breeder chickens	49
	Abstract	50
	Introduction	51
	Data	54
	Model Development	57
	Case Studies	59
	Model 1: Time dependence	59
	Model 2: Species dependence	62
	Model 3: Time and species dependence	64

Model 4: ST perseverance	67
Model 5: Chicken dependence	69
Model 6: Chicken and species dependence	70
Model 7: Chicken and density dependence	72
Discussion	76
4 Mechanisms of biodiversity between <i>Campylobacter</i> sequence types in a flock of broiler-breeder chickens	83
Abstract	84
Introduction	85
Methods	89
Data	89
Competition Matrix	92
Patch-Occupancy Model	95
Results	98
Discussion	100
5 Can good broiler flock welfare prevent colonization by <i>Campylobacter</i>?	109
Abstract	110
Introduction	111
Methods	113
Data	113
Data Cleaning	116
Logistic Regression	117
Bayesian Network	119
Discretisation	123
Banned Arcs	124
Results	124
Logistic Regression	124
Bayesian Network	126
Discussion	129
6 Discussion	135
Summary of Chapters	135
Further Work	139
Conclusions	142

Appendices

A	Supplementary Material for Chapter 2	147
	A1 - Model boundedness	148
	A2 - Simulated annealing	151
B	Supplementary Material for Chapter 3	153
	B1 - Bayesian statistics	154
	B2 - Model simulations of flock health	156
C	Supplementary Material for Chapter 4	159
	C1 - Patch-occupancy model pseudo-code	160
D	Supplementary Material for Chapter 5	161
	D1 - Logistic regression: maximal model to minimal adequate model . .	162
	D2 - Bayesian network structure with BDE scoring	169
	D3 - <i>Campylobacter</i> conditional probability tables	171
	Bibliography	177

1

Introduction

Contents

<i>Campylobacter</i> and chickens	3
Control attempts to date	4
Thesis outline	9

Approximately one in three people will experience food poisoning caused by *Campylobacter* in their lifetime, based upon the latest estimates from the Food Standards Agency (FSA) (Holland and Mahmoudzadeh, 2020). At best, such an infection is characterised by severe stomach cramps, diarrhoea, fever, and vomiting for several days (Locht and Krogfelt, 2002), while roughly 3,500 cases each year result in hospitalisation, and in some cases, death (Daniel et al., 2020). Even upon recovery, campylobacteriosis has been associated with a multitude of subsequent sequelae. Cases of campylobacteriosis leave the host ~ 100 times more susceptible to the auto-immune disorder Guillain-Barré syndrome, and evidence has been found of increased risk of reactive arthritis (Hannu et al., 2002), bacteremia (Pigrau et al., 1997), and inflammatory bowel disease (Zautner et al., 2014). This considerable public health burden was initially estimated to present a £50 million annual direct cost to the UK economy (Tam and O’Brien, 2016), but after factoring for the indirect costs such as loss of earnings and business disruption, recent models instead

approximate the true cost at £712 million per year (Daniel et al., 2020).

While *Campylobacter* had previously been known to infect animals, the importance of *Campylobacter* as a human pathogen was first noted by Skirrow (1977), whereupon confirmed instances of campylobacteriosis began to be recorded in the UK. *Campylobacter* spp. are one of the sixty listed ‘notifiable organisms’ in the UK, with medical practitioners legally required to report confirmed cases, to ensure that potential outbreaks can be responded to quickly. While several species of *Campylobacter* have been identified over the years, 90% of infections in humans are attributed to *Campylobacter jejuni* (Public Health England, 2018b), with *Campylobacter coli* comprising most of the remaining cases. While the number of confirmed cases has fluctuated over the years, coinciding with various societal and agricultural changes (Strachan, Rotariu, et al., 2013), the number of yearly cases shows no signs of slowing (Strachan and Forbes, 2010; Public Health England, 2018a).

Global research into tackling this health crisis is severely lacking in mathematical modelling approaches. Such theoretical frameworks are a crucial tool in epidemiological study, enabling researchers to test understandings of underlying disease dynamics by comparing predictions with observed data (Huppert and Katriel, 2013). A great deal of *Campylobacter* prevalence data has been gathered over the last 40 years, and yet few theoretical tool sets have been developed to meaningfully query this data. This thesis addresses this paucity by presenting a variety of different mathematical approaches to specific epidemiological questions as to the underlying dynamics of *Campylobacter* colonisation in chickens. I present previously unreported results detailing the nature of the host-bacteria interaction, using flexible frameworks that can be repurposed by other researchers throughout the field.

***Campylobacter* and chickens**

Many sources have been attributed to campylobacteriosis over the years, yet by far the most commonly identified transmission route is poultry, responsible for as many as 80% of infections (Ackerley, 2016; ACMSF, 2018; Domingues et al., 2012; Pires et al., 2010). In 2014, Public Health England (PHE) began a year-long survey of supermarket chicken carcasses in the UK, testing over 4,000 whole chickens across the country. This survey revealed that 73.3% of chickens tested positive for *Campylobacter*, and 6.8% of samples tested positive on just the outer packaging alone (Jorgensen, Madden, et al., 2015). This study received widespread media attention, highlighting the extent to which *Campylobacter* has dominated the commercial poultry industry.

While there has been some evidence of cross-contamination events at the slaughterhouse, where flocks that test negative for *Campylobacter* at the farm are later found to be colonised after processing (Herman et al., 2003), there is a vast amount of literature suggesting that the majority of broilers (chickens grown for their meat) arrive at the slaughterhouse already colonised (Rivoal et al., 1999; Evans and Sayers, 2000). Indeed, in a study covering 88% of the UK’s poultry slaughter throughput, Powell et al. (2012) demonstrated that 75.8% (303/400) of incoming broiler samples tested positive for *Campylobacter* compared to 87.3% (349/400) outgoing broiler carcasses. Much like in human incidence, the majority of these positive samples were *Campylobacter jejuni*, with the rest being *Campylobacter coli*. In light of this, the current advice given by the Advisory Committee on the Microbiological Safety of Food (ACMSF), a committee of independent experts advising the UK government, is that *Campylobacter* “... cannot be controlled in the poultry slaughterhouse and encouragement and support should be given to the UK poultry industry to continue to maintain high levels of on-farm biosecurity and to continue to explore other on-farm interventions...” (ACMSF, 2018).

Modern-day industrial chicken farming sees broilers grown in flocks of 10-to-100

thousand birds (Sheppard, 2004) in large houses, with roughly 20 million birds slaughtered every week in the UK (DEFRA, 2020). Over the years, genetic selection and the development of high-protein feed has caused the modern-day meat bird to be four times the size of the broilers of the 1950s (Zuidhof et al., 2014). Broilers are grown for roughly five or six weeks before slaughter (though free-range and organic birds are required to be reared for at least 56 and 81 days respectively (Council of European Union, 2008)), in densely packed houses, where communal feeders and drinkers are placed throughout.

It may perhaps be no surprise that, under such housing conditions, when *Campylobacter* are introduced to a flock, they spread quickly and aggressively, until the entire flock is colonised. From the introduction of one inoculated broiler, it can take just a single week for an entire flock to be colonised (Stern, Cox, Musgrove, et al., 2001). The bacteria are spread via the faecal-oral route, whereby a colonised bird excretes the bacteria in its faecal and cecal droppings, which is then ingested by the surrounding birds. Despite its significant impact on human health (Black et al., 1988), *Campylobacter* were long considered commensal in chickens, however more recent studies have begun to unearth clearer signs of pathology such as inflammatory responses (Wigley, 2015; S. Humphrey et al., 2014), adding concerns regarding broiler welfare to the escalating need to bring this health crisis under better control.

Control attempts to date

In response to the increased research attention and growing need for *Campylobacter* intervention strategies, the ACMSF released a report in 2005 outlining their formal advice to the FSA on how to control *Campylobacter* outbreaks. Upon consultation with science and industry experts, the report concluded that biosecurity measures such as improved cleaning of houses, the implementation of ‘boot dips’ (a sanitising solution for farmers to step in before entering the broiler houses), hand-washing, and crate-washing should be prioritised immediately (ACMSF, 2005). The primary motivating evidence for this was seemingly drawn from the far lower incidence of

Campylobacter observed in Scandinavian farms, and the far higher standard of biosecurity observed on these farms. The report also drew attention to the potential impact of the practice of ‘thinning’ on triggering *Campylobacter* colonisation. Thinning, or, ‘partial depopulation events’, is the management practice whereby part of a flock is removed during the growing period. As broiler farms in the EU are not allowed to house birds at densities greater than 39kg/m² (Council of European Union, 2007), the most efficient practice for farmers is to grow a great number of birds when they are lighter, newly-hatched, chicks, and as the stocking limit is approached, to then remove the smaller, less-efficient females, to create more space for the faster-growing, large-breasted males. The hand-removal of these birds had begun to appear in studies as a major contamination event (Hald, Rattenborg, et al., 2001), likely due to the immense breach of biosecurity regulation during these hand-catching depopulation events. While the 2005 ACMSF report prioritised the need to improve biosecurity, it was also recommended that research should begin to explore how the thinning practice could be improved, and if it would be possible to ban the practice.

Later in 2005/early 2006, an exhaustive, systematic literature review was undertaken by Adkin et al. (2006). While the array of research into *Campylobacter* prevalence in broiler flocks was rapidly growing, the majority of studies were small, directed investigations of varying flock sizes into specific management practices. As such, Adkin et al.’s review employed a methodology to compare and contrast the findings of all studies up to May 2003. The conclusions of the review were broadly in line with the previous advice of the ACMSF, general biosecurity was identified as a strong contributing factor; however, thinning was overwhelmingly implicated as the most important aspect of poultry management contributing to *Campylobacter* colonisation on broiler farms. The review also highlighted other factors as having little impact on *Campylobacter* colonisation, such as variation in feed, water, and litter; however, the majority of factors considered by the literature to that point showed a lack of consensus on their possible importance. Most notable was a

significant amount of disagreement and contradiction on a multitude of potential point-of-entry mechanisms by which *Campylobacter* may first enter a flock. Earlier (in 2003), *Salmonella* had been shown to be transmitted vertically, from parent bird to chick, prompting a multitude of studies investigating similar mechanisms in *Campylobacter* (Newell and Fearnley, 2003). Research on this mechanism of spread was firmly divided, with different studies finding both positive and negative association between *Campylobacter* prevalence and: vertical transmission; insect/rodent-borne transmission; wild bird contamination; and flock carry over. Adkin et al. (2006) concluded that future research should prioritise further study of the most promising avenues of interest, namely biosecurity and thinning.

It is hard to ascertain the extent to which farm management was impacted by these reports; however, no reduction in human disease was observed in the following years (Public Health England, 2018a), with recorded instances of campylobacteriosis increasing year-on-year up to 2011. In 2010, jointly agreed between government and industry, the FSA published a target to reduce *Campylobacter* prevalence in broiler meat over the next five years (FSA, 2010). Instead of attempting to reduce the number of colonised birds at farm-level, the target was to reduce the enumeration of bacteria on bird carcasses. The formal target was to reduce the amount of chicken carcasses carrying more than 1,000 colony forming units (CFU) per gram, from a baseline of 27% in 2008 to 10% by 2015. This was to be measured at the end of the slaughterhouse chain, post-chill. Despite the recommendation of the previous ACMSF report to prioritise on-farm biosecurity, the FSA's approach urged that, while updated biosecurity procedures should be adhered to, the primary focus would be on improving procedures in the slaughterhouse.

In 2011, a systematic review by Newell, Elvers, et al. (2011) again summarised the literature, investigating the progress achieved since the previous review of Adkin et al. (2006). They found that while the majority of research had continued investigating the impact of farm biosecurity, alternate preventative measures had

also begun to be considered, such as probiotics, feeding regimens, and competitive exclusion. While the general consensus of the literature was that management practice could have a significant impact on preventing *Campylobacter* colonisation, Newell, Elvers, et al. (2011) concluded that the experimental design of much of the primary research into mechanisms of introduction was questionable. It was well reported that *Campylobacter* rapidly colonised a flock after introduction, but many studies would only test for the presence of *Campylobacter* at the beginning and end of the growth period, thus struggling to observe and hypothesise events that may have triggered an initial introduction. The methods of testing were themselves brought into question, as *Campylobacter* had been shown to struggle to survive in the environment outside of their hosts, and was difficult to capture via the culture methods used to date (Ridley et al., 2008).

The review was, however, conclusive in highlighting that a clear pattern of seasonality in the colonisation of flocks had been identified, with broiler flocks across Europe more frequently found to be *Campylobacter* positive in the warmer summer months, compared to the winter. The review also labelled vertical transmission an unlikely source of initial colonisation, referencing several studies that were unable to identify *Campylobacter* from the eggs of broilers, even using PCR testing. They also showed that thinning as a practice continued to be associated with *Campylobacter* colonisation; however, this review did highlight that the practice was not a definitive cause of introduction of *Campylobacter* into a flock. New research instead suggested the possibility that the significant stress caused to broilers by the thinning process, due to disruption and feed withdrawal, could itself expedite *Campylobacter* colonisation by causing significant change to the microbiome as a result of increased cortisol (T. Humphrey, 2006). The overall conclusion of the review was that little new information had been found in the last five years. They suggested that efforts be focused towards the point at which *Campylobacter* was first introduced to a flock (at odds with the previously announced FSA target), and, while recommending that biosecurity measures continue to be pursued, expressed

concern as to whether maintaining such methods on farms was a sustainable goal. As such, they concluded that additional and complementary interventions would need to be more reliably implemented to limit the colonisation of flocks.

Adding to this growing concern as to the realistic impact of improved biosecurity measures, a review by Hermans et al. (2011) concluded that biosecurity measures alone would not be sufficient to control incidence of *Campylobacter*. Mirroring the concerns of Newell, Elvers, et al. (2011), they expressed concern as to the usefulness of repeated biosecurity studies, arguing that further understanding of the mechanisms of the broiler immune system was urgently needed, and that fundamental new research was required to provide a meaningful reduction in human incidence of campylobacteriosis.

The FSA-set target of <10% high-bacterial-load chicken carcasses was not met (Jorgensen, Madden, et al., 2015), and laboratory-reported cases of human campylobacteriosis decreased only slightly between 2011 and 2016 (Public Health England, 2018a). While prevention was not attained, this period saw increase in knowledge and understanding of the relationship between *Campylobacter* and broiler. The review by Awad, Hess, et al. (2018) explored these in detail. The notion of the ‘commensal *Campylobacter*’ was firmly rejected, initiated by evidence of a clear immune response mounted by broilers to *Campylobacter* colonisation (Han, Willer, L. Li, et al., 2017; Pielsticker et al., 2016). Research into a wealth of indicators has begun to show the wide network of factors influenced by (and influencing) the introduction of *Campylobacter*. Colonised flocks were found to have significantly different microbiome diversity to uncolonised flocks (Thibodeau et al., 2015; Awad, Mann, et al., 2016), which in turn was linked with decreased bird weight (Gharib et al., 2012), footpad lesions (Alpigiani et al., 2017), and even bird behaviour (Colles, Cain, et al., 2016). This wealth of recent discoveries has revealed the burden of *Campylobacter* in broiler flocks stretches beyond being a purely food-based biosecurity challenge. Instead the crisis should be tackled as a so-called ‘One

Health’ challenge, requiring the collaborative effort of human medical expertise, veterinary authorities, researchers, policymakers, public health authorities, and industry, working systematically to reduce the epidemic of campylobacteriosis cases in humans (Gölz et al., 2014).

To this end, it has become apparent through exploring the literature how the majority of *Campylobacter* research is primarily comprised of on-farm microbiological surveys. When compared with the research undertaken in tackling a variety of other zoonotic diseases, such as tuberculosis, rabies, or bird flu, there is an astonishing paucity of theoretical investigation into the driving dynamics and mechanisms of *Campylobacter*. Indeed, until 2019, the only studies to have considered a mathematical approach to understanding *Campylobacter* colonisation of broiler flocks were a simplified susceptible-infected (SI) model by Van Gerwe et al. (2005), and a quantitative risk assessment model by Hartnett et al. (2001). Mathematical models are a crucial tool in driving updated public health policy, and a variety of theoretical frameworks are needed to be used, and importantly, understood, by the field at large to help derive new hypotheses, mechanistic understandings, and analyse data (Kretzschmar, 2020). It is this distinct lack of mathematical tools available for the study of *Campylobacter* that this thesis was constructed to address.

Thesis outline

Each thesis chapter is presented as a stand-alone publication, presenting a novel mathematical modelling approach, designed to investigate either a specific mechanism or data set.

Chapter 2 addresses one of the most glaring omissions from the literature - a bespoke system of differential equations capturing the dynamics of *Campylobacter* colonising a flock of broilers. Starting from an initial set of ordinary differential equations (ODEs) capturing the microbiological interactions occurring within the gut of a broiler, this system is then methodically expanded to a stochastic framework

describing the dynamics of multiple *Campylobacter* sequence types (STs) across a whole flock of broilers. The model is fit to available time-series data on the proliferation of *Campylobacter* across time, and is then used to explore multiple dynamical questions that could not previously be considered. The effects of inter-strain competition, stocking density, and competitive exclusion are explored, and a variance-based sensitivity analysis is performed to highlight which factors are most likely to prevent flock-wide colonisation from occurring. The model is introduced and explained gradually, with the aim that this framework may be used by other researchers to trial their own hypotheses and explore their own data. Indeed many of the conclusions of later chapters can be expressed and displayed through this initial dynamical model, and is gradually being recognised by the wider field.

Chapter 3 uses the most robust and expansive data set of *Campylobacter* dynamics available to date, a year-long survey of specific *Campylobacter* STs and their appearance in specific birds, first presented by Colles, McCarthy, Bliss, et al. (2015). Demonstrating how bespoke mathematical frameworks can unveil far deeper effects than descriptive analyses alone, I present a series of Bayesian models to explore the transition effects of different STs between different broilers. I present remarkable results indicating that the host response of just one bird to bacterial challenge can have significant impact on the colonisation state of the entire flock. Evidence is displayed of ‘super shedders’, broilers who routinely excrete a diverse number of *Campylobacter* STs, and that these super shedders act as important points of *Campylobacter* proliferation, growth, and expansion.

Chapter 4 uses the same data as the previous chapter to investigate the competitive hierarchies displayed by the multitude of STs, through a combination of both competitive matrix models, and patch-occupancy models (developed throughout the wider ecological literature). I show that the diverse assortment of STs displayed a strongly hierarchical structure, with certain frequently-isolated STs being particularly capable of replacing previously observed STs within host birds. The competitive

relationships observed therefore indicated that demographic differences in mortality between STs must exist to drive the vast diversity of STs observed, and yet no significant signs of such demographic differences could be identified. This suggests that, instead of being simply a commensal carrier for *Campylobacter*, the specific immune responses of individual host broilers play a decisive role in the dynamics of *Campylobacter* itself (in agreement with the recent literature and previous chapters).

Finally, Chapter 5 expands on the recent call for One Health approaches in flock management, by investigating the inter-relationships of multiple management and welfare measures alongside the emergence of *Campylobacter* in broiler flocks. Using data from multiple commercial broiler flocks, I investigate what factors may directly influence the likelihood of *Campylobacter* colonisation by using and comparing both traditional logistic regression models, with more sophisticated Bayesian network models. Comparing the insights and advantages of both methodologies, I ask, “can good broiler welfare prevent *Campylobacter* colonisation?”, and demonstrate that the two are conditionally dependent. Namely, that while poor welfare does not directly cause colonisation by *Campylobacter*, the two are both simultaneously influenced by flock management decisions.

2

A mathematical model of *Campylobacter* dynamics within a broiler flock

Contents

Abstract	14
Introduction	15
Modelling Frameworks	18
Deterministic Model	18
Stochastic Model	23
Model Validation	26
Simulations	29
Staggered Strain Infection	29
Stochastic Model - One strain in one broiler	30
Stochastic Model - One strain in multiple broilers	33
Stochastic Model - Five strains in multiple broilers	34
Sensitivity Analysis	36
Discussion	41

Published in *Frontiers in Microbiology* as Rawson, Dawkins, et al. (2019)

DOI: <https://doi.org/10.3389/fmicb.2019.01940>

Abstract

Globally, the bacterial genus *Campylobacter* is one of the leading causes of human gastroenteritis, with its primary route of infection being through poultry meat. The application of biosecurity measures is currently limited by a lack of understanding of the transmission dynamics within a flock. Our work is the first to undertake a mathematical modelling approach to *Campylobacter* population dynamics within a flock of broilers (chickens bred specifically for meat). A system of stochastic differential equations is used to model the routes of infection between co-housed birds. The presented model displays the strong correlation between housing density and *Campylobacter* prevalence, and shows how stochastic variation is the driving factor determining which strains of *Campylobacter* will emerge first within a flock. The model also shows how the system will rapidly select for phenotypic advantages, to quickly eliminate demographically-weaker strains. A global sensitivity analysis is performed, highlighting that the growth and death rate of other native bacterial species likely contributes the greatest to preventing flock colonisation, presenting a promising approach to hypothesising new methods of combatting disease transmission.

Introduction

Campylobacter is recognised as the leading cause of human gastroenteritis in the developed world (Ghareeb, Awad, Mohnl, Schatzmayr, et al., 2013). While several transmission routes have been noted over the years (Nauta et al., 2005), poultry meat has been overwhelmingly attributed as the leading route of ingestion for humans (EFSA Panel on Biological Hazards (BIOHAZ), 2011). An ongoing study by Public Health England has highlighted the extent to which *Campylobacter* spp. have dominated our commercial poultry. Seventy-three percent of supermarket chicken carcasses were found to contain *Campylobacter* and seven percent of the outer packaging was similarly contaminated (Jorgensen, Madden, et al., 2015). An estimated 450,000 people across the United Kingdom are infected every year, with ten percent of these infections resulting in hospitalisation (Strachan and Forbes, 2010). The immediate impact of infection is rarely fatal in the developed world, characterised by stomach cramps and diarrhoea, however the resulting sequelae, while rare, are far more serious. Campylobacteriosis leaves the host ~ 100 times more likely to develop the auto-immune disorder Guillain-Barré syndrome (McCarthy and Giesecke, 2001).

While the bacteria provoke an aggressive response in human hosts, the most common species, *Campylobacter jejuni*, is considered commensal within its most common host, broiler chickens. The term ‘broiler’ refers to any chicken bred and raised specifically for meat production. Once *Campylobacter* is present in a flock, full colonisation of all birds occurs very rapidly (Evans and Sayers, 2000). From the introduction of one infected bird, it can take only a single week for an entire flock to become infected (Stern, Cox, Musgrove, et al., 2001). The bacteria are spread via the faecal-oral route. After becoming infected, the newly-infected host broiler spends a brief period in a non-infectious incubation period, before excreting the bacteria in its faecal and cecal matter. Surrounding susceptible broilers are then exposed to this by ingesting the surrounding feed and water (Shanker et al., 1990). While the direct cause of introduction to the flock is uncertain, an exhaustive review

by Adkin et al. (2006) considered that horizontal transmission is by far the most likely route, primarily being brought into a susceptible flock from some other source on the farm, such as the enclosures of other farm animals. This is as opposed to vertical transmission from breeder flocks, which are themselves often fully colonized by *Campylobacter* spp.. Nevertheless, there may be a combination of both routes of entry into a flock, which deserves greater consideration.

Campylobacter is rarely observed to colonise the gut of very young chickens (0 to 2 weeks of age) (Newell and Wagenaar, 2000). This is theorised to be the result of a supply of innate maternal antibodies acquired during a pre-laying period. This immunity has been shown to have significant bactericidal properties (Sahin, Zhang, et al., 2001).

Despite numerous intervention measures being trialled and employed on farms, little impact has been seen in reducing *Campylobacter* incidence (Hermans et al., 2011). This is due in part to the aggressive rate of proliferation once *Campylobacter* has entered a flock, coupled with persisting uncertainty in the exact route of primary infection. Specifically designed prevention methods are also marred by genetic variation and plasticity of *Campylobacter* spp. (Tresse et al., 2017).

Of increasing concern is the growing trend of antimicrobial resistance in campylobacteriosis outbreaks. Roughly ninety percent of the antibiotics applied in agriculture are used only to promote growth or as prophylactic agents, as opposed to being used to treat infection (Khachatourians, 1998). This overzealous use has been a major contributing factor to the continuing spread of antibiotic resistance. Ge et al. (2003) conducted a study showing that ninety-four percent of tested raw chicken samples were resistant to at least one of seven antibiotics being tested, fifty-four percent of which showed resistance to erythromycin, the antibiotic most commonly used to treat campylobacteriosis. These antimicrobial-resistant strains cause more prolonged and severe illness in humans (Travers and Michael, 2002) and create a

scenario where *in-vitro* susceptibility testing may be necessary before any drugs may be prescribed.

Despite a wealth of empirical investigations, there is a lack of knowledge synthesising these empirical findings through theoretical modelling frameworks. Only two studies have considered a theoretical approach to understanding *Campylobacter* spp. colonisation of broiler flocks; Van Gerwe et al. (2005) and Hartnett et al. (2001), who built a basic susceptible-infected (SI) model and a probabilistic model, respectively. Both frameworks only consider a model on the scale of a flock through basic susceptible-infected interactions. These approaches are not sophisticated enough to develop any meaningful theories on *Campylobacter* dynamics, as they do not represent or convey any specific interbacterial actions by *Campylobacter* populations. The lack of modelling approaches is likely due in part to the inherent challenges of mathematically simulating a gut microbiome. Over 100 different bacterial genera have been isolated from the intestines of chickens (Pan and Z. Yu, 2014), all with a range of individual ecological interactions with one-another. Questions must then be asked regarding how to simulate the temporal and spatial impact of gut motility on the development of a microbial community. Despite these challenges, simplified models of stochastic differential equations have proved effective in capturing the often frenetic bacterial population dynamics within the gut (Wiles et al., 2016).

Here, we introduce a framework of stochastic differential equations that captures the basic interactions that are known to be observed within the broiler gut. Using this framework we simulate the propagation of multiple strains of *Campylobacter* through multiple birds in a flock. In the analysis presented below we observe key dynamical behaviour commonly observed through experimentation, which can now be mechanistically explained using this theoretical framework. The theoretical insights derived from this model can be used to refine current hypotheses regarding *Campylobacter* transmission and inform future experimental and control efforts. The model will likely be of use to experimentalists and risk assessors

in theorising the impact of potential new disease prevention methods on the bacterial transmission dynamics.

Modelling Frameworks

Deterministic Model

Before presenting the stochastic differential equation framework, we begin by introducing the underlying deterministic core of the framework and the particular interactions modelled. Consider four variables to describe the bacterial populations within a broiler's digestive tract. C , the proportion of a single bird's gut flora made up of *Campylobacter*. B , the proportion of the gut flora made up of other bacterial species competing for space and resources. P , the proportion of the gut containing host defence peptides (HDPs) (this may also be interpreted as other plausible forms of host autoimmune response). Lastly, M , the proportion of the gut containing innate maternal antibodies. These all take values ranging such that $0 \leq C, B, P, M \leq 1$. The set of ODEs describing the dynamics follows:

$$\frac{dC}{dt} = r_1 C (1 - (C + \alpha_1 B)) - \gamma CP - d_1 C - \beta CB - \sigma CM, \quad (2.1)$$

$$\frac{dB}{dt} = r_2 B (1 - (B + \alpha_2 C)) - d_2 B, \quad (2.2)$$

$$\frac{dP}{dt} = \xi CP (1 - P) - d_3 P, \quad (2.3)$$

$$\frac{dM}{dt} = -d_4 M. \quad (2.4)$$

All rate constants are defined below in Table 2.1. Model boundedness is shown in Appendix A1. The first term ($r_1 C (1 - (C + \alpha_1 B))$) in equation (2.1) describes the logistic growth of *Campylobacter* to a relative carrying capacity of 1, while in competition with other bacteria B . Competition for resources is the key to success within the gut. *Campylobacter* is known to be an effective coloniser (Stahl et al., 2012), as it is very effective at drawing zinc and iron from its environment (Giella and DiRita, 2012; Raines et al., 2016). The second term (γCP) in equation (2.1) models the inhibitory effect of host defence peptides, P . These peptides are created

in response to challenge by *Campylobacter*, as shown by Cawthraw et al. (1994). The third term (d_1C) of equation (2.1) simply describes the natural death rate of *Campylobacter*. The fourth term (βCB) simulates an important interbacterial interaction; that some of the most abundant competing bacteria in the microbiome have an inhibitory effect on *Campylobacter* (Schoeni and Doyle, 1992). The final term (σCM) of equation (2.1) represents the strong bactericidal abilities of the bird's maternal antibodies. All chickens hatch with an initial supply of antibodies that depletes over time, gone by about three weeks of age (Sahin, Zhang, et al., 2001) (most broilers are slaughtered at five or six weeks of age, however some organic and free-range flocks are slaughtered at approximately eight weeks). These antibodies have a strong inhibitory effect on *Campylobacter*, and many studies are unable to detect *Campylobacter* (by culture methods) in birds under 2 weeks of age under commercial conditions (Sahin, Kassem, et al., 2015). However, forced inoculation of high quantities of *Campylobacter* soon after hatching can still result in expression of the bacteria (Welkos, 1984).

Equations (2.2), (2.3), and (2.4) follow a similar logic to equation (2.1). Other bacteria, B , grow in competition with *Campylobacter* to a carrying capacity. Defence peptides, P , grow in response to *Campylobacter* expression (not in competition for resources), and the population of maternal antibodies, M , does not grow. All variables decay at a rate proportional to their respective populations.

Note that the above model could be reduced by amalgamating terms in equations (2.1) and (2.2), however we choose to keep these separate to (i) keep biological processes clearly defined, and (ii) make further model development and sensitivity analyses clearer.

Ignoring the trivial cases of complete domination by either C or B , the basic dynamical behaviour observed for this simplified model is illustrated in Figure 2.1. Notably, *Campylobacter* is absent from the microbiome until the maternal antibody

Table 2.1: Model parameters and baseline values. Descriptions for all parameter values appearing in the final stochastic model, equations (2.14) - (2.18). Baseline values are given, used for model validation and simulation case studies. Values were calculated by using simulated annealing to identify a parameter set that best fits the experimental data of Achen et al. (1998) (Figure 2.2). See Appendix A2 for an explanation of simulated annealing. $^*\Omega$ value is dependent on the experiment specifics for model validation, but flock case studies consider a flock of 400 chickens, and an Ω value of 200,000.

Expression	Description	Value
r_{C_j}	Growth rate for <i>Campylobacter</i> strain j .	0.3009
r_2	Growth rate for other bacteria (B).	0.1407
α_1	<i>Campylobacter</i> competition coefficient.	0.9744
α_2	Other bacteria competition coefficient.	1
γ_{C_j}	Rate of inhibition by host defence peptides (P) on <i>Campylobacter</i> strain j .	0.6358
ξ_j	Rate of host defence peptide growth in response to <i>Campylobacter</i> strain j .	0.7411
b	Rate of broiler shedding <i>Campylobacter</i> into the environment, E_j .	10
Ω	Total environmental carrying capacity of <i>Campylobacter</i> .	200,000*
d_{C_j}	Death rate of <i>Campylobacter</i> strain j .	0.0185
d_2	Death rate of other bacteria.	0.0212
d_3	Decay rate of host defence peptides.	0.0463
d_4	Decay rate of maternal antibodies.	0.0046
d_5	Death rate of <i>Campylobacter</i> in the environment.	0.05
β_{C_j}	Rate of inhibition by other bacteria on <i>Campylobacter</i> strain j .	0.0276
σ_{C_j}	Rate of inhibition by maternal antibodies on <i>Campylobacter</i> strain j .	0.0661
η_{C_j}	Scaling factor applied to stochastic <i>Campylobacter</i> growth in the gut.	0.01
η_{BC_j}	Scaling factor applied to stochastic <i>Campylobacter</i> inhibition by other competing bacteria.	0.0847
η_2	Scaling factor applied to stochastic competing bacteria (B) growth.	0.01
η_3	Scaling factor applied to stochastic host defence peptide (P) growth.	0.01
η_4	Scaling factor applied to stochastic maternal antibody (M) decay.	0.01
η_5	Scaling factor applied to stochastic <i>Campylobacter</i> growth in the environment.	0.01

population has been exhausted. At this point a sudden, temporary, surge in the population of *Campylobacter* is observed. This phenomena is due to the very low population of HDPs, caused by the strong effect of the initial maternal antibodies. The HDP population then quickly rises to meet this sudden challenge, bringing the *Campylobacter* population back to a lower level in an oscillating manner, where it eventually reaches a steady-state equilibrium. This behaviour is commonly observed in experimental studies (Achen et al., 1998; Newell and Fearnley, 2003).

From this simple core of four equations we adapt the model to allow for N unique strains of *Campylobacter*, by describing each strain as a separate variable. Equation (2.1) is repeated for each individual strain, while altering the growth rate terms to reflect the fact that all strains will also be in competition with one another. This alteration is represented by the following set of ODEs:

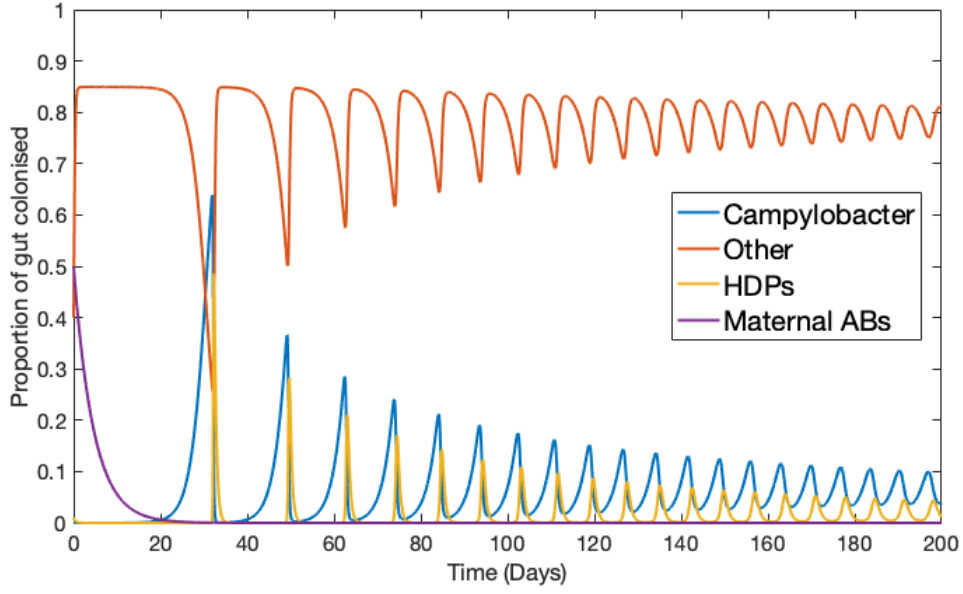


Figure 2.1: Deterministic model for one chicken. An example of the typical dynamical behaviour observed for simulations of equations (2.1) - (2.4). Parameters defined in Table 2.1.

$$\begin{aligned} \frac{dC_j}{dt} = & r_{C_j} C_j \left(1 - \left(\sum_{j=1}^N C_j + \alpha_1 B \right) \right) - \gamma_{C_j} C_j P - d_{C_j} C_j \\ & - \beta_{C_j} C_j B - \sigma_{C_j} C_j M, \end{aligned} \quad (2.5)$$

$$\frac{dB}{dt} = r_2 B \left(1 - \left(B + \alpha_2 \sum_{j=1}^N C_j \right) \right) - d_2 B, \quad (2.6)$$

$$\frac{dP}{dt} = \sum_{j=1}^N \xi_j C_j P (1 - P) - d_3 P, \quad (2.7)$$

$$\frac{dM}{dt} = -d_4 M. \quad (2.8)$$

Here C_j represents the j^{th} strain of *Campylobacter*, where $j \in \{1, 2, \dots, N\}$, and N is the total number of strains. As such, this adjusted model is composed of $N + 3$ variables. The next alteration is to allow for multiple birds and the ability for *Campylobacter* to move from one bird to another. This is done by repeating the $N + 3$ equations presented in equations (2.5)-(2.8) for each bird, and introducing new variables to display the saturation of *Campylobacter* strains in the shared living space.

As such, the newly adjusted model to describe the population dynamics of N strains of *Campylobacter* within L broilers, is written as,

$$\begin{aligned} \frac{dC_{ij}}{dt} = & r_{C_j} C_{ij} \left(1 - \left(\sum_{j=1}^N C_{ij} + \alpha_1 B_i \right) \right) - \gamma_{C_j} C_{ij} P_i - d_{C_j} C_{ij} \\ & - \beta_{C_j} C_{ij} B_i - \sigma_{C_j} C_{ij} M_i + a \frac{E_j}{\Omega}, \end{aligned} \quad (2.9)$$

$$\frac{dB_i}{dt} = r_2 B_i \left(1 - \left(B_i + \alpha_2 \sum_{j=1}^N C_{ij} \right) \right) - d_2 B_i, \quad (2.10)$$

$$\frac{dP_i}{dt} = \sum_{j=1}^N \xi_j C_{ij} P_i (1 - P_i) - d_3 P_i, \quad (2.11)$$

$$\frac{dM_i}{dt} = -d_4 M_i, \quad (2.12)$$

$$\frac{dE_j}{dt} = \sum_{i=1}^L b C_{ij} \left(1 - \frac{E_j}{\Omega} \right) - d_5 E_j. \quad (2.13)$$

Here then, C_{ij} represents the proportion of the i^{th} broiler's gut bacteria which is composed of *Campylobacter* strain j . B_i is the proportion of the i^{th} broiler's gut bacteria made up of other bacterial species competing for space and resources. P_i , the proportion of the i^{th} broiler's gut containing host defence peptides. M_i is the proportion of the i^{th} broiler's gut containing innate maternal antibodies. Here $i \in \{1, 2, \dots, L\}$, where L is the total number of broilers. E_j represents the amount of *Campylobacter* strain j that is currently in the flock's enclosed living space. We assume a living space of fixed size shared by all broilers. As such, Ω represents this total size, or carrying capacity for strains. The first term in equation (2.13) shows that the amount of strain j in the environment is increased by being shed from birds that are already infected with strain j at a rate b . Note from the final term $\left(a \frac{E_j}{\Omega}\right)$ in equation (2.9) that birds may then ingest strain j from the environment at a rate $\frac{a}{\Omega}$. This route of infection simulates the faecal-oral route of infection, but may be interpreted as some other intermittent transmission stage between birds. As such we do not remove *Campylobacter* from the environment (E_j) upon an ingestion event, as the possibility of further environmental contamination is not yet understood and

may indeed depend on the specific route of infection. The model is now composed of $L(N + 3) + N$ equations, for N strains of *Campylobacter*, and L individual broilers.

Stochastic Model

While several important biological phenomena can be discovered and better understood with the model in its current, deterministic, form, there are key reasons to pursue a stochastic framework. First, having one variable alone to represent the multitudes of bacterial species that make up the constantly-evolving gut microbiome is, of course, a significant simplification. In practice, these other bacterial species competing with *Campylobacter* will be constantly changing, both in resurgences of population and in how they interact with *Campylobacter*. Adding stochastic elements to these populations and interactions is a small step towards capturing some of this more unpredictable behaviour. Indeed the biomass of *Campylobacter* measurable in faecal and cecal matter has been observed to fluctuate widely (Morishita et al., 1997; Achen et al., 1998). Secondly, the density dependent assumptions made when formulating the initial deterministic model, that is that interaction rates are directly proportional to the variable populations, are assumptions that break down for smaller populations. The simulations undertaken often display bacterial populations at very small quantities, especially in the initial period dominated by maternal antibodies. A stochastic system behaves very differently under these circumstances and means that the model is more likely to display cases of strain extinction, a phenomena that the deterministic model cannot capture. Indeed, the very nature of *Campylobacter* infections is one that is often described in the language of probability. The all-or-nothing nature of flock infections means that we often must ask what measures can reduce the likelihoods of infections, rather than the magnitude. Through a stochastic framework we explore multiple realisations of potential outcomes, and investigate reducing the likelihood of *Campylobacter* colonisation in a flock of broilers. The case studies presented below highlight the need for a stochastic modelling approach to accurately capture the multi-strain dynamics

of *Campylobacter* within the gut. This approach allows for the simulation of turnover and resurgence of dominant strains, an experimentally observed behaviour (Colles and Maiden, 2014) that cannot be captured by a deterministic system.

For the stochastic framework, equations (2.9)-(2.13) are adjusted to the following set of stochastic differential equations,

$$\begin{aligned} dC_{ij} = & \left[r_{C_j} C_{ij} \left(1 - \left(\sum_{j=1}^N C_{ij} + \alpha_1 B_i \right) \right) - \gamma_{C_j} C_{ij} P_i - d_{C_j} C_{ij} \right. \\ & \left. - \beta_{C_j} C_{ij} B_i - \sigma_{C_j} C_{ij} M_i + a(E_j) \right] dt \\ & + [\eta_{C_j} C_{ij} + \lambda_j(t) - \eta_{BC_j} C_{ij} B_i] dW_t, \end{aligned} \quad (2.14)$$

$$dB_i = \left[r_2 B_i \left(1 - \left(B_i + \alpha_2 \sum_{j=1}^N C_{ij} \right) \right) - d_2 B_i \right] dt + [\eta_2 B_i] dW_t, \quad (2.15)$$

$$dP_i = \left[\sum_{j=1}^N \xi_j C_{ij} P_i (1 - P_i) - d_3 P_i \right] dt + [\eta_3 P_i] dW_t, \quad (2.16)$$

$$dM_i = [-d_4 M_i] dt + [\eta_4 M_i] dW_t, \quad (2.17)$$

$$dE_j = \left[\sum_{i=1}^L b C_{ij} \left(1 - \frac{E_j}{\Omega} \right) - d_5 E_j \right] dt + [\eta_5 E_j] dW_t, \quad (2.18)$$

where $\lambda_j(t)$ is defined by;

$$\lambda_j(t) = \begin{cases} 0, & \text{if } C_{ij}(t) = 0. \\ 2.873 \times 10^{-4}, & \text{otherwise.} \end{cases}$$

and where $a(E_j)$ is defined by;

$$a(E_j) = \begin{cases} 0.015, & \text{if } X < \frac{E_j}{\Omega} \text{ for random variable } X \sim \mathcal{U}(0, 1). \\ 0, & \text{otherwise.} \end{cases}$$

Here W_t denotes a Wiener process (standard Brownian motion process). These stochastic increments are scaled by the respective population size and constants η_2 through to η_5 . These constants dictate the variance of their respective Wiener processes, defining the range of stochasticity attributed to the growth rate of their

respective variables. The changes and additions shown in equation (2.14) warrant further explanation. The sixth term ($a(E_j)$) in equation (2.14) (the last of the deterministic terms), has been changed from a constant rate of ingestion from the environment, as seen in equation (2.9), to instead have ingestion modelled by a chance to ingest *Campylobacter* depending on the amount of that strain in the environment, E_j . The greater E_j is, the more likely it is for ingestion to occur.

The eighth term ($\lambda_j(t)$) in equation (2.14) is a stochastic term independent of the population of C_{ij} . This is introduced to allow for the possibility of extinction events, should the population of C_{ij} reach a particularly low threshold. This threshold is decided by the value taken by $\lambda_j(t)$, in this case 2.873×10^{-4} . As with all other parameter values displayed in Table 2.1, the values used in the expressions $\lambda_j(t)$ and $a(E_j)$ are calculated through model validation against the studies presented below. Finally, the ninth term of equation (2.14) considers stochasticity surrounding the interactions between C_{ij} and the competing bacteria B_i . This term allows for instances when the particular biodiversity and spatial structure of the gut microbiome may be more inhibitory towards *Campylobacter*, or perhaps actually assisting its growth instead.

Several interesting dynamical behaviours can be observed using this model, which are highlighted through some specific question-led case studies. Table 2.1 defines all parameters presented in the final stochastic model ((2.14) - (2.18)) as well as a baseline of parameter values that were used in model validation against real world data sets (presented below). The model is constructed to an arbitrary timescale, however the parameter values given in Table 2.1 ensure that multiple oscillations in the *Campylobacter* population can be observed in the below case studies, a phenomena observed in the lifespan of broilers (Morishita et al., 1997).

Note that throughout we have chosen to use a *Campylobacter* competition coefficient of $\alpha_1 = 0.9744 < 1$. This choice is justified in that bacterial populations can inhabit

multiple intestinal niches that cannot be colonised by other competing bacteria. Indeed competitive exclusion therapies have been far less effective in tackling *Campylobacter* compared to other foodborne illnesses such as *Salmonella* (Stern, Cox, Bailey, et al., 2001). The deterministic model is solved using the `ode45` solver, a fifth-order Runge-Kutta method in Matlab. The stochastic model is solved numerically using the Euler-Maruyama method (Higham, 2001) with $N = 2^{14}$ timesteps, also programmed in Matlab. The code used to produce all figures presented is available at: <https://osf.io/b3duc/>.

We also note that, while the model is general enough that it is not specific to any one species of *Campylobacter*, two of the three datasets that were used to tune our model parameter values were from studies unique to the most common species, *Campylobacter jejuni* (Achen et al., 1998; Stern, Cox, Musgrove, et al., 2001). As such we may expect slightly different parameter values to be applicable for other species of *Campylobacter*. However, from testing the model, the results we present below are ones that are generally observed for a range of parameter values, and as such are relevant findings to the whole genus.

Model Validation

We test our model by comparing its predictions against three experimental studies on *Campylobacter* expression and spread. Firstly, we consider the work of Achen et al. (1998). Achen et al. (1998) performed an experiment with twenty-four broilers, who were kept in individual, isolated wire-bottomed cages. Birds were confirmed as free of *Campylobacter* before being inoculated with a *C. jejuni* suspension. A cloacal swab was then obtained from each bird every day for forty-two days, to monitor whether or not each bird was shedding *Campylobacter*. Figure 2.2 shows their experimental results alongside the predictions made by our model.

Specifically, the blue line represents the modal value of the percentage of the 24 birds shedding across a thousand simulations, with error bars depicting the standard

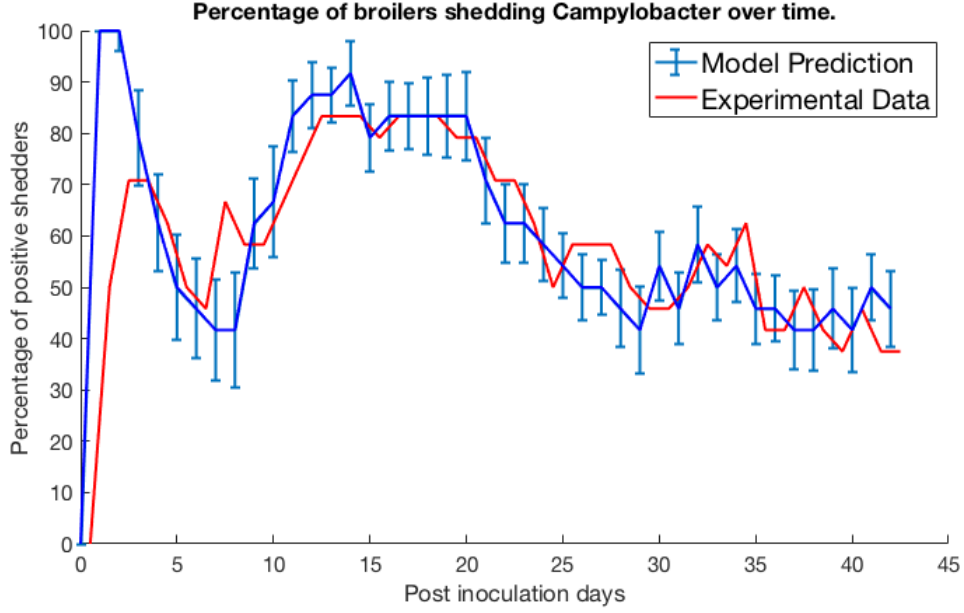


Figure 2.2: Model validation against data of Achen et al. (1998). The percentage of a group of isolated broilers shedding *Campylobacter* across several weeks following inoculation.

deviation across these simulations. Achen et al. (1998) also reports how most birds would shift from phases of positive shedding to negative shedding, a phenomena also captured by the oscillating behaviour displayed by the model. Sampling via culture methods like those performed in this experiment is prone to false-negative results for samples with very low quantities of *Campylobacter* (Acke et al., 2009). Therefore, for this model validation, we considered a broiler as being clear of *Campylobacter* if its proportion of *Campylobacter* (variable C) was below 0.005. This was considered a more accurate measure to correspond with the experimental data. While our model was constructed to an arbitrary timescale, comparing to this real-world data set it was found that our time axis is best rescaled by a factor of 0.021 to align with the measure of days used in these studies. This corrected timescale is used for all subsequent case studies within this paper.

Secondly, we consider the experiment conducted by Stern, Cox, Musgrove, et al. (2001). Multiple separate pens were prepared, each containing 70 broilers, all free of *Campylobacter*. A *Campylobacter*-positive seeder bird was then added to

the flock. Different pens had seeder birds introduced at different points in time. 3, 5, and 7 days after a seeder bird was introduced, a sample of chickens were tested for *Campylobacter* to estimate the percentage of the flock that was currently *Campylobacter*-positive. We plot our model predictions against the experimental data of Stern, Cox, Musgrove, et al. (2001) below in Figure 2.3. To match the housing density of the experiment, a value of $\Omega = 45,369$ was used for the model. An error band is plotted around our model prediction displaying the standard deviation of values across 100 simulations.

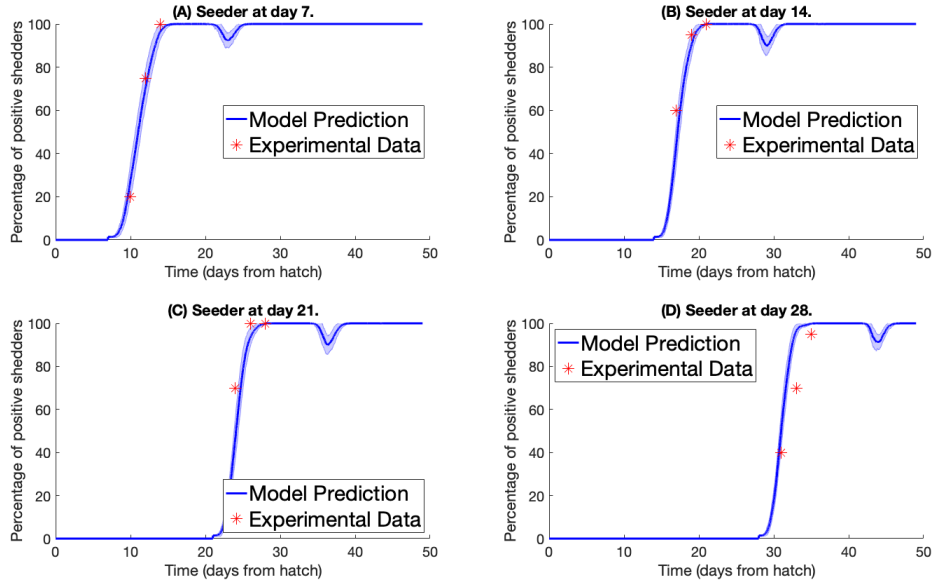


Figure 2.3: Model validation against data of Stern, Cox, Musgrove, et al. (2001). The percentage of a flock of broilers shedding *Campylobacter* across several weeks after introduction of a *Campylobacter*-positive seeder bird at (A) seven days, (B) fourteen days, (C) twenty one days, (D) twenty eight days.

Lastly we simulated the experiment performed by Van Gerwe et al. (2005). Four flocks of 400 birds were set up in individual enclosures from day of hatch. Four birds in each flock were then inoculated with a *Campylobacter* suspension and returned to the flock. Birds were then sampled from each flock throughout the next few weeks to record the percentage of flock infection. Figure 2.4 plots their experimental data against our model prediction. For the experiments shown in

Figure 2.4A and Figure 2.4B, the four seeder birds were inoculated at day of hatch, and chickens were sampled by cloacal swabbing. For the experiments shown in Figure 2.4C and Figure 2.4D, the seeder birds were inoculated one day after hatch, and the flock was analysed by collecting fresh fecal samples.

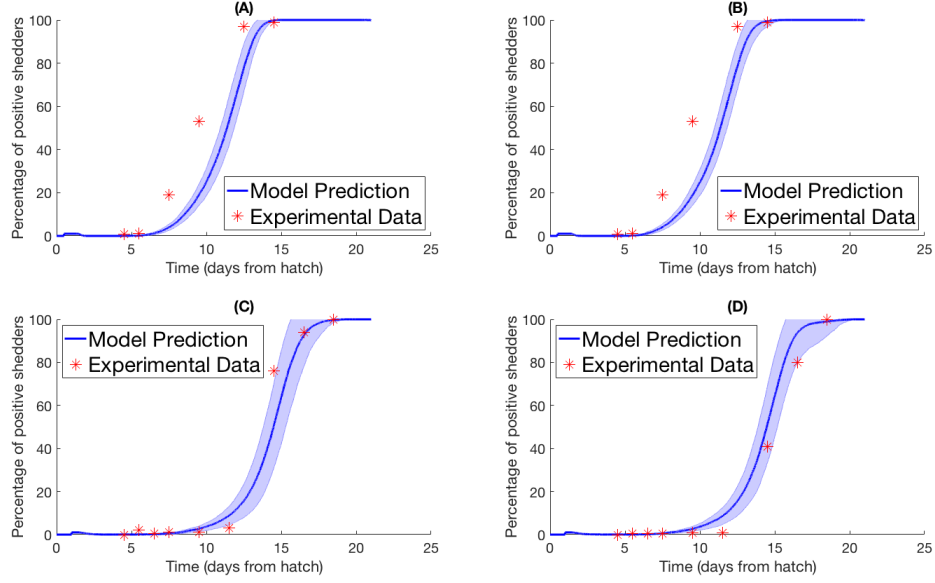


Figure 2.4: Model validation against data of Van Gerwe et al. (2005). The percentage of a flock of broilers shedding *Campylobacter* across several weeks after introduction of a *Campylobacter*-positive seeder bird. (A)/(B) Seeder bird introduced at day of hatch, samples collected via cloacal swab, (C)/(D) seeder bird introduced one day after hatch, samples collected via fresh fecal droppings.

Simulations

We now use a series of (simulated) case studies to investigate key dynamical behaviours and predictions from the model.

Staggered Strain Infection

In this first example, the deterministic model for multiple strains in one broiler (equations (2.5) - (2.8)) is considered. Five strains of *Campylobacter* within one chicken are simulated, all with the exact same respective rate constants as shown in Table 2.1. Figure 2.5A shows the results when all five strains are introduced at

$t = 0$ days, with the same initial inoculation amount of $C_i(0) = 0.0001$. Figure 2.5B shows instead when each strain is introduced in intervals of $t = 5$ days. Therefore only strain 1 is introduced at $t = 0$ days, strain 2 is introduced at $t = 5$ days and so on until finally strain 5 is introduced at $t = 20$ days. In both cases the other three variables are initialised at $B(0) = 0.4$, $P(0) = 0.01$ and $M(0) = 0.5$.

While the maternal antibodies (M) are not plotted on these figures, they approach 0 at approximately $t = 20$ days, as can be seen by the following surge in *Campylobacter* populations following this point in Figure 2.5. While, unsurprisingly, all strains perform identically in Figure 2.5A (where strains are initialised at the same point in time), a more curious dynamic is observed in Figure 2.5B. The strain that performs best and exists at the highest proportion in the staggered release example is strain 2, the second strain to be introduced. The reason for this is that strain 1, present at $t = 0$ days, is initially suppressed by the maternal antibodies (parameter M), reducing the proportion of strain 1. As a result, when strain 2 is introduced, it is able to capitalise on the severely reduced amount of strain 1, and the reduced amount of maternal antibodies, to quickly grow and dominate the competitive space. Strain 2's increased presence then puts future strains at a disadvantage as it has already had the opportunity to establish itself within the gut. These results suggest that dominant *Campylobacter* strains can prevent new strains from taking hold. Moreover, there is an optimal point in time for inoculation to occur for a strain to become dominant, as shown in Figure 2.5B where strain 2 is consistently occupying a higher proportion of the gut than other strains.

Stochastic Model - One strain in one broiler

The stochastic model (equations (2.14) - (2.17)) is run to simulate one strain of *Campylobacter* within one broiler. In this scenario, we ignore the environmental variable E (equation (2.18)), as its input is negligible for only one broiler. The rate constants are kept at the same values as used previously, defined in Table 2.1, with the additions of the stochastic variance scaling rate constants, parameters that limit

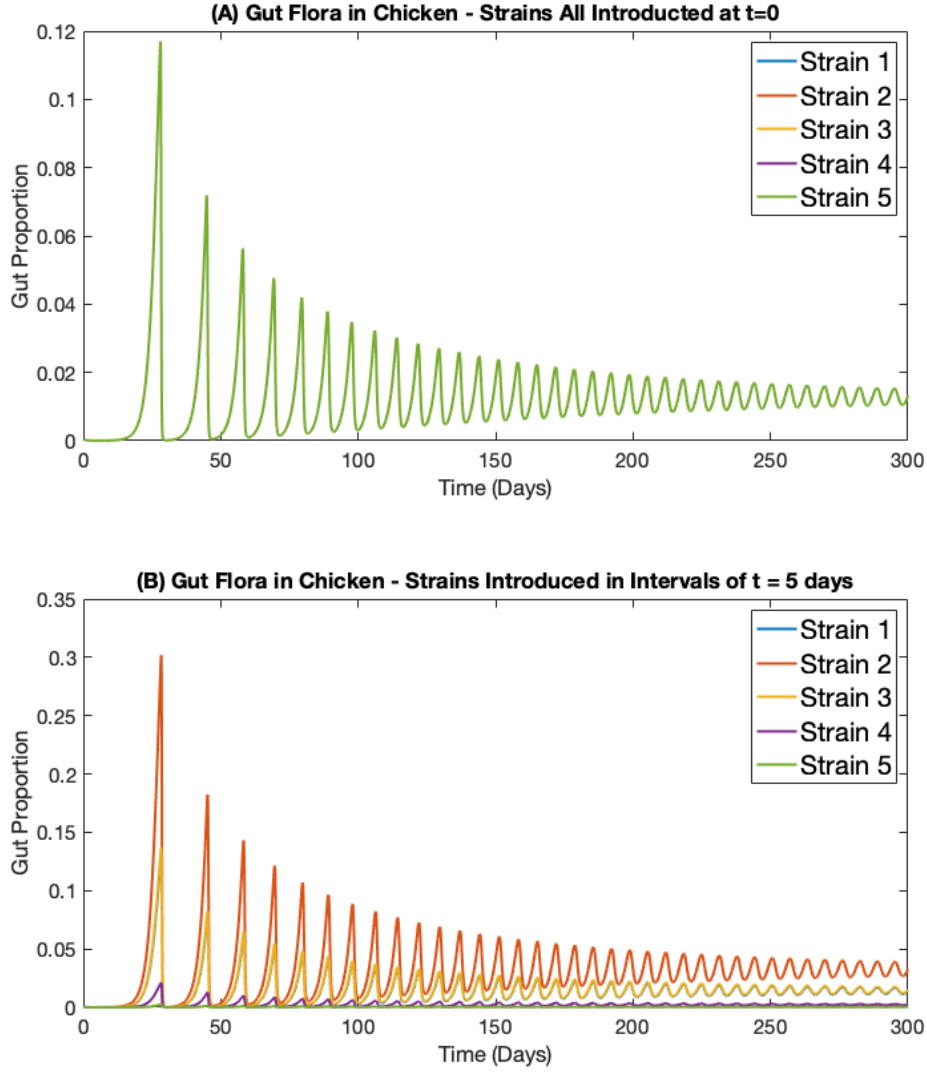


Figure 2.5: Simulations of multiple *Campylobacter* strains within one broiler. Population growth of five strains of *Campylobacter* within one broiler that are (A) all introduced at $t = 0$ days (B) introduced in intervals of $t = 5$ days. Strains are initialised at $C_j(t) = 0.0001$ at their respective time of introduction. Other variables are initialised at $B(0) = 0.4$, $P(0) = 0.01$ and $M(0) = 0.5$. Note that the single green line in plot (A) is due to overlap, all five strains exhibit the exact same dynamical behaviour, as would be expected.

the variance of the stochastic additions. These are set as $\eta_{C_j} = \eta_2 = \eta_3 = \eta_4 = 0.01$, and $\eta_{BC_j} = 0.0847$. η_{BC_j} is set higher than the other stochastic rate constants to capture the greater unpredictability surrounding these bacterial interactions. Four different realisations of this model are presented in Figure 2.6, all initialised at $C(0) = 0.02$, $B(0) = 0.4$, $P(0) = 0.01$, $M(0) = 0.5$.

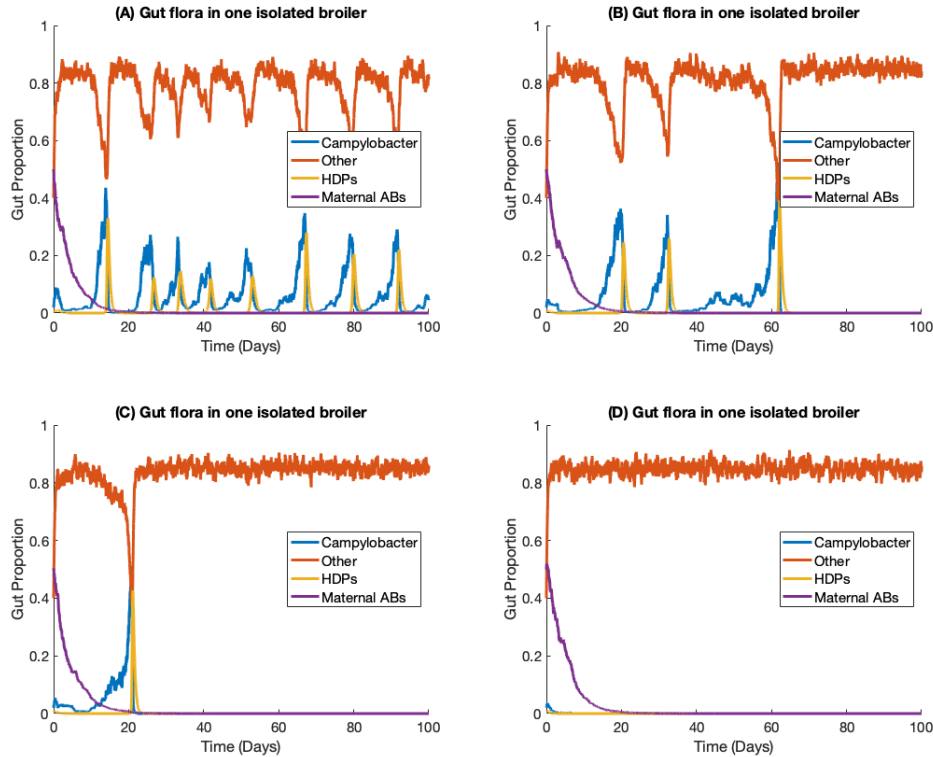


Figure 2.6: Stochastic simulations of one *Campylobacter* strain within one broiler. Four different realisations of a stochastic model simulating one strain of *Campylobacter* within one isolated broiler.

Empirical studies measuring the amount of *Campylobacter* in the faecal matter of isolated broilers have shown a spectrum of results (Achen et al., 1998). Some broilers display sustained high populations, others express initial peaks followed by great reduction and potentially later resurgence, and sometimes extinction cases are observed. All these dynamical behaviours can be observed in different realisations of this model (Figure 2.6). Figure 2.6A shows an instance where a broiler is consistently infected and shedding into the environment, unable to effectively clear the *Campylobacter* from its system. Figure 2.6B instead shows an instance where a broiler has multiple periods of high infection and shedding, before being able to clear the infection. Figure 2.6C shows an instance where after one initial peak in *Campylobacter* expression, a broiler is able to quickly clear infection. Finally, Figure

2.6D shows an instance where the broiler successfully clears *Campylobacter* at the initial point of inoculation. All these realisations are run with the same parameters given in Table 2.1, demonstrating the benefit of a stochastic framework being able to better capture the more diverse range of possible events. This broad array of dynamical profiles is not observed in commercial broiler flocks, a phenomena that is demonstrated in the following case study.

Stochastic Model - One strain in multiple broilers

The previous scenario is now extended to consider multiple broilers. Figure 2.7 presents the results for one *Campylobacter* strain in a flock of 400 broilers. We use the parameter values stated in Table 2.1. The total size of the enclosure, or the carrying capacity of E , is set at $\Omega = 200,000$. This value is considered in cm^2 , and so with 400 broilers, translates to 500cm^2 per broiler. EU directive 2007/43/CE states that broilers may never be stocked at more than $39\text{kg}/\text{m}^2$ (Council of European Union, 2007). Assuming a targeted bird weight of 1.5kg , this translates to 385cm^2 per bird. This simulation models slightly more space allowed to each bird than the limit. The death rate of *Campylobacter* in the environment is set at $d_5 = 0.05$, higher than the death rate within a broiler as, despite their many survival mechanisms (Murphy et al., 2006) *Campylobacter* is susceptible to many exterior environmental stresses (Park, 2002) and is exceptionally fragile outside of its host. The simulation began with no *Campylobacter* in the surrounding environment ($E(0) = 0$) and the other initial conditions are set the same as for the previous example, with the exception that two of the 400 broilers start with an initial condition of $C_1(0) = C_2(0) = 0.02$, while the others are initialised without any *Campylobacter*. These results are shown in Figure 2.7.

While birds who are not initialised with *Campylobacter* become infected at a slightly later time, the dynamical behaviour is very similar across all birds in the flock. Multiple realisations do not display the broader spectrum of behaviour observed in the one broiler case (Figure 2.6). The implication is that housing a greater

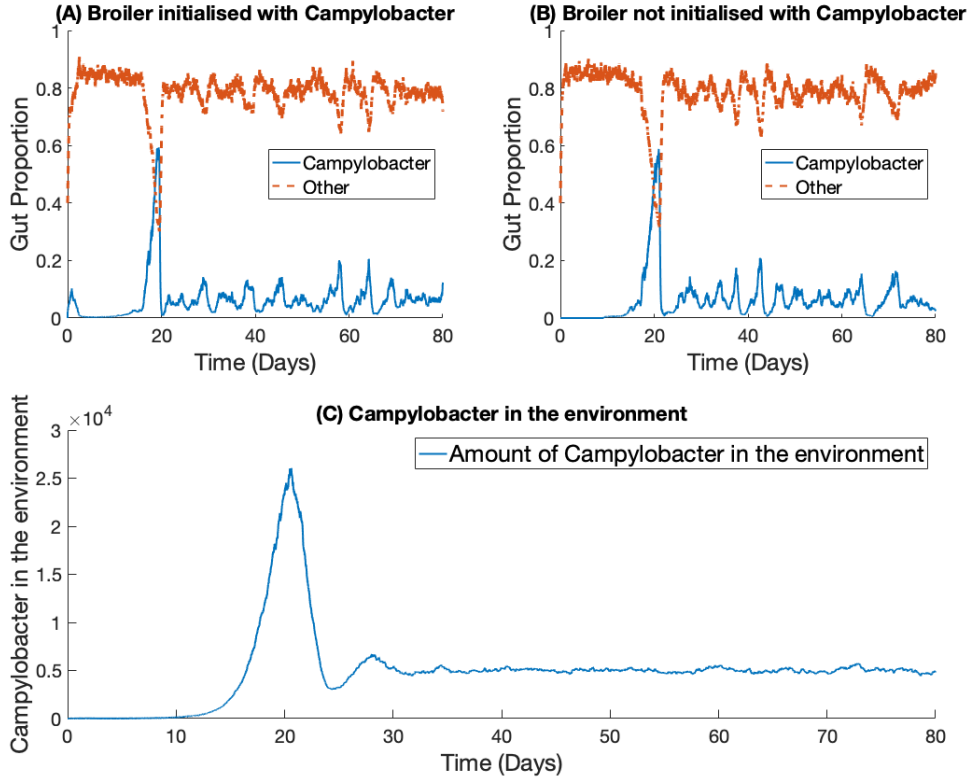


Figure 2.7: Stochastic simulations of one *Campylobacter* strain within multiple broilers. The proportion of a broiler’s gut containing *Campylobacter* for (A) a broiler in the flock initialised with a small proportion of *Campylobacter* (B) a broiler in the flock initialised with no *Campylobacter*. (C) shows how much of the environment (total size of 200,000) contains *Campylobacter*. This is variable E in the model.

number of birds causes more homogeneous dynamical behaviour, and indeed the wide variety of *Campylobacter* expression seen in the isolated bird experiments of Achen et al. (1998) is not so commonly observed in experiments with group-housed birds (Van Gerwe et al., 2005).

Stochastic Model - Five strains in multiple broilers

We extend the previous scenario to investigate dynamics of multiple strains. Five strains of competing *Campylobacter* are modelled within the same flock of 400 birds. The same constants are used as in the previous scenario, with each strain having identical rate constants. One key difference is that all broilers are initialised without any *Campylobacter*, instead an initial amount is present in the environment.

Each strain of *Campylobacter* in the environment is initialised at $E_1(0) = E_2(0) = E_3(0) = E_4(0) = E_5(0) = 100$. The results of this simulation are shown in Figure 2.8.

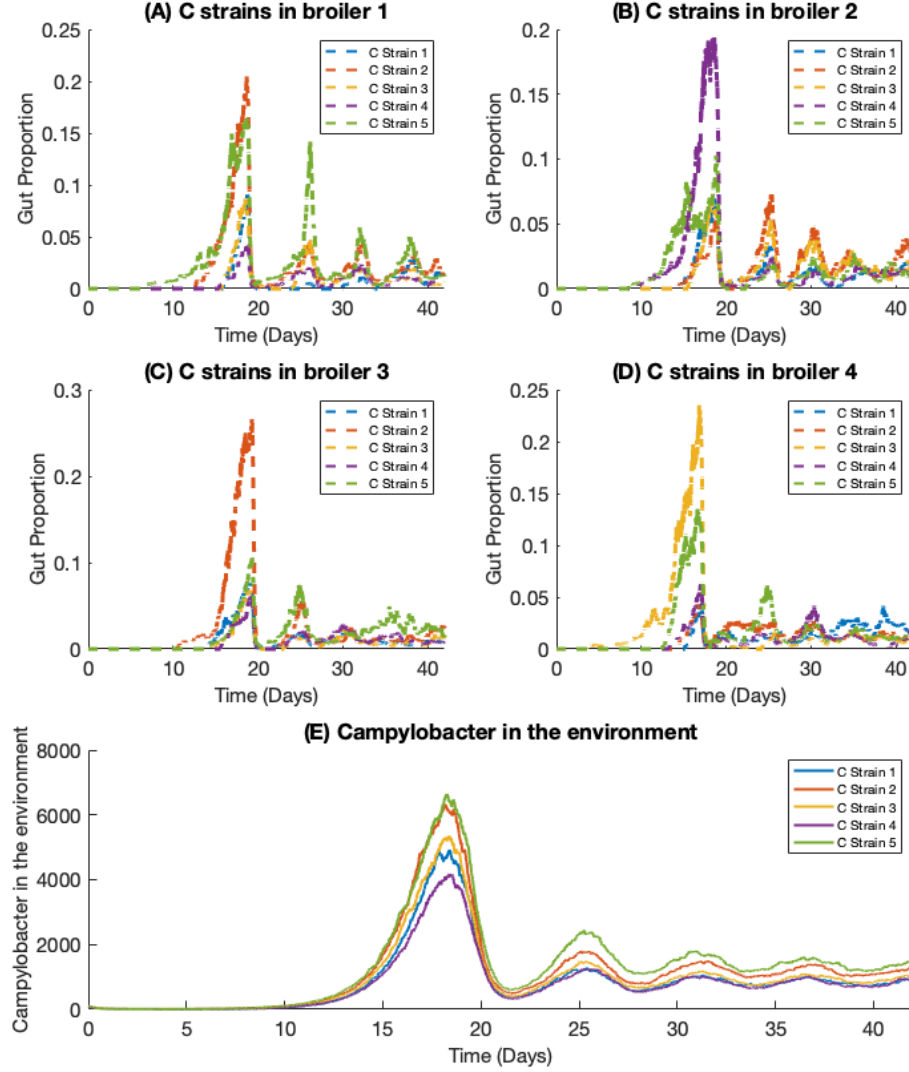


Figure 2.8: Stochastic simulations of multiple *Campylobacter* strains within multiple broilers. The proportion of four different broilers' microbiomes that contain five strains of *Campylobacter*. All birds are within the same flock. (E) shows how much of the environment (total size of 200,000) contains the five strains of *Campylobacter*. These are variables E_j in the model.

On average, all strains perform equally well across the flock, as shown in Figure 2.8E. All strains slowly converge to roughly equal amounts in the environment,

reflecting an equal presence on average across all birds in the flock. However, when observing the *Campylobacter* proportions within individual broilers, one or two strains will tend to dominate early on in colonising a broiler's gut, which can in turn prevent other strains from taking hold (seen most clearly in Figure 2.8D). This dynamical behaviour was first observed in our deterministic simulations (see Figure 2.5B), however unlike in the deterministic case, stochastic events can cause dominant *Campylobacter* strains to reduce in population, presenting an opportunity for a different strain to establish itself.

This phenomena is more clearly seen if the timescale of the simulation is extended, as illustrated in Figure 2.9. Although the average population of strains across the flock is equal, the stochastic model shows that a single strain of *Campylobacter* tends to dominate the gut of individual broilers at any one time. Although there are brief periods where strains exist in equal amounts, eventually the balance shifts again to longer periods of dominance by one or perhaps two strains.

Disadvantaged strains of *Campylobacter* are quickly eliminated. Figure 2.10 shows the results for a simulation where strain 4's growth rate, r_{C_4} , is reduced from 0.3009 to 0.295, and strain 5's growth rate, r_{C_5} , is reduced to 0.29. Strains 1, 2 and 3 are kept with a growth rate of 0.3009. As Figure 2.10 shows, the weaker strains are unable to outcompete the other three and are quickly eliminated. Changing other demographic parameters of a strain achieves a similar result of driving a strain to extinction, the phenomenon is not unique to only altering the growth rate. Making only very small reductions to the growth rate can result in a strain surviving at a lower average population size, although this may only be due to the time needed for extinction to occur being too long to observe in these simulations.

Sensitivity Analysis

A powerful use of this model is to conduct a robust sensitivity analysis to identify the parameters of greatest impact in causing *Campylobacter* colonisation. We

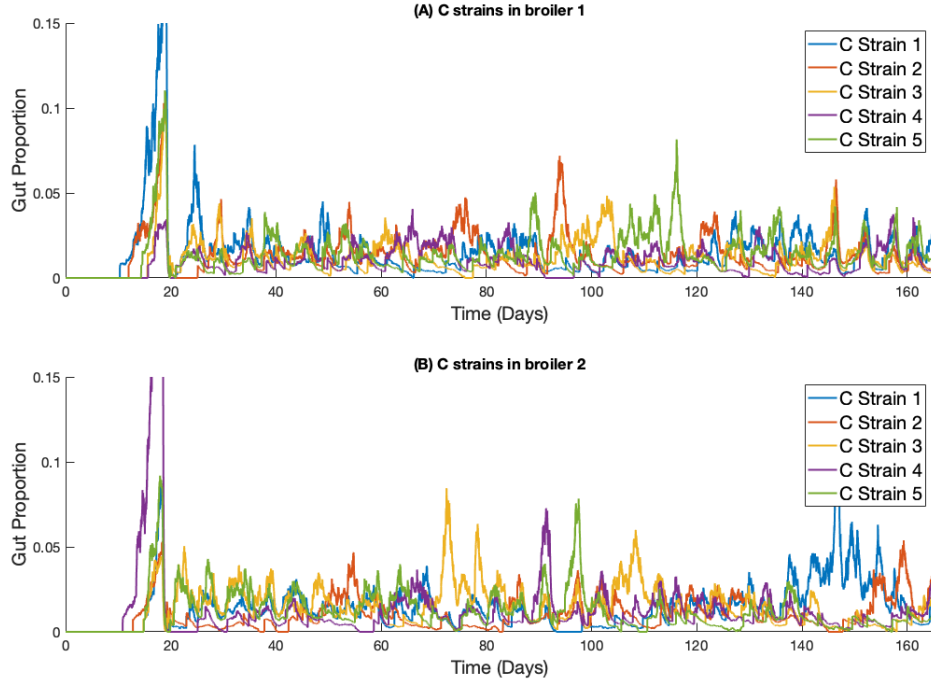


Figure 2.9: Stochastic simulations of multiple *Campylobacter* strains within multiple broilers across a greater timescale. The proportion of two different broilers' microbiomes that contain five identical strains of *Campylobacter*.

adopt a variance-based analysis of the model, and investigate the likelihood of flocks remaining free of *Campylobacter* based on a random assignment of parameter values.

We consider the case of a flock of broilers infected with a single strain of *Campylobacter*, the scenario shown in Figure 2.7. Model parameters are sampled randomly from a uniform range, and the model is run multiple times for these values. We then record how many of these stochastic runs resulted in the flock successfully eliminating *Campylobacter* infections, before drawing a new random sample of parameters values and repeating as necessary. Eventually we finish with a final data set which we display an example of below in Figure 2.11.

As such, the most “important” parameters will be the ones which exhibit a strong trend in their scatter plot. A seemingly randomly distributed scatter plot would indicate a parameter value which has little impact on our output. To report

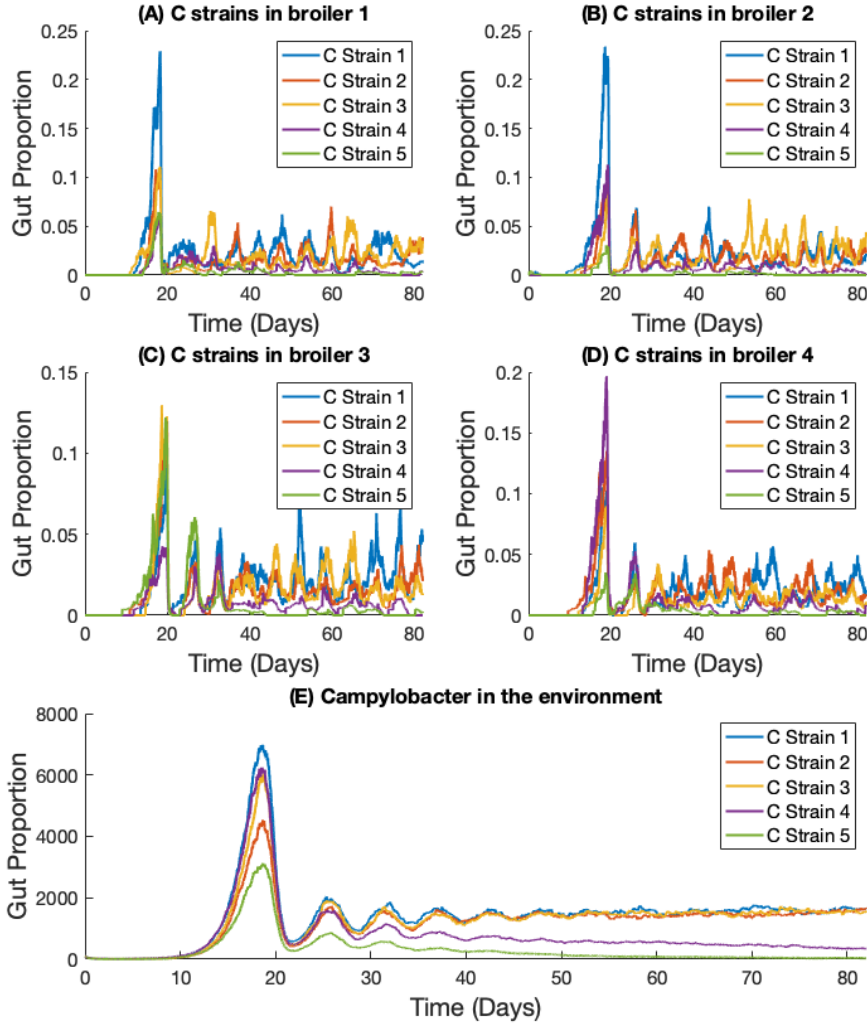


Figure 2.10: Stochastic simulations of multiple *Campylobacter* strains, differing in growth rates, within multiple broilers. The proportion of four different broilers' microbiomes that contain five strains of *Campylobacter*. Strain 4 has had its growth rate reduced from 0.3009 to 0.295 and strain 5 has had its growth rate reduced to 0.29. Strains 1, 2 and 3 have a growth rate of 0.3009. **(E)** shows how much of the environment (total size of 200,000) contains the five strains of *Campylobacter*. These are variables E_j in the model.

more accurately this measure, we use the first-order sensitivity index, S_i , and the total effect index, S_{T_i} , defined as:

$$S_i = \frac{V(E(Y|X_i))}{V(Y)}, \quad S_{T_i} = \frac{E(V(Y|X_{\sim i}))}{V(Y)},$$

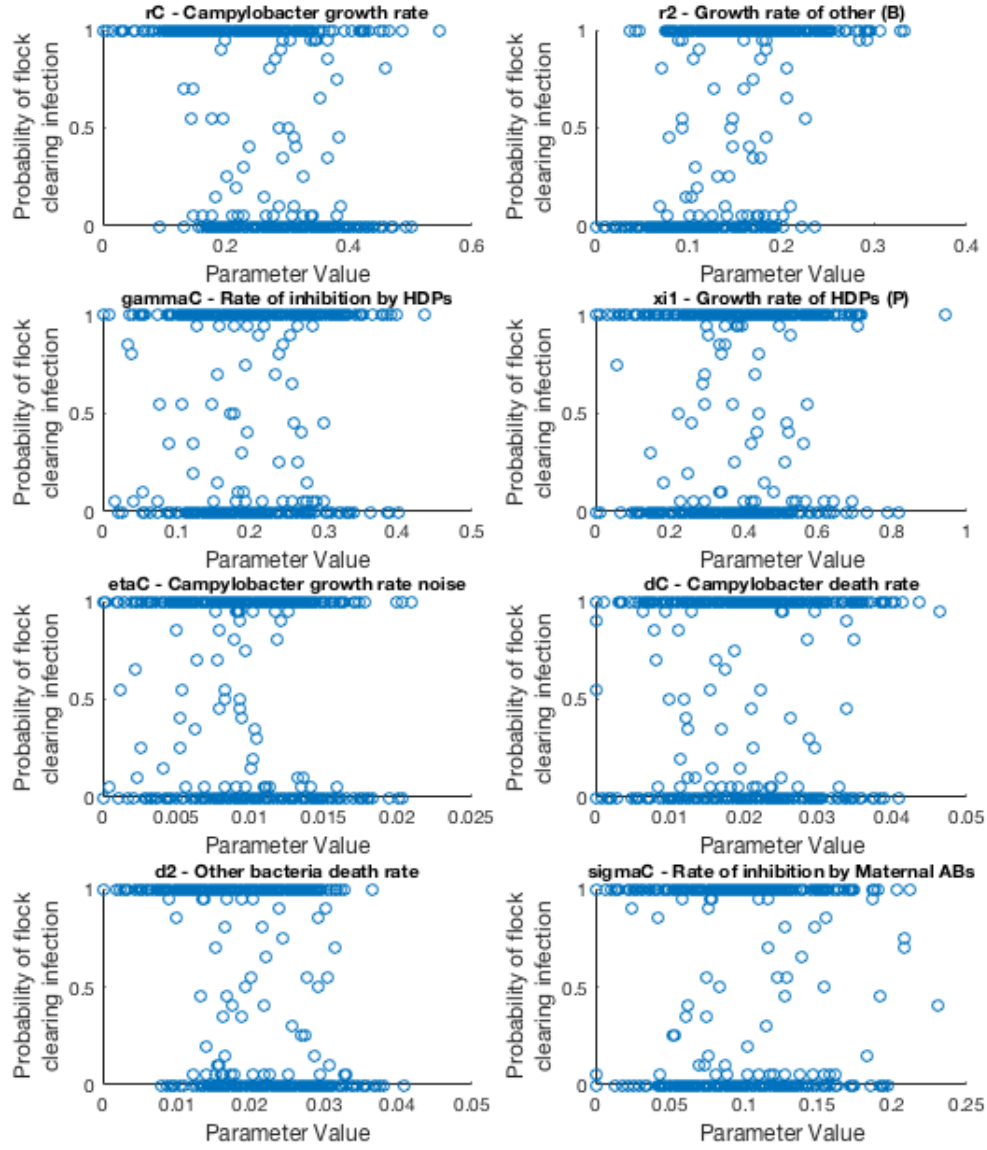


Figure 2.11: Scatter plots displaying probability of a flock clearing *Campylobacter* infection against randomly sampled parameter values. Each scatter plot depicts the results for a specific parameter value. Probability is calculated by running the model for a sampled parameter set twenty times, and recording how many of those runs resulted in the flock not becoming infected with *Campylobacter*.

where X_i denotes parameter i , and Y denotes the model output. $X_{\sim i}$ denotes the vector of all factors but X_i . $V(\cdot)$ denotes the variance, and $E(\cdot)$ the expectation. Specifically $E(A|B)$ denotes the expectation of variable A when B is held fixed. In

short S_i will measure the changes observed in the output when parameter X_i is kept fixed, while S_{T_i} measures the changes to the output when all other parameters are kept fixed. A full derivation and explanation can be found in Saltelli et al. (2008). In short, both are values that range from zero to one, that explain the impact of a parameter on the model output. The higher the value, the more “important” the parameter is. S_{T_i} is considered a stronger metric, as it also considers the higher-order impact of a parameter, whereas S_i only considers the immediate first-order impact. As such S_i would be a sufficient measure for a linear model, but for a more complex model such as the one presented in this paper, S_{T_i} can better reveal the impact that each parameter plays. An initial sensitivity analysis was run for twenty parameters with 2,000 parameter set samples, drawn from a quasi-random Sobol set (Saltelli et al., 2008). The results of this analysis are displayed in Table 2.2, and the code used to produce them is available to access at: <https://osf.io/b3duc/>.

Specifically, our objective function will run the stochastic model for a flock of chickens with the random parameter set drawn. If this model run results in no *Campylobacter* being present in the flock, it is considered to have successfully eliminated infection. The model is run twenty times with this parameter set, and the proportion of these twenty runs that results in an elimination of *Campylobacter* is the final output value, the ‘probability of flock clearing infection’.

Note that some of the values in Table 2.2 are negative, despite S_i and S_{T_i} being limited to being between zero and one. This is due to the computational error in estimating the value, however the ordering of parameters for these particular runs will not be affected by this error. Table 2.2 shows that the S_{T_i} values associated with most parameters ranges between 0 and 0.1. The “most important” parameters however have a wider spread of associated S_{T_i} values.

The main result from this analysis is that the growth, death and inhibition rates of the other bacteria present in a broiler’s gut (parameters r_2 , d_2 and β_C) collectively

Table 2.2: Sensitivity analysis of parameters in a stochastic model for one *Campylobacter* strain in a flock of broilers. The first-order sensitivity index and total effect index is given for a sensitivity analysis of 2,000 runs for 20 parameters. The output function considered is the probability of *Campylobacter* going extinct within the flock based on the given parameter set.

S_i	Parameter	S_{T_i}	Parameter
-0.0112	η_3	-0.0031	σ_C
-0.0054	η_{BC}	-0.0026	η_4
-0.0051	a	-0.0016	η_5
-0.0017	η_C	-0.0011	d_4
-0.0011	η_2	0.0056	b
0.0009	η_5	0.0065	a
0.0020	d_4	0.0114	γ_C
0.0025	b	0.0122	ξ
0.0038	η_4	0.0224	η_C
0.0054	ξ	0.0228	η_{BC}
0.0062	σ_C	0.0378	η_3
0.0077	γ_C	0.0389	Ω
0.0085	Ω	0.0417	η_2
0.0096	d_5	0.0543	d_5
0.0125	β_C	0.1380	d_C
0.0441	d_C	0.1925	β_C
0.0459	r_C	0.2088	r_C
0.1131	d_3	0.3117	d_3
0.1622	r_2	0.4418	d_2
0.1900	d_2	0.4789	r_2

carry the largest impact in eliminating *Campylobacter* from a flock. As such, we can begin to consider which preventative methods could best take advantage of this heightened sensitivity.

Discussion

Here, we have investigated the dynamics of *Campylobacter* across a range of model applications. Our framework reveals several key dynamics of microbial interaction that explain many experimentally observed phenomena. This presents promising new approaches to understanding and tackling this bacteria.

First, the most apparent prediction is that the *Campylobacter* population is

successfully suppressed by the innate maternal antibodies (an experimentally observed phenomenon (P. L. Connerton et al., 2018)), until these antibodies are eventually removed from the system. At this point an initial surge in the population of *Campylobacter* is observed, before it comes to rest at a lower level, reaching an equilibrium with the broiler's immune-response. This can be seen in all of the above figures, but most clearly in Figure 2.1. This initial surge creates an interesting opportunity for certain strains of *Campylobacter* to emerge as an early dominating strain. Figure 2.5B shows that, due to the antibacterial properties of a broiler's maternal antibodies, any strains that infect a broiler early on in its lifespan will be heavily inhibited. This creates a brief window at the point in which maternal antibodies have depleted, whereby any new strain introduced is observed to quickly colonise and dominate the gut flora, suppressing other strains. This hypothesis has been verified experimentally (P. L. Connerton et al., 2018).

The proposition of damped oscillations between *Campylobacter* population size and the host's immune-response is better reinforced by observations that host antibody populations will also oscillate in birds infected with *Campylobacter* (Cawthraw et al., 1994). This basic interaction has been experimentally observed by Achen et al. (1998), with a high degree of variability between birds. This variability is better captured by the stochastic model, as shown in Figure 2.6. Indeed, many birds in Achen et al.'s (1998) study are shown to clear *Campylobacter* successfully from their system, a result rarely observed on commercial broiler farms. Likewise this result was only observed in the model case of individual, isolated broilers (see Figure 2.6D).

Most important is the mechanism observed in Figure 2.7, where the broad spectrum of oscillatory behaviour observed within a broiler is greatly reduced in a large flock of birds. Indeed the vast examples of individual dynamics observed in Figure 2.6, large oscillations and perhaps extinctions, are completely replaced by the same, homogenised dynamics seen within flock-reared birds in Figures 2.7A and 2.7B, as the populations of *Campylobacter* within each bird are consistently

reinforced by the amount of *Campylobacter* in the environment. The wealth of experiments in monitoring flock *Campylobacter* expression for varying flock sizes means this effect can be observed taking place across multiple experiments of different flock magnitudes and densities. Morishita et al. (1997) measured the amount of *Campylobacter* in a flock of thirty birds in a sizeable pen. This flock was small enough to observe oscillating behaviour in the prevalence of *Campylobacter*, and yet there do not appear to be any clear cases of birds being able to clear the bacteria from their system. Stern, Cox, Musgrove, et al. (2001) experimented with flocks of 70 birds at a density of 15.4 birds/m². A small cyclic pattern is observable in their results but there are clearly far higher incidence rates. Lastly, Van Gerwe et al. (2005) studied flocks of 400 birds housed at 20 birds/m² (the same density considered in the above flock modelling), where now no cyclic patterns can be observed, and all birds quickly reach a constant state of *Campylobacter* expression. This effect is seen in Figure 2.7, and almost always observed in commercial farms (Stern, Clavero, et al., 1995; Evans and Sayers, 2000). Our work presented here is the first, to our knowledge, to be able to propose a mechanistic explanation for this observed effect, namely that the housing density of reared flocks is correlated to *Campylobacter* prevalence.

This dynamic, whereby broilers are consistently infected with *Campylobacter* due to highly contaminated living space, can also explain the observed phenomena whereby broiler breeder flocks (flocks kept for the breeding of meat birds) display a consistently lower *Campylobacter* prevalence rate than commercial broiler flocks (Colles, McCarthy, Layton, et al., 2011). Breeder birds will regularly move between periods of testing positive and negative for *Campylobacter*, inconsistently with the state of other birds in the flock, unlike the much younger birds grown for meat which remain consistently positive. Our case studies suggest that this may be due to the lower stocking density afforded to breeder birds, as it would appear the route of infection between breeder birds is weaker than that between broilers. Our sensitivity analysis however also highlighted that the gut flora can have a strong

impact on the survival of *Campylobacter*. The differences in diet and management practice for breeder birds likely results in a different variety of bacterial colonies to broilers, which could also be a cause of the differences seen between breeders and broilers in *Campylobacter* expression.

Additionally, we note that over these case studies we have seen that the population dynamics are unaffected by the initial method of inoculation. There is no clear consensus yet on whether flocks are initially infected through horizontal or vertical transmission, and our model predictions show that this may not be possible to determine from flock infection dynamics. Case studies showed no difference between initialisation with an infected environment, or an infected number of broilers. More specifically, Figure 2.7 shows clearly how the dynamic profile of a broiler that is initialised with *Campylobacter*, is not significantly different from that of a bird which is infected through the environment, representing the effect of vertical and horizontal transmission respectively.

Over time, our model shows strains of equal fitness will tend to settle at equal levels of prevalence on average across a flock (Figure 2.8E), a result that has also been shown experimentally (Colles and Maiden, 2014; Colles, McCarthy, Bliss, et al., 2015). However, it is very common for an individual broiler to have only one or two dominant strains against far smaller proportions of other strains (Colles, Preston, et al., 2019), as our model represents in Figures 2.8A - 2.8D and Figure 2.9. Our results show that this effect is most prominently seen early on in the chicken's lifespan, where usually only one strain will be present during the initial population surge of *Campylobacter*. Colles, Preston, et al. (2019) shows that a greater diversity of strains are observable later on in a broiler's lifespan, but usually at a far lower prevalence compared to a dominating strain. Evidently, when one strain is already well-established within a chicken's gut, it is difficult for new competing strains to grow. This is due to the broiler already having a heightened level of immune response

(P) due to the currently present strain. In the deterministic case, later strains would never be able to establish themselves as much as strains that were earlier to arrive (Figure 2.5B). However, in the stochastic model, there is the potential for a stochastic event to reduce the population of the currently dominating strain, and increase the population of a less-established strain.

Across the whole flock, weaker strains can be quickly outcompeted by other strains. Figure 2.10 shows two weaker strains (strains with lower growth rates) attempting to survive within a flock, even having a slight population peak at the optimal point of strain introduction, before eventually being forced to extinction by the other three strains. Parameter variation showed that reducing a strain's demographic parameters by a very small amount can allow it to persist still in the flock at a smaller average population than the others, but the majority of realisations would always end with weaker strains becoming extinct. Clearly this shows an environment where genetic dominance is very quickly selected for.

These results have considerable implications for biosecurity. While smaller flocks may have a very real opportunity to be protected from *Campylobacter* invasions, *Campylobacter* prevalence is far more stable in larger commercial flocks, and our model shows it to be exceptionally difficult to remove. Efforts can be made to prevent initial inoculations, but once a bacterial presence is established, it may be all but impossible to remove from a flock. Considerable improvements to biosecurity have been made in recent years, but very little impact has been observed in this having reduced *Campylobacter* incidence (Hermans et al., 2011). These measures do not reduce the speed of proliferation of the bacteria, and our results suggest that better attention to bird health is likely to have a greater effect on preventing flock infection.

This model's greatest strength is its lack of overarching assumptions. We model only the most basic bacterial interactions, all supported and verified through experimental work. Our stochastic system is capable of exhibiting a plethora of

interesting dynamical interactions based on just a few known biological interactions. In moving forward with this work, the model can be used to theorise optimal methods by which to decrease the likelihood of *Campylobacter* colonisation, and begin collaborative efforts in better explaining the evolving genetic diversity of this bacteria.

One area in which the model is admittedly lacking currently, is that it does not represent the physiological changes that occur as a bird grows. Broilers have been genetically selected over the many decades to grow excessively fast, which has been shown to have numerous concerning implications for their health (Buzala et al., 2015). This is likely to then result in differences to their auto-immune capabilities over time. More pertinently, the gut flora of a chicken is known to change and develop as the birds age (Lu et al., 2003), suggesting varying degrees of inter-bacterial uncertainty.

Our sensitivity analysis gives great insight into the optimal routes of infection prevention. Table 2.2 clearly shows that bolstering the growth rate and inhibition capabilities of the other bacteria populating a broiler's gut is the best way to force extinction of *Campylobacter*, primarily through suppressing *Campylobacter* at its initial appearance in a system, before it has the opportunity to propagate. As such, the sensitivity analysis suggests further exploration and experimentation into the impact of factors which would affect the gut flora of a broiler. Probiotics are a clear way of impacting the microflora (Mountzouris et al., 2007) and have shown some effect in studies into their impact on *Campylobacter* expression (Santini et al., 2010). Equally, the stressors linked with stocking density have been shown to affect the gut microflora by Guardia et al. (2011). Burkholder et al. (2008) have shown that feed withdrawal and heat stress can considerably alter and limit the gut microflora. These highlight that general bird health and welfare can be equally strong factors in determining the values of r_2 , d_2 and β_C ; some of the parameters highlighted as most "important" by the sensitivity analysis. Table 2.2 also however highlights

the importance of parameters d_3 , the death rate of host defence peptides. This parameter has been shown to be strongly affected by stressors such as overcrowding (Gomes et al., 2014). As such, this result would lend further support to giving greater care to the health and welfare of broilers, as the resulting improvement to host defence peptide production would have a positive impact on helping prevent *Campylobacter* colonisation.

These caveats notwithstanding, the model presented is capable of explaining a wealth of experimentally observed *Campylobacter* population dynamics, further elucidating an urgent public health risk. We have used our framework to investigate multiple strain interactions, to understand better the spread of genotypes across a flock. Finally, we were able to use the model to highlight the factors most responsible for causing infection. Looking forward, this work can be used to understand better observed differences in colonisation dynamics between different farms and indeed countries, and further our goal of minimising public exposure to this dangerous pathogen.

Code Availability

The code used to produce all figures presented is available at: <https://osf.io/b3duc/>.

3

A mathematical modelling approach to uncover factors influencing the spread of *Campylobacter* in a flock of broiler-breeder chickens

Contents

Abstract	50
Introduction	51
Data	54
Model Development	57
Case Studies	59
Model 1: Time dependence	59
Model 2: Species dependence	62
Model 3: Time and species dependence	64
Model 4: ST perseverance	67
Model 5: Chicken dependence	69
Model 6: Chicken and species dependence	70
Model 7: Chicken and density dependence	72
Discussion	76

Published in *Frontiers in Microbiology* as Rawson, Paton, et al. (2020).

DOI: <https://doi.org/10.3389/fmicb.2020.576646>

Abstract

Despite continued efforts to improve biosecurity protocols, *Campylobacter* continues to be detected in the majority of commercial chicken flocks across Europe. Using an extensive data set of *Campylobacter* prevalence within a chicken breeder flock for over a year, multiple Bayesian models are presented to explore the dynamics of the spread of *Campylobacter* in response to seasonal variation, species-specificity, bird health, and total colonisation prevalence. These models indicated that birds within the flock varied greatly in their response to bacterial challenge, and that this phenomena had a large impact in the overall prevalence of different species of *Campylobacter*. *Campylobacter jejuni* appeared more frequently in the summer, while *Campylobacter coli* persisted for a longer duration, amplified by the most susceptible birds in the flock. Our study suggests that strains of *Campylobacter* that appear most frequently likely possess no demographic advantage, but are instead amplified due to the health of the birds that ingest it.

Introduction

Poultry meat has been decisively implicated as a leading infection route for campylobacteriosis in humans (EFSA Panel on Biological Hazards (BIOHAZ), 2011). With an estimated 450,000 cases a year in the UK, approximately ten percent of which result in hospitalisation (Strachan and Forbes, 2010), *Campylobacter* presents an important public health challenge, and an estimated £50 million annual economic burden to the UK (Tam and O’Brien, 2016). An investigation by Public Health England indicated the extent to which *Campylobacter* spp. dominated the commercial poultry industry: seventy-three percent of supermarket chicken carcasses were found to contain *Campylobacter* and seven percent of the outer packaging was similarly contaminated (Jorgensen, Madden, et al., 2015). Consequently, preventing the spread of the bacteria to humans by reducing the number of colonised broiler flocks, i.e. chickens grown specifically for meat, at slaughter is as an urgent endeavour (D. J. Wilson et al., 2008).

Current attempts at tackling *Campylobacter* colonisation have largely focused around on-farm biosecurity measures; however, little impact in reducing incidence has been demonstrated (Hermans et al., 2011). This is predominantly due to the aggressive rate of proliferation once *Campylobacter* has entered a flock, and further complicated by uncertainty in the exact route of primary colonisation. Specifically designed prevention methods are also marred by high genetic diversity of *Campylobacter* spp. (Tresse et al., 2017).

Campylobacter has historically been considered to be a gastrointestinal commensal of chickens, but recent studies suggest that, at least in some circumstances, they are pathogenic (Wigley, 2015; S. Humphrey et al., 2014). We here refer to birds being ‘colonised’ and *Campylobacter* spp. ‘shed’ where they could be detected from samples. Once an initial bird has become colonised by *Campylobacter*, colonisation of the rest of the flock occurs very rapidly, usually within the course of a week (Evans and Sayers, 2000; Shreeve et al., 2000; Stern, Cox, Musgrove, et al., 2001).

The bacteria are spread via the faecal-oral route. After becoming newly-colonised, the host broiler spends a brief period in a non-infectious incubation period, before excreting the bacteria in its faecal and caecal matter. Surrounding susceptible broilers are then exposed to this via coprophagy (Shanker et al., 1990).

Understanding of the spread of *Campylobacter* is hindered by incomplete understanding of the transmission dynamics of the bacteria at farm level. Multiple strains of *Campylobacter* can simultaneously inhabit broiler flocks (Höök et al., 2005), with some strains appearing to dominate the flock at different times (Kudirkiene et al., 2010; De Cesare et al., 2008). It has been suggested that these dynamical behaviours are driven by the appearance of demographically superior strains that outcompete other strains (Calderón-Gómez et al., 2009) within the gut. However, another study suggests that strains are lost or transmitted randomly, regardless of their genotypic differences (Grant et al., 2005). Indeed, recent mathematical modelling approaches have demonstrated that stochastic simulations can effectively capture the broad dynamical differences between strains of equal demographic ability (Rawson, Dawkins, et al., 2019).

An area of more recent study is the role played by ‘super-shedders’ (birds who consistently shed high amounts of *Campylobacter* in their faeces) in the transmission dynamics of *Campylobacter* within a flock. The impact of ‘super-shedders’ has been well documented as a key factor in the rapid spread of *Salmonella* throughout chicken flocks (Menanteau et al., 2018; Gopinath et al., 2012), but the impact on the dynamics of *Campylobacter* spread within broiler flocks is not well-studied. These ‘super-shedders’ have been found experimentally to have fewer circulating heterophilic cells, but this does not appear to be a genetically acquired trait, nor the result of differences in adaptive immunity (Barrow et al., 2004). Variation in *Campylobacter* shedding in broiler flocks has been observed in an experimental study measuring *Campylobacter* prevalence (Achen et al., 1998), and it is reasonable to assume that this could have implications for the transmission dynamics within a

flock.

Some factors affecting transmission are well-reported, if incompletely understood. The effect of seasonal variation on both the carriage rate, and number of *Campylobacter* found in the caeca, of colonised chickens has been noted (Wallace et al., 1997), with an increase often observed in the spring or summer. The exact timing of these peaks, however, varies within and among countries (Kovats et al., 2005), and experimental work is not always able to detect such an effect (T. Humphrey et al., 1993). Less well-investigated is the impact of different species of *Campylobacter* competing within a flock. *C. jejuni*, the most common species, has been found in approximately 90% of British chicken flocks, compared with *C. coli* appearing in 10% of flocks (Jorgensen, Ellis-Iversen, et al., 2011). This ratio has been reported by other studies in broiler flocks (Bull, Allen, et al., 2006), with species rarely both being simultaneously present. It is not understood whether this is due to established strains suppressing new strains from emerging, demographic differences, or the short-lifespan of commercial broiler flocks not providing enough time for multiple species to colonise a flock. Under laboratory conditions, *C. coli* has been shown to have lower growth rates, motility, and invasiveness than *C. jejuni* (Aroori et al., 2013), potentially explaining its rarer appearance in chicken flocks. There is also some suggestion that *C. coli* is more commonly isolated from older, free-range, birds (Colles, T. A. Jones, et al., 2008).

This study explores the impact of multiple factors on the transmission of multiple sequence types (STs) of *Campylobacter* amongst individual birds within a flock across 51 weeks. A broiler breeder flock, i.e. the parents of broiler/meat birds, was studied rather than a broiler flock, since the production period of around a year, compared to a small number of weeks for broiler birds, allows much greater potential to study the interaction of different *Campylobacter* strains over time. When interpreting the results, however, it should be noted that broiler breeder flocks will differ from broiler flocks with respect to, for example, host genetics, age, feed, and

flock density. The majority of the *Campylobacter* genotypes (STs), however, have been isolated from other chicken flocks, most typically housed commercial broiler flocks (Colles, McCarthy, Bliss, et al., 2015).

We use a robust data set from Colles, McCarthy, Bliss, et al. (2015), which is currently the best available for monitoring the *Campylobacter* strain dynamics amongst individual birds within a commercially reared flock across 51 weeks. Through a Bayesian modelling approach we show the range of receptiveness to colonisation throughout the flock, and highlight the role that more-susceptible, ‘super-shedder’, birds play in driving disease. The impact of seasonal variation is also investigated, and specific attention is given to differences between species of *Campylobacter*, so as to understand how certain strains persist at higher levels throughout the flock. Seven exploratory models are presented, each investigating a specific research question, analysing the transition probabilities at both a flock-wide, and individual level.

A Bayesian approach is considered for this study due to the methodology’s innate strengths in analysing incomplete data (Dorazio, 2016), and enabling efficient inference of missing data. Numerical computations were carried out using the Just Another Gibbs Sampler (JAGS) program (Plummer, 2007), a Markov chain Monte Carlo (MCMC) sampling program utilising Gibbs sampling.

Data

The field data used for this study was originally presented in Colles, McCarthy, Bliss, et al. (2015). Within a flock of 500 broiler breeders, 200 birds were labelled with leg-rings and monitored for a total of 51 weeks. Each week, roughly 75 unique birds were picked at random from the labelled 200, and a swab was taken of the cloacal opening. Some small variation is seen around this sample size due to human error. These swabs were then tested for the presence of *Campylobacter* through standard culture methods, and positive samples were then genotyped by multi-locus

sequence typing (MLST) of seven house-keeping genes, enabling the sequence type (ST) and species of the *Campylobacter* isolate to be specified. Further experimental details can be found in the original publication.

As such we build a dataset providing information on real-time evolution of *Campylobacter* prevalence and diversity throughout the flock. This is shown below in Figure 3.1, with all positive samples classified by species of *Campylobacter*.

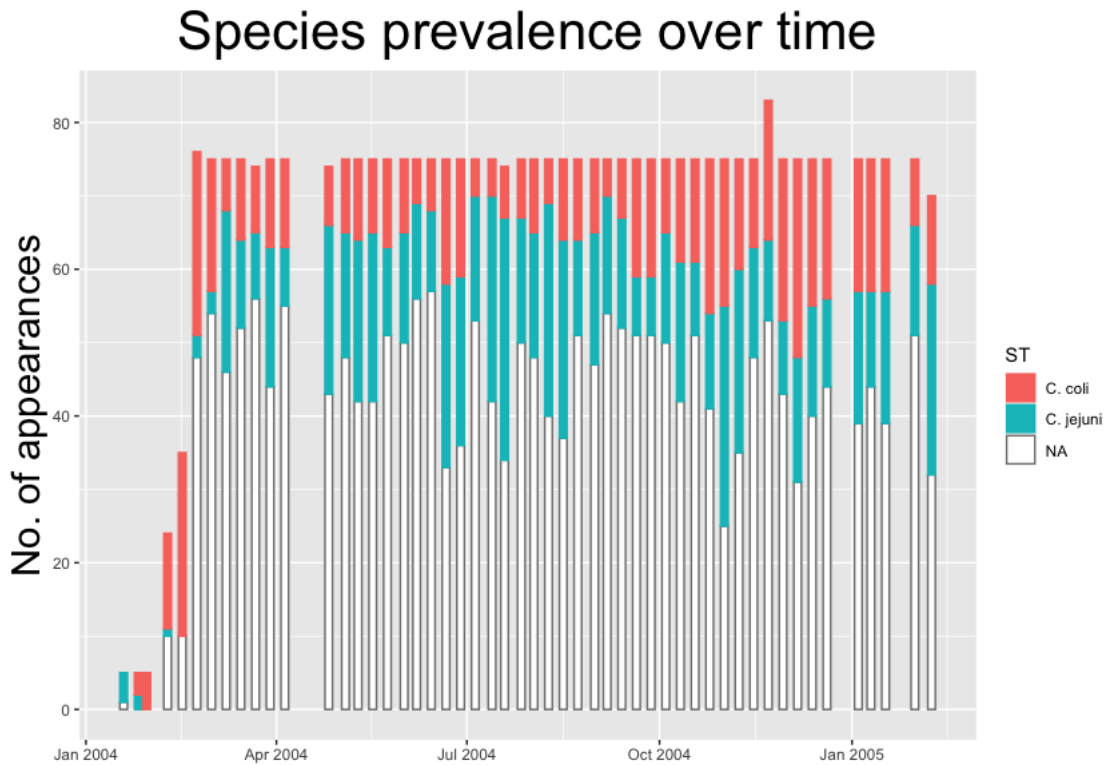


Figure 3.1: Histogram showing the count of positive and negative samples from a breeder flock for different species of *Campylobacter*. White ‘NA’ counts represent samples that were negative for *Campylobacter*.

Within each species, multiple STs are recorded. In Figures 3.2 and 3.3 below we plot the five week moving averages of total positive samples for each species. Beneath each point we plot a histogram showing how this average is split between the competing STs.

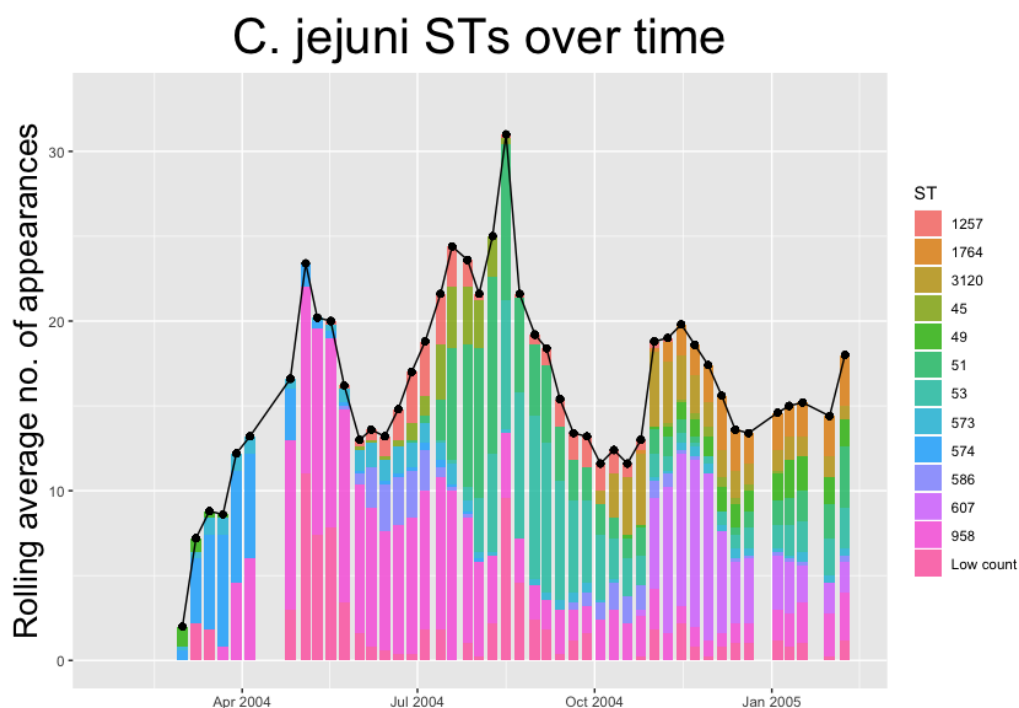


Figure 3.2: The five-week rolling average number of positive samples for *Campylobacter jejuni*, with both the total number and separate ST averages. STs that appear less than twenty times throughout the entire experiment are amalgamated into a group "Low Count".

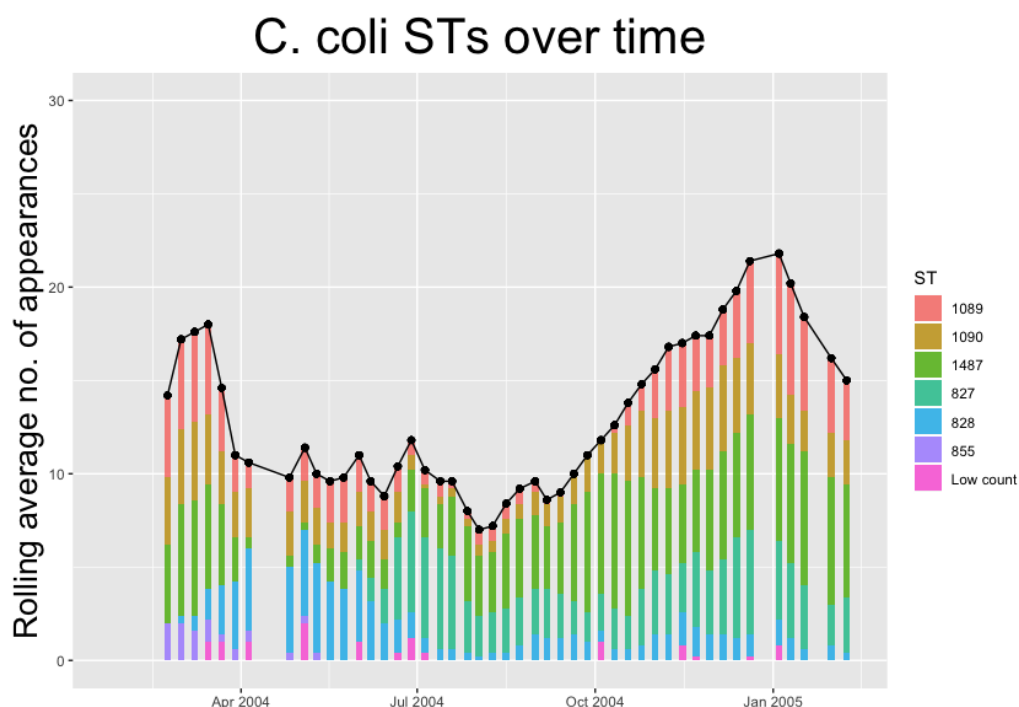


Figure 3.3: The five-week rolling average number of positive samples for *Campylobacter coli*, with both the total number and separate ST averages. STs that appear less than ten times throughout the entire experiment are amalgamated into a group "Low Count".

We notice from Figures 3.2 and 3.3 that there are more unique STs of *C. jejuni* than *C. coli*, despite both species existing at roughly equal levels. We also see that *C. jejuni* incidence appears to peak in the summer, around the August period, coinciding with a dip in the population of *C. coli* STs. Within each species we can observe that different ST populations grow and shrink across the study period. For example, within Figure 3.2 we see that the summer peak is dominated by the prevalence of ST 51 and 53, however by November/December, this population shrinks, and instead ST 607 rapidly increases in population.

Figures 3.2 and 3.3 effectively illustrate the key research questions tackled by this study. Namely, why do some STs seem to exist at higher quantities and persevere better than other STs which may die out? Do the dynamical behaviours of species and STs correlate to any particular trait? We investigate what mechanisms are dynamically driving these observed differences through querying the probability of chickens transitioning from different states of colonisation using a series of Bayesian models presented below.

Model Development

In this section we discuss the general methodology behind all of our models. A general step-by-step process to model formulation is also presented in Box 3.1. Each model begins by classifying each of the datapoints into certain state labels. For example, at the simplest level each reading can be classified as either “State 1: Uncolonised” or “State 2: Colonised”. Other models may use more states to further distinguish colonisations by species or ST. After doing this, we are able to convey this classification data in the form of a matrix $S[c, t]$ where $c \in \{1, 2, \dots, 200\}$ is the index denoting which chicken is considered, and $t \in \{1, 2, \dots, 51\}$ is the index denoting which week is considered. Therefore each element of S will be a number conveying the state classification of that particular data point. For example, $S[3, 7] = 1$, would indicate that on week 7, chicken number 3 was classified as state 1: uncolonised. Because only 75 of the 200 chickens were tested at random each

week, many of these matrix elements are undefined, and as such are marked as ‘NA’.

Once the matrix is defined, each model uses a Bayesian process to find the transition probabilities between these states. Formally we seek the matrix π , where $\pi_{i,j} = P(S[m, n] = j \mid S[m, n - 1] = i)$, for every $m \in \{1, 2, \dots, 200\}$ and $n \in \{2, 3, \dots, 51\}$. In short, $\pi_{i,j}$ is the probability that a chicken moves from state i to state j across a week. The exact choice of how to formulate the expressions is where our models vary, as different formulations are able to investigate different relationships governing these transition probabilities. For example, at the simplest level, we could define

$$\pi_{1,1} = \alpha_1, \tag{3.1}$$

$$\pi_{1,2} = 1 - \pi_{1,1},$$

$$\pi_{2,1} = \alpha_2,$$

$$\pi_{2,2} = 1 - \pi_{2,1},$$

where we seek to find the values $\alpha_1 \in [0, 1]$ and $\alpha_2 \in [0, 1]$ that best fit the data S . Note that we have bounded $\pi_{i,j}$ between 0 and 1, as each value represents a probability. Likewise each row of π must sum to 1, as these probabilities cover all transition possibilities. In the example of equations (3.1) above, when starting from state 1, one can transition to state 2 ($\pi_{1,2}$), or remain in state 1 ($\pi_{1,1}$), hence $\pi_{1,1} + \pi_{1,2} = 1$. Different models below will use more complex definitions for π to explore the impact of time, density dependence, and chicken health on transitions between different states.

A Bayesian statistical model provides a way to iteratively deduce parameters of interest in regards to given data. The process is derived from Bayes’ theorem:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}, \tag{3.2}$$

where θ is the parameter/s we wish to discover, and D is the data provided. In short, equation (3.2) reads that when starting from an initial, **prior**, belief in what values

θ may take ($P(\theta)$), one may obtain an updated, **posterior**, probability distribution for these possible values given some provided data ($P(\theta|D)$). A more thorough introduction to Bayesian modelling is provided in Appendix B1. In our case, the parameters we seek, θ , are the ones used in our definition of π , such as α_1 and α_2 in the example above. The data, D , we use is the matrix S .

Box 3.1 - Model construction process

1. Decide state classifications.

Choose how data should be classified, and construct matrix S containing all state classifications for each data point.

2. Decide formulation of transition matrix.

Choose how model will define transition probabilities and dependencies.

3. Run Bayesian model.

Define prior probability distributions for model parameters. Program and run Bayesian model using JAGS, to acquire a posterior probability distribution for all model parameters defined in step 2.

4. Assess convergence.

Investigate model output to assure posterior distribution is well-constructed and has converged.

5. Present results.

Plot the transition probabilities, $\pi_{i,j}$, and interpret the results.

Below we present a series of case studies presenting our different models and their results. All models were run using JAGS (Plummer, 2007) from within R using the `run.jags` package (Plummer et al., 2016). All code used for the following models is made available at <https://osf.io/m5yua/>.

Case Studies

Model 1: Time dependence

Our first model investigates how time affects the transition probabilities between states. Following the process outlined in Box 3.1, we choose to initially classify our

data as one of two states: “state 1: uncolonised” and “state 2: colonised”.

To assess how the transition probabilities vary through time we must ensure that we define our transition probabilities such that they depend on time. One way would be to adapt equations (3.1) above such that $\pi_{1,1}$ was a function of $\alpha + \beta t$. However, this would impose structure upon the transition probabilities, enforcing them to change linearly with time. Ideally a model formulation should allow as much freedom as possible to fit to the data. As such, we shall instead construct π as a three-dimensional array. In essence this means that each time period can be described by its own transition matrix. Formally we write this as,

$$\pi_{1,2,t} = \text{ilogit}(\alpha_1 + C_1[t]), \quad (3.3)$$

$$\pi_{1,1,t} = 1 - \pi_{1,2,t},$$

$$\pi_{2,1,t} = \text{ilogit}(\alpha_2 + C_2[t]),$$

$$\pi_{2,2,t} = 1 - \pi_{2,1,t},$$

for $t \in \{1, 2, \dots, 51\}$. Here $\text{ilogit}()$ is the inverse logit function defined by $\text{ilogit}(x) = \frac{e^x}{1+e^x}$. This function is bounded between 0 and 1, scaling the argument so that our probabilities, $\pi_{i,j,t}$ remain correctly bounded. The underlying theory is that we assume there is some mean probability for $\pi_{i,j,t}$ across all t . These mean probabilities are described by α_1 and α_2 . We then assume that, for each t , there is some “correction term” away from the mean unique to each week. These correction terms are captured by $C_1[t]$ and $C_2[t]$ for each t .

Now that we have decided on our model formulation, we move to step 3 and run the model to find the posterior distributions for α_1 , α_2 , C_1 and C_2 . First we define our prior probability distributions for each of the model parameters. This distribution represents our initial assumptions on what value our variables may take, and is often informed by expert opinion. Since we do not have any initial assumptions on what values our variables may take, we use wide noninformative priors. For α_1 and α_2 we choose a prior distribution of $U(0, 25)$ for each, a uniform

distribution between 0 and 25. For C_1 and C_2 , we wish each element of these vectors to be a small perturbation away from the mean of α_1 or α_2 . As such, we would ideally have these elements drawn from a normal distribution with mean 0, and some, yet to be determined, standard deviation. This represents a hierarchical model formulation (discussed further in Appendix B1), where we instead define priors on the two standard deviations for these two normal distributions associated with C_1 and C_2 . Following the advice of Gelman (2006) for noninformative improper priors, we use a uniform distribution between 0 and 50 for the prior distribution of each of these standard deviation parameters. The model was then run using two chains, with a burn-in period of 5,000 iterations, and then a final sample of 25,000 iterations to build the posterior distributions.

Convergence was considered well-achieved via investigation of the trace plots of the chains, the effective sample size (ESS) and Monte Carlo Standard Error (MCSE) of the variables. The Gelman-Rubin statistic, or ‘shrink factor’, is the most commonly used metric for convergence, with a value close to 1 signifying effective convergence. Heuristically, any shrink factor below 1.1 is considered by Kruschke (2014) to signify sufficient convergence. The presented model run resulted in a multivariate potential scale reduction factor (mpsrf) of 1.0059.

The results for this model are presented below in Figure 3.4. The median values of the transition probabilities for (3.4A) $\pi_{1,1,t}$, (3.4B) $\pi_{1,2,t}$, (3.4C) $\pi_{2,1,t}$, (3.4D) $\pi_{2,2,t}$ are plotted, and a linear regression is fit to these outputs using the `lm` function in R. Fitting a general additive model (GAM) to these median values revealed that there was no significant model fit for higher order models, hence only linear regression fits are displayed. The probability of transitioning from state 1 (plots 3.4A and 3.4B) was not significantly correlated against time (t-test, $p = 0.135$), however transitions from state 2 (plots 3.4C and 3.4D) against time were statistically significant (t-test, $p < 0.01$).

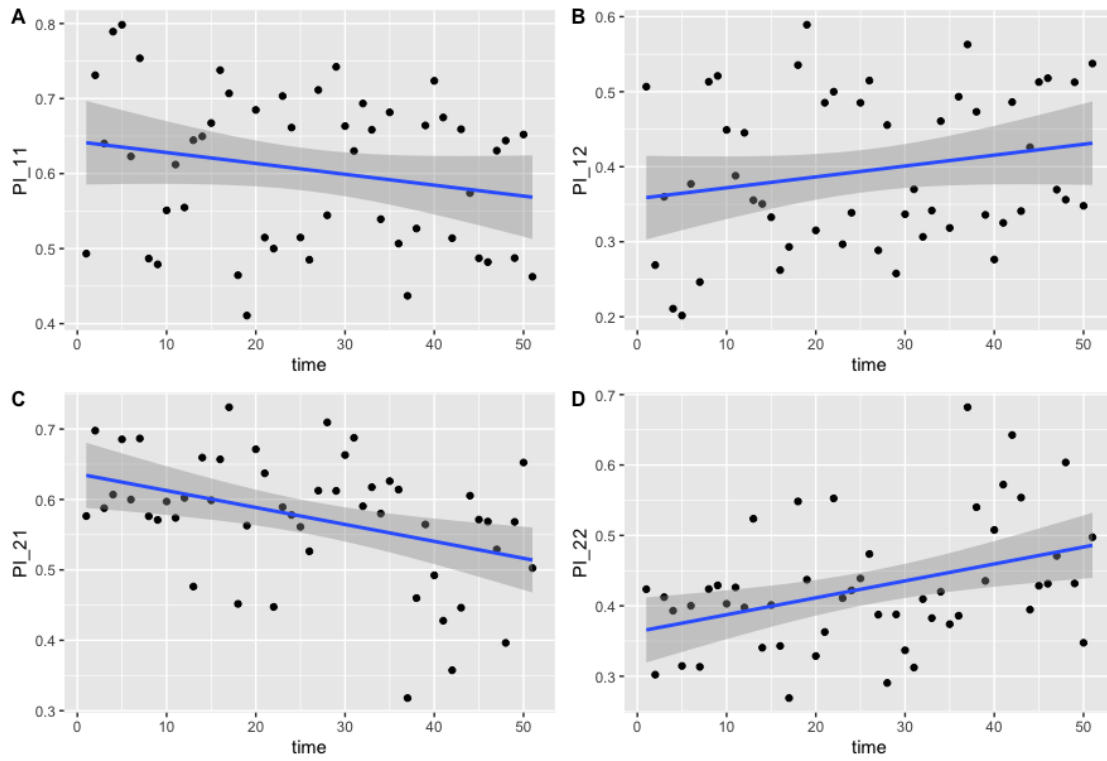


Figure 3.4: Transition probabilities between two states, ‘uncolonised’ and ‘colonised’. Plots show (A) $\pi_{1,1,t}$, (B) $\pi_{1,2,t}$, (C) $\pi_{2,1,t}$ and (D) $\pi_{2,2,t}$ against time. Each point is the calculated transition probability for that time point. Also plotted is a linear regression against these points in blue, with a shaded region depicting the 95% confidence interval of the regression. (C) and (D) are significant ($p < 0.01$).

These findings suggest that, as time progresses, colonised chickens become more likely to remain colonised, and similarly become less likely to clear such a colonisation. Figures 3.4C and 3.4D show that, at the start of the experiment, colonised chickens would be more likely to clear the colonisation the following week, but by the end of the experiment this had reduced to a probability of roughly 50%.

Model 2: Species dependence

For the next model we investigate transition differences between the two species present in the study; *C. jejuni* and *C. coli*. As such, this time we classify our data as belonging to one of three states; ‘state 1: uncolonised’, ‘state 2: colonised by *C. jejuni*’ and ‘state 3: colonised by *C. coli*’. Therefore our transition matrix will be of size 3×3 . We define each row of the transition matrix by a 3-variable

Dirichlet distribution (the multivariate generalisation of the Beta distribution), ensuring each row sums to 1. As such, we infer the transition probabilities directly, using prior distributions of

$$(\pi_{1,1}, \pi_{1,2}, \pi_{1,3}) \sim \text{Dirichlet}(1, 1, 1),$$

$$(\pi_{2,1}, \pi_{2,2}, \pi_{2,3}) \sim \text{Dirichlet}(1, 1, 1),$$

$$(\pi_{3,1}, \pi_{3,2}, \pi_{3,3}) \sim \text{Dirichlet}(1, 1, 1).$$

The model was run with two chains and an initial burn-in period of 5,000 iterations. Posterior distributions were built from a sample of 10,000 iterations. Convergence was once again well-achieved with a mpsrf of 1.0035. The results are plotted below in Figure 3.5. Results show slight variations between species across the

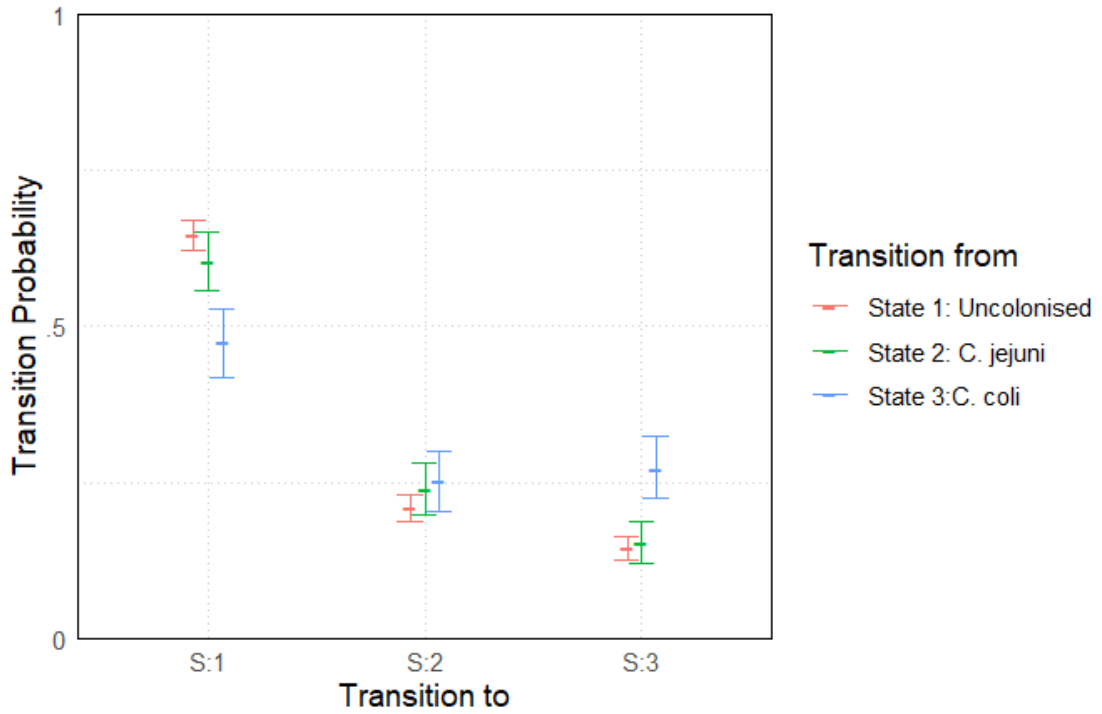


Figure 3.5: Transition probabilities between three states, ‘uncolonised’, ‘colonised by *C. jejuni*’ and ‘colonised by *C. coli*’. Plots show the median values of the posterior distributions and the 95% highest density intervals (HDIs).

entire experiment. General transition probabilities from each state are very similar, however one can note that a chicken is more likely to be colonised by *C. coli* when transitioning from a state of already being colonised by *C. coli*. We also see

that a chicken colonised by *C. coli* is less likely to transition to being uncolonised than a chicken colonised by *C. jejuni*.

Model 3: Time and species dependence

We now combine the previous two models together, to investigate how the transitions between species alter across time. We once again therefore classify our data into three categories, as per the previous model.

We will be constructing a three-dimensional array once again for our transition probabilities, with each time period being described by a separate 3×3 transition matrix. To ensure each row of these matrices sums to 1, we start by framing the transition probabilities as an unbounded array p , before scaling these into our final array π . p is defined as

$$\begin{aligned}
 p_{1,1,t} &= \exp(\alpha_1), & p_{1,2,t} &= \exp(\alpha_2 + C_1[t]), & (3.4) \\
 p_{1,3,t} &= \exp(\alpha_3 + C_2[t]), & p_{2,1,t} &= \exp(\alpha_4), \\
 p_{2,2,t} &= \exp(\alpha_5 + C_3[t]), & p_{2,3,t} &= \exp(\alpha_6 + C_4[t]), \\
 p_{3,1,t} &= \exp(\alpha_7), & p_{3,2,t} &= \exp(\alpha_8 + C_5[t]), \\
 p_{3,3,t} &= \exp(\alpha_9 + C_6[t]).
 \end{aligned}$$

Like in our initial model, our α parameters will describe the average transition value across time, with the C parameters describing a small perturbation away from this mean. The exponential function ensures that values of $p_{i,j,t}$ will be positive. C values only need to be implemented on two probabilities in each row, as we will next scale these so that each row sums to 1, meaning that two free correction terms are sufficient

to describe the distribution of the row. Our scaling is then performed like so,

$$\begin{aligned}
 \pi_{1,1,t} &= \frac{p_{1,1,t}}{p_{1,1,t} + p_{1,2,t} + p_{1,3,t}}, & \pi_{1,2,t} &= \frac{p_{1,2,t}}{p_{1,1,t} + p_{1,2,t} + p_{1,3,t}}, \\
 \pi_{1,3,t} &= \frac{p_{1,3,t}}{p_{1,1,t} + p_{1,2,t} + p_{1,3,t}}, & \pi_{2,1,t} &= \frac{p_{2,1,t}}{p_{2,1,t} + p_{2,2,t} + p_{2,3,t}}, \\
 \pi_{2,2,t} &= \frac{p_{2,2,t}}{p_{2,1,t} + p_{2,2,t} + p_{2,3,t}}, & \pi_{2,3,t} &= \frac{p_{2,3,t}}{p_{2,1,t} + p_{2,2,t} + p_{2,3,t}}, \\
 \pi_{3,1,t} &= \frac{p_{3,1,t}}{p_{3,1,t} + p_{3,2,t} + p_{3,3,t}}, & \pi_{3,2,t} &= \frac{p_{3,2,t}}{p_{3,1,t} + p_{3,2,t} + p_{3,3,t}}, \\
 \pi_{3,3,t} &= \frac{p_{3,3,t}}{p_{3,1,t} + p_{3,2,t} + p_{3,3,t}}.
 \end{aligned} \tag{3.5}$$

We choose priors of $N(0, 1000)$ for all our α values (normal distributions with mean 0 and standard deviation 1000). Like the first model, we shall construct a hierarchical dependency such that our $C_i[t]$ are all drawn from a normal distribution for each t . Motivated by the correlation observed in the first model, we actually set these six C_i terms to all be drawn from a 6-variable multivariate normal distribution, with mean $(0, 0, 0, 0, 0, 0)$ and a covariance matrix as our parameter to be defined. JAGS requires the input of a precision matrix (the inverse of the covariance matrix) for its formulation of the multivariate normal distribution, so we set a prior distribution on the precision matrix of $\text{Wishart}(I_6, 6)$, where I_6 is the 6×6 identity matrix.

The model was run with two chains for an initial burn-in period of 5,000 iterations, and then a posterior distribution was built from a sample of 250,000 iterations, thinned at a rate of 1 in 5, meaning only 1 in every 5 iterations was used for the posterior distribution so as to reduce autocorrelation. Results are plotted below in Figure 3.6.

Of the nine transition probabilities presented, five were found to be statistically significant for correlation against time when a linear regression was applied: $\pi_{1,3,t}$, $\pi_{2,1,t}$, $\pi_{2,3,t}$, $\pi_{3,1,t}$ and $\pi_{3,3,t}$ (t-tests, $p < 0.0005$, $p < 0.05$, $p < 0.05$, $p < 0.0005$ and $p < 0.0005$ respectively). We once again see that transitions to a state of uncolonisation reduce over time, however, whereas Model 1 reported overall

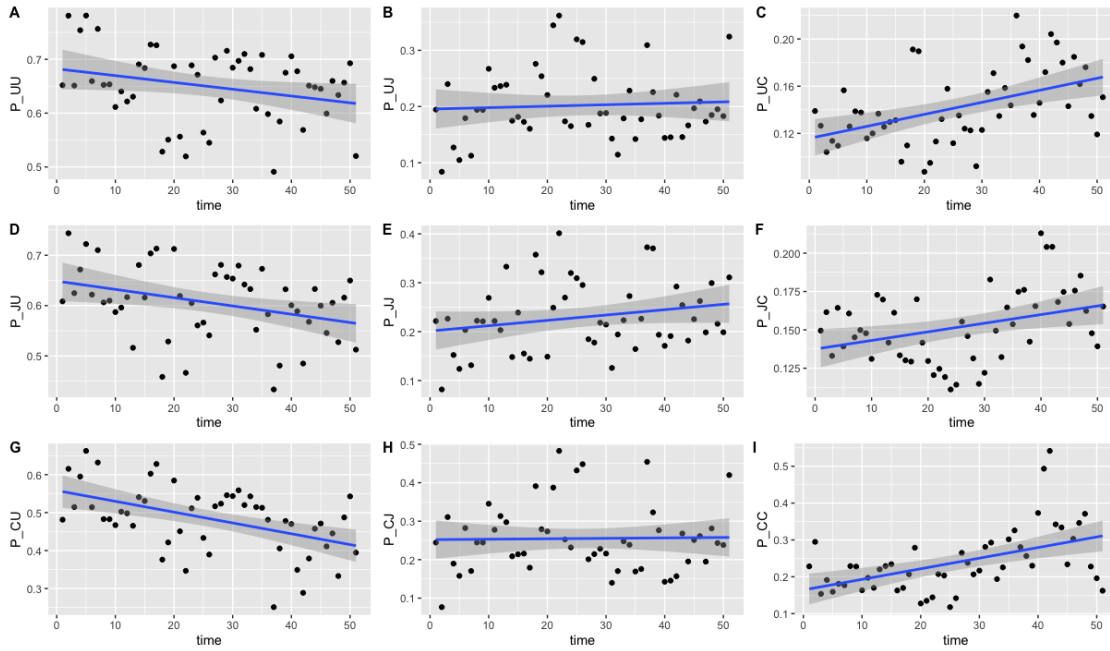


Figure 3.6: Transition probabilities between three states, ‘uncolonised’, ‘colonised by *C. jejuni*’ and ‘colonised by *C. coli*’. Plots show (A) $\pi_{1,1,t}$, (B) $\pi_{1,2,t}$, (C) $\pi_{1,3,t}$, (D) $\pi_{2,1,t}$, (E) $\pi_{2,2,t}$, (F) $\pi_{2,3,t}$, (G) $\pi_{3,1,t}$, (H) $\pi_{3,2,t}$ and (I) $\pi_{3,3,t}$ against time. Each point is the calculated transition probability for that time point. Also plotted is a linear regression against these points in blue, with a shaded region depicting the 95% confidence interval of the regression. Five transition probabilities were found to be statistically significant for correlation against time: $\pi_{1,3,t}$, $\pi_{2,1,t}$, $\pi_{2,3,t}$, $\pi_{3,1,t}$ and $\pi_{3,3,t}$ (t-tests, $p < 0.0005$, $p < 0.05$, $p < 0.05$, $p < 0.0005$ and $p < 0.0005$ respectively).

transitions to a state of colonisation increasing, Model 3 shows that only transitions to colonisation by *C. coli* increase over time. Given the spread of the data in Figure 3.6, we also tested for statistical significance against a quadratic regression. A quadratic fit would be a strong argument for the existence of seasonal variation, by capturing a difference in the middle of the time series as the time axis moves to summer, before returning to winter. Recall again that this time period plotted is in weeks from February 2004 to February 2005. Only one transition probability was found to fit a quadratic regression with statistical significance, the transition from colonisation by *C. jejuni* to *C. coli*, $\pi_{2,3,t}$ (t-test, $p < 0.05$). This quadratic regression is presented in Figure 3.7 below. This would correlate with the behaviour observed in Figures 3.2 and 3.3, whereby *C. jejuni* appears to be most prevalent in the summer, and *C. coli* most prevalent in the winter. (Similarly to Model 1,

fitting a general additive model (GAM) to these median values revealed that there was no significant model fit for higher order models, hence only linear regression fits, and the one quadratic regression fit, are displayed.)

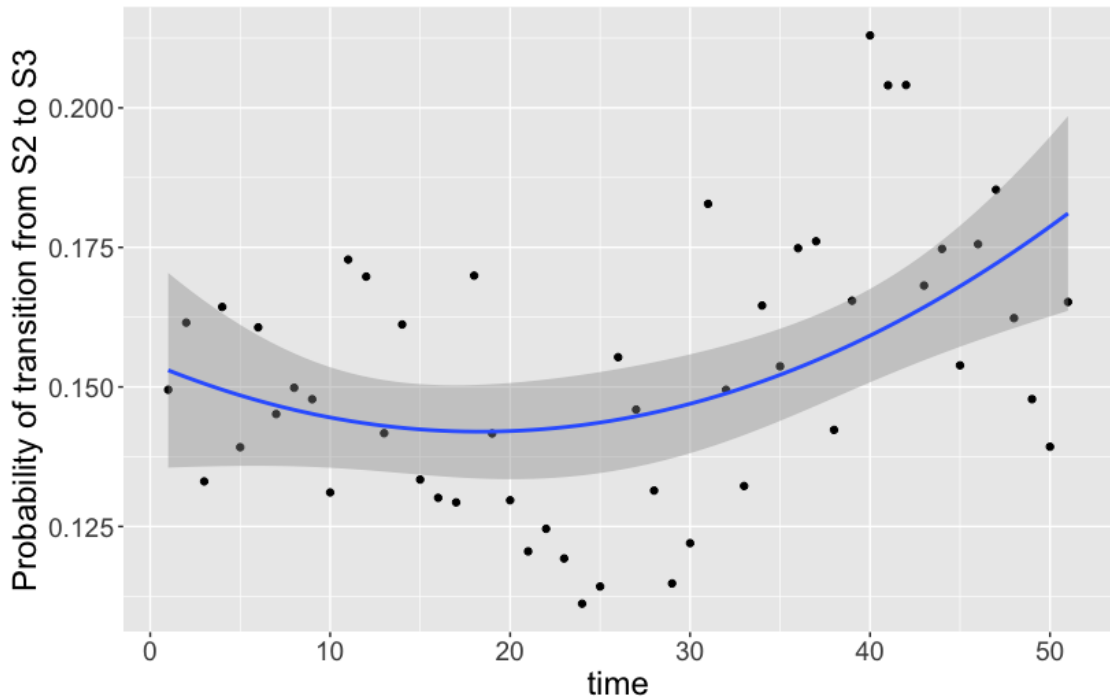


Figure 3.7: Transition probabilities between state 2 ‘colonised by *C. jejuni*’ and state 3 ‘colonised by *C. coli*’ against time. Each point is the calculated transition probability for that time point. Also plotted is a quadratic regression against these points in blue, with a shaded region depicting the 95% confidence interval. The transition probability was found to be statistically significant for correlation against time (t-test, $p < 0.05$).

Model 4: ST perseverance

For this model, we extend Model 2 to now capture species-specific ST perseverance within a chicken. To do this, we re-classify the data into five different states: ‘S1: uncolonised’, ‘S2: new *C. jejuni* ST’, ‘S3: same *C. jejuni* ST as previous week’, ‘S4: new *C. coli* ST’ and ‘S5: same *C. coli* ST as previous week’. To further clarify the meaning of state 2 and state 4, we mean a ST of either *C. jejuni* or *C. coli* that was not present in the previous week for the chicken in question. For example, if one chicken had the following colonisation data for ten days: {“Uncolonised”, “Colonised

by *C. coli* ST 1089”, “Colonised by *C. coli* ST 1090”, “Colonised by *C. coli* ST 1090”, “NA”, “Colonised by *C. coli* ST 1090”, “Colonised by *C. jejuni* ST 958”, “Colonised by *C. jejuni* ST 958”, “Colonised by *C. jejuni* ST 1257”, “Uncolonised”}, then this row of ten would be classified as $\{1, 4, 4, 5, \text{NA}, 4, 2, 3, 2, 1\}$. Because, by definition, one can only transition to state 3 from state 2 or state 3, we can fix $\pi_{1,3} = \pi_{4,3} = \pi_{5,3} = 0$, and likewise for transitions to state 5: $\pi_{1,5} = \pi_{2,5} = \pi_{3,5} = 0$. The non-zero transition probabilities can then be calculated by drawing each row from a 3 or 4 variable Dirichlet distribution. Formally we set a prior on each row of,

$$\begin{aligned} (\pi_{1,1}, \pi_{1,2}, \pi_{1,4}) &\sim \text{Dirichlet}(1, 1, 1), \\ (\pi_{2,1}, \pi_{2,2}, \pi_{2,3}, \pi_{2,4}) &\sim \text{Dirichlet}(1, 1, 1, 1), \\ (\pi_{3,1}, \pi_{3,2}, \pi_{3,3}, \pi_{3,4}) &\sim \text{Dirichlet}(1, 1, 1, 1), \\ (\pi_{4,1}, \pi_{4,2}, \pi_{4,4}, \pi_{4,5}) &\sim \text{Dirichlet}(1, 1, 1, 1), \\ (\pi_{5,1}, \pi_{5,2}, \pi_{5,4}, \pi_{5,5}) &\sim \text{Dirichlet}(1, 1, 1, 1). \end{aligned} \tag{3.6}$$

The model was run with 2 chains for a burn-in period of 5,000 iterations before building posteriors from a final sample of 10,000 iterations. Chains were well-mixed and convergence well-achieved with an mpsrf of 1.0037. Results are plotted below in Figure 3.8.

The most notable difference is seen in the perseverance of *C. coli* STs compared to *C. jejuni* STs. Comparing columns 3 and 5 of Figure 3.8, we see that, a colonisation by a new ST of either *C. coli* or *C. jejuni* have a roughly equal chance of persevering to the next week. However, once a ST has carried over for one week, *C. coli* colonisations are then considerably more likely to further persist for later weeks. In fact, a repeated instance of colonisation by a *C. coli* ST (state 5) is more likely to continue in subsequent weeks than to transition to any other state (seen by comparing the pink lines in Figure 3.8). Comparing also columns 2 and 3 of Figure 3.8, we see that transitions to colonisations by new *C. coli* / *C. jejuni* STs are roughly comparable, meaning that the primary difference we observe between the two species is in perseverance as opposed to infectivity.

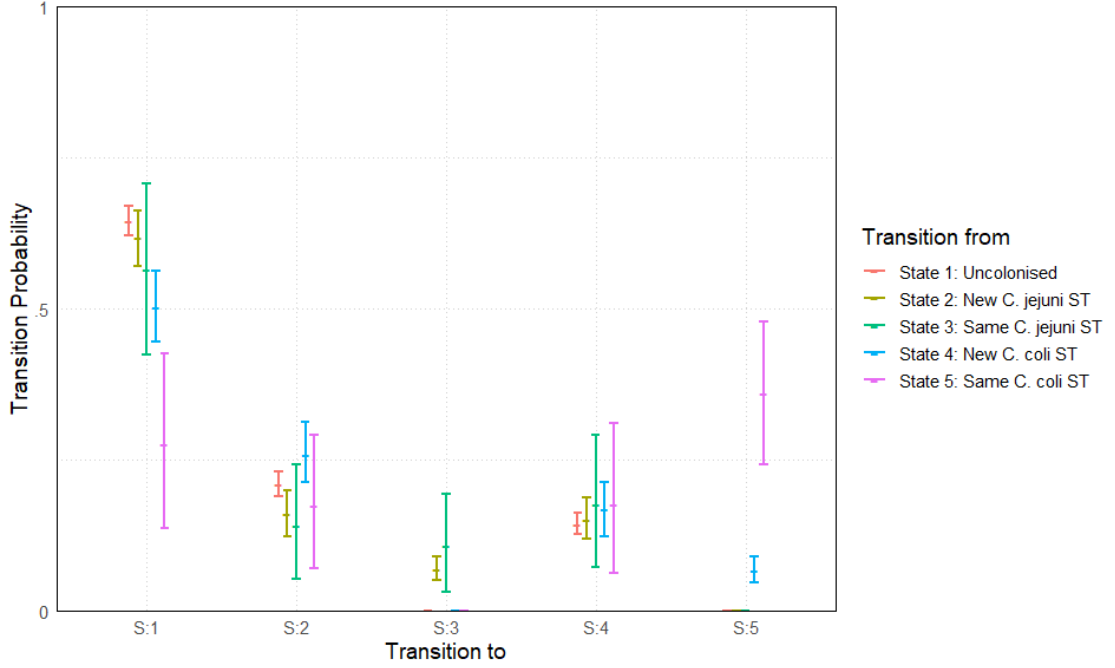


Figure 3.8: Transition probabilities between five states, ‘uncolonised’, ‘colonised by a new *C. jejuni* ST’, ‘colonised by the same *C. jejuni* ST as previously’, ‘colonised by a new *C. coli* ST’ and ‘colonised by the same *C. coli* ST as previously’. Plots show the median values of the posterior distributions and the 95% highest density intervals (HDIs).

Model 5: Chicken dependence

Whereas Model 1 considered how transition probabilities vary across time, we now consider how transition probabilities vary across different chickens. We follow a very similar framework to Model 1, beginning by classifying all data as one of two states: ‘S1: uncolonised’ or ‘S2: colonised’. We then, like Model 1, consider some average transition probability that each chicken is close to, and then consider some small “correction term” unique to each chicken, which may make them more or less likely to transition to a certain state. Formally, we write,

$$\pi_{1,2,t} = \text{ilogit}(\alpha_1 + C_1[c]), \quad (3.7)$$

$$\pi_{1,1,t} = 1 - \pi_{1,2,t},$$

$$\pi_{2,1,t} = \text{ilogit}(\alpha_2 + C_2[c]),$$

$$\pi_{2,2,t} = 1 - \pi_{2,1,t},$$

for $c \in \{1, 2, \dots, 200\}$. We set a noninformative prior distribution for α_1 and α_2 of $N(0, 1000)$. Our chicken correction terms, $C_1[c]$ and $C_2[c]$, are each drawn from a two-variable multivariate normal distribution for each c , with mean $(0, 0)$ and covariance matrix to be calculated. Like described in Model 3, we therefore set a prior distribution on the precision matrix for this multi-variate normal distribution of $\text{Wishart}(I_2, 2)$, where I_2 is the 2×2 identity matrix.

The model was run with two chains for an initial burn-in period of 20,000 iterations, before posteriors were then constructed from a sample of 50,000 iterations. Convergence was well-achieved, with all chains well-mixed and all parameters sampled with a high ESS and $\text{MCSE} < 0.01$. The mpsrf was unable to be calculated due to the high number of stochastic nodes, however there were no signs to suggest invalid convergence.

Upon calculating our transition probabilities for each bird, we plot the values for $\pi_{1,2}$ against the value of $\pi_{2,1}$ for each bird and investigate the correlation. Figure 3.9 shows these results overlaid with a contour of the associated multivariate normal distribution, indicating the probability density of the transition probabilities for the flock.

The strong linear relation observed reveals the presence of distinct sub-groups within the flock of birds who are colonised often, and those who are colonised very rarely.

Model 6: Chicken and species dependence

We now alter the previous model to consider the differences in transition between species of *Campylobacter* across all birds. As such, the data is instead classified into the three states: ‘state 1: uncolonised’, ‘state 2: colonised by *C. jejuni*’ and ‘state 3: colonised by *C. coli*’. This model is formulated the same way as in Model 3 above. The transition probabilities follow the same structure as equations (3.4) and (3.5), except that our correction terms $C_i[c]$ are corrections for each chicken

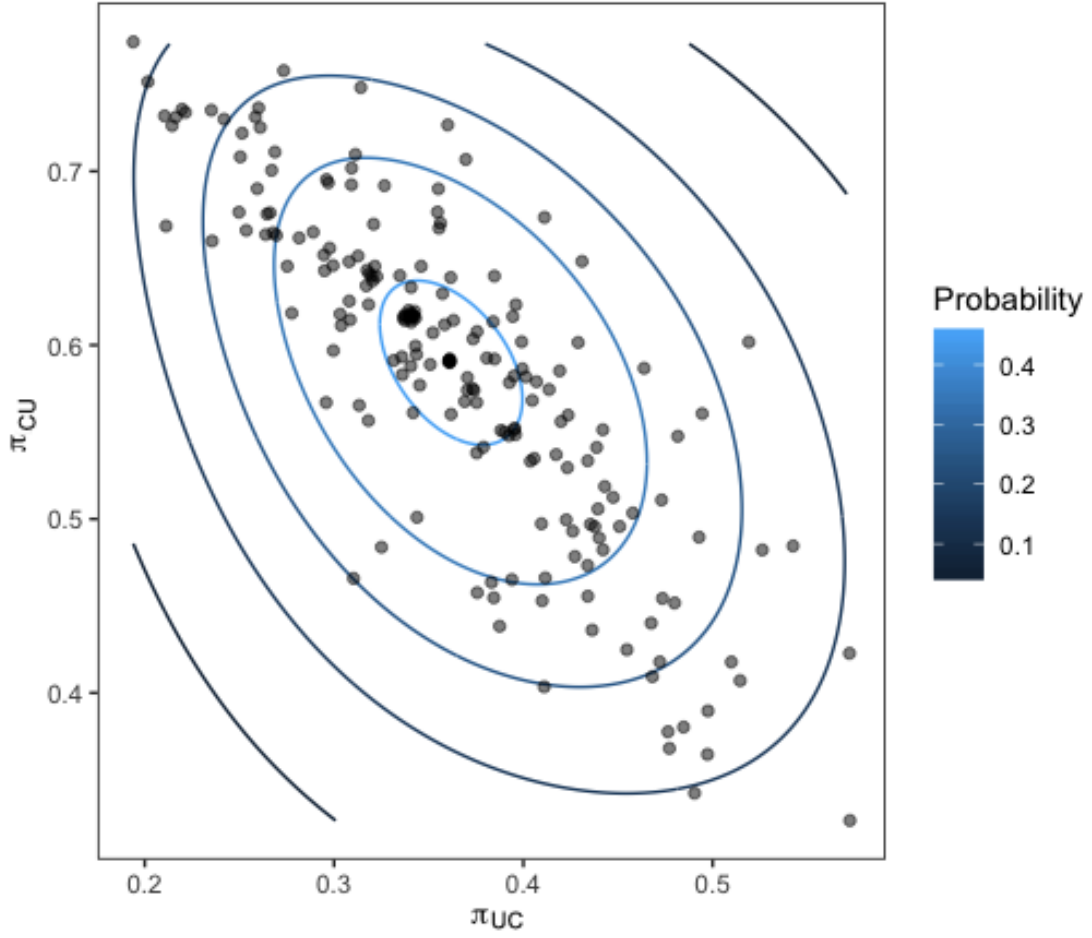


Figure 3.9: Transition probabilities for each bird in the flock from a state of being colonised to uncolonised (y-axis) against the transition probability from uncolonised to colonised (x-axis). Contours show the fit of a multivariate normal distribution to the output.

in the flock ($c \in \{1, 2, \dots, 200\}$) as opposed to each time step. As such we craft a 3×3 transition matrix for each chicken. A prior distribution of $N(0, 1000)$ is used for each α_i parameter, and the six chicken correction terms, $C_i[c]$ are drawn from a six-variate multivariate normal distribution for each c , with mean $(0, 0, 0, 0, 0, 0)$ and a precision matrix as a parameter to find. The prior distribution for this precision matrix is $\text{Wishart}(I_6, 6)$, where I_6 is the 6×6 identity matrix.

The model was run with two chains for an initial burn-in period of 10,000 iterations, before posterior distributions were constructed from a sample of 50,000 iterations,

thinned at a rate of 1 in 25, meaning only one iteration was kept in every 25.

The idea of this model is to assess how bird variation affects the transition of each species of *Campylobacter*. The previous model revealed the existence of variation in bird resistance to colonisation throughout the flock. Figure 3.10 below plots the result of multiple transition probabilities against one-another. Each point on the graphs represents the transition probabilities for a specific chicken. Plots 3.10A to 3.10C use $\pi_{1,1,c}$, the transition from uncolonised to uncolonised as the y-axis. This acts as a rough metric for “bird resilience to colonisation”, as the more resistant birds are more likely to continue being uncolonised. As such, plots 3.10A to 3.10C depict how transitions related to each species vary according to host bird susceptibility. Plot 3.10D uses $\pi_{3,3,c}$, the transition from *C. coli* to *C. coli* as the y-axis, to compare how the perseverance of *C. coli* affects the colonising ability of *C. jejuni*. Linear regressions are fit to all plots in Figure 3.10, and all were found to be statistically significant (t-test, $p < 0.0001$).

Interestingly, the gradients of all the shifting transition probabilities are different between species, confirming that, indeed, the transition probabilities of each species varies differently across chickens. We see that the probability of a species persisting unsurprisingly increases as bird susceptibility increases, but curiously our linear regressions for each species overlap. This result indicates that, in the more resilient birds, *C. coli* is less likely to persevere than *C. jejuni* colonisations, however the inverse is seen in the more susceptible birds.

Model 7: Chicken and density dependence

This model builds on Model 5 by now considering how transition probabilities are affected by the number of total colonisations in the previous week. *Campylobacter* is known to be transmitted via the faecal-oral route between chickens, so it seems likely that a higher density of colonisations one week will cause an increased number

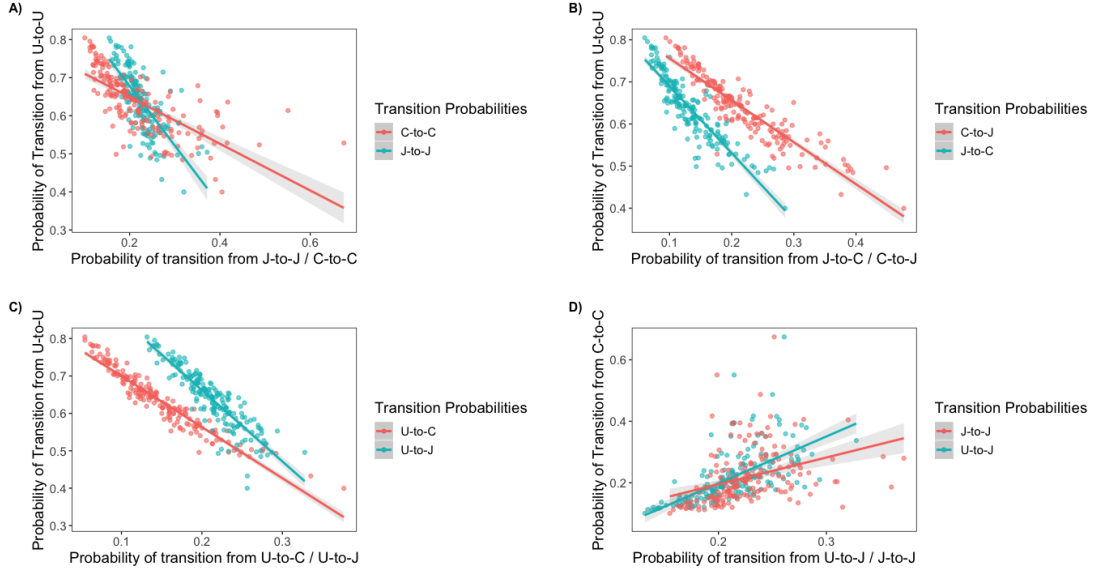


Figure 3.10: Transition probabilities for a three state system. In these plots, ‘U’ refers to being uncolonised, ‘J’ refers to colonisation by *C. jejuni* and ‘C’ refers to colonisation by *C. coli*. Each of the points is the transition probability for a specific bird within the flock. Linear regression fits are plotted with a shaded region representing the 95% confidence intervals of the regression. All regressions were statistically significant (t-test $p < 0.0001$). (A) The transition probabilities of J-to-J and C-to-C against the transition probability of U-to-U. (B) The transition probabilities of C-to-J and J-to-C against the transition probability of U-to-U. (C) The transition probabilities of U-to-J and U-to-C against the transition probability of U-to-U. (D) The transition probabilities of J-to-J and U-to-J against the transition probability of C-to-C.

of colonisations the following week. We classify our data into two states, uncolonised and colonised.

The model formulation is then as follows,

$$\begin{aligned}
 \pi_{1,2,c,t} &= \text{ilogit} \left(\alpha_1 + C_1[c] + \beta_1 \left(\sum_{i=1}^{51} \frac{S[i,t] - 1}{N_t} \right) \right), \\
 \pi_{1,1,c,t} &= 1 - \pi_{1,2,c,t}, \\
 \pi_{2,1,c,t} &= \text{ilogit} \left(\alpha_2 + C_2[c] + \beta_2 \left(\sum_{i=1}^{51} \frac{S[i,t] - 1}{N_t} \right) \right), \\
 \pi_{2,2,c,t} &= 1 - \pi_{2,1,c,t},
 \end{aligned} \tag{3.8}$$

where N_t is the number of birds that data is available for at time t . Here, as with

previous models, α_i represents some mean transition probability that all birds are clustered around, and $C_i[c]$ represents the slight correction for each bird c . Recall that the matrix S is populated by elements ‘1’ denoting uncolonised and ‘2’ denoting colonised. Therefore the expression $S[i, t] - 1$ for every i and t shifts this to instead be captured as ‘0’ signifying uncolonised, and ‘1’ signifying colonised. Therefore, the expression $\sum_{i=1}^{51} (S[i, t] - 1)$ will be a tally of exactly how many birds are recorded as being colonised at time t . Therefore, the expression $\sum_{i=1}^{51} \frac{S[i, t] - 1}{N_t}$ conveys the exact proportion of how many birds are currently colonised. Note the use of N_t as, for most weeks 75 birds are recorded for every t , however, as can be seen in Figure 3.1, occasionally a few more or less were recorded each week. Note however, that during the Bayesian modelling process, values for each element of S can be imputed in the process, meaning that we can choose to measure our density dependence using either just the provided data, or also the imputed data. There are merits to both approaches, and so results are included for both below. Here then, β_i represents the strength of the density dependent effect.

The model was initialised with prior distributions of $N(0, 1000)$ for all α_i and β_i parameters. The chicken corrections terms $C_i[c]$ were, like above, drawn from a multivariate normal distribution of mean $(0, 0)$ whose precision matrix we seek. The precision matrix was initialised with a prior distribution of $\text{Wishart}(I_2, 2)$ where I_2 is the 2×2 identity matrix. The model was run with two chains for an initial burn-in period of 6,000 iterations and then posterior distributions built from a sample of 25,000 iterations. This was done twice with two variations of the model: one where density dependence is calculated from provided data, and one with the addition of imputed data. The posterior distributions of our model parameters were used to simulate the transition probabilities for each flock across a full range of total flock prevalences. i.e. using the median values for α_i , β_i , C_i and the precision matrix, we are able to build the functions

$$\pi_{1,2,c} = \text{ilogit}(\alpha_1 + C_1[c] + \beta_1 D,) \quad (3.9)$$

$$\pi_{2,1,c} = \text{ilogit}(\alpha_2 + C_2[c] + \beta_2 D,) \quad (3.10)$$

for any value $D \in [0, 1]$, for each chicken c . The results of these functions for both the imputed and non-imputed density models are presented below in Figure 3.11. The data only records flock colonisations proportions ranging from 0.1818 to 0.6667, so dotted lines are placed in Figure 3.11 to show the range beyond which the result was further imputed.

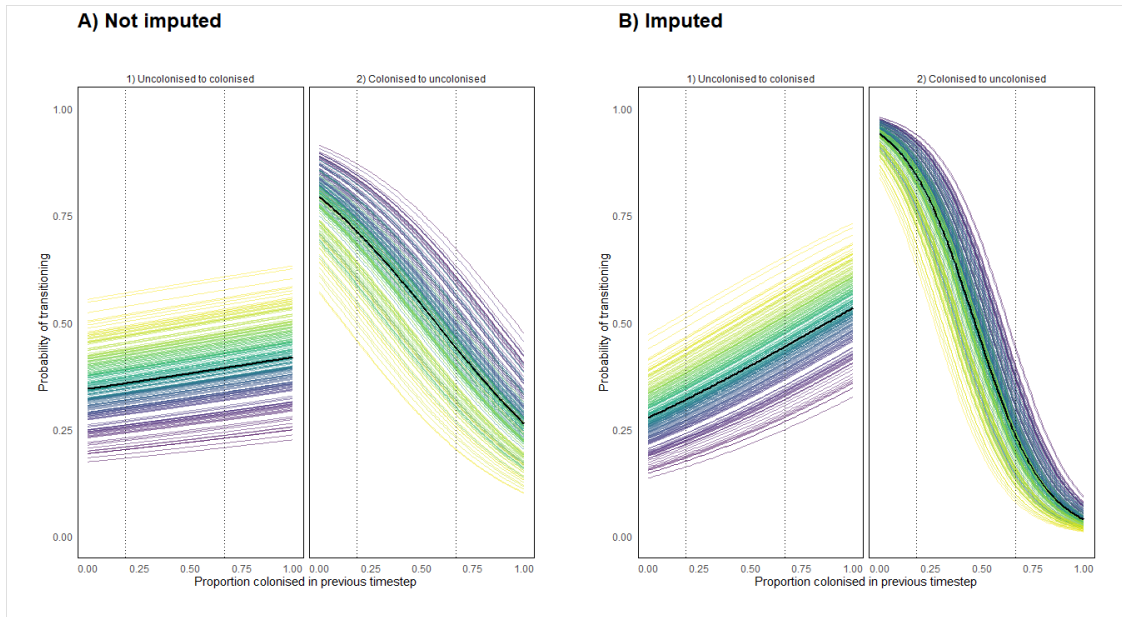


Figure 3.11: Transition probabilities from a state of uncolonisation (by *Campylobacter*) to colonisation, and from colonisation to uncolonisation using a density dependent model programmed using (A) recorded data (B) recorded and imputed data. Each coloured line represents the transition probabilities for a single chicken, with a black line depicting the flock mean. Dotted lines show the region for which data was available for such a flock colonisation proportion.

Importantly Figure 3.11 confirms that density dependence is apparent within the flock. This was an important result to capture to reinforce the findings of Model 6. It confirms that birds are influenced by the colonisation prevalence of

the flock, suggesting that the more resilient birds truly are less likely to become colonised, as opposed to just never becoming exposed to particularly virulent STs. Of interest here is that the probability of clearing colonisation (transitioning to uncolonised) is affected far more by flock prevalence proportion than the probability of becoming colonised.

Discussion

An improved understanding of the transmission dynamics of *Campylobacter* among and within poultry flocks in the commercial environment is essential for the design of effective intervention strategies to reduce levels of human disease. Here, these dynamics were evaluated within a flock of broiler-breeder chickens through a series of seven models, each constructed to investigate and answer a specific question. The analyses demonstrated the extent to which data can capture and describe multiple underlying dynamical behaviours, when queried with modelling approaches.

A number of the analyses were consistent with the existence of a ‘*Campylobacter* super-shedder’ state within the flock, with other ‘resilient’ birds that rarely or never transmit *Campylobacter* (Models 1, 5, 6, and 7). These resilient birds persisted even with high levels of colonisation amongst the remainder of the flock (Model 7). We specifically define super-shedders in this instance to be chickens who have a far higher likelihood of becoming colonised by *Campylobacter*, and a far lower likelihood of clearing such colonisation, compared to the flock average, i.e. the chickens populating the lower-right quadrant of Figure 3.9. Likewise we define ‘resilient birds’ as those with a lower likelihood of being colonised, and a greater likelihood of clearing any colonised *Campylobacter* STs (the upper-left quadrant of Figure 3.9). We likewise refer The existence of super-shedders is well-documented for *Salmonella* colonisation in chickens (Menanteau et al., 2018; Gopinath et al., 2012) and a small number of studies indicate varied shedding levels of *Campylobacter* amongst cattle (Rapp et al., 2012); however, although the models presented here were developed from field data, there is otherwise a paucity of published evidence

for a *Campylobacter* super-shedder status in chickens. This observation warrants further investigation, as the concept of individual colonisation-resistant birds within a flock raises the potential of alternative approaches to controlling *Campylobacter* colonisation in chicken flocks, e.g. those based on feed and/or probiotics. Small-scale experiments have shown that probiotic treatment can reduce *Campylobacter* bacterial load in individual birds (Ghareeb, Awad, Mohnl, Porta, et al., 2012; Willis and Reid, 2008), but the wider impact on flock transmission dynamics is poorly understood. Numerical modelling approaches have, however, highlighted the growth rate of competing bacteria as the most powerful factor in reducing the spread of *Campylobacter* (Rawson, Dawkins, et al., 2019).

The models further indicated that individual bird status was more important than *Campylobacter* strain in determining the dynamics of flock colonisation (Models 6 and 7) and, to our knowledge, this is the first time this has been shown. Most (26/39, 66.7%) of the *Campylobacter* sequence types (STs) observed have been isolated from other chickens (Colles, McCarthy, Bliss, et al., 2015), which is consistent with them being competent at colonising chicken hosts, although there is still potential for competition, for example over space or metabolic advantage. Whilst *Campylobacter* is generally considered to be a commensal of the chicken gut, there is evidence that in some breeds there is a prolonged inflammatory response, damage to the gut mucosa, and diarrhoea (S. Humphrey et al., 2014). Our results are consistent with the colonisation of a chicken flock by *Campylobacter* being a multi-factorial process, and the health and welfare of individual birds should be considered alongside *Campylobacter* strain type. The immune response of chickens has been shown to be impacted by welfare measures such as stocking density (Guardia et al., 2011; Gomes et al., 2014), food withdrawal, and heat stress (Burkholder et al., 2008). Consequently these results suggest an additional incentive to uphold good bird welfare, as a small subpopulation of susceptible birds can ultimately have a large impact on the colonisation status of the whole flock. The birds were weighed on two occasions through the study, but no correlation was found between weight and

Campylobacter shedding rate.

Having confirmed the existence of variation in bird transition probabilities, the possible impact of such variation on the proliferation of *Campylobacter* STs was investigated. Using a previously-published stochastic differential equation model of *Campylobacter* population dynamics within a broiler flock (Rawson, Dawkins, et al., 2019), two variant scenarios were explored: (i) one simulating a homogenous flock of chickens; and (ii) another simulating variation in immune response. The simulations demonstrated that demographically equal strains of *Campylobacter* could be sustained at broadly different levels across the flock as a consequence of varying bird immune response (Appendix B2). This is a random process, in that whichever strain is initially acquired by a super-shedder is then shed in large amounts into the environment, increasing the likelihood of colonising other birds in the flock. This result implies that the observation that some STs persist at higher levels than others in the flock, is likely due to the variation in bird transition probabilities, as opposed to phenotypic differences between STs. For example, ST 958 may appear more than ST 45 (Figure 3.2), not because it has a competitive advantage, but because it was initially ingested by super-shedders. Indeed, upon examination of the first recorded appearance of specific STs, the STs that would appear most frequently throughout the experiment were first observed in the most susceptible birds. Likewise the STs that appeared to die out were first observed in the more resilient birds; however, as only 75 out of 200 birds were sampled each week the exact date of when a ST first occurred cannot be determined.

With respect to the two species, there was evidence that *Campylobacter coli* was shed by individual birds more consistently over time compared to *Campylobacter jejuni*, with some indication that *C. coli* was more prevalent in the winter and *C. jejuni* more prevalent in summer (Models 2, 3, and 4). These results should be interpreted with caution. Firstly, there is no corroborating evidence in the literature to support such seasonal difference between species, and secondly, although the

number of birds from which *Campylobacter* was detected was variable from week to week, single colony picks (i.e. one *Campylobacter* strain per bird) were used, meaning the two variables were not entirely independent in this instance.

Human incidence of campylobacteriosis has been shown to vary in a repeated pattern each year (Nylen et al., 2002), which numerous studies have correlated with a similar pattern observed in broiler house colonisation rates (Jore et al., 2010; Patrick et al., 2004; Kapperud et al., 1993), an artifact disputed by other studies (T. Humphrey et al., 1993). Despite this, there was no effect of seasonality on *Campylobacter* spp. shedding rate detected by the modelling approaches here (Models 1 and 3, Figures 3.4 and 3.6), or in the original study examining local environmental variables (Colles, McCarthy, Bliss, et al., 2015). This lack of seasonality could be due to the different housing conditions and diet provisions between broiler and breeder flocks (Leeson and Summers, 2010). Breeder flocks have also been shown to shed smaller amounts of *Campylobacter* than commercial broilers (Cox, Stern, et al., 2002).

The age of the flock has been shown to be an important factor associated with increasing *Campylobacter* strain diversity. *C. coli* are more commonly isolated from older broiler flocks and may be over-represented in this broiler-breeder flock in comparison to commercially housed broilers, who live for typically 5 to 10 weeks before slaughter. The finding of increased colonisation duration as time progressed (Model 3), primarily by *C. coli*, was most likely due to the increased flock prevalence resulting in a positive feedback loop, whereby more *Campylobacter* is being shed into the environment by colonised birds, and then further ingested by the other birds in the flock before they are able to clear colonisation. Biologically, there is some indication that *C. coli* differs from *C. jejuni* based upon its genomic structuring, with the majority of farm animal related *C. coli* isolates grouping into the large ST-828 clonal complex, and *C. jejuni* isolates forming around 40 complexes. The shorter colonisation periods by *C. jejuni* (Models 4 and 6, Figures 3.8 and 3.10) potentially reflects a stronger immune response from the host leading to control or

clearance of STs, although this did not preclude colonisation by other *C. jejuni* (or *C. coli*) STs. Due to the inherent nature of sampling, it is not possible to know if strains were cleared altogether, or if multiple *Campylobacter* strains colonised birds simultaneously, although the modelling methods used in this study were chosen to take this uncertainty into account as far as possible. These results further highlight the importance of individual bird responses in determining flock-wide prevalence, and further ecological competition hierarchy modelling could be applied to the same data to assess the impact this phenomenon has on ST population stability across time. If structured hierarchical competition could be verified, this would support our contention that flock-wide defensive strategies should be conducted at an individual bird level.

We also stress the importance of our final model in not just investigating the weekly bacterial prevalence turnover, but how it further substantiates the importance of bird-to-bird transmission. Without this, one could argue from our earlier results that more resilient birds were simply the ones who did not ingest a more invasive ST. Figure 3.11 shows the influence of flock colonisation proportion on transition probabilities. Most notably we see that the transition from uncolonised to colonised is affected less by total colonisation prevalence than the transition from colonised to uncolonised. This means that in a highly colonised flock, uncolonised birds still have a possibility to not become colonised, while those who are already colonised will be far less likely to then clear their colonisation. This would likely be caused by the immune system of currently uncolonised birds being just as likely as previously to prevent an initial colonisation, but currently colonised birds will be more likely to add to their current bacterial load by ingesting more *Campylobacter* and reduce their likelihood of recovery.

In conclusion, these analyses have highlighted the diversity of individual bird response to bacterial challenge, and how this range of responses can be a key driver of *Campylobacter* prevalence dynamics. It is now important to find an

observable metric that correlates with the resilience of a bird to colonisation. If it were possible to identify ‘super-shedder’ birds on the farm, targeted interventions could be instituted to improve their response to bacterial challenge, or to better inform industry of how to raise broiler flocks with low rates of shedding within a flock. Such super-shedders will amplify the prevalence of *Campylobacter* within a flock, leading to rapid colonisation. Now that we have highlighted the critical role that individual bird status plays, future work must elucidate how one may act to help prevent the emergence of super-shedders within the flock.

Code Availability

The code used to produce all figures presented is available at: <https://osf.io/m5yua/>.

4

Mechanisms of biodiversity between *Campylobacter* sequence types in a flock of broiler-breeder chickens

Contents

Abstract	84
Introduction	85
Methods	89
Data	89
Competition Matrix	92
Patch-Occupancy Model	95
Results	98
Discussion	100

Under review in *Vet. Microbiol.* Preprint available as Rawson, Colles, Terry, et al. (2021).

DOI: <https://doi.org/10.1101/2021.04.14.439797>

Abstract

A long-term study of *Campylobacter* sequence types was used to investigate the competitive framework of the *Campylobacter* metacommunity, and understand how multiple sequence types simultaneously co-occur in a flock of chickens. A combination of matrix and patch-occupancy models were used to estimate parameters describing the competition, transmission, and mortality of each sequence type. It was found that *Campylobacter* sequence types form a strong hierarchical framework within a flock of chickens, and occupied a broad spectrum of transmission-mortality trade-offs. Upon further investigation of how biodiversity is thus maintained within the flock, it was found that the demographic capabilities of *Campylobacter*, such as mortality and transmission, could not explain the broad biodiversity of sequence types seen, suggesting that external factors such as host-bird health and seasonality are important elements in maintaining biodiversity of *Campylobacter* sequence types.

Introduction

Campylobacter are one of the most frequent causes of food poisoning in the UK (Blaser, 1997; Strachan and Forbes, 2010), presenting an estimated £50 million direct economic burden to the UK (Tam and O'Brien, 2016). The most commonly identified route of transmission to humans is via poultry meat (EFSA Panel on Biological Hazards (BIOHAZ), 2011), with seventy three percent of UK supermarket chicken carcasses shown to carry the bacteria (Jorgensen, Madden, et al., 2015). Whereas some foodborne pathogens, such as *Salmonella*, have been shown to proliferate primarily at the slaughterhouse (Heyndrickx et al., 2002), *Campylobacter* instead emerge and spread rapidly at the farm level (Evans and Sayers, 2000; Powell et al., 2012). As a result, limiting the spread of *Campylobacter* within poultry farms has been one of the primary goals of the Food Standards Agency (FSA) across the last ten years (FSA, 2010), where attempts to date have focused on biosecurity measures (Newell, Elvers, et al., 2011; Sibanda et al., 2018), such as employing anti-bacterial 'boot dips' at the entrance to chicken houses, and greater stress placed on farmers to practise consistent hand-washing and facility cleaning. Since *Campylobacter* have been shown to spread from a single bird, to an entire flock, in as little as one week (Stern, Cox, Musgrove, et al., 2001), the thinking behind such prevention methods is to minimise the chance of the bacteria entering the flock in the first instance. The focus on such measures has not resulted in decreased incidence, either proving ineffective (Fraser et al., 2010; Hermans et al., 2011; Wales et al., 2019), or not being correctly implemented, prompting calls for greater study into the ecology of this microbe (Sibanda et al., 2018; Kretzschmar, 2020), in the hope of gaining insight into how it can be controlled.

Different strains of *Campylobacter* are commonly categorised by sequence type (ST); genotyping samples by multi-locus sequence typing (MLST) of seven house-keeping genes (Dingle et al., 2001; Maiden et al., 2013). Broiler flocks (birds grown for their meat) are grown for only a short time, ranging from roughly five weeks for standard flocks, to 12 weeks for organic flocks (Council of European Union, 2008).

Yet despite this short window of time available for *Campylobacter* to colonise a flock, multiple STs are commonly observed simultaneously within a broiler flock (Colles, K. Jones, et al., 2003; Lydekaitienė and Kudirkienė, 2020; Wieczorek et al., 2020). For multiple STs to co-occur within a flock for several weeks implies the presence of regulatory mechanisms driving the sustained biodiversity within the flock, that have not yet been identified, let alone studied in depth.

Understanding the inter-strain competition mechanisms amongst different strains of *Campylobacter* can both aid understanding of the host-pathogen relationship, but also presents new opportunities in disease control. Understanding how certain STs may be excluded from colonising a flock by pre-established STs creates the opportunity for manipulation of these dynamics to reduce the incidence of certain STs. Strains of *Campylobacter* are known to vary in their pathogenic potential (Hofreuter et al., 2006), with some strains particularly effective at cell invasion (Hu and Kopecko, 1999). Introducing competitively superior strains into a transmission source presents a way to ensure that particularly pathogenic strains are unable to establish via competitive exclusion, as has been demonstrated in experimental studies (Chen and Stern, 2001). Alternatively, an understanding of these competitive frameworks presents the possibility for the use of live vaccine candidates, whereby bacterial strains that have been weakened can be used to trigger an immune response and limit pathogenic strains (Nothaft et al., 2016). While promising results in such vaccine candidates have begun to appear (Meunier et al., 2017), reliable effectiveness of these approaches requires knowledge of the underlying population dynamics. As of yet, such dynamics are not properly understood (Coward et al., 2008).

Understanding of these mechanisms is further exacerbated due to the fact that the exact route of entry into the flock is still uncertain. While it is generally considered that horizontal transmission is the most likely source of flock infections (Adkin et al., 2006), with STs carrying over from other locations on a farm, there still exists evidence of some infections caused due to vertical transmission (Cox, Richardson,

et al., 2012) and wild bird crossover (Colles, T. A. Jones, et al., 2008). The possibility of multiple points of entry for *Campylobacter* to enter a flock would explain the inability for improved biosecurity alone to reduce colonisation incidence, and may even suggest that stopping colonisation outright may be a fruitless endeavour, further supporting the need to utilise the manipulation of competitive hierarchies within the host microbiome; if the bacteria cannot be kept out of the farms, perhaps it can yet be kept out of the birds.

Investigations into the varying prevalence of specific STs have shown, experimentally (Colles, Preston, et al., 2019) and numerically (Rawson, Dawkins, et al., 2019), that a multitude of STs can be isolated from a chicken at any given time. Yet, within this pattern of co-occurrence, only one specific ST will usually be seen to dominate the gut, being isolated in far greater proportions than its co-colonisers. Through this mechanism, a diverse mix of STs can exist within a flock of many chickens, each carrying their own cohort of STs, and each with their own resident dominant strain. This observation constitutes a metacommunity (Leibold et al., 2004) of STs. A metacommunity is defined as a system where small communities interact with one-another, and influence the dynamics within each individual community. In our instances, the competing STs within a single host chicken can be thought of as a community, with multiple STs competing for dominance within one chicken, yet the dynamics within each individual chicken influences neighbouring chickens, resulting in a level of flock-wide dynamics as well. By utilising various mathematical frameworks from the wider ecological literature, we can begin to uncover how STs can co-exist within the flock, and to then ascertain what dynamic properties cause some newly introduced STs to die out, and others to persist.

To investigate this dynamic behaviour, this study utilises two mathematical modeling approaches to query the data from a long-term broiler-breeder flock prevalence study by Colles, McCarthy, Bliss, et al. (2015), which reports the STs isolated from individual birds within a flock across a year. A competition matrix model, such as

that outlined by Ulrich et al. (2014), is used to estimate a global competition matrix, detailing the competitive outcomes of pairwise competition between STs. This matrix quantifies the likelihood of specific competitive outcomes, namely if some STs will always outcompete some other STs, or whether such competitive outcomes can have unpredictable results. More importantly, they also provide insight into the competitive hierarchy seen within the broiler microbiome, whether that be a highly structured hierarchy, whereby dominant STs will always outcompete lesser-able STs in a gradually decreasing order of competitive advantages, or perhaps instead a system of intransitive competition. Intransitive competition, or ‘rock-paper-scissors’ competition, instead is defined as a system whereby loops are observed in the rank of competitive outcomes, for example if ST A outcompetes ST B, ST B outcompetes ST C, and ST C then outcompetes ST A (Soliveres and Allan, 2018). We refer to this cyclic relationship as an intransitive triad. In such a system, there can be frequent turnover of competing organisms, as no one entity is necessarily globally superior. Intransitive competition has been shown to have far-reaching implications for ecological stability and biodiversity, enabling species coexistence (Laird and Schamp, 2006), promoting biodiversity (Grilli et al., 2017), and enabling species cooperation (Hibbing et al., 2010).

Building on this, we then use the estimated competition matrix within a discrete-time patch-occupancy model to simulate and explore the broader dynamics of how STs move between birds in a flock, displace one another, and capitalise on the niches presented by uncolonised birds. Patch-occupancy models simplify a system to a series of ‘patches’, be it spatial units or, in our case, individual chickens, where each patch can be occupied by only one organism at a time, in our case, we consider the dominant ST of *Campylobacter*. The turnover in occupation by different organisms is captured by a series of probabilistic transition mechanisms, which have had great success in demonstrating persistence within metacommunities (Sutherland et al., 2014), due to minimising the assumptions placed upon the population dynamics of the system. The mechanisms that allow for sustained biodiversity in metapopulation

models have been shown to primarily be the demographic factors of transmission and mortality of competing species (May and Nowak, 1994; Hanski and Gyllenberg, 1997). i.e. how well a bacteria can invade a host, and how well it can remain there. In our case, we consider transmission as a measure of how many subsequent chickens will likely be challenged by the established ST in a host bird in the following timestep, the outcome of such a challenge is then decided by the previously estimated competition matrix. Bacterial mortality meanwhile is considered as the probability that a dominant ST will die out in the subsequent timestep, leaving the host bird susceptible to a new invading ST (not to be confused with bird mortality). By building a simulation of the system from which the data was gathered, we estimate these two specific parameters for each ST, and examine how these vary between STs and how they correlate with the observed frequency of each ST.

By presenting quantified estimates into the growth, spread, and competitive ability of each individual ST, we are able to provide insight into how STs of *Campylobacter* interact with one another, both within a host chicken, and within a flock as a whole.

Methods

Data

In the original study, a flock of 500 broiler breeders was monitored, with 200 birds labelled with leg-rings and monitored for a total of 51 weeks. Each week, cloacal swabs were taken from a random selection of 75 of the labeled birds, and tested for the presence of *Campylobacter* through standard culture methods. Positive samples were then genotyped (MLST), enabling the ST and species of the *Campylobacter* isolate to be specified. Note that, while multiple STs can occupy a host-bird simultaneously, it is frequently observed, experimentally (Colles, Preston, et al., 2019) and theoretically (Rawson, Dawkins, et al., 2019), that a single ST will broadly dominate the gut at any given time. Hence the sole ST recorded from a positive bird is a reflection of which STs are most dominantly expressed at that timepoint. Furthermore, these dominant STs in a host bird will dominate for roughly

a week before being replaced by a competitor (Rawson, Dawkins, et al., 2019). 39 distinct STs of varying prevalence were observed across the year within the flock, 25 of *Campylobacter jejuni* and 14 of *Campylobacter coli*. 19 of these STs appear very rarely, with less than ten total appearances in the data. Due to this limited number of data, meaningful conclusions as to their competitive abilities cannot be given, and as such we do not consider these STs in our analysis, considering only the 20 STs for which more than ten instances of occurrence were recorded in the data. An example layout of a small portion of this data is presented in Figure 4.1, and the total prevalence of STs over time is displayed in Figure 4.2. Negative samples are not shown in Figure 4.2, as this data is not used for the competition matrix model. Further experimental details can be found in the original publication (Colles, McCarthy, Bliss, et al., 2015).

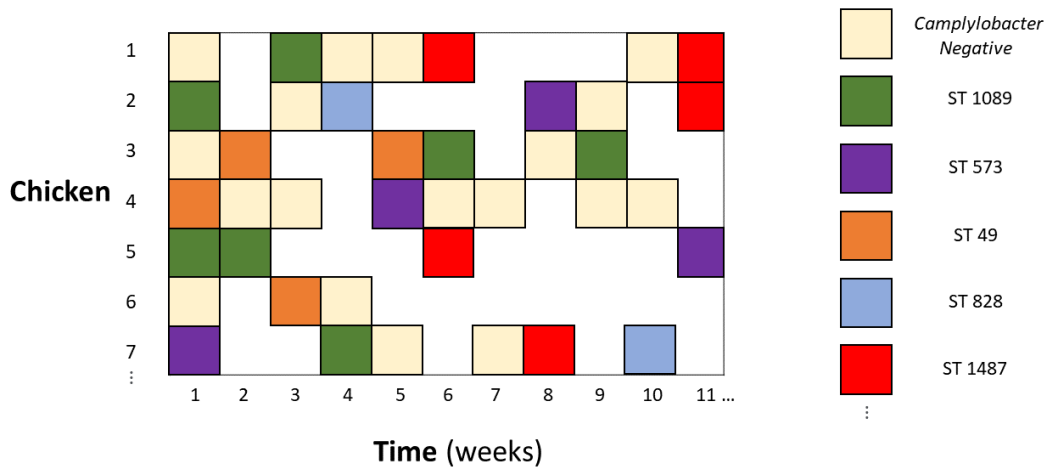


Figure 4.1: Example portion of the ST prevalence data. From a total flock of 500 broiler breeders, 200 were labelled with leg-rings. These 200 are captured in the rows of the data frame. Each week 75 of these birds were tested for the presence of *Campylobacter* for 51 weeks (columns). Birds were marked as either free from *Campylobacter* (marked in tan), or, if found to be *Campylobacter* positive, the sequence type (ST) of the bacteria was recorded. Blank white spaces indicate where a bird was not tested for that particular week. The whole data set comprises 200 rows, 51 columns, and captures 39 distinct STs.

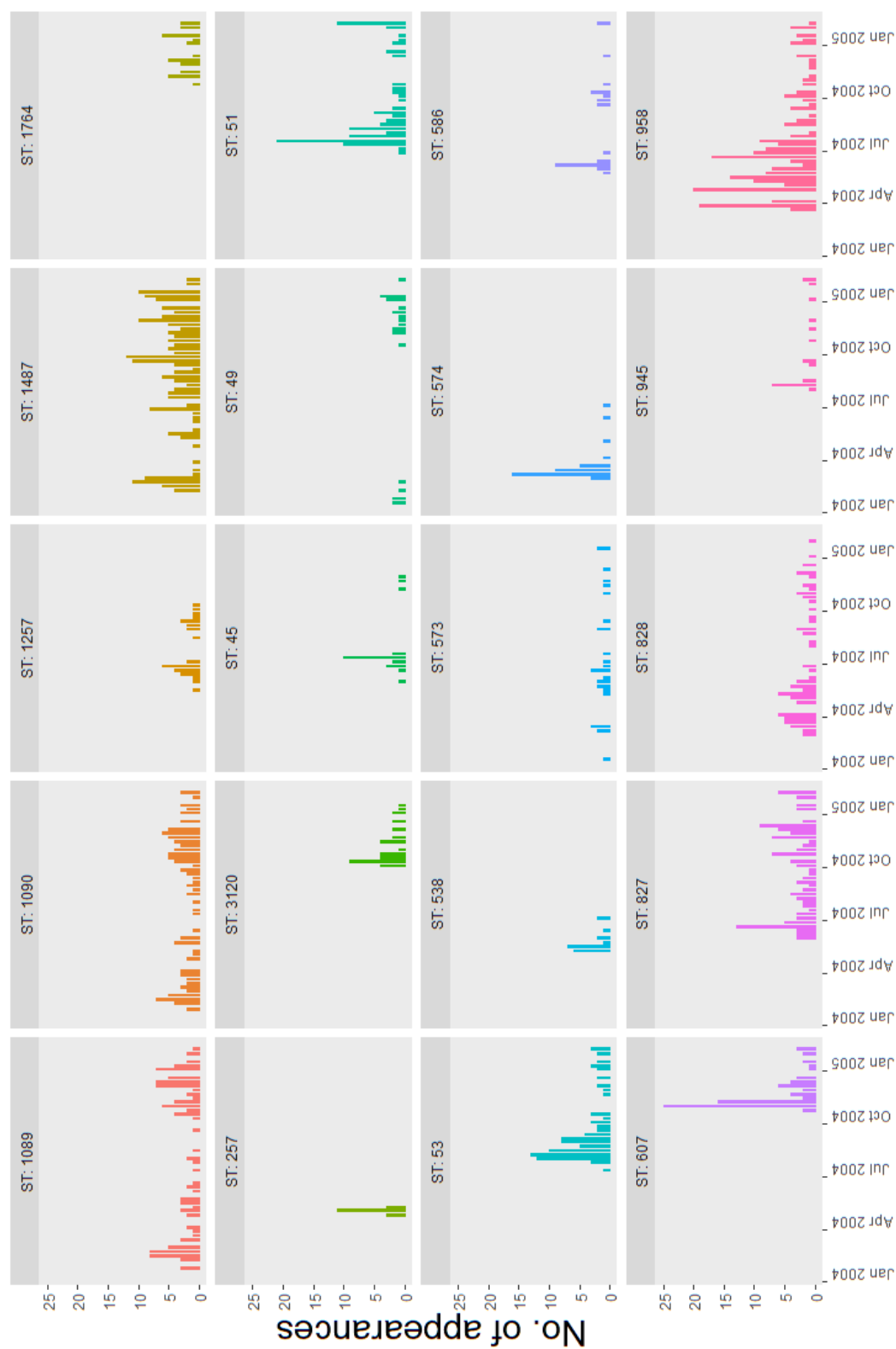


Figure 4.2: Histograms of ST prevalence across the study period reported in Colles, McCarthy, Bliss, et al. (2015). STs with less than ten appearances are not displayed.

Competition Matrix

We first estimate a competition matrix, detailing all pairwise competition outcomes between all STs. Formally, we define that, for a system of n STs, the competition matrix, C , is an $n \times n$ square matrix where element $C_{i,j}$ represents the probability that ST i outcompetes ST j in a pairwise competition. By definition, the diagonal elements of C are equal to 1, and $C_{i,j} = 1 - C_{j,i}$.

By using the time series abundance data of all STs throughout the flock, as shown in Figure 4.2, one may back-infer the pairwise competitive strengths between all STs within the flock. Based upon the methods outlined by Ulrich et al. (2014), this competition matrix may be estimated by first inferring a transition matrix, P : an $n \times n$ square matrix where $P_{i,j}$ represents the probability that a chicken colonised by ST j is instead colonised by ST i in the next time period. Note that this matrix P is not the same as the competition matrix C , as the observed transitions could represent the result of multiple sequential competitions between STs - the replacing ST has not only outcompeted the present occupant, but also all other incoming STs.

To estimate this transition matrix, P , consider an $n \times 51$ frequency matrix A , where element $A_{i,t}$ denotes the number of chickens that ST i was isolated from at time t , and where n is the number of distinct STs. This matrix is directly built from our data, where element $A_{i,t}$ can be seen as the ‘No. of appearances’ of ST i in week t from Figure 4.2. This frequency matrix is then related to our transition matrix, P , via the equation;

$$PA_t = A_{t+1} \quad (4.1)$$

where A_t is a column vector of n elements, reporting the abundance of all STs at time t . This provides a method by which to estimate P by choosing the matrix P that best fits equation (4.1). Homogeneous mixing of STs is assumed, however,

another assumption is made in equation (4.1) that all STs are present and are capable of appearing at each time point. This is not representative of biological reality. We see from Figure 4.2 that some STs do not appear in the flock until later in the experiment, and while it could plausibly be being outcompeted in every prior instance, it is more plausible that the ST has simply not yet infected the flock. As such, we adapt equation (4.1) by also implementing a binary-filled ‘presence’ matrix Z , an $n \times 51$ matrix, where element $Z_{i,j}$ is either 0 or 1, denoting whether or not ST i is present in the flock at time j . i.e. when a ST is not observed within a flock in a particular week, we do not consider its impact on that week’s transition dynamics.

If ST i is isolated in the data at time t , we mark it as present in matrix Z for times t through to $t + 3$, to account for the possibility of a ST being reduced to low levels, not captured in the data. This three week window was determined from our earlier numerical simulations (Rawson, Dawkins, et al., 2019), showing the average duration for which a low-level “non-dominant” ST might persist within a host. We rewrite equation (4.1) as:

$$(PA_t) \odot Z_{t+1} = A_{t+1} \quad (4.2)$$

where Z_{t+1} is the $(t + 1)^{\text{th}}$ column of Z , and $\odot$ is the Hadamard (element-wise) product. In essence, Z simply acts as a switching mechanism, to switch off the possibility of transitions to a ST that has not yet emerged. This approach carries multiple benefits. Primarily, the transition matrix now represents the transition probabilities for a flock where all STs are present simultaneously. This inference allows more of the dataset to be utilised, without having to divide our experimental data into multiple regions of different sized matrix calculations. A possible limitation to this approach is that it allows inference of competitive outcomes between STs that do not appear at the same time in the original dataset. i.e. it can infer based on the growth abilities of a ST at a later time how it would fare against a ST from an earlier time. While this inference is useful, these limited instances are not experimentally verifiable. As such, we do not display these few “assumed”

competitive strengths in our results, to avoid confusion.

Once the best fitting P to equation (4.2) has been found, we may use this P to estimate the associated competition matrix C . Ulrich et al. (2014) presents such a methodology whereby, assuming homogeneous mixing, the transition matrix P and the competition matrix C are linked by the relationship:

$$\begin{aligned} P_{i,j} &= P(1, \dots, n) [j \rightarrow i] \\ &= \frac{1}{n-1} C_{i,j} + \frac{1}{n-1} \sum_{k=1, k \neq i,j}^n C_{j,k} P(1, \dots, k-1, k+1, \dots, n) [j \rightarrow i] \end{aligned} \quad (4.3)$$

for $i \neq j$, and

$$P_{i,i} = \prod_{k=1, k \neq i}^n C_{i,k} \quad (4.4)$$

where the range of summation in (4.4) is calculated across the subset considered in the notation $P(1, \dots, n)$. Heuristically, one considers the transition probabilities as the proportional outcomes of all possible competitive interactions. In a four-species system, equations (4.3) and (4.4) would define:

$$P_{i,j} = \frac{1}{3} C_{i,j} + \frac{1}{3} \left(\frac{1}{2} C_{i,j} C_{j,k} + \frac{1}{2} C_{i,j} C_{j,k} C_{j,l} \right) + \frac{1}{3} \left(\frac{1}{2} C_{i,j} C_{j,l} + \frac{1}{2} C_{i,j} C_{j,l} C_{j,k} \right).$$

In small systems, the probability of successful transition for each ST could be directly calculated as the proportional outcome of all possible competitive interactions as given in equations (4.3) and (4.4). However, for our system of 20 STs this is computationally impossible, as the size of equation (4.3) will rapidly balloon for such a large system. Instead we therefore used the approximation approach of Ulrich et al. (2014):

$$P_{i,j} \approx \frac{1}{m-1} \sum_{k=0}^{n-2} \left(\frac{\prod_{l=1}^n C_{j,l}}{C_{j,j} C_{j,i}} \right)^{\frac{k}{n-2}}. \quad (4.5)$$

This approximation was found to estimate a randomly drawn 20×20 test competition matrix with a mean value error < 0.001 .

The above methodology allows us to choose a trial competition matrix, C , convert this to a transition matrix, P , via equation (4.5), and then evaluate how well this transition matrix simulates the observed data, A , via equation (4.2). All that is required now is an approach by which to find the “best” competition matrix C . As such, we estimate the competition matrix C using the above equations within a Bayesian framework, using the Just Another Gibbs Sampler (JAGS) program (Plummer, 2007), a Markov chain Monte Carlo (MCMC) sampling program utilising Gibbs sampling. Specifically the model was called and analysed within R by using the `rjags` package (Plummer et al., 2016). We considered wide, uninformative, uniform priors on the elements of C . Convergence was considered well-achieved, with every element of C ’s posterior distribution displaying a potential scale reduction factor (PSRF) < 1.03 , and a Monte Carlo standard error (MCSE) less than 5% of the standard deviation of the sample. The code used is made available at <https://osf.io/3rd4e/>.

Lastly, we quantify the amount of intransitivity observed from the best-fit competition matrix C . While many metrics of measuring intransitivity have been proposed (Feng et al., 2020), the most suitable is generally considered to be Kendall and Babington Smith’s d_s (Kendall and B. B. Smith, 1940); a measure of the proportion of three-species intransitive loops found within the competition matrix. i.e. we measure the number of cyclical intransitive triads seen in the competition matrix, and divide this by the total number of possible triads for a competition matrix of that size.

Patch-Occupancy Model

The estimated competition matrix gives insight into the interactions between different *Campylobacter* STs; however, it cannot by itself answer our questions as to how biodiversity of STs is maintained within the flock. The previous

metacommunity modelling studies of May and Nowak (1994) and Hanski and Gyllenberg (1997) have demonstrated that persistence can be largely managed by differences between the colonising ability and mortality of competing organisms. As such, we estimate parameters describing the colonising ability and mortality for each of our 20 considered STs. Figure 4.2 shows that some STs occur with increased frequency compared to other present STs. For example, STs 1487 and 573 both seem to persist within the flock throughout the entire recorded experimental duration, and yet ST 573 is observed in far fewer birds throughout this time. We hypothesise that differences in the demographic parameters between these STs may explain the differences in the underlying population dynamics.

A patch-occupancy model was designed to simulate the experimental data as closely as possible. In this instance, the patches considered are the 500 chickens that make up the flock, and the STs of *Campylobacter* present are the occupying entities.

A 500×51 matrix is initialised, where each row denotes a specific chicken in a flock, and each column a time-point (a week), so as to replicate the data structure shown in Figure 4.1. Element (i, t) thus records which ST, if any, has colonised chicken i at time t . The first column is initialised to match the proportion of STs recorded in the first week of the dataset in Figure 4.2. Each time-step is then simulated in turn, to iteratively generate the subsequent 50 columns. In each timestep, each established ST may be removed for the following timestep with probability, μ_i , the ST-specific mortality parameter that we seek to estimate. STs that persist to the next timestep then have the opportunity to infect other chickens. The number of other chickens that are challenged by this resident ST is drawn from a Poisson distribution, $\text{Pois}(\lambda_i)$, where λ_i is a ST-specific parameter. Borrowing from the parlance of the ecological literature, we refer to this parameter as the average ‘propagules released’ by ST i . If a challenged chicken is currently uncolonised by *Campylobacter*, they then become colonised by the invading ST. If a challenged

chicken is currently colonised by a different ST, this is treated as a competitive event, whereby the winner of the pairwise competition will be the occupying ST for the following timestep, and the loser is removed. If a chicken is challenged by multiple STs in the same time-step, then a further weighted competitive event is considered between the challenging STs to decide the host ST for the following timestep. This outcome is decided by the probabilities estimated in our previous model, given by the matrix C .

When new STs appeared for the first time in the experimental data, they are directly introduced into the patch-occupancy model at the proportion and time-step they were first observed. One exception is made for ST 49, which was unobserved for so long in the experimental data, that two specific introduction events were allowed. Appendix C1 outlines the pseudo-code detailing this model structure. The model was programmed in R and the code is available at <https://osf.io/3rd4e/>.

Considering transition events on the weekly timescale provided in the original data is considered valid based upon theoretical modelling work showing that dominant STs in a host bird will dominate for roughly one week before being replaced by a competitor (Rawson, Dawkins, et al., 2019). Much like our previous model, this provides a framework whereby a trial solution of μ_i and λ_i for each ST i can be used, and the resulting ST population dynamics can be compared against the population dynamics observed in the original data. We wish to find the values of μ_i and λ_i that best capture the patterns seen in Figure 4.2. We score a trial solution by comparing the relative proportions of ST frequency at each time-step with the proportions shown in the original data. The specific iterative framework as outlined in Appendix C1 cannot be integrated into a Bayesian system, so we instead utilise machine learning techniques to seek the optimum solution.

We first find an estimate for the average parameter values across all STs to use as an initial trial solution for each individual ST. We collapse the data to a binary state

of either *Campylobacter*-positive or *Campylobacter*-negative, and use simulated annealing to find the average μ and λ values that best simulate the data, using a scoring function defined by the absolute difference between the infection proportions in every column and every row between the model data and experimental data. This is so that the algorithm selects the parameters that also capture the frequency with which chickens may transition from being *Campylobacter*-positive to *Campylobacter*-negative. This provided a best-fit solution of $\mu = 0.7$, and $\lambda = 3.2$. These values were then used as initialisation points for each ST-specific parameter set (μ_i, λ_i) , which are then iteratively adapted using genetic algorithm approaches to find the best-fit solution.

Genetic algorithms, so named for their inspiration by natural selection, generate “mutations” of the initial trial solutions, and the resulting mutations which best describe the data will in turn inform the next generation of trial solutions. A genetic algorithm of population size 200 was run for 100 iterations, using the $(0.7, 3.2)$ estimate as a suggested population element for each specific ST.

Results

Figure 4.3 shows the pair-wise competition values for all STs. STs that do not naturally co-occur during the experiment have been represented with a grey box, as meaningful conclusions as to their competitive interactions cannot be drawn. The matrix has been re-ordered to maximise the number of values >0.5 in the upper-diagonal, thus showing the identified hierarchy.

A strong hierarchical structure can be observed, with STs at the top of the matrix mostly outcompeting all STs below them. Some intransitive loops can be seen within the matrix however, for example ST 607, which is able to outcompete some STs higher up the hierarchy. When uniformly sampling the missing values of the matrix shown in Figure 4.3, an average of 125 intransitive triads are recorded for the competition network, compared to a hypothetical maximum of 330 for a

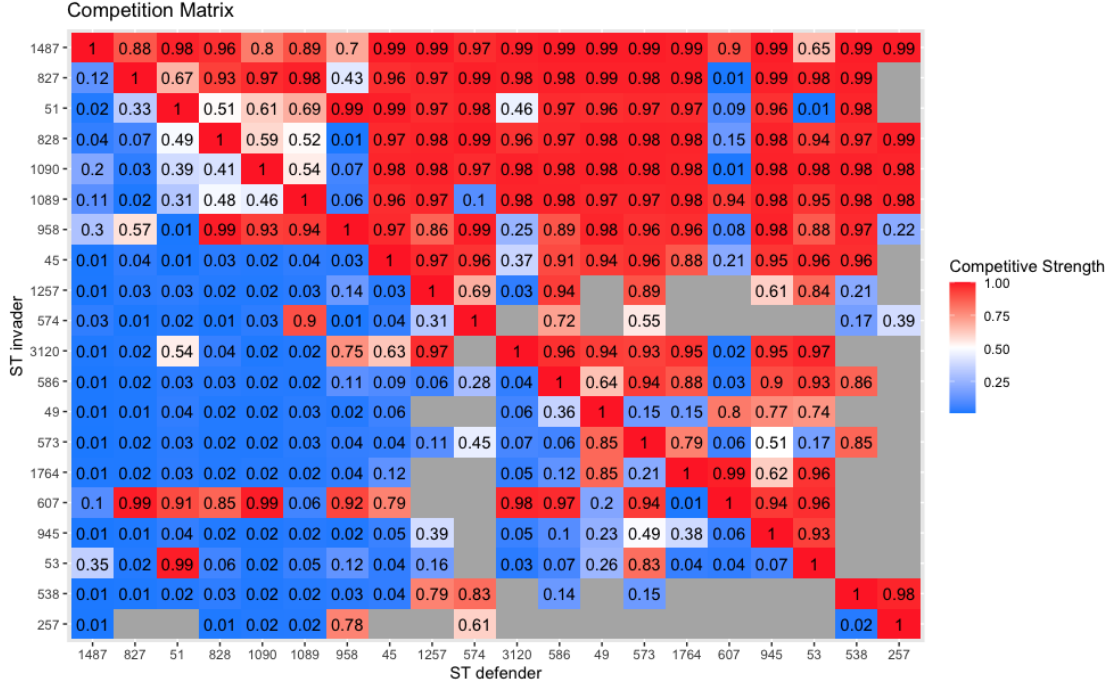


Figure 4.3: Matrix of pairwise competition strengths between *Campylobacter* STs. Element (i,j) depicts the probability that ST i outcompetes ST j in a pairwise competition. Empty grey boxes depict cases where two STs do not coexist during the experiment, thus their competitive relationship cannot be estimated. Rows are ordered to maximise the number of values >0.5 above the diagonal. The structure reveals a strong competitive hierarchy, with the strongest competitors at the top of the matrix.

(complete) 20×20 matrix, resulting in an intransitivity score of $d_s = 0.379$ (Kendall and Babington Smith's d_s (Kendall and B. B. Smith, 1940)). In comparison, on sampling 100,000 random 20×20 competition matrices, the lowest number of intransitive triads generated was 196, hence our observation of only 125 triads supports a system of significant hierarchical competition.

The competition matrix shown in Figure 4.3 is then utilised within the patch-occupancy model to estimate ST-specific transmission and mortality parameters. These parameters are displayed below in Figure 4.4. Mortality (μ) we define as the probability that an established ST will die out from its host bird naturally from one week to the next. To capture ST-specific transmission effects we report the average propagules released (λ), the average number of other chickens that an occupying ST will challenge for the following timestep, with the outcome of these challenges

decided by the above competition matrix.

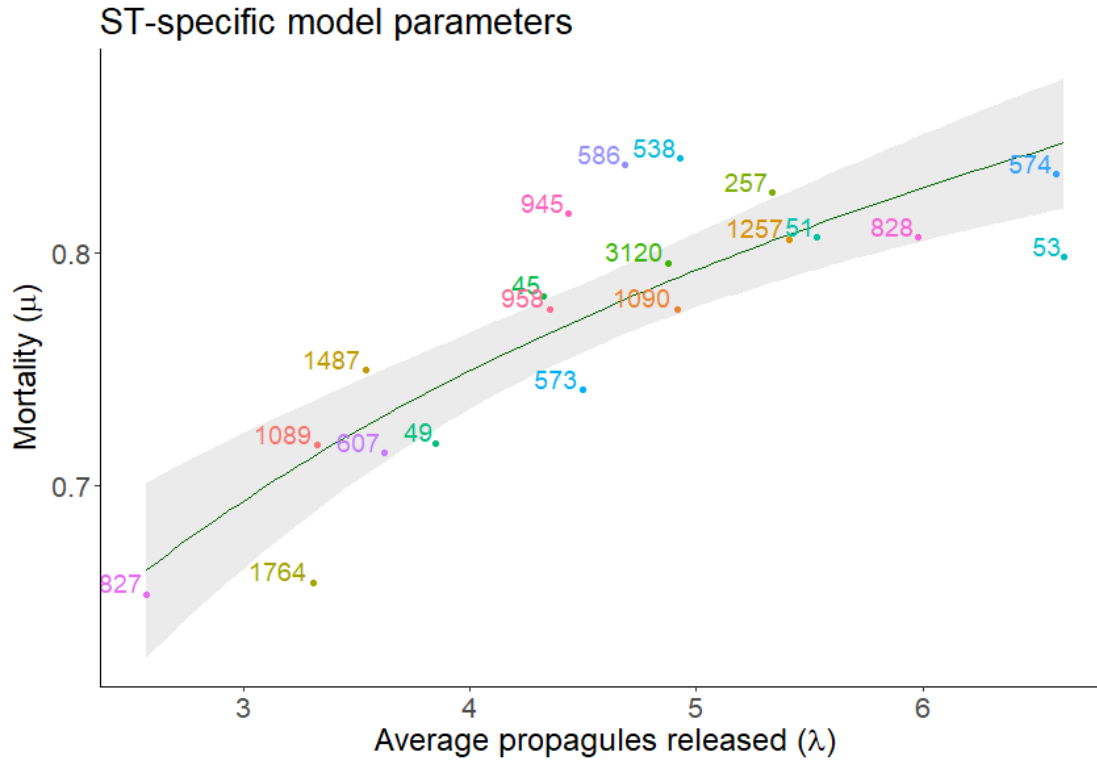


Figure 4.4: ST-specific model parameters for patch-occupancy model. Mortality (μ) depicts the probability that a ST dies out from one time point (a week) to the next. If the ST does not vacate a host, it releases propagules that challenge other host chickens. The average number of chickens challenged (λ) is a model parameter depicted on the x -axis. The green line displays the statistically significant (t-test, $p < 0.0001$) logarithmic regression between the two variables. We see a positive trend whereby higher mortality is compensated by a greater number of propagules being released.

The positive logarithmic trend (t-test, $p < 0.0001$) shows a relationship whereby STs with a higher mortality (they die out more frequently) can maintain their presence in the flock by being able to colonise more chickens.

Discussion

Here we have investigated the ecological drivers maintaining *Campylobacter* diversity within chicken flocks. By quantifying competition, transmission, and mortality parameters through two mathematical frameworks, we have highlighted the demographic differences between *Campylobacter* sequence types, and shown that the

metacommunity of STs operates within a strict competitive hierarchy, with some STs capable of outcompeting other STs, and hence replacing them as the dominant strain within host birds.

The competition matrix shown in Figure 4.3 effectively disproves the hypothesis that ST diversity may have been maintained by a system of intransitive competition, as very few intransitive triads were found within the system. Intransitive loops have been shown to theoretically support coexistence of many competing organisms, dependent on growth rate differences and intransitive cycle length (Gallien et al., 2017), and such effects have been demonstrated in small plant communities (Soliveres, Maestre, et al., 2015). Despite the wealth of theoretical work surrounding the impact of intransitive competition, real-world evidence of such systems is lacking. An experimental study searching for such effects across five different taxonomic groups by Soliveres, Lehmann, et al. (2018) was unable to find strong evidence of intransitivity in any of their studies other than experiments with mosses, and found zero intransitive triads in their bacterial experiment. As such, our inability to identify clear signs of intransitivity is unsurprising. Only STs 53, 607, and 3120 showed clear evidence of being able to outcompete STs higher in the hierarchy. All three of these STs appear to have remained prevalent in the flock from their point of entry to the end of the experiment time, possibly suggesting that STs that are able to form an intransitive loop may be more capable of invading and persisting in the flock.

Within this competitive hierarchy, we also show that the magnitude of the respective competitive probabilities are relatively large. In the the upper diagonal of Figure 4.3 values are greater than 0.9, suggesting that the competitively superior STs not only outcompete a vast number of other STs, but that they outcompete these other STs decisively, winning competitive interactions over 90% of the time in most instances. This is in line with experimental studies of inter-strain competition, with El-Shibiny, P. L. Connerton, et al. (2007) demonstrating how a strain of *Campylobacter* is not only able to outcompete a multitude of other strains, but to do so repeatedly in

multiple experiments.

This evidence of a clear competitive hierarchy further stresses how specific mechanisms must underpin the observed maintenance of biodiversity of *Campylobacter* STs. Under such competitive conditions, biodiversity of a metacommunity has been shown to be feasibly maintained by trade-offs between transmission and mortality (May and Nowak, 1994; Hanski and Ovaskainen, 2003; Bonsall et al., 2004; Scheffer and van Nes, 2006). Under such a system, the co-occurrence of multiple STs can be explained by competitively strong STs displaying high mortality rates, namely that after replacing a resident ST, they naturally die out from the host quickly. Alternatively, their transmission ability may be compromised such that, although they may be very effective competitors, they are unable to proliferate as fast as other STs, and thus may not challenge a high number of other chickens from one week to the next. Likewise, a competitively weak ST, such as ST 53 in Figure 4.3, may not be able to withstand competition from incoming STs, but is able to persist in the flock by challenging a higher number of chickens each week (high number of propagules released), and surviving within these host birds for a longer period of time (low mortality). The patch-occupancy model presented was designed to specifically quantify these mortality and mean propagule release parameters, and are presented in Figure 4.4.

Figure 4.4 shows that all STs can be placed somewhere within a life-history trade-off. In general, STs displaying high mortality, may persist in the environment by releasing a higher number of average propagules, and vice-versa. May and Nowak (1994) theoretically showed that for a newly emerging organism to successfully invade a metacommunity, and to then persist, they need to fill a yet unrepresented area of this transmission-mortality spectrum. i.e. to persist, they need to have no close neighbours in the plot of Figure 4.4. This may be demonstrated by STs 827 and 53. Both STs can be seen from Figure 4.2 to appear within the flock midway through the time span, and to then successfully persist through to the end of the

experiment. Both of these STs can also be seen from Figure 4.4 to be outliers on the transmission-mortality spectrum, with ST 827 having the lowest mortality of all observed STs, and ST 53 having the highest number of mean propagules released. As a further interesting contrast, competitive ability does not appear to have influenced this, as ST 53 is one of the weakest competitors in the metacommunity, and ST 827 is one of the strongest, as shown in Figure 4.3.

However, this mechanism alone has historically been unable to account for the vast amount of sustained biodiversity observed in nature. Building on the theoretical findings of May and Nowak (1994), Bonsall et al. (2004) demonstrated that species within a hierarchical competition structure, competing for the same resource, may co-exist by clustering into ‘life-history guilds’. Competitively strong species may simultaneously co-exist by sharing similar demographic parameters. At the same time, competitively weaker species will also persist in the environment, by also sharing similar demographic capabilities with one another. Scheffer and van Nes (2006) highlighted the same result, concluding that newly emergent species would only persist in the environment if either (i) they were significantly competitively superior to all existing species, or (ii) if they were similar enough to existing species, both competitively and demographically, so as to exist within this particular life-history guild niche. Our results however do not show evidence of such ecological guilds.

Figure 4.4 shows that, while STs do form a life-history trade-off, STs appear in a broadly even distribution across this mortality-propagule trade-off. Furthermore, some STs that appear to be demographically similar vary greatly in their competitive ability and respective population dynamics. From Figure 4.2, we can broadly delineate STs by four distinct dynamic profiles: a ST may either persist in a flock or die out, and it may exist at high-frequency or low-frequency. It was assumed that one could characterise these four distinct dynamic profiles by their competition, average propagule release, and mortality parameters, and yet no such pattern has

been found in this study. For example, the STs 257, 574, 45, and 1257, could all be characterised as appearing in high frequency, before then dying-out. Yet despite these similar dynamical behaviours, all STs place broadly across the competition-propagule-mortality spectrum, with no common trends in their placement. Likewise, STs 586, 573, and 945, could all be categorised as persisting in the flock, though recovered at low frequency, and yet all three STs are found in broadly different placements in Figures 4.3 and 4.4. In general, STs that appear in high frequency appear to correlate with higher competitive potential in Figure 4.3, though no such trend can be associated with persistence.

Since these STs do not demonstrate the guild-assemblage ‘clumping’ structure in Figure 4.4 (shown by Bonsall et al. (2004) to be necessary for biodiversity maintenance in this instance), it suggests that some other mechanism must be enabling the co-occurrence and persistence of *Campylobacter* STs. Based upon the broader wealth of investigations into *Campylobacter* dynamics, we can posit three potential hypotheses driving these clearly seen differences in population dynamics between STs:

(i) Host-bird variability. It has been shown in numerous patch-occupancy systems that patch quality (meaning that some patches are ‘easier’ to colonise than others) can have a tremendous impact on the overall population dynamics, having even greater impact than differences between how patches are connected (Donahue et al., 2003; Poniowski et al., 2018). D. W. Yu and H. B. Wilson (2001) theoretically showed that while differences in life-history trade-offs were necessary for co-existence, significant heterogeneity in patch quality or density was necessary to support a large number of species. Such patch variation also made it possible for newly emergent species to persist even if the species was inferior in both competitive and colonisation ability. In our context, variation in patch quality and density would translate to host birds varying in their response to bacterial challenge, with some chickens ‘easier’ to colonise than others. Indeed, through Bayesian transition models

we have shown using this same data set in Rawson, Paton, et al. (2020) that a flock contains a mixture of birds that are highly resilient to bacterial challenge, and highly susceptible birds that operate as ‘super shedders’. These super shedders are consistently being colonised by a variety of *Campylobacter* STs with high turnover. Poor individual bird health and welfare has been previously shown to correlate with a reduced immune response, with measures such as stocking density (Guardia et al., 2011; Gomes et al., 2014), food withdrawal, and heat stress (Burkholder et al., 2008) all contributing to increased *Campylobacter* colonisation. D. W. Yu and H. B. Wilson (2001) directly showed that the host-bird variation seen in Rawson, Paton, et al. (2020) removes the need for newly emerging STs to be sufficiently similar to persist in the flock. This further supports the idea that the proliferation of *Campylobacter* in a flock is influenced primarily by the individual birds.

(ii) Seasonal variation. Broiler flock colonisation by *Campylobacter* has been well-documented to follow a seasonal trend (Patrick et al., 2004; Jore et al., 2010), with flocks more likely to become colonised in the warmer summer months than the winter. The data behind this modelling study was gathered over 51 weeks, January 2004 to January 2005, so would plausibly have been impacted by seasonal variation. The original study examining the impact of local environmental variables on the data set we have considered (Colles, McCarthy, Bliss, et al., 2015) (and subsequent Bayesian transition analyses (Rawson, Paton, et al., 2020)), were unable to identify any temporal trend within the total *Campylobacter* prevalence, however the *Campylobacter coli* STs did appear less frequently during the summer. It is thus plausible that seasonal variation may have impacted the population dynamics of the occupying STs in the flock via some yet-unidentified mechanism. An example of this may be seen by comparing the population dynamics of STs 53 and 574. Both STs occupy a similar placement in the propagule-mortality spectrum of Figure 4.4, and yet, despite ST 574 being more competitively able than ST 53, ST 574 does not persist in the flock, while ST 53 does. One possible explanation for this is that ST 53 first appeared within the flock in July, while ST 574 appeared in February.

(iii) Stochasticity. While our patch-occupancy model is a probabilistic one, the mechanisms by which a metacommunity of *Campylobacter* STs persist is determined by a number of random events. The events of a ST first entering the flock, chickens ingesting colonised faeces, and of then establishing themselves within the microbiome all encompass a wide number of stochastic events which could change the resulting population dynamics. Coward et al. (2008) showed that attempts to replicate population dynamics of *Campylobacter* within broilers were largely unsuccessful, even in the most simple cases of just two competing strains. They posited that this was likely due to “founder effects”, small variations in population level at first inoculation which could have large consequences for the flock-wide population dynamics. We have previously shown this effect through a series of stochastic differential equations in Rawson, Dawkins, et al. (2019), whereby a variety of overall population dynamics can be observed dependent on stochastic events when the population of a *Campylobacter* ST is very low. Likewise, upon running the patch-occupancy model for the estimated parameters presented in the results, some STs would persist in some actualisations, but not others. Thus, attempting to characterise some dynamical profiles by mortality and transmission parameters, may not be possible as the data displays only one dynamic outcome of many possible ones.

One important caveat to this work must be stressed. Since broiler flocks are slaughtered anywhere from 5 to 11 weeks of age, longitudinal studies into the *Campylobacter* population dynamics are not possible, birds are not alive long enough for us to observe long-term dynamics from which to extract parameter estimates. As such, this experimental data was gathered from a flock of broiler-breeders, the birds that lay the eggs that become broiler flocks. As we have discussed above, host bird factors may have significant implications for the overarching population dynamics of the microbiome, meaning that these estimated parameters could plausibly be different in commercial broiler flocks. Broiler and breeder flocks are kept under slightly different housing conditions and diet provisions (Leeson and Summers, 2010),

and breeder flocks have also been shown to shed smaller amounts of *Campylobacter* than commercial broilers (Cox, Stern, et al., 2002). Since this study has focused on investigating *Campylobacter*-specific factors, our conclusions remain relevant to commercial broiler flocks, namely that the population dynamics remain deeply susceptible to impact from a variety of factors, such as season and host bird health.

The primary finding of this work highlights how the life-history trade-offs we have identified fail to provide an explanation for the persistence and co-occurrence of multiple *Campylobacter* STs. This further supports the notion that suppressing and controlling colonisation by *Campylobacter* cannot be achieved through biosecurity alone, and reflects calls for a ‘One Health’ (Gözl et al., 2014) approach, whereby further understanding is needed of how *Campylobacter* and broilers interact and affect each other. We have shown that demographic advantages alone cannot determine which STs of *Campylobacter* will come to dominate a flock of chickens, and that it may instead come down to a ST being in the right place at the right time, or rather, the right chicken in the right season.

Code Availability

The code used to produce all figures and tables presented is available at: osf.io/3rd4e/.

5

Can good broiler flock welfare prevent colonization by *Campylobacter*?

Contents

Abstract	110
Introduction	111
Methods	113
Data	113
Data Cleaning	116
Logistic Regression	117
Bayesian Network	119
Discretisation	123
Banned Arcs	124
Results	124
Logistic Regression	124
Bayesian Network	126
Discussion	129

Under review in *Poultry Science*. Preprint available as Rawson, Colles, A. L. Smith, et al. (2021).

DOI: <https://doi.org/10.1101/2021.03.30.437710>

Abstract

Using data on rearing and welfare metrics of multiple commercial broiler flocks from the last ten years, we investigate how welfare measures such as hock burn, mortality, weight, and pododermatitis, among others, impact the likelihood of a flock becoming colonized by *Campylobacter*. Using both logistic regression and Bayesian networks, we show that, while some welfare metrics were weakly related to *Campylobacter* colonization, evidence could not be found to suggest that these metrics actively exacerbated *Campylobacter* colonization, rather that they were both symptoms of the same underlying cause. Instead, observed dependency on the management of the flock suggested that yet-undiscovered differences in rearing practice were the principal cause of both poor bird welfare and increased risk of *Campylobacter*, suggesting that action can be taken to improve both these factors simultaneously.

Introduction

For several years campylobacteriosis has been the most frequently observed zoonotic disease in humans throughout the EU (Westrell et al., 2009), with poultry meat identified as a leading infection route (EFSA Panel on Biological Hazards (BIOHAZ), 2011). This acute form of food poisoning, characterised by diarrhoea, fever, and abdominal pain, is estimated to affect 450,000 individuals a year in the UK, approximately ten percent of which result in hospitalisation (Strachan and Forbes, 2010). An investigation by Public Health England into the extent of *Campylobacter* within the poultry industry revealed that seventy-three percent of supermarket chicken carcasses were found to contain *Campylobacter* and seven percent of the outer packaging was similarly contaminated (Jorgensen, Madden, et al., 2015). This considerable public health burden posed by *Campylobacter* spp. represents an estimated £50 million annual economic cost to the UK (Tam and O'Brien, 2016).

Given the extent to which *Campylobacter* dominates commercial chicken flocks, attempting to reduce the proliferation of the pathogen at farm level would have significant impacts in reducing disease incidence in humans. Once *Campylobacter* is first identified within a broiler flock (chickens grown specifically for their meat), colonization of all birds occurs very rapidly (Evans and Sayers, 2000). In experimental studies, it can take only a single week for an entire flock to become infected following the introduction of a single infected bird (Stern, Cox, Musgrove, et al., 2001). This speed of proliferation makes identifying the initial point of entry of *Campylobacter* into a flock challenging, and has resulted in a focus on preventative measures.

To date, the poultry industry has largely focused upon on-farm biosecurity measures (Fraser et al., 2010; Gibbens et al., 2001), such as boot-dips and improved cleaning of housing. However, little impact in reducing incidence has been achieved with these measures (Hermans et al., 2011). As such, research has instead turned to a broad array of preventative measures (Ghareeb, Awad, Mohnl, Schatzmayr, et al.,

2013), such as treatment of food and water (Peh et al., 2020), probiotics (Saint-Cyr et al., 2016), and bacteriophage therapy (El-Shibiny, Scott, et al., 2009). Such measures have thus far had mixed, and at times contradictory, success.

One area of research still greatly overlooked is the role of bird welfare in the emergence of *Campylobacter* within a flock, both as a potential indicator of *Campylobacter* colonization, and as a driving factor. *Campylobacter* spp. were long considered to be commensal within broiler chickens, but recent studies have begun to suggest they may be pathogenic under some circumstances (S. Humphrey et al., 2014; Wigley, 2015). Some welfare measures in the past have been observed to correlate with changes in the gut microbiota and immune response of birds, such as stocking density (Guardia et al., 2011; Gomes et al., 2014), food withdrawal, and heat stress (Burkholder et al., 2008). More directly, digital dermatitis (Bull, Thomas, et al., 2008), lesions on the footpad, and arthritis (Alpigiani et al., 2017) have been shown to be strong predictors of *Campylobacter* prevalence, further supporting findings that flock movement patterns and behaviour can also accurately predict *Campylobacter* prevalence (Colles, Cain, et al., 2016). Lawes et al., 2012 have shown bird age, bird mortality and depopulation events to be significant risk factors for *Campylobacter* colonisation at slaughter. Our own previous mathematical modelling studies have highlighted the potential for stocking density (Rawson, Dawkins, et al., 2019) to impact the population dynamics of *Campylobacter* within a flock, and have also shown that the colonization status of an entire flock is greatly impacted by the most susceptible birds within the flock (Rawson, Paton, et al., 2020), suggesting that attention to individual birds must not be overlooked.

This study investigates the relationship between multiple welfare indicators on *Campylobacter* prevalence in flocks using two different mathematical modelling approaches. We firstly employ a logistic regression analysis to test for direct relationships between *Campylobacter* colonization and predictor variables, such as weight, mortality, and hock burn incidence. While this methodology has long served

as a useful tool for highlighting potential relationships between variables, it cannot elucidate the exact mechanism of such a relationship, nor how these relationships interact with one another. We combine our logistic regression with a Bayesian network analysis to demonstrate the network of conditional dependencies between variables, to investigate more precisely how variables affect and impact each other. In combination with the logistic regression analysis, we are able to posit where welfare directly increases the likelihood of *Campylobacter* colonization, or to what extent infection by this bacteria is a symptom of the same root cause.

The greatest challenges to welfare-focused studies is ensuring a broad collection of data from varied sources, and using easily reproducible metrics. Studies utilising welfare concepts such as the ‘Welfare Quality®’ (De Jong et al., 2016) or the ‘five freedoms’ (Iannetti et al., 2020) are useful, but can be difficult to recreate due to differences in individual assessment. To this end, this study uses data spanning six years from multiple farms, logging reproducible metrics, such as temperature, flock parent age, pododermatitis rates, and flock size, amongst others.

Methods

Data

Data was provided across six years (2010 to 2015) from multiple farms throughout the UK. Each data point represents a flock of broilers, listing multiple welfare parameters and rearing information, as well as a measure of whether the flock tested positive for *Campylobacter*. All variables measured for flocks are detailed and defined below:

- ***Company*** - A two-factor categorical variable, depicting whether the flock is overseen by company “1” or “2”. This variable will also therefore capture differences in company-specific rearing methodologies not represented by our current list of predictor variables.

- **Farm** - A categorical variable, further delineating the *Company* measure, detailing which farm the flock was located at, so as to investigate trends unique to certain locations.
- **Number placed** - A numerical variable describing how many broilers made up the flock. While modelling studies have primarily implicated stocking density as a high *Campylobacter* risk factor, the total flock population may also increase the likelihood of initial flock inoculation (Rawson, Dawkins, et al., 2019).
- **Date placed** - The date the flock was first placed into the house. *Campylobacter* is well reported to show seasonal trends, with the warmer, summer months seeing flocks test positive for *Campylobacter* more frequently (Nylen et al., 2002; Djennad et al., 2019).
- **Breed** - A three-factor variable describing the breed of broilers grown. Two commercial breeds of broiler were investigated, with flocks comprised of either: Breed A, Breed B, or a mixture of Breed A & B. Both breeds refer to the breeding companies, each with many different genetic lines of broiler. Host-bird genetics have been shown to impact *Campylobacter* prevalence (Stern, Meinersmann, et al., 1990; Babacan et al., 2020; Psifidi et al., 2021), hence the consideration of the genetic line of the flock.
- **Number of parent flocks** - The number of parent flocks the broiler flock was sourced from. While the possibility of vertical transmission of *Campylobacter* is still debated, the hypothesis is that a greater number of parent flocks could increase the number of *Campylobacter* sequence types (and thus phenotypic specialisations) that a flock is exposed to at hatch (Petersen et al., 2001).
- **Mean parent age** - The average age (in weeks) of all parent flocks sourced from. Parent age has been shown to impact egg weight and embryo weight (Shanawany, 1984), and thus could potentially impact the general health of the chick.

- **7/14/21/28/35/Total mortality percentage** - Six different variables, describing the percentage of the flock that had died after x days.
- **Pododermatitis percentage** - What percentage of the flock suffered from pododermatitis; inflammation and ulcers on the footpad and toes. This was measured post-mortem by abattoir staff.
- **Hock burn percentage** - What percentage of the flock suffered from hock burn; areas where ammonia from the waste of other birds has burned through the skin of the leg. This was measured post-mortem by abattoir staff.
- **7/14/21/28/35/Final day weight** - Six variables showing the mean weight of the flock, in grams, at weekly intervals.
- **Maximum/minimum temperature** - A variable describing the maximum and minimum recorded external temperature, in degrees centigrade, for the time the flock was housed, as sourced from historical records from the Met Office for the nearest weather station.
- **Campylobacter 21/28/35 days** - A two-factor variable depicting whether a flock was found to be positive or negative for *Campylobacter* after 21/28/35 days. This was sampled via fabric boot swabs in the flock house at 21/28/35 days. In addition, fresh faecal samples were collected concurrently on day 28. *Campylobacter* prevalence was then tested for in all samples via culture methods. Full details of this methodology are given in Colles, Cain, et al. (2016).

A total of 212 flocks were monitored, however not all variables could be measured for all flocks due to the practical difficulties in obtaining all measures from farms. As such there is some degree of missing data across all variables, most notably that only 149 of these flocks have a final record of *Campylobacter* infection status. Before incorporating this data into a mathematical model, we must consider the detail of data available given the absence of some variables for some flocks. To ensure the maximum number of flocks are able to be used in model fitting, a balance must be

found between filtering out variables to increase data availability, while not overly limiting the number of variables investigated. We detail these decisions below.

Data Cleaning

Before beginning the regression analysis, we clean and simplify our data to aid interpretation. The *Campylobacter* variables across time points 21, 28, and 35 days were simplified to a single variable that reads as positive if a flock was recorded positive on any of the three dates recorded, and negative if the flock was reported negative on all of the measured dates provided. This was to increase the data availability, as some flocks were only measured on certain dates. There were six instances of a flock being recorded as negative after previously testing positive. These six instances were cases where the faecal samples taken on day 28 tested positive, but the boot swab on day 35 tested negative. It was considered appropriate to rely on the more targeted faecal sample for these six cases.

The *Date placed* variable was converted to a 4-level factor variable, denoting the season that the flock was reared in. This was done as date is known to have a non-linear effect on *Campylobacter* prevalence (Jorgensen, Ellis-Iversen, et al., 2011), with incidence in both flocks and humans more frequently observed in the UK summer compared to the winter (Louis et al., 2005). It is this effect that we wish to investigate as opposed to variation between years. Season classification is partitioned by the dates December 1st, March 1st, June 1st, and September 1st, aligning with the meteorological seasons, which more accurately capture temperature variation than the astronomical seasons classification.

Regression analysis requires that the explanatory variables be independent of the response variable (and each other) otherwise predictive power is weakened across all dependent descriptor variables. In some cases, parameters of the linear model then become indeterminate due to the high degree of multicollinearity. For example, the *7/14/21/28/35/Total mortality percentage* variables are, as expected, all highly

correlated with one another, hence we use only the *28-day mortality percentage* measure, as this is the one that most data was available for. We do the same for the average bird weight variables. Likewise, the *Company* variable was removed for the logistic regression, as it is heavily correlated with the *Farm* variable (companies do not share farms), however the *Farm* variable was also then found to have very strong correlation with the *Number placed* variable. For this reason the *Farm* variable is also removed, as *Number placed* is a preferred metric of interest. Similarly, we use only the *Minimum temperature*, and not the *Maximum temperature*, or the *Date placed*, as these three are strongly correlated. By reducing the number of model predictors, the generalised variance-inflation factors (GVIF) (Fox and Monette, 1992) of all variables are less than 3, far less than the commonly-used threshold of 10.

Finally, the data was filtered to remove any flocks with missing values for the explanatory variables under consideration. 84 data points remained for the final mathematical model. Flocks with missing data were later utilised for the parameter learning stage. A summary table of all variables considered in the final model is presented in Tables 5.1 and 5.2.

Logistic Regression

Multiple logistic regression is an adaptation of multiple linear regression for instances where the response variable of interest is a two-factor binary output ($Y \in \{0, 1\}$), in our case where a flock is either *Campylobacter* negative or positive. A multiple linear regression model structures the response variable, Y , as a linear predictor of a set of explanatory variables, X_i , like so;

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n,$$

for n variables, and where β_i are the coefficients to be determined. A logistic regression instead models $p = P(Y = 1)$, the probability that $Y = 1$, as:

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n, \quad (5.1)$$

Table 5.1: Factor variable summaries

Variable	Factor level	Total	<i>Campylobacter</i> positive	<i>Campylobacter</i> negative
<i>Company</i>	1	49	22	27
	2	35	20	15
<i>Farm</i> *	C1-F1	49	22	27
	C2-F1	15	7	8
	C2-F2	11	7	4
	C2-F3	9	6	3
<i>Breed</i>	A	32	20	12
	B	48	21	27
	A & B	4	1	3
<i>No. Parents</i>	1	39	23	16
	2	25	13	12
	3	15	5	10
	4	5	1	4
<i>Date placed</i>	Spring	24	19	5
	Summer	18	17	1
	Autumn	16	1	15
	Winter	26	5	21

* Company 1 has only one farm ‘C1-F1’. Company 2 has three farms; ‘C2-F1’, ‘C2-F2’, ‘C2-F3’.

Table 5.2: Continuous variable summaries

Variable	Mean	Standard Deviation
<i>Number placed</i>	27,639	7,283
<i>Mean parent age</i>	39.13	9.82
<i>28-day mortality percentage</i>	3.87	1.40
<i>Pododermatitis percent</i>	57.62	28.48
<i>Hock burn percent</i>	20.00	18.95
<i>28-day average bird weight</i>	1419.6	83.2
<i>Minimum temperature</i>	6.59	3.70

where $\text{logit}()$ is the log-odds ratio $\text{logit}(p) = \log \frac{p}{1-p}$, which ensures that p is bounded between 0 and 1. To model the impact of factor variables with m levels, we use treatment contrasts; $m - 1$ distinct descriptor variables within the model. For example, consider a simplified model which investigated the impact of breed alone on the probability of a flock being colonized by *Campylobacter* (p). *Breed* has three factor levels; ‘Breed A’, ‘Breed B’, and ‘Breed A & B’, and therefore the logistic regression model would be:

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2,$$

where Breed A, is represented by $X_1 = X_2 = 0$, Breed B by $X_1 = 1, X_2 = 0$, and the mixture of Breed A & B by $X_2 = 1, X_1 = 0$.

Nine explanatory variables were used for the final maximal logistic regression fit: *Number placed*, *Breed*, *Mean parent age*, *Number of parent flocks*, *28-day mortality percentage*, *Pododermatitis percentage*, *Hock burn percentage*, *28-day average weight*, and *Minimum temperature*. After initially fitting the maximal model of nine explanatory variables, a step wise simplification is then performed, removing the least significant term iteratively to finally reach the minimal adequate model: a model composed of only statistically significant explanatory variables. The model was fit using the `glm` package in R, which fits the model via iteratively reweighted least squares (IWLS). All code is made freely available at osf.io/pb62g/.

Bayesian Network

Bayesian networks are probabilistic graphical models that display the network of conditional dependencies between a collection of variables. Each variable in the model is visually represented as a node, with directed edges, called ‘arcs’, between nodes representing a directly dependent relationship. $A \rightarrow B$ indicates that B depends on A . Since arcs are directed, there is a cause-and-effect (from-and-to) relationship between variables. A node with an arc directed towards another node is called a ‘parent’ node to the respective ‘child’ node. Each node’s output is

then explicitly detailed by a probability distribution that is dependent on any and all parent variables. This highlights the two greatest strengths of Bayesian networks as tools to investigate relationships between variables: firstly, the Markov property imposed by the network of conditional dependencies means that the global probability distribution of the system can be expressed as a far smaller product of dependent probabilities. As such, a large and complicated probabilistic system can be simplified by knowledge of how some variables do or do not influence one another. Secondly, these types of models provide a straightforward way of visually conveying how certain explanatory variables influence (or do not influence) each other, something that would otherwise require the analysis of a large variety of logistic regression models, and could easily overlook certain dependencies. As a result of this architecture, “cycles” are by definition not allowed within a Bayesian network, meaning a path cannot be drawn from any node back to itself. Such a structure is called a directed acyclic graph (DAG). We provide a short example below to understand how such networks are calculated, but greater insight can be found in Nagarajan et al. (2013).

Calculating a Bayesian network model is separated into two tasks. Firstly, structural learning: learning the network model of dependencies (i.e. identifying all arcs in the system), followed by parameter learning: finding the specific parameters of probability distributions linking parent to child nodes. Consider an example of a dataset of three discrete variables in a broiler flock we wish to explore: *Mortality* (low, average, high), *Age* (young, adult, old), and *Feather condition* (good, average, poor). We start by learning the structure between these three variables. Many algorithms and approaches exist for finding the structure of a Bayesian network (Bouchaala et al., 2010), however within this paper we utilise the hill-climbing algorithm (Bouckaert, 1995), a score-based structure learning algorithm. The algorithm starts with a randomly chosen graph (though usually the empty graph made up of no arcs), and calculates a network score that ascertains how effectively such a graph describes the data. It then iteratively adds, removes and reverses one

arc at a time, altering the global probability distribution via the introduction (or removal) of a dependency, selecting the alteration that increases the network score the most. This process is then repeated until no further improvement can be found. Multiple network scores can be used, but we use the Bayesian information criterion (BIC) (Bhat and Kumar, 2010), a variation on the traditional likelihood function. After using this algorithm on our example data, we discover the “best” network as being the network of two arcs shown in Figure 5.1.

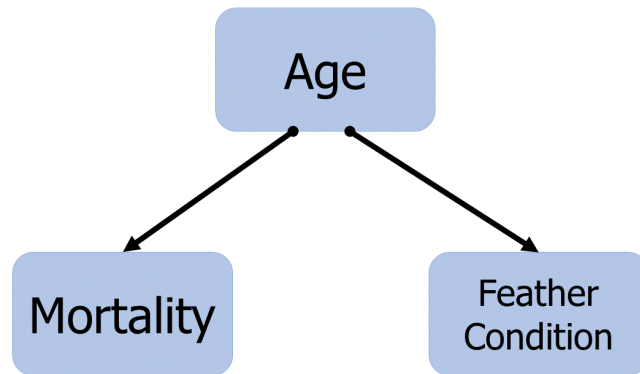


Figure 5.1: Bayesian network for the example problem posed. Two conditionally dependent relationships are found, from *Age* to *Mortality* and from *Age* to *Feather condition*. This example relationship was demonstrated by Comin et al. (2019).

We see from Figure 5.1 that *Age* is a parent variable to both *Mortality* and *Feather condition*. This indicates that, from this imagined example data, *Age* directly informs the mortality rate of a bird and the feather condition of the bird (this result was directly demonstrated by Comin et al. (2019)). An important insight gained from this network analysis would be that *Mortality* and *Feather condition* would likely be found to be correlated via a logistic regression analysis. However, *Feather condition* itself does not affect *Mortality*, rather they are both impacted by the same direct cause; *Age*. This illustrative example shows the objective of the Bayesian network approach to our particular question; what **causes** *Campylobacter* to colonize a flock, rather than just what is correlated with *Campylobacter* colonization. Another advantage of such a model, is that inference can be made even with missing data.

The network of Figure 5.1 presents a structure whereby the mortality of a bird can be predicted with data on their feather condition, as this gives important indication of what the age of the bird may be. In Bayesian terms, this informs our prior belief as to the age of the bird, thus impacting our posterior belief as to the mortality of the bird. In contrast, the logistic regression approach would require an assumption on the age of such an individual, usually the mean of the training data, but no such requirement exists for Bayesian networks. Note, however, that if the age of a broiler is known, the prediction of their relative mortality rate is not improved by further information on their feather condition, as mortality is found to be predicted by age alone.

Note also from Figure 5.1 the mathematical advantage of such a network for expressing the joint probability distribution of the system. By definition the arcs indicate that $P(\text{Age}, \text{Mortality}, \text{Feather Cond.})$ can be expressed as

$$P(\text{Age}, \text{Mortality}, \text{Feather Cond.}) = P(\text{Age})P(\text{Mortality}|\text{Age})P(\text{Feather Cond.}|\text{Age}).$$

Since each variable has three factor levels, this reduces a distribution of 27 (3^3) parameters, to 21, where each arc indicates that the child variable is modeled by a multinomial distribution dependent on the parent variable.

Indeed, for the second step, parameter estimation, we treat each node as being described by a multinomial distribution, and fit these using two separate well-known techniques, the maximum likelihood estimator (MLE), and a nested Bayesian approach, using uninformative uniform priors. See Appendix B1 for a brief introduction to Bayesian statistical inference.

A further benefit to a Bayesian network model is that we do not need to test for multicollinearity, which required us to remove several variables from consideration in the logistic regression, as structure learning specifically investigates these inter-variable correlations. As such we are able to include *Company*, *Farm*, and *Date*

placed within our Bayesian network model. We also include the *7-day bird weight*, and *7-day mortality percentage* variables, alongside the 28-day measures, to serve both as a sanity check (we would expect these two variables to be linked with their 28-day counterparts), but also to increase the predictive power of the model, so inference could be made on the *Campylobacter* status of a flock from the 7-day as well as 28-day measures. This decision did however reduce the number of available training data from 84 to 81.

All of these introduced methodologies are implemented using the **bnlearn** package in R (Scutari, 2009), and all code used in the model analysis is provided at osf.io/pb62g.

Discretisation

While we have displayed the many inherent strengths of Bayesian network models, one considerable weakness is the implementation of models consisting of both discrete and continuous variables. While methodologies exist for the assessment of such “hybrid” Bayesian networks (Scutari and Denis, 2014), the approaches are considerably more computationally demanding, and require a greater amount of data to give a robust fit to a Bayesian network. Given the comparatively smaller size of our training data ($n = 84$), we instead take the commonly used route of discretisation, whereby our continuous variables are converted into discrete bins. Of the many approaches to discretisation, a wide-ranging comparison by Kohavi and Sahami (1996) found the best approach to be the supervised, entropy-based Minimal Description Length (MDL) (Fayyad and Irani, 1993) method, whereby each variable is discretised based upon its informative potential on a variable of interest. This approach was undertaken on our data, in relation to the *Campylobacter* variable, using the **FSelectorRcpp** package in R. However, only *Minimum temperature* was found to be able to discretised in such a way (foreshadowing our later results). As such, we instead used a quantile binning (equal-frequency) approach, to separate out our continuous variables into three bins of equal frequency, and confirming

against the histograms for each variable that no obvious separation was missed. These bin intervals are provided in Table 5.3.

Table 5.3: Discretisation intervals of continuous variables for the Bayesian network model

Variable	Bin 1 Intervals	Bin 2 Intervals	Bin 3 Intervals
<i>Number placed</i>	[11770, 22000]	(22000, 33503.3]	(33503.3, 34650]
<i>Mean parent age</i>	[25, 32]	(32, 44.78]	(44.78, 58]
<i>7-day mortality percentage</i>	[0.65, 1.21]	(1.21, 1.81]	(1.81, 7.26]
<i>28-day mortality percentage</i>	[1.97, 3.06667]	(3.06667, 4.05667]	(4.05667, 9.61]
<i>Pododermatitis percent</i>	[1, 42.6667]	(42.6667, 76]	(76, 95]
<i>Hock burn percent</i>	[0, 10]	(10, 20]	(20, 90]
<i>7-day average bird weight</i>	[144, 170.667]	(170.667, 181]	(181, 213]
<i>28-day average bird weight</i>	[1138, 1388.33]	(1388.33, 1475.33]	(1475.33, 1565]
<i>Minimum temperature</i>	[1.3, 4]	(4, 8.7]	(8.7, 13.8]

Banned Arcs

To both aid the structure learning process, and to disallow erroneous network structures, we also introduce a list of banned arcs, defining all arcs which are not to be considered by the algorithm, based on logical reasoning. For example, we do not allow any arcs directed towards the *Company* variable, as this is clearly not affected by any other variables. While the company that a flock belongs to may in turn affect the mean parent bird age for example, it is illogical to say that the mean parent bird age could affect which company the flock is managed by. *Company* is a variable that is predetermined before the flock even hatches, and as such cannot be influenced by factors that occur during the lifespan of the flock. A full list of these banned illogical arcs is provided with all associated code in the online repository.

Results

Logistic Regression

The results of the logistic regression for the minimal adequate model are presented in Table 5.4, alongside a variety of model evaluation metrics. Appendix D1 shows the analysis of the original maximal model comprised of all explanatory variables,

Table 5.4: Logistic regression analysis of the minimal adequate model for 84 broiler flocks using the `glm` function in R.

Predictor	β	SE β	Wald-test z -score	p	e^β (odds ratio)
(Intercept)	-1.490	0.621	-2.398	0.0165	NA
<i>Breed</i> (1 = B, 0 = Other)	-1.474	0.628	-2.348	0.0189	0.225
<i>Hock burn percentage</i>	-0.041	0.0189	-2.165	0.0304	0.960
<i>Minimum temperature</i>	0.469	0.105	4.474	$\ll$ 0.0001	1.599
Test			χ^2	p	
Overall model evaluation					
Likelihood ratio test ¹			35.972	7.59×10^{-8}	
Likelihood ratio test ²			7.776	0.557	
Goodness of fit					
McFadden's R^2	0.309				
Cox & Snell's R^2	0.348				

¹ Compared against null model.² Compared against maximal model.

and describes the reduction steps taken to reach the minimal adequate model.

Three variables were found to be statistically significant in relation to the *Campylobacter* status of a flock via the Wald-test: *Breed*, *Hock burn percentage*, and *Minimum temperature*. Note that for the minimum adequate model, while Breed B flocks were found to be statistically significantly different to Breed A birds with relation to *Campylobacter* incidence, the mixed breed flocks were not found to differ from Breed A flocks. As such, the ‘Breed A’ and ‘Breed A & B’ flocks were collapsed into one variable for the minimal adequate model. Table 5.4 shows that flocks of Breed B birds were found to be 0.225 times as likely to test positive for *Campylobacter* than any other flock. *Hock burn percentage* was, unintuitively, found to have a negative correlation with *Campylobacter* colonization. *Minimum temperature* was very strongly correlated, with an odds ratio showing that an increase of 1 degree to the minimum recorded temperature corresponded with a flock being 1.599 times more likely to test positive for *Campylobacter*. The generalised variance-inflation factors (GVIF) (Fox and Monette, 1992) of all variables in the minimal

adequate model was less than 2, and all variables of the maximal model (Appendix D1) had a GVIF of less than 3, far less than the commonly-used threshold of 10.

Bayesian Network

Figure 5.2 displays the final global network structure, fit using the hill-climbing algorithm, and with networks scored via BIC. This was run using the `bnlearn` package in R. The strength of individual arcs (as measured by BIC) is represented by arrow-width in Figure 5.2. Table 5.5 also explicitly provides these arc strength scores.

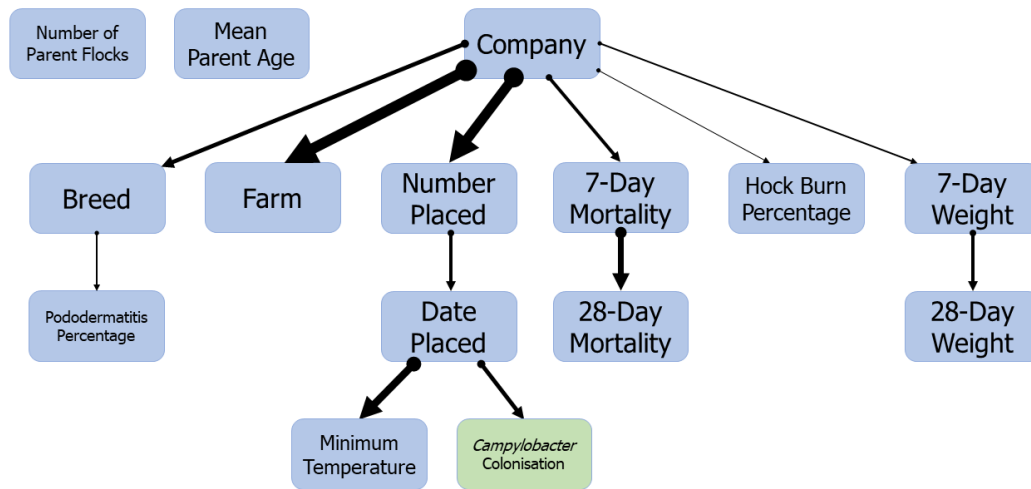


Figure 5.2: Bayesian network structure showing the interrelationships between multiple welfare and rearing practice factors in a flock of broilers. *Campylobacter* colonization is directly impacted by the season the flock is grown in. Structure was learned using a hill-climbing algorithm, and sampled networks scored using the Bayesian information criterion (BIC). Arrow-width indicates arc strength as scored by BIC, the values of which are given in Table 5.5.

To test the significance of the fit structure, structure learning was also performed with a tabu search algorithm, and by introducing random network restarts into the hill-climbing algorithm (10, 100, and 1000 random restarts were all performed), all of which resulted in the same network structure. We also performed a hill-climbing structure learning algorithm using the logarithm of the Bayesian Dirichlet equivalent score (BDE) (Castelo and Siebes, 2000), as opposed to the BIC, a Bayesian-based

score equivalent to the Dirichlet posterior density (and initialised with uniform priors). This scoring metric resulted in a very similar network structure which we present in Appendix D2. The only differences were that, (i) *Hock burn percentage* no longer had *Company* as a parent node, meaning it was unconnected to any other node. (ii) *Minimum temperature* had an additional arc from itself to *Campylobacter* colonization, suggesting that *Campylobacter* could also be impacted by temperature variation throughout the season; and finally, (iii) an additional arc was introduced from *Breed* to *Number placed*, simply representing that flocks of Breed B birds were larger in size than flocks of Breed A birds.

Table 5.5: Arc strengths of the Bayesian network shown in Figure 5.2. Arc strength is measured by Bayesian information criterion (BIC), where a lower value indicates a stronger link.

Parent	Child	Arc Strength
<i>Company</i>	<i>Breed</i>	-16.656
<i>Company</i>	<i>Farm</i>	-47.756
<i>Company</i>	<i>Number placed</i>	-43.069
<i>Company</i>	<i>7-day mortality percentage</i>	-11.975
<i>Company</i>	<i>Hock burn percentage</i>	-0.234
<i>Company</i>	<i>7-day weight</i>	-3.644
<i>Breed</i>	<i>Pododermatitis percentage</i>	-2.236
<i>Number placed</i>	<i>Date placed</i>	-6.272
<i>7-day mortality percentage</i>	<i>28-day mortality percentage</i>	-20.876
<i>7-day weight</i>	<i>28-day weight</i>	-6.031
<i>Date placed</i>	<i>Minimum temperature</i>	-29.040
<i>Date placed</i>	<i>Campylobacter colonization</i>	-17.246

Other results to be noted from Figure 5.2 is that neither the number of different parent flocks that a broiler flock was born from, nor the mean age of these parent was found to have any correlation to any other variable. *Pododermatitis* was interestingly found to be influenced by the *Breed* of broiler comprising the flock. We also see that many variables are directly influenced by the *Company* variable, suggesting that many observed differences are due to yet-unidentified differences in management practice between the two companies.

Table 5.6: Conditional probability table for *Campylobacter* colonization status. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference				
	<i>Date placed</i>			
<i>Campylobacter</i> colonization	Spring	Summer	Autumn	Winter
Negative	0.273	0.071	0.878	0.802
Positive	0.727	0.929	0.122	0.198

Maximum Likelihood Estimator				
	<i>Date placed</i>			
<i>Campylobacter</i> colonization	Spring	Summer	Autumn	Winter
Negative	0.217	0.063	0.937	0.808
Positive	0.783	0.937	0.063	0.192

Figure 5.2 shows that the season (*Date placed*) in which a flock is reared is the sole parent node to *Campylobacter* status. This means that *Date placed* alone captures the uncertainty and probabilistic distribution of whether or not a flock is likely to test positive for *Campylobacter*. This means that while data on the number of birds in the flock (*Number placed*) can inform whether or not a flock is *Campylobacter* positive, this data is superfluous when one has knowledge of the *Date placed*. The conditional probability table for *Campylobacter* colonization is given in Table 5.6. These model parameters can be fit either via maximum likelihood estimators (MLE) or through Bayesian inference. Model parameters via both methods are provided in Table 5.6. Note that one advantage of the Bayesian inference method is that this approach can learn parameters from data containing missing values. Hence while the MLE parameters are fit from the 84 data points used in structure learning, the Bayesian inference method uses 114 data points, incorporating those that were removed from structure learning due to missing values.

Conditional probability tables for *Campylobacter* dependent on all other variables, assuming the absence of data on any other variable, are provided in Appendix D3.

Discussion

Here, through a combination of both logistic regression and Bayesian network analysis, we have investigated the interrelationships between a selection of welfare and rearing practice explanatory variables for multiple commercial broiler flocks. At the inception of this work, our hypothesis was that poor welfare indicators such as low weight and hock burn, among others, would result in an increased risk of colonization by *Campylobacter* due to poor health compromising the immune response of birds in the flock (T. Humphrey, 2006). Social stress (Mohamed and Hanson, 1980), heat stress (Burkholder et al., 2008), and overcrowding stress (Gomes et al., 2014), have all been shown to increase susceptibility to disease in chickens by compromising the immune response (Hirakawa et al., 2020; Heckert et al., 2002), and in many cases have been correlated with increased risk of colonization with *Salmonella* (Alhenaky et al., 2017; Gomes et al., 2014). As such, it was initially assumed that similar measures may increase incidence of *Campylobacter* in broiler flocks. While our work has revealed some level of correlation between poor welfare metrics and *Campylobacter* incidence (see the conditional probabilities of Appendix D3), these relationships were not found to be statistically significant via a logistic regression model, and our Bayesian network model suggests that poor bird welfare, as judged by the measures used here, is not in fact a cause of increased *Campylobacter* colonization. Despite this, our model reveals many yet-unconsidered relationships between rearing variables, provides evidence against multiple existing hypotheses, and highlights multiple promising new lines of enquiry towards identifying the source of *Campylobacter* colonization in commercial poultry flocks.

Our logistic regression analysis, shown in Table 5.4, identified three statistically significant explanatory variables; *Breed*, *Minimum temperature* and *Hock burn percentage*, with p values of 0.0189, 7.69×10^{-6} , and 0.0304 respectively. Seasonal

variation in *Campylobacter* incidence has long been observed in broiler flocks (Louis et al., 2005; Jorgensen, Ellis-Iversen, et al., 2011), with minimum/maximum temperature and sunshine hours significantly correlated with both the incidence and total bacterial load found in chicken flocks (Wallace et al., 1997). The warmer summer months see greater *Campylobacter* prevalence, yet despite the large body of research confirming this phenomenon, the precise mechanism for this increase remains unclear. While the growth rate of *Campylobacter* is found to vary in relation to temperature (Doyle and Roman, 1981), the minimum temperature required for *Campylobacter* growth is estimated to be around 30 degrees centigrade, somewhat precluding the impact of UK seasonal temperatures. Previous studies have suggested that the seasonal increase of flies (Hald, Skovgård, et al., 2004; Hald, Sommer, et al., 2007), rodents (Meerburg and Kijlstra, 2007), and wild birds (Colles, T. A. Jones, et al., 2008) as vectors of *Campylobacter* transmission may be responsible, while seasonal patterns in country-wide clonal complex incidence potentially point to genetic adaptation to seasonal trends (Jorgensen, Ellis-Iversen, et al., 2011). Investigating this trend in human incidence of campylobacteriosis, Djennad et al. (2019) conducted a rigorous statistical assessment of spatial and weather factors, concluding that the correlation between incidence and temperature was “likely to be indirect”. Our above results reach the same conclusion for broiler colonization. While our logistic regression shows the strong correlation between temperature and *Campylobacter* colonization, our Bayesian network analysis shows in Figure 5.2 that the two variables are conditionally independent upon the date placed, i.e. the correlation is indirect.

Footpad dermatitis, commonly referred to as ‘hock burn’, the dark discolouration and ulceration of the lower leg of birds, was also found to be statistically significantly correlated with *Campylobacter* prevalence, however this relationship was curiously found to be negatively correlated. These painful lesions are considered a sign of poor bird welfare, usually caused by litter unsuitably saturated with chicken waste. As such, the suggestion that more instances of hock burn in a flock are linked with less

cases of *Campylobacter* is surprising, considering that the bacteria are transmitted via the faecal-oral route. One hypothesis is that the presence of *Campylobacter* may in turn limit colonization of the flock by more pathogenic bacteria that could more easily trigger diarrhoea within a host-bird, thus impacting the litter quality and the resulting development of hock burn. Alternatively this relationship may be an artifact of how the *Hock burn percentage* variable was defined. Namely, it was recorded as the cross-sectional prevalence of any signs of hockburn in the flock (Dawkins, Roberts, et al., 2017). In short, it is a measure of how many birds showed signs of hock burn, and not a measure of the extremity of these burns. Bull, Thomas, et al. (2008) observed this same effect, whereby the flock-wide presence of hock burn was generally higher in *Campylobacter* negative flocks, however the number of birds in the flock rejected from consumption due to extreme cases of hock burn was positively correlated with rates of *Campylobacter* colonization. Figure 5.2 also concludes that this correlation between *Campylobacter* colonization and hock burn prevalence is conditionally independent upon the managing company.

The Bayesian network structure displayed in Figure 5.2 reveals a wide variety of insight into the various interrelationships of the included variables. Firstly we see that the number of parent flocks a broiler flock is sourced from, and the mean age of these parent flocks, had no meaningful impact on association with any other variable. The feasibility of vertical transmission of *Campylobacter* from parent to broiler flock is still frequently discussed in the literature, and the inclusion of this variable was based upon the hypothesis that a greater number of parent flocks may challenge a broiler flock with a greater genotypic variety of *Campylobacter* isolates (Petersen et al., 2001). Parent age has also been shown to influence egg weight and embryo weight of chicks (Shanawany, 1984). Given the potential importance of maternal antibodies in suppressing *Campylobacter* in the first few weeks of age (Rawson, Dawkins, et al., 2019), parent age could potentially impact the likelihood of *Campylobacter* colonization. Our results however indicate that factors relating to the parent flock have no effect on any of the metrics considered in this study.

The logistic regression analysis of the minimum adequate model found statistical significance in the relationship between *Breed* and *Campylobacter* colonization, where flocks of Breed A birds were more frequently observed to become colonized than Breed B birds. Caffrey et al. (2021) recently identified a correlation between breed and *Campylobacter*, with flocks comprised of Cobb birds, or a mixture of Cobb and Ross birds, 4.75 times more likely to test positive for fluoroquinolone resistant *Campylobacter jejuni* than flocks comprised of just Ross birds. Further to this, Cobb birds have been found to be more frequently colonized by *Campylobacter* than Hubbard birds by Babacan et al. (2020), however they were unable to separate this association from other rearing factors such as age-of-slaughter. Our Bayesian network analysis, similar to the hock burn conclusions, was unable to detect any direct arc of causation between *Breed* and *Campylobacter* colonization, suggesting that the breed of chicken is indicative of the company managing the flock, and not necessarily an indicator of a breed-specific susceptibility. Host-bird genetics have however previously been shown to cause differences in host-resistance to *Campylobacter* challenge (Stern, Meinersmann, et al., 1990; X. Li et al., 2008; Connell et al., 2013), with such resistances shown to be inheritable under experimental conditions (Boyd et al., 2005). Further linking breed and welfare measures, S. Humphrey et al. (2014) found that faster-growing breeds of broiler showed evidence of prolonged inflammation in the intestines in response to *Campylobacter jejuni*, suggesting that the impact of breed is yet a plausible route of further study. An interesting relationship observed in Figure 5.2 was the implication of *Breed* as a determinant of the prevalence of pododermatitis, with flocks of Breed A birds more frequently displaying heavy incidence of pododermatitis. No study to our knowledge has directly investigated this supposed relationship in broilers, however one study in turkeys found no correlation between breed and pododermatitis (S. Clark et al., 2002). Pododermatitis has previously been shown to be associated with a poor-nutrient diet (Nagaraj et al., 2007), hence the hypothesis that this factor could correlate with general gut health and/or the composition of the gut microbiome.

The primary conclusion of our work, as shown in Figure 5.2, was that our network of variables was closely related by yet-unobserved factors concealed within the *Company* variable. *Company* was found to be a parent variable to six factors; *Breed*, *Farm*, *Number placed*, *7-day mortality*, *Hock burn percentage*, and *7-day weight*. This indicates that these six factors significantly vary, due to which of the two companies considered within this study they are managed by. This suggests that choices made within the complex decision network relating to the rearing of these flocks, encompassing factors such as diet, water provision, housing, thinning protocols, cleaning regimens, antibiotic usage, and stocking density among others (Sibanda et al., 2018), will have the significant potential to both decrease incidence of *Campylobacter* and may simultaneously improve the welfare of the flock. While disappointing to not ascertain the primary root cause of increased *Campylobacter* prevalence from within our considered set of variables, the work has revealed a key network of dependencies within commonly recorded and studied metrics. While far from the first study to examine the contributions of multiple health factors towards *Campylobacter* colonization (T. Humphrey et al., 1993; Rushton et al., 2009; Babacan et al., 2020; Frosth et al., 2020), our work is the first, to our knowledge, to utilise the powerful methodologies underlying Bayesian network analysis in studying the spread of *Campylobacter*. Such approaches, in combination with more traditional logistic regression analyses, greatly increase the descriptive power of gathered datasets, and it is our hope that this work will help expedite their adoption throughout the field of *Campylobacter* risk management. Bayesian networks have had some early success already in specifically implicating welfare measures with specific housing variables (Comin et al., 2019), we now further our attempts to identify the variables that exacerbate the spread of *Campylobacter*.

This study illustrates the need to investigate, more thoroughly, management decisions in the broiler industry, so as to reduce *Campylobacter* incidence whilst

improving bird health and welfare, to provide the consumer with a better product whilst reducing the impact of campylobacteriosis on human health.

Code Availability

The code used to produce all figures and tables presented is available at: osf.io/pb62g/.

6

Discussion

This thesis has presented four distinct mathematical approaches to understanding the spread of *Campylobacter* in chicken flocks. Previously published data was used within these new frameworks to answer specific questions relating to *Campylobacter* dynamics, with my results displaying the various relationships between host-bird health and the proliferation and diversity of *Campylobacter* strains within the flock. These methods are presented with the intention of providing frameworks that can be adapted, applied to future works and previously collated *Campylobacter* data, and to stimulate the uptake and consideration of modelling approaches to tackling the public health burden of *Campylobacter*.

In this chapter I summarise the methods and findings of the previous chapters, before suggesting further routes of work to pursue. Finally, I present my closing conclusions and how these conclusions may be incorporated into commercial practice.

Summary of Chapters

Chapter 1 provided an introduction to the public health crisis posed by *Campylobacter*. I discussed the extent to which the poultry industry is particularly burdened by *Campylobacter*, and why broilers are thought to be such a significant reservoir of the bacteria. I then provided a summary of the research landscape

to date, with several extensive reviews signposted throughout the years (Adkin et al., 2006; Newell, Elvers, et al., 2011; Hermans et al., 2011; Awad, Hess, et al., 2018). Government-led initiatives to reduce *Campylobacter* incidence had primarily focused on biosecurity measures up to 2015, attempting to increase the hygiene quality in farms to prevent *Campylobacter* from first entering a flock. The focus on such measures did not result in decreased incidence, either proving ineffective, or not being correctly implemented, leading to an increased call for fundamental research into the ecology of *Campylobacter*, and greater consideration of the wider associations between *Campylobacter* and host-chicken health and welfare, a so-called ‘One Health’ approach (Gölz et al., 2014). The broad aim of this thesis was to answer such a call.

In Chapter 2, I presented the first significant mathematical model of *Campylobacter* dynamics in a flock of broiler chickens. A system of stochastic differential equations (SDEs) was constructed to simulate the dynamics of *Campylobacter* in the gut, its competitive effects with other native bacteria, and its suppression by early-life maternal antibodies. Such a system was then expanded from a single bird to simulate the gut dynamics of an entire flock of broilers, with *Campylobacter* strains moving from one host to another via the shared environment. Comparing model predictions against multiple datasets presented over the last 30 years, I found that my model could successfully predict a variety of observed dynamical behaviours, such as that birds housed in isolation showed a time-dependent likelihood of successfully clearing *Campylobacter* infection. This probability of *Campylobacter* dying out from the gut was then shown to decrease as the number of birds housed in the environment was increased. Expanding the model to consider a multitude of *Campylobacter* strains, I showed that demographically disadvantaged strains, such as those with reduced growth rates, could be quickly removed from a flock via competitive exclusion. Furthermore, while multiple strains could be isolated simultaneously from a flock, within individual chickens the gut would be broadly dominated by a single strain of *Campylobacter* at any given time, with the resident dominant strain changing on a

roughly weekly basis. This notable difference between flock-wide, and single-bird dynamics was experimentally shown by Colles, Preston, et al. (2019) shortly prior to the publication of my theoretical findings. Lastly, I conducted a variance-based sensitivity analysis to determine which model parameters had the greatest impact in preventing (or expediting) the colonisation of a flock. The demographic parameters of the other native gut bacteria were found to have the greatest impact on preventing *Campylobacter* colonisation, suggesting that competitive exclusion strategies could be a beneficial further avenue of pursuit.

Wishing to explore the interplay between *Campylobacter* and broiler further, Chapter 3 presented a series of Bayesian models to investigate a large and informative previously published dataset, reporting how a community of *Campylobacter* STs was transmitted between individual birds in a flock over time. By investigating how ST-specific dynamics were impacted by time, species, and host-birds, I discovered the existence of ‘super-shedder’ birds within the flock. While some chickens would prove especially resilient to colonisation by *Campylobacter* for the duration of the experiment, other birds would be consistently colonised by a wide range of STs with frequent turnover in the dominant ST. While the role of ‘super shedders’ had been discussed in the transmission of *Salmonella* (Menanteau et al., 2018), this was the first study to my knowledge to demonstrate the existence, and more importantly, the dynamical impact of such birds within a flock colonised by *Campylobacter*. My findings here, namely that individual bird status is a more important driver of *Campylobacter* ST prevalence than the demographic capabilities of the ST itself, were then further corroborated by simulations using my previous model presented in Chapter 2 (shown in Appendix B2). This work further demonstrated how mathematical tools can be used to test dynamical hypotheses in this host-bacteria system.

To investigate further how multiple STs are able to co-occur and persist within a flock, Chapter 4 used the same dataset as in Chapter 3 to determine if demographic

factors alone could explain the maintenance of the observed levels of *Campylobacter* biodiversity, and to evaluate the competitive relationship between STs. Using a competition matrix model, the community of *Campylobacter* STs was shown to exist within a competitive hierarchy, with certain STs repeatedly being able to outcompete other STs, replacing them in host birds. A patch-occupancy model was then constructed to quantify both the transmission and mortality of each ST, and to investigate further whether observed differences in ST dynamics could be explained by these two parameters. While some STs were observed to occur in greater frequency, or for longer durations, no pattern could be drawn between ST demographics and their respective population dynamics. This suggests that environmental variables, namely, the host chicken occupied, were the most likely mechanism by which biodiversity was maintained within a flock, further supporting the findings of Chapter 3.

Lastly, as findings from previous chapters (Chapters 2-4) repeatedly implicated bird health with increased risk of *Campylobacter* colonisation, Chapter 5 used data reporting multiple rearing and welfare factors for a multitude of commercial broiler flocks to investigate the causal links between such variables. By combining and comparing the use of a logistic regression model and a Bayesian network model, correlations and causal dependencies were demonstrated between flock colonisation and environmental and welfare factors such as season, hock burn incidence, bird weight, breed, and parent flocks, among others. While the logistic regression was able to demonstrate statistical significance between the breed of the flock, hock burn incidence, and the external temperature with *Campylobacter* colonisation, the more sophisticated Bayesian network analysis was able to provide greater insight into how these correlations are structured. I showed that, while such a correlation could be drawn between variables, *Campylobacter* prevalence was conditionally dependent on the majority of other welfare factors associated with the rearing company. Namely, that some yet-unidentified decisions in rearing or management practice were simultaneously negatively influencing bird welfare and

increasing *Campylobacter* incidence. *Campylobacter* incidence and bird welfare did not directly affect one another, but rather, were highlighted in this chapter as symptoms of the same root cause.

Further Work

As each chapter introduces a new modelling framework, each individual chapter can be expanded, utilised, and queried in many novel ways. Here I list some of the most pertinent developments that may now be borne from these studies, but do not doubt that researchers in both academia and industry will continue to raise yet-unconsidered questions that these models will be able to inform.

- Solution existence and stability:** Chapter 2 introduced an ODE and SDE framework with which to simulate multiple strains of *Campylobacter* in a flock of multiple broilers. While I presented a variety of case studies to demonstrate the informative potential of the model, the wider analysis of model stability and solution existence requirements was not performed for this thesis. While Chapter 4 investigated the data-driven differences in transmission and mortality between strains of *Campylobacter*, performing a stability analysis of either the deterministic or stochastic model of Chapter 2 would provide robust mathematical results on what determines whether a newly-emergent strain will persist in a flock or not. In short, one may seek the qualifying conditions of the gut microbiome required for a new strain to achieve a stable, non-zero, equilibrium state. Fan et al. (2015) list a multitude of methodologies by which such a study could be conducted; Young et al. (2015) performed such a study using a simplified model of *Salmonella* dynamics in an unspecified host organism, demonstrating conditions under which *Salmonella* may or may not outcompete the native gut flora. Similar studies could now be undertaken for *Campylobacter*.
- Optimal control solutions to *Campylobacter* prevention:** Using a system of differential equations, one may optimise a certain state variable

by dynamically changing certain control variables. Using the methodology laid out in Pontryagin's maximum principle (Lenhart and Workman, 2007), my mathematical framework presented in Chapter 2 may be used to answer a variety of optimisation questions, subject to a variety of constraints. For example, one may investigate the optimal combination of a variety of improved housing measures to minimise *Campylobacter* prevalence, subject to economic constraint. Or, one may query how to optimally reduce the likelihood of *Campylobacter* colonisation, subject to constraints regarding stocking density and lifespan. Such approaches would be beneficial to both investigating plausible new intervention methods, and informing optimal rearing practices.

- **Identification and bacterial load of super-shedders:** Chapter 3 is the first study to my knowledge that identifies the presence, and importance, of super-shedders in the spread of *Campylobacter* throughout a broiler flock. I have shown that some birds in a flock will consistently be at greater risk of colonisation by *Campylobacter*, and appear to influence the colonisation status of other birds in the flock. Following on from this, experimental investigation is urgently required both to quantify the frequency and bacterial load of shedding from these super-shedder birds, and then more importantly, to uncover ways by which these birds may be identified within a flock. If, as my studies suggest, the resilience of a flock to colonisation is only as great as their most susceptible bird, then methods by which to identify, and then improve the health of, such super-shedder birds presents a vital tool in preventing flock-wide colonisation. We stress however that such findings should be further investigated alongside maintaining good biosecurity, not as a replacement. Again, such a study has already been conducted verifying the significance of super-shedder broilers in the spread of *Salmonella* (Menanteau et al., 2018), and similar studies should be undertaken for *Campylobacter*.
- **The ecology of *Campylobacter* in flocks:** The maintenance of such biodiversity of *Campylobacter* STs, as presented in Chapter 4, is particularly

curious given the strict competitive hierarchy observed. Results from Chapter 4 reveal that such persistent co-occurrence cannot be explained by ST-specific factors alone, and that variation in bird susceptibility must drive this effect. While the model of Chapter 2 is able to demonstrate the dynamic impact of differences in the microbiome of host-birds, demographic advantage is still slowly selected for within this system at a greater rate than is seemingly observed from the data of Chapters 3 and 4. The ecological literature is vast, with numerous theoretical frameworks and explanations for the maintenance of biodiversity in specific ecosystems. While the depths of this research could not be fully explored during this thesis, it is likely that many such mechanistic hypotheses for the observed biodiversity of *Campylobacter* can be further investigated within these ecological paradigms.

- **Identification of key management practice variables:** Chapter 5 considered a variety of rearing practice and welfare variables, investigating the causal relationships between such variables. It was found that the majority of factors considered were conditionally dependent on one-another via the ‘Company’ variable. Namely, that some variable not considered or recorded in this study, but one that varies between companies, was responsible for both increased risk of *Campylobacter* colonisation, but also increased incidence of measures of poor bird welfare. Further investigation into such factors is required. One potential variable to consider next would be feed nutrition and delivery, as work by Han, Willer, Pielsticker, et al. (2016) has shown that the nutrient-rich feed given to longer-lived layer birds was correlated with reduced *Campylobacter* gut populations compared to the high-energy broiler feed. Furthermore, the majority of the variables reported in Chapter 5 can be readily obtained and recorded by farm management during rearing. Negotiating the provision of such data with management companies would provide the necessary data each year with which to strengthen the conclusions and links within my Bayesian network analysis.

Conclusions

All the work within this thesis supports an ultimate conclusion; that increased focus and research into the health of individual broilers within the flock presents a great opportunity by which to limit the spread of *Campylobacter*. Mechanistically, this may involve improving the host-chicken immune response (Chapter 2), enabling a reduction in the number of ‘super-shedder’ birds that expedite total flock colonisation (Chapter 3), limiting the number of *Campylobacter* STs supported simultaneously (Chapter 4), or that observable changes to health and welfare may predict the emergence of *Campylobacter* without the need for time-consuming sample collection and testing (Chapter 5).

It is one thing to make such a conclusion, and another to see it meaningfully enacted. The rearing of chickens has been heavily industrialised, with flocks grown in the tens of thousands simultaneously, with an overall economic net margin of 3 pence to be made per bird (Sheppard, 2004). Meaningful change to the *Campylobacter* load of commercial poultry can only be made possible through the close collaboration of science and industry. Clear communication is required between the two to demonstrate the commercial incentive of ensuring good broiler health and welfare (Dawkins, 2017), but equally to understand how plausible intervention can be achieved. While further research on commercial broiler flocks is still required, the research and development of monitoring technologies to identify where management intervention is necessary are of key importance to enact my conclusions. While the pathogenic potential of *Campylobacter* to broilers has been identified (S. Humphrey et al., 2014), these health impacts are difficult to observe. Technologies such as the optical flow analysis presented by Dawkins, H.-J. Lee, et al. (2009) are an example of an interface by which farmers may be alerted to differences in flock movement patterns, which have been shown to predict more accurately both colonisation by *Campylobacter* (Colles, Cain, et al., 2016) and future mortality and hock burn (Dawkins, Wang, et al., 2021). As I have concluded, the development and application of such technologies will be necessary to know when to respond to the health needs

of broilers.

Moreover, I have demonstrated the wide array of insight that can be drawn through the application of various mathematical techniques to answering questions relating to the proliferation of *Campylobacter* through broiler flocks. This in itself likely did not need explicitly demonstrating, as mathematical models have been integral to the understanding and research of other foodborne pathogens such as *Salmonella* (Mastroeni et al., 2009; Carrasco et al., 2012) and *E. coli* (Murray and Howard, 2019; Ongeng et al., 2014). Rather, it is surprising that so few mathematical modelling studies of *Campylobacter* exist to date. Not only have I demonstrated the unique insights that can be provided by such approaches, but one will notice that no new experimental studies were performed as part of this thesis, and that all conclusions were borne out of data spanning back as late as 1998 (Achen et al., 1998). There have been countless flock-wide population studies of *Campylobacter* over the last several decades, and this data has been greatly underutilised to date. Further utilisation of such mechanistic and empirically-derived theoretical methodologies will enable the synthesis and testing of new hypotheses, the simulation of cellular behaviour, and the validation of observed phenomena. The development of descriptive mathematical models is inherently modular and iterative in nature (Stelling, 2004). It is my hope that the work presented in this thesis will act as the first steps towards a robust systems-level understanding of *Campylobacter* in poultry flocks, and the food chain at large.

Appendices

A

Supplementary Material for Chapter 2

Contents

A1 - Model boundedness	148
A2 - Simulated annealing	151

A1 - Model boundedness

The basic deterministic core of the model is:

$$\frac{dC}{dt} = r_1 C (1 - (C + \alpha_1 B)) - \gamma CP - d_1 C - \beta CB - \sigma CM, \quad (\text{A.1})$$

$$\frac{dB}{dt} = r_2 B (1 - (B + \alpha_2 C)) - d_2 B, \quad (\text{A.2})$$

$$\frac{dP}{dt} = \xi CP(1 - P) - d_3 P, \quad (\text{A.3})$$

$$\frac{dM}{dt} = -d_4 M. \quad (\text{A.4})$$

We can show that C and B are bounded between 0 and 1. If $C = 0$, then $\frac{dC}{dt} = 0$. Likewise if $B = 0$, then $\frac{dB}{dt} = 0$. Hence neither variable is capable of decreasing below 0, and they are bounded below at 0.

If $C = 1$, then equation (A.1) becomes:

$$\left. \frac{dC}{dt} \right|_{C=1} = r_1(-\alpha_1 B) - (\gamma P + d_1 + \beta B + \sigma M).$$

Since all parameters are positive, and we have shown that $B > 0$, we can conclude that $\left. \frac{dC}{dt} \right|_{C=1} < 0$. Therefore C cannot exceed one, and is bounded above. We can likewise show the same for B .

Likewise equation (A.3) shows that when $P = 0$, $\frac{dP}{dt} = 0$. When $P = 1$, $\frac{dP}{dt} = -d_3$. Hence P is also bounded between 0 and 1.

M is clearly bounded as it is only decreasing down to 0, and we initialise it at a value less than 1.

There is however some issue in our formulation of the stochastic system, which we present again here,

$$\begin{aligned}
dC_{ij} = & \left[r_{C_j} C_{ij} \left(1 - \left(\sum_{j=1}^N C_{ij} + \alpha_1 B_i \right) \right) - \gamma_{C_j} C_{ij} P_i - d_{C_j} C_{ij} \right. \\
& \left. - \beta_{C_j} C_{ij} B_i - \sigma_{C_j} C_{ij} M_i + a(E_j) \right] dt \\
& + [\eta_{C_j} C_{ij} + \lambda_j(t) - \eta_{BC_j} C_{ij} B_i] dW_t,
\end{aligned} \tag{A.5}$$

$$dB_i = \left[r_2 B_i \left(1 - \left(B_i + \alpha_2 \sum_{j=1}^N C_{ij} \right) \right) - d_2 B_i \right] dt + [\eta_2 B_i] dW_t, \tag{A.6}$$

$$dP_i = \left[\sum_{j=1}^N \xi_j C_{ij} P_i (1 - P_i) - d_3 P_i \right] dt + [\eta_3 P_i] dW_t, \tag{A.7}$$

$$dM_i = [-d_4 M_i] dt + [\eta_4 M_i] dW_t, \tag{A.8}$$

$$dE_j = \left[\sum_{i=1}^L b C_{ij} \left(1 - \frac{E_j}{\Omega} \right) - d_5 E_j \right] dt + [\eta_5 E_j] dW_t, \tag{A.9}$$

where $\lambda_j(t)$ is defined by;

$$\lambda_j(t) = \begin{cases} 0, & \text{if } C_{ij}(t) = 0. \\ 2.873 \times 10^{-4}, & \text{otherwise.} \end{cases}$$

and where $a(E_j)$ is defined by;

$$a(E_j) = \begin{cases} 0.015, & \text{if } X < \frac{E_j}{\Omega} \text{ for random variable } X \sim \mathcal{U}(0, 1). \\ 0, & \text{otherwise.} \end{cases}$$

Having terms independent of their respective variables is necessary to ensure that the model may simulate new infections and extinction events; however, this does leave the model analytically unbounded. To fix this issue, when the model was solved numerically we used min and max functions to ensure that the model was fixed between 0 and 1. To express the model in a more rigid mathematical framework, we can include the Kronecker delta function.

The Kronecker delta is defined as:

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j. \\ 1, & \text{if } i = j. \end{cases}$$

As such, we can better define the model by changing equations (A.5) to (A.9) to,

$$\begin{aligned} dC_{ij} = & \left[r_{C_j} C_{ij} \left(1 - \frac{\sum_{j=1}^N C_{ij} + \alpha_1 B_i}{K} \right) - \gamma_{C_j} C_{ij} P_i - d_{C_j} C_{ij} \right. \\ & \left. - \beta_{C_j} C_{ij} B_i - \sigma_{C_j} C_{ij} M_i + (1 - \delta_{(1)(C)}) a(E_j) \right] dt \\ & + \left[(1 - \delta_{(1)(C)})(1 - \delta_{(0)(C)})(\eta_{C_j} C_{ij} + \lambda_j(t) - \eta_{BC_j} C_{ij} B_i) \right] dW_t, \end{aligned} \quad (\text{A.10})$$

$$dB_i = \left[r_2 B_i \left(1 - \frac{B_i + \alpha_2 \sum_{j=1}^N C_{ij}}{K} \right) - d_2 B_i \right] dt + \left[(1 - \delta_{(1)(B)}) \eta_2 B_i \right] dW_t, \quad (\text{A.11})$$

$$dP_i = \left[\sum_{j=1}^N \xi_j C_{ij} (1 - P_i) - d_3 P_i \right] dt + \left[(1 - \delta_{(1)(P)}) \eta_3 P_i \right] dW_t, \quad (\text{A.12})$$

$$dM_i = [-d_4 M_i] dt + \left[(1 - \delta_{(1)(M)}) \eta_4 M_i \right] dW_t, \quad (\text{A.13})$$

$$dE_j = \left[\sum_{i=1}^L b C_{ij} \left(1 - \frac{E_j}{\Omega} \right) - d_5 E_j \right] dt + \left[(1 - \delta_{(1)(E)}) \eta_5 E_j \right] dW_t, \quad (\text{A.14})$$

where $\lambda_j(t)$ is defined by;

$$\lambda_j(t) = \begin{cases} 0, & \text{if } C_{ij}(t) = 0. \\ 2.873 \times 10^{-4}, & \text{otherwise.} \end{cases}$$

and where $a(E_j)$ is defined by;

$$a(E_j) = \begin{cases} 0.015, & \text{if } X < \frac{E_j}{\Omega} \text{ for random variable } X \sim \mathcal{U}(0, 1). \\ 0, & \text{otherwise.} \end{cases}$$

This alteration was not included in the main manuscript for the sake of readability.

A2 - Simulated annealing

Inspired by the process of annealing in metallurgy, simulated annealing is a process of finding a global extrema (minimum or maximum value) of a function. The most common methods of hill-climb algorithms or gradient-descent are incredibly efficient, but for cases where there are many local extrema, these methods can become easily stuck in local extrema, and are unable to detect the “true” minimum or maximum value of the function.

Simulated annealing tackles this problem by taking a probabilistic approach. Consider a minimisation problem. Other methods will consider neighbouring points and move there if the output value is lower. Simulated annealing, however, will consider a neighbouring point and instead have a *probability* of moving to this point. This probability is dependent on the value at the two points being considered, and a time-dependent variable T , that slowly decreases to 0 as the algorithm proceeds. The probability of moving to a neighbouring position will be higher if the new position gives a desired lower value, and will also be higher the greater the value of T , i.e. the earlier it is in the process’ runtime.

As time progresses, and T converges to 0, the probability of moving position also slowly converges to 0. By having the possibility of moving to a “poorer choice” of neighbour, this method searches a greater amount of the parameter space. This method was ideal for use in this project, as it was applied to minimising a cost function with a large quantity of local minima, across a very large parameter space. The cost function in our case was the distance between the model predictions and experimental data presented in Figure 2.2. As such, the method was able to provide an accurate parameter set to use for the ensuing case studies. The drawbacks of the method are its substantially longer runtimes, which increases exponentially for multi-variate functions.

B

Supplementary Material for Chapter 3

Contents

B1 - Bayesian statistics	154
B2 - Model simulations of flock health	156

B1 - Bayesian statistics

This brief section aims to convey the basic principles of Bayesian statistics, and familiarise the reader with the terminology that is used throughout the thesis. For an in-depth explanation, I recommend the text by Kruschke (2014).

Bayesian statistics is derived wholly from the relationship defined by Bayes' theorem,

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}. \quad (\text{B.1})$$

If we consider θ as some statistical parameter we wish to infer, and D as some data informing the parameter, then equation (B.1) expresses that the probability distribution for our value of θ , given our dataset ($P(\theta|D)$), is proportional to the **likelihood** of such data ($P(D|\theta)$) multiplied by the probability distribution of θ free of any data ($P(\theta)$).

Spoken plainly, one starts with a **prior** probabilistic understanding of the values θ , often informed by expert opinion, and by utilising relevant data, D , we update our belief in the values θ may take, producing a new **posterior** distribution. Mnemonically, if we wished to calculate the probability that a flipped coin will land heads up, we may have a **prior** belief that the coin is fair. However, upon observing a data set of 5 coin flips, all of which produced heads, we may update our **posterior** belief to reflect that the coin may be biased.

The analytical difficulty in this calculation lies in computing $P(D) = \int P(D|\theta)P(\theta)d\theta$, which is often near impossible for realistically complex models. Fortunately modern computing power enables us to efficiently estimate our posterior distributions through algorithms such as Gibbs sampling and other Metropolis-Hastings schemes.

Hierarchical systems represent multi-variable models where some parameters depend on other parameters. Returning to the example of a coin flip, say the probability

of heads (θ) is dependent on the factory in which the coin was minted. The probability that a coin was from a certain factory (ω) will then inform our value of (θ). Expressed mathematically, equation (B.1) now becomes:

$$\begin{aligned} P(\theta, \omega|D) &= \frac{P(D|\theta, \omega)P(\theta, \omega)}{P(D)} \\ &= \frac{P(D|\theta, \omega)P(\theta|\omega)P(\omega)}{P(D)}. \end{aligned} \tag{B.2}$$

This means that a prior distribution is only required for ω , as this distribution will directly inform our **conditional prior** of θ , via our model formulation. As such, when provided with data on coin flips from multiple coins from different factories, we obtain a posterior probability distribution of which factory a coin has come from, and the resulting probability of a coin flip resulting in heads. This structure of conditional independence means that data relating specifically to one parameter can still help inform the posterior of all other dependent variables, a key advantage of Bayesian inference.

B2 - Model simulations of flock health

After confirming from Model 5 the existence of variation in bird transition probabilities, we asked what the impact of this variation could be on the proliferation of *Campylobacter* STs. Using a previously published stochastic differential equation model of *Campylobacter* population dynamics within a broiler flock (Rawson, Dawkins, et al., 2019) we simulated two variant scenarios. Figure B.1 displays a case study of the spread of five demographically identical strains of *Campylobacter* within a flock of 400 demographically identical broilers. Figure B.1 shows that, as expected, all strains perform equally well and are equally represented in the amount being shed into the environment. Figure B.2 instead shows the same model of five demographically identical strains of *Campylobacter* within a flock of 400 birds whose strength of immune response is drawn from a normal distribution centred around the value used for Figure 3.8. Figure B.2E shows how five demographically equal strains can be sustained at broadly different levels across the flock due only to variation in bird immune response. This is caused by random chance, in that whichever strain is initially picked up by a super-shedder, such as the one shown in Figure B.2D then sheds large amounts of that strain of *Campylobacter* into the environment, increasing the likelihood of then infecting other birds in the flock. This result greatly implies that the results shown in the data, whereby some STs seem to persist at higher levels than others in the flock, is likely due to the variation in bird transition probabilities, as opposed to phenotypic differences between STs.

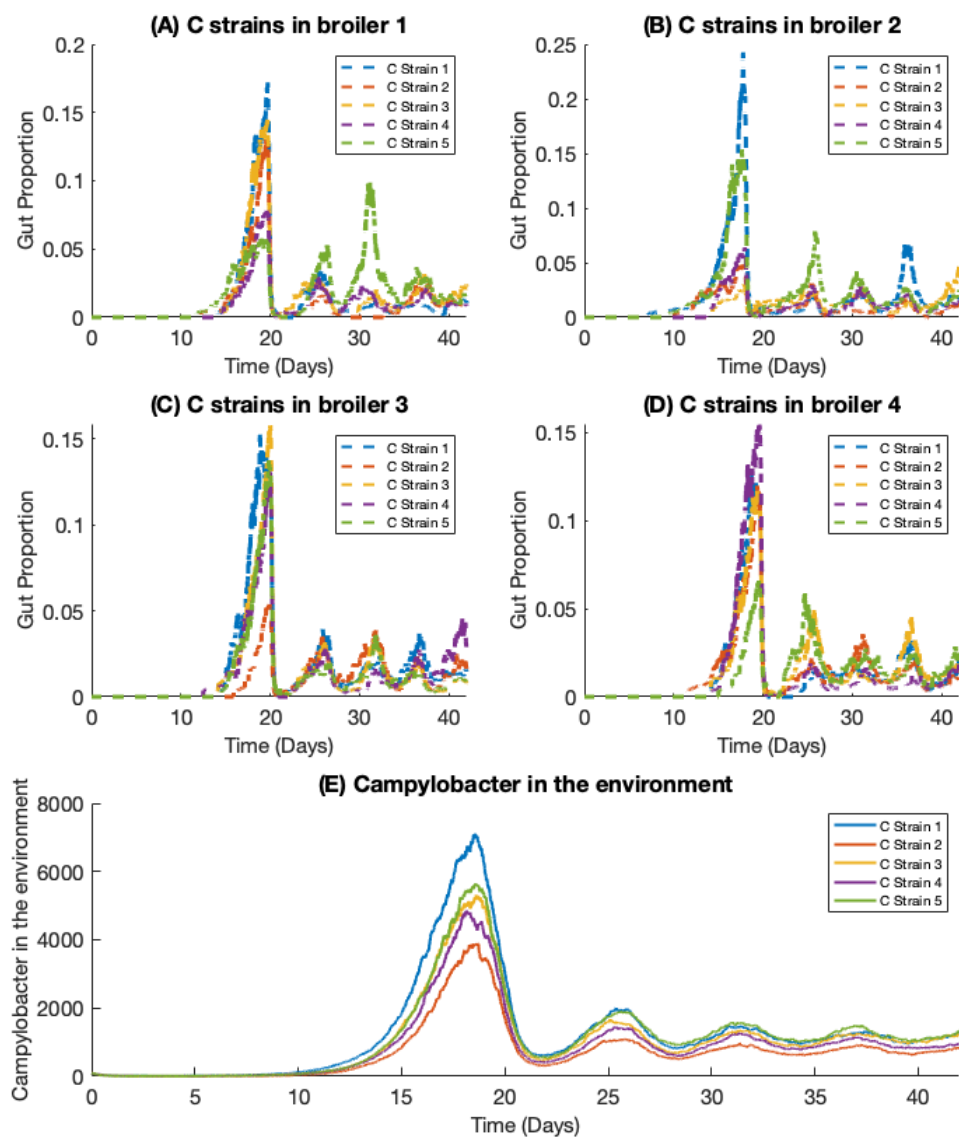


Figure B.1: Dynamic behaviour of five identical strains of *Campylobacter* in a flock of identical broilers. (A) - (D) shows the population within the gut of individual broilers, while (E) displays the amount of *Campylobacter* in the environment, an expression of the average amount throughout the flock.

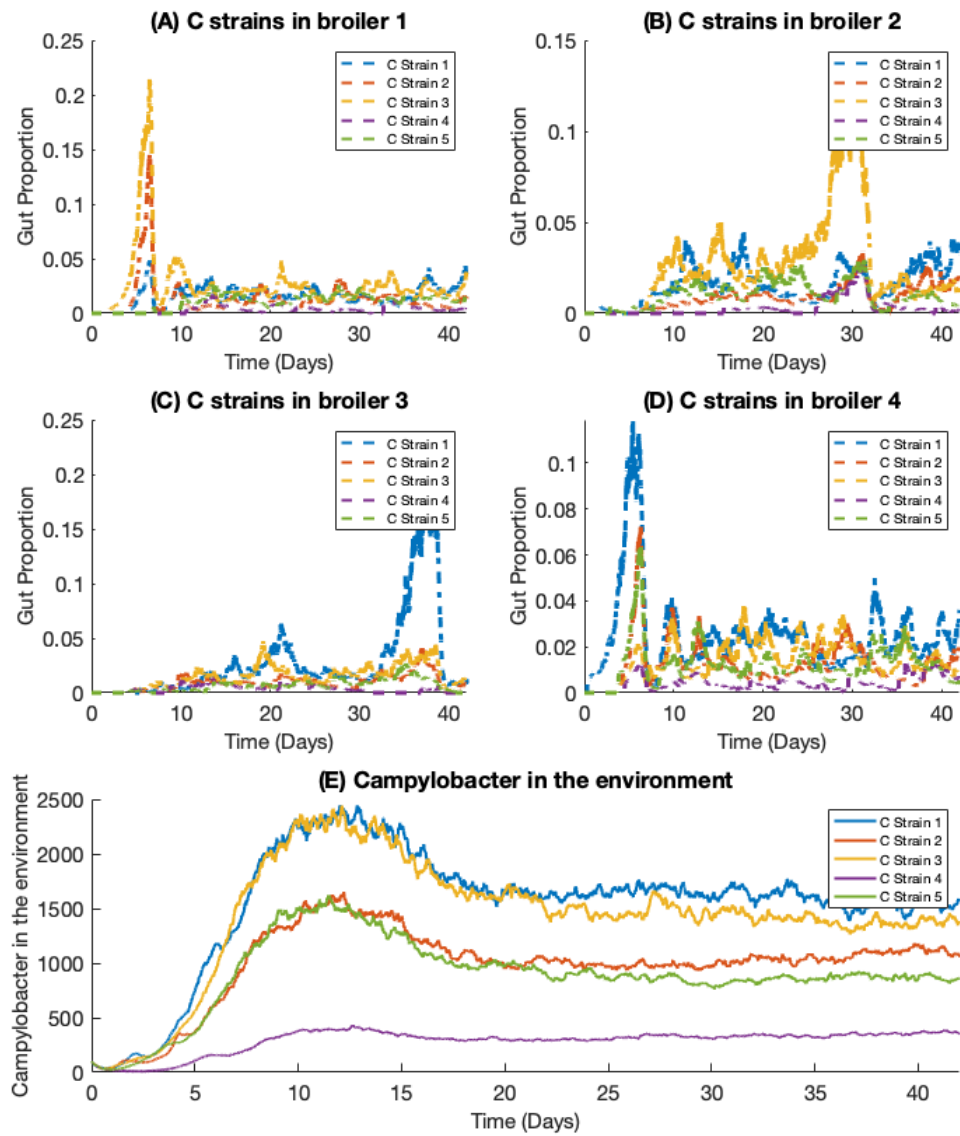


Figure B.2: Dynamic behaviour of five identical strains of *Campylobacter* in a flock of broilers of varying susceptibility to infection. (A) - (D) shows the population within the gut of individual broilers, while (E) displays the amount of *Campylobacter* in the environment, an expression of the average amount throughout the flock.

C

Supplementary Material for Chapter 4

Contents

C1 - Patch-occupancy model pseudo-code	160
--	-----

C1 - Patch-occupancy model pseudo-code

Algorithm 1: Patch-occupancy model pseudo code

```

1  Initialise chickens with STs in proportion to first timestep in data.
2  for every timestep do
3      Prepare placeholder vector for current timestep, equal to previous
        timestep.
4      for every chicken do
5          if chicken currently colonised then
6              Record currently occupying ST  $s$ .
7              Draw random number  $x$  from uniform distribution  $U(0, 1)$ .
8              if  $x < \mu_s$  then
9                  Remove ST  $s$  from chicken in placeholder vector.
10             else
11                 ST  $s$  will challenge other chickens:
12                 Draw random number  $y$  from  $\text{Pois}(\lambda_s)$ .
13                 Add  $y$  to a running tally,  $Y_s$ , of how many other chickens
                    will be challenged by ST  $s$ .
14             end
15         end
16     end
17 end
18 for every ST,  $s$  do
19     for  $j \leftarrow 1$  to  $Y_s$  do
20         Randomly select a chicken,  $c$ , to be challenged by ST  $s$ .
21         if  $c$  not colonised then
22             Chicken  $c$  is now colonised by ST  $s$  in placeholder vector.
23         else
24             Record currently occupying ST,  $r$ .
25             Draw  $z$  from uniform distribution  $U(0, 1)$ .
26             if  $z < C_{s,r}$  then
27                 ST  $s$  replaces ST  $r$  in placeholder vector.
28             else
29                 ST  $r$  remains in chicken  $c$  in placeholder vector.
30             end
31         end
32     end
33 end
34 end
35 end
36 Placeholder vector is assigned as frequency vector for current timestep.
37 end

```

D

Supplementary Material for Chapter 5

Contents

D1 - Logistic regression: maximal model to minimal adequate model	162
D2 - Bayesian network structure with BDE scoring	169
D3 - <i>Campylobacter</i> conditional probability tables	171

D1 - Logistic regression: maximal model to minimal adequate model

First we present the logistic regression analysis of the maximal model:

Table D.1: Logistic regression analysis of the maximal model.

Predictor	β	SE β	Wald-test z-score	<i>p</i>
(Intercept)	8.958	7.00	1.28	0.2006
<i>Number placed</i>	-5.435×10^{-5}	6.928×10^{-5}	-0.785	0.433
<i>Breed</i> (1 = B, 0 = A)	-1.510	0.788	-1.916	0.0554
<i>Breed</i> (1 = A & B, 0 = A)	-1.213	1.815	-0.668	0.504
<i>Mean parent age</i>	0.0255	0.0406	0.629	0.529
<i>No. parent flocks</i> (1 = 2 flocks, 0 = 1 flock)	-0.456	0.755	-0.604	0.546
<i>No. parent flocks</i> (1 = 3 flocks, 0 = 1 flock)	-0.930	0.846	-1.1	0.271
<i>No. parent flocks</i> (1 = 4 flocks, 0 = 1 flock)	-1.372	1.519	-0.903	0.367
<i>28-day mortality percentage</i>	-0.339	0.318	-1.065	0.287
<i>Pododermatitis percentage</i>	-0.016	0.015	-1.142	0.253
<i>Hock burn percentage</i>	-0.0531	0.0240	-2.215	0.0268
<i>28-day average weight</i>	-0.0050	0.0047	-1.070	0.284
<i>Minimum temperature</i>	0.512	0.135	3.809	0.0001
Test			χ^2	<i>p</i>
Overall model evaluation *				
Likelihood ratio test			43.748	1.685×10^{-5}
Goodness of fit test				
Hosmer-Lemeshow			7.9779	0.436
McFadden's R^2	0.376			
Cox & Snell's R^2	0.406			

* Compared against null model.

We then iteratively remove the least significant variables, until only significant variables remain. Firstly we remove *Mean parent age*.

Table D.2: Logistic regression analysis of the maximal model with *Mean parent age* removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	8.887	6.863	1.295	0.195
<i>Number placed</i>	-7.137×10^{-5}	6.412×10^{-5}	-1.113	0.195
<i>Breed</i> (1 = B, 0 = A)	-1.382	0.755	-1.832	0.067
<i>Breed</i> (1 = A & B, 0 = A)	-1.155	1.806	-0.639	0.523
<i>No. parent flocks</i> (1 = 2 flocks, 0 = 1 flock)	-0.461	0.749	-0.616	0.538
<i>No. parent flocks</i> (1 = 3 flocks, 0 = 1 flock)	-0.973	0.850	-1.144	0.253
<i>No. parent flocks</i> (1 = 4 flocks, 0 = 1 flock)	-1.426	1.556	-0.916	0.360
<i>28-day mortality percentage</i>	-0.392	0.312	-1.257	0.209
<i>Pododermatitis percentage</i>	-0.017	0.015	-1.160	0.246
<i>Hock burn percentage</i>	-0.0527	0.0239	-2.201	0.0277
<i>28-day average weight</i>	-0.0038	0.0041	-0.917	0.359
<i>Minimum temperature</i>	0.497	0.128	3.883	0.0001
Test			χ^2	p
Overall model evaluation				
Likelihood ratio test ¹			47.16	2.014×10^{-6}
Likelihood ratio test ²			0.4033	0.5254
Goodness of fit test				
Hosmer-Lemeshow			11.131	0.1944
McFadden's R^2	0.372			
Cox & Snell's R^2	0.403			

¹ Compared against null model.

² Compared against maximal model.

Next we remove *Number of parent flocks*.

Table D.3: Logistic regression analysis of the maximal model with *Number of parent flocks* removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	10.63	6.715	1.583	0.113
<i>Number placed</i>	-8.973×10^{-5}	6.650×10^{-5}	-1.349	0.1773
<i>Breed</i> (1 = B, 0 = A)	-1.542	0.744	-2.072	0.0382
<i>Breed</i> (1 = A & B, 0 = A)	-1.408	1.602	-0.879	0.3794
<i>28-day mortality percentage</i>	-0.383	0.289	-1.325	0.185
<i>Pododermatitis percentage</i>	-0.017	0.014	-1.248	0.212
<i>Hock burn percentage</i>	-0.0477	0.0237	-2.013	0.044
<i>28-day average weight</i>	-0.0050	0.0039	-1.265	0.2058
<i>Minimum temperature</i>	0.506	0.130	3.893	0.0001
Test			χ^2	p
Overall model evaluation				
Likelihood ratio test ¹			41.44	1.728×10^{-6}
Likelihood ratio test ²			2.314	0.678
Goodness of fit test				
Hosmer-Lemeshow			27.57	0.0005
McFadden's R^2	0.356			
Cox & Snell's R^2	0.389			

¹ Compared against null model.

² Compared against maximal model.

Next we remove *Pododermatitis* percentage.

Table D.4: Logistic regression analysis of the maximal model with *Pododermatitis* percentage removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	8.086	6.309	1.282	0.200
<i>Number placed</i>	-6.77×10^{-5}	6.029×10^{-5}	-1.123	0.2615
<i>Breed</i> (1 = B, 0 = A)	-1.501	0.735	-2.043	0.0410
<i>Breed</i> (1 = A & B, 0 = A)	-0.647	1.456	-0.444	0.657
<i>28-day mortality percentage</i>	-0.366	0.283	-1.295	0.195
<i>Hock burn percentage</i>	-0.0519	0.0228	-2.283	0.0224
<i>28-day average weight</i>	-0.0046	0.0038	-1.186	0.2356
<i>Minimum temperature</i>	0.540	0.129	4.192	0.00003
Test			χ^2	p
Overall model evaluation				
Likelihood ratio test ¹			39.76	1.398×10^{-6}
Likelihood ratio test ²			3.986	0.551
Goodness of fit test				
Hosmer-Lemeshow			25.464	0.00130
McFadden's R^2	0.341			
Cox & Snell's R^2	0.377			

¹ Compared against null model.

² Compared against maximal model.

Next we remove *Number placed*.

Table D.5: Logistic regression analysis of the maximal model with *Number placed* removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	4.983	5.434	0.917	0.359
<i>Breed</i> (1 = B, 0 = A)	-1.825	0.697	-2.617	0.0089
<i>Breed</i> (1 = A & B, 0 = A)	-1.267	1.358	-0.933	0.351
<i>28-day mortality percentage</i>	-0.228	0.238	-0.956	0.339
<i>Hock burn percentage</i>	-0.0513	0.0211	-2.435	0.0149
<i>28-day average weight</i>	-0.0037	0.0037	-1.016	0.310
<i>Minimum temperature</i>	0.491	0.113	4.341	0.00001
Test			χ^2	p
Overall model evaluation				
Likelihood ratio test ¹			38.512	8.921×10^{-7}
Likelihood ratio test ²			5.237	0.514
Goodness of fit test				
Hosmer-Lemeshow			22.82	0.0036
McFadden's R^2	0.331			
Cox & Snell's R^2	0.368			

¹ Compared against null model.

² Compared against maximal model.

Next we remove *28-day mortality percentage*.

Table D.6: Logistic regression analysis of the maximal model with *28-day mortality percentage* removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	3.835	5.266	0.728	0.466
<i>Breed</i> (1 = B, 0 = A)	-1.639	0.652	-2.513	0.0120
<i>Breed</i> (1 = A & B, 0 = A)	-1.133	1.406	-0.806	0.420
<i>Hock burn percentage</i>	-0.045	0.019	-2.33	0.0197
<i>28-day average weight</i>	-0.0036	0.0037	-0.978	0.328
<i>Minimum temperature</i>	0.471	0.108	4.344	0.00001
Test			χ^2	p
Overall model evaluation				
Likelihood ratio test ¹			37.59	4.553×10^{-7}
Likelihood ratio test ²			6.155	0.522
Goodness of fit test				
Hosmer-Lemeshow			23.16	0.0032
McFadden's R^2	0.323			
Cox & Snell's R^2	0.361			

¹ Compared against null model.

² Compared against maximal model.

Next we remove *28-day average weight*.

Table D.7: Logistic regression analysis of the maximal model with *28-day average weight* removed.

Predictor	β	SE β	Wald-test z-score	p
(Intercept)	-1.29	0.669	-1.935	0.0529
<i>Breed</i> (1 = B, 0 = A)	-1.596	0.650	-2.455	0.0141
<i>Breed</i> (1 = A & B, 0 = A)	-1.030	1.341	-0.768	0.443
<i>Hock burn percentage</i>	-0.043	0.019	-2.238	0.0252
<i>Minimum temperature</i>	0.464	0.106	4.382	0.00001
Test	χ^2			p
Overall model evaluation				
Likelihood ratio test ¹	36.611			2.166×10^{-7}
Likelihood ratio test ²	7.137			0.5219
Goodness of fit test				
Hosmer-Lemeshow	16.699			0.033
McFadden's R^2	0.314			
Cox & Snell's R^2	0.353			

¹ Compared against null model.

² Compared against maximal model.

The only remaining non-significant variable remains within the *Breed* variable. Table D.7 shows that the factor ‘Breed A & B’ is not statistically significantly different in its predictive potential from Breed A birds. As such, we collapse the ‘Breed A’ and ‘Breed A & B’ factors together, to produce the final minimal adequate model as provided in Table 5.4 of the manuscript.

D2 - Bayesian network structure with BDE scoring

Here we present the best-fit network structure when the hill-climbing algorithm is used with BDE scoring instead of BIC scoring.

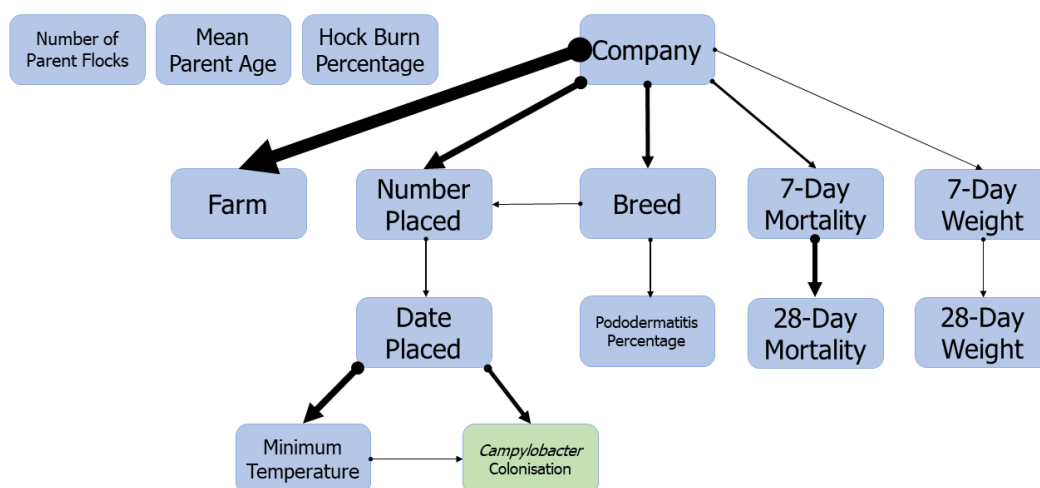


Figure D.1: Bayesian network structure showing the interrelationships between multiple welfare and rearing practice factors in a flock of broilers. *Campylobacter* colonization is directly impacted by the season the flock is grown in. Structure was learned using a hill-climbing algorithm, and sampled networks scored using the Bayesian Dirichlet equivalent score (BDE). Arrow-width indicates arc strength as scored by BDE.

We provide the exact arc strengths as scored by BDE below.

Table D.8: Arc strengths of the Bayesian network shown in Figure D.1. Arc strength is measured by Bayesian Dirichlet equivalent score (BDE), where a lower value indicates a stronger link.

Parent	Child	Arc Strength
Company	Farm	-52.501334
Company	Number placed	-28.237221
Date placed	Minimum temperature	-31.034381
7-day Mortality Percentage	28-day Mortality Percentage	-20.568572
Company	Breed	-17.193869
Date placed	<i>Campylobacter</i> colonization	-18.176669
Company	7-day Mortality Percentage	-11.413902
Number placed	Date placed	-5.470563
Minimum temperature	<i>Campylobacter</i> colonization	-4.876092
Breed	Number placed	-4.271554
7-day average weight	28-day average weight	-3.898640
Breed	Pododermatitis percentage	-2.886749
Company	7-day average weight	-2.621290

D3 - *Campylobacter* conditional probability tables

The network structure shown in Figure 5.2 displays that the date placed alone captures the probabilistic distribution of whether or not a flock is colonized by *Campylobacter*. However, in the absence of data on the date placed, other predictor variables can inform our expectations of whether or not a flock will be *Campylobacter* positive. The following tables provide these conditional probabilities for *Campylobacter* colonization under the assumption that no data is known other than the variable displayed. The best fit parameters via both Bayesian inference and MLE are given. The Bayesian estimates are built from a larger dataset of 114 entries, 33 of which contain some missing data. We do not provide tables upon mean parent age or number of parent flocks, as these variables were found to be unassociated.

Table D.9: Conditional probability table for *Campylobacter* colonization status, when data is only available on the Company variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference		
	Company	
<i>Campylobacter</i> colonization	1	2
Negative	0.508	0.379
Positive	0.492	0.621

Maximum Likelihood Estimator		
	Company	
<i>Campylobacter</i> colonization	1	2
Negative	0.576	0.430
Positive	0.424	0.570

Table D.10: Conditional probability table for *Campylobacter* colonization status, when data is only available on the Farm variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference				
	Farm			
<i>Campylobacter</i> colonization	C1F1	C2F1	C2F2	C2F3
Negative	0.505	0.386	0.387	0.388
Positive	0.495	0.614	0.613	0.612

Maximum Likelihood Estimator				
	Farm			
<i>Campylobacter</i> colonization	C1F1	C2F1	C2F2	C2F3
Negative	0.576	0.430	0.430	0.430
Positive	0.424	0.570	0.570	0.570

Table D.11: Conditional probability table for *Campylobacter* colonization status, when data is only available on the ‘number placed’ variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
	Number placed		
<i>Campylobacter</i> colonization	[11770,22000]	(22000,33503.3]	(33503.3,34650]
Negative	0.362	0.533	0.533
Positive	0.638	0.467	0.467

Maximum Likelihood Estimator			
	Number placed		
<i>Campylobacter</i> colonization	[11770,22000]	(22000,33503.3]	(33503.3,34650]
Negative	0.417	0.616	0.544
Positive	0.583	0.384	0.456

Table D.12: Conditional probability table for *Campylobacter* colonization status, when data is only available on the breed variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
Breed			
<i>Campylobacter</i> colonization	B	A	A & B
Negative	0.471	0.410	0.479
Positive	0.529	0.590	0.521

Maximum Likelihood Estimator			
Breed			
<i>Campylobacter</i> colonization	B	A	A & B
Negative	0.555	0.454	0.576
Positive	0.445	0.546	0.424

Table D.13: Conditional probability table for *Campylobacter* colonization status, when data is only available on the 7-day mortality percentage variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
7-Day Mortality Percentage			
<i>Campylobacter</i> colonization	[0.65,1.21]	(1.21,1.81]	(1.81,7.26]
Negative	0.496	0.444	0.406
Positive	0.504	0.556	0.594

Maximum Likelihood Estimator			
7-Day Mortality Percentage			
<i>Campylobacter</i> colonization	[0.65,1.21]	(1.21,1.81]	(1.81,7.26]
Negative	0.571	0.517	0.468
Positive	0.429	0.483	0.532

Table D.14: Conditional probability table for *Campylobacter* colonization status, when data is only available on the 28-day mortality percentage variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
28-Day Mortality			
<i>Campylobacter</i> colonization	[1.97,3.06667]	(3.06667,4.05667]	(4.05667,9.61]
Negative	0.480	0.445	0.419
Positive	0.520	0.555	0.581
Maximum Likelihood Estimator			
28-Day Mortality			
<i>Campylobacter</i> colonization	[1.97,3.06667]	(3.06667,4.05667]	(4.05667,9.61]
Negative	0.557	0.518	0.481
Positive	0.443	0.482	0.519

Table D.15: Conditional probability table for *Campylobacter* colonization status, when data is only available on the 7-day average weight variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
7-Day Average Weight			
<i>Campylobacter</i> colonization	[144,170.667]	(170.667,181]	(181,213]
Negative	0.481	0.453	0.418
Positive	0.519	0.547	0.582
Maximum Likelihood Estimator			
7-Day Average Weight			
<i>Campylobacter</i> colonization	[144,170.667]	(170.667,181]	(181,213]
Negative	0.555	0.522	0.479
Positive	0.445	0.478	0.521

Table D.16: Conditional probability table for *Campylobacter* colonization status, when data is only available on the 28-day weight variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
28-Day Average Weight			
<i>Campylobacter</i> colonization	[1138,1388.33]	(1388.33,1475.33]	(1475.33,1565]
Negative	0.459	0.451	0.433
Positive	0.541	0.549	0.567

Maximum Likelihood Estimator			
28-Day Average Weight			
<i>Campylobacter</i> colonization	[1138,1388.33]	(1388.33,1475.33]	(1475.33,1565]
Negative	0.538	0.522	0.496
Positive	0.462	0.478	0.504

Table D.17: Conditional probability table for *Campylobacter* colonization status, when data is only available on the hock burn percentage variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
Hock Burn Percentage			
<i>Campylobacter</i> colonization	[0,10]	(10,20]	(20,90]
Negative	0.438	0.447	0.452
Positive	0.562	0.553	0.548

Maximum Likelihood Estimator			
Hock Burn Percentage			
<i>Campylobacter</i> colonization	[0,10]	(10,20]	(20,90]
Negative	0.490	0.520	0.548
Positive	0.510	0.480	0.452

Table D.18: Conditional probability table for *Campylobacter* colonization status, when data is only available on the pododermatitis percentage variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
Pododermatitis Percentage			
<i>Campylobacter</i> colonization	[1,42.6667]	(42.6667,76]	(76,95]
Negative	0.448	0.457	0.436
Positive	0.552	0.543	0.564

Maximum Likelihood Estimator			
Pododermatitis Percentage			
<i>Campylobacter</i> colonization	[1,42.6667]	(42.6667,76]	(76,95]
Negative	0.524	0.537	0.490
Positive	0.476	0.463	0.510

Table D.19: Conditional probability table for *Campylobacter* colonization status, when data is only available on the minimum temperature variable. We present values calculated by Bayesian inference using uniform priors, and an equivalent sample size of 10. Below that we present values calculated via a maximum likelihood estimator.

Bayesian Inference			
Minimum Temperature			
<i>Campylobacter</i> colonization	[1.3,4]	(4,8.7]	(8.7,13.8]
Negative	0.452	0.442	0.416
Positive	0.548	0.558	0.584

Maximum Likelihood Estimator			
Minimum Temperature			
<i>Campylobacter</i> colonization	[1.3,4]	(4,8.7]	(8.7,13.8]
Negative	0.694	0.465	0.313
Positive	0.306	0.535	0.687

Bibliography

- Achen, M., Morishita, T. Y., and Ley, E. C. (1998). “Shedding and colonization of *Campylobacter jejuni* in broilers from day-of-hatch to slaughter age”. In: *Avian Diseases*, pp. 732–737.
- Acke, E. et al. (2009). “A comparison of different culture methods for the recovery of *Campylobacter* species from pets”. In: *Zoonoses and Public Health* 56.9-10, pp. 490–495.
- Ackerley, L. (2016). “The key role of hygiene and cleanliness in the domestic environment”. In: *Perspectives in Public Health* 136.4, pp. 210–212.
- ACMSF (2005). *Second Report of the Advisory Committee on the Microbiological Safety of Food (ACMSF) dealing with Campylobacter*. Advisory Committee on the Microbiological Safety of Food (ACMSF). URL: https://acmsf.food.gov.uk/sites/default/files/mnt/drupal_data/sources/files/multimedia/pdfs/acmsfcampylobacter.pdf.
- ACMSF (2018). *Third Report of the Advisory Committee on the Microbiological Safety of Food (ACMSF) dealing with Campylobacter*. Advisory Committee on the Microbiological Safety of Food (ACMSF). URL: https://acmsf.food.gov.uk/sites/default/files/annex_b_campylobacter_report.pdf.
- Adkin, A. et al. (2006). “Use of a systematic review to assist the development of *Campylobacter* control strategies in broilers”. In: *Journal of Applied Microbiology* 100.2, pp. 306–315.
- Alhenaky, A. et al. (2017). “The effect of heat stress on intestinal integrity and *Salmonella* invasion in broiler birds”. In: *Journal of Thermal Biology* 70, pp. 9–14.
- Alpigiani, I. et al. (2017). “Associations between animal welfare indicators and *Campylobacter* spp. in broiler chickens under commercial settings: A case study”. In: *Preventive Veterinary Medicine* 147, pp. 186–193.
- Aroori, S. V., Cogan, T. A., and Humphrey, T. J. (2013). “The effect of growth temperature on the pathogenicity of *Campylobacter*”. In: *Current Microbiology* 67.3, pp. 333–340.
- Awad, W. A., Hess, C., and Hess, M. (2018). “Re-thinking the chicken–*Campylobacter jejuni* interaction: a review”. In: *Avian Pathology* 47.4, pp. 352–363.
- Awad, W. A., Mann, E., et al. (2016). “Age-related differences in the luminal and mucosa-associated gut microbiome of broiler chickens and shifts associated with *Campylobacter jejuni* infection”. In: *Frontiers in Cellular and Infection Microbiology* 6, p. 154.
- Babacan, O. et al. (2020). “Factors affecting the species of *Campylobacter* colonizing chickens reared for meat”. In: *Journal of Applied Microbiology* 129.4, pp. 1071–1078.
- Barrow, P. et al. (2004). “Faecal shedding and intestinal colonization of *Salmonella enterica* in in-bred chickens: the effect of host-genetic background”. In: *Epidemiology & Infection* 132.1, pp. 117–126.

- Bhat, H. S. and Kumar, N. (2010). "On the derivation of the Bayesian Information Criterion". In: *School of Natural Sciences, University of California* 99.
- Black, R. E. et al. (1988). "Experimental *Campylobacter jejuni* infection in humans". In: *Journal of Infectious Diseases* 157.3, pp. 472–479.
- Blaser, M. J. (1997). "Epidemiologic and clinical features of *Campylobacter jejuni* infections". In: *Journal of Infectious Diseases* 176.Supplement_2, S103–S105.
- Bonsall, M. B., Jansen, V. A., and Hassell, M. P. (2004). "Life history trade-offs assemble ecological guilds". In: *Science* 306.5693, pp. 111–114.
- Bouchaala, L. et al. (2010). "Improving algorithms for structure learning in Bayesian Networks using a new implicit score". In: *Expert Systems with Applications* 37.7, pp. 5470–5475.
- Bouckaert, R. R. (1995). "Bayesian belief networks: from construction to inference". PhD thesis.
- Boyd, Y. et al. (2005). "Host genes affect intestinal colonisation of newly hatched chickens by *Campylobacter jejuni*". In: *Immunogenetics* 57.3, pp. 248–253.
- Bull, S. A., Allen, V., et al. (2006). "Sources of *Campylobacter* spp. colonizing housed broiler flocks during rearing". In: *Applied Environmental Microbiology* 72.1, pp. 645–652.
- Bull, S. A., Thomas, A., et al. (2008). "Flock health indicators and *Campylobacter* spp. in commercial housed broilers reared in Great Britain". In: *Applied and Environmental Microbiology* 74.17, pp. 5408–5413.
- Burkholder, K. et al. (2008). "Influence of stressors on normal intestinal microbiota, intestinal morphology, and susceptibility to *Salmonella enteritidis* colonization in broilers". In: *Poultry Science* 87.9, pp. 1734–1741.
- Buzala, M., Janicki, B., and Czarnecki, R. (2015). "Consequences of different growth rates in broiler breeder and layer hens on embryogenesis, metabolism and metabolic rate: a review". In: *Poultry Science* 94.4, pp. 728–733.
- Caffrey, N. et al. (2021). "A cross-sectional study of the prevalence factors associated with fluoroquinolone resistant *Campylobacter jejuni* in broiler flocks in Canada". In: *Preventive Veterinary Medicine* 186, p. 105164.
- Calderón-Gómez, L. I. et al. (2009). "Potential use of characterised hyper-colonising strain(s) of *Campylobacter jejuni* to reduce circulation of environmental strains in commercial poultry". In: *Veterinary Microbiology* 134.3-4, pp. 353–361.
- Carrasco, E. et al. (2012). "A review on growth/no growth *Salmonella* models". In: *Food Research International* 47.1, pp. 90–99.
- Castelo, R. and Siebes, A. (2000). "Priors on network structures. Biasing the search for Bayesian networks". In: *International Journal of Approximate Reasoning* 24.1, pp. 39–57.
- Cawthraw, S. et al. (1994). "Isotype, specificity, and kinetics of systemic and mucosal antibodies to *Campylobacter jejuni* antigens, including flagellin, during experimental oral infections of chickens". In: *Avian Diseases*, pp. 341–349.
- Chen, H.-C. and Stern, N. J. (2001). "Competitive exclusion of heterologous *Campylobacter* spp. in chicks". In: *Applied and Environmental Microbiology* 67.2, pp. 848–851.
- Clark, S. et al. (2002). "Pododermatitis in turkeys". In: *Avian Diseases* 46.4, pp. 1038–1044.

- Colles, F. M., Cain, R. J., et al. (2016). “Monitoring chicken flock behaviour provides early warning of infection by human pathogen *Campylobacter*”. In: *Proceedings of the Royal Society B: Biological Sciences* 283.1822, p. 20152323.
- Colles, F. M., Jones, K., et al. (2003). “Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment”. In: *Applied and Environmental Microbiology* 69.12, pp. 7409–7413.
- Colles, F. M., Jones, T. A., et al. (2008). “*Campylobacter* infection of broiler chickens in a free-range environment”. In: *Environmental Microbiology* 10.8, pp. 2042–2050.
- Colles, F. M. and Maiden, M. C. (2014). “Exploring the population dynamics of multiple *Campylobacter* strains in poultry”. In: *Campylobacter Ecology and Evolution*, p. 213.
- Colles, F. M., McCarthy, N. D., Bliss, C. M., et al. (2015). “The long-term dynamics of *Campylobacter* colonizing a free-range broiler breeder flock: an observational study”. In: *Environmental Microbiology* 17.4, pp. 938–946.
- Colles, F. M., McCarthy, N. D., Layton, R., et al. (2011). “The prevalence of *Campylobacter* amongst a free-range broiler breeder flock was primarily affected by flock age”. In: *PLOS One* 6.12, e22825.
- Colles, F. M., Preston, S. G., et al. (2019). “Parallel sequencing of *porA* reveals a complex pattern of *Campylobacter* genotypes that differs between broiler and broiler breeder chickens”. In: *Scientific Reports* 9.1, p. 6204.
- Comin, A. et al. (2019). “Revealing the structure of the associations between housing system, facilities, management and welfare of commercial laying hens using Additive Bayesian Networks”. In: *Preventive Veterinary Medicine* 164, pp. 23–32.
- Connell, S. et al. (2013). “Genome-wide association analysis of avian resistance to *Campylobacter jejuni* colonization identifies risk locus spanning the *CDH13* gene”. In: *G3: Genes, Genomes, Genetics* 3.5, pp. 881–890.
- Connerton, P. L. et al. (2018). “The effect of the timing of exposure to *Campylobacter jejuni* on the gut microbiome and inflammatory responses of broiler chickens”. In: *Microbiome* 6.1, p. 88.
- Council of European Union (2007). *Council Directive 2007/43/EC of 28 June 2007 laying down minimum rules for the protection of chickens kept for meat production [2007] OJ L182/19*.
<https://eur-lex.europa.eu/eli/dir/2007/43/oj>.
- Council of European Union (2008). *Commission Regulation (EC) No 889/2008 of 5 September 2008 laying down detailed rules for the implementation of Council Regulation (EC) No 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and control*.
<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02008R0889-20210101>.
- Coward, C. et al. (2008). “Competing isogenic *Campylobacter* strains exhibit variable population structures in vivo”. In: *Applied and Environmental Microbiology* 74.12, pp. 3857–3867.
- Cox, N. A., Richardson, L., et al. (2012). “Evidence for horizontal and vertical transmission in *Campylobacter* passage from hen to her progeny”. In: *Journal of Food Protection* 75.10, pp. 1896–1902.
- Cox, N. A., Stern, N. J., et al. (2002). “Prevalence and level of *Campylobacter* in commercial broiler breeders (parents) and broilers”. In: *Journal of Applied Poultry Research* 11.2, pp. 187–190.

- Daniel, N. et al. (2020). *The Burden of Foodborne Disease in the UK 2018*. Food Standards Agency. URL: https://www.food.gov.uk/sites/default/files/media/document/the-burden-of-foodborne-disease-in-the-uk_0.pdf.
- Dawkins, M. S. (2017). “Animal welfare and efficient farming: is conflict inevitable?” In: *Animal Production Science* 57.2, pp. 201–208.
- Dawkins, M. S., Lee, H.-J., et al. (2009). “Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait”. In: *Applied Animal Behaviour Science* 119.3-4, pp. 203–209.
- Dawkins, M. S., Roberts, S., et al. (2017). “Early warning of footpad dermatitis and hockburn in broiler chicken flocks using optical flow, bodyweight and water consumption”. In: *Veterinary Record* 180.20, pp. 499–499.
- Dawkins, M. S., Wang, L., et al. (2021). “Optical flow, behaviour and broiler chicken welfare in the UK and Switzerland”. In: *Applied Animal Behaviour Science* 234, p. 105180.
- De Cesare, A. et al. (2008). “Genotypic and phenotypic diversity within three *Campylobacter* populations isolated from broiler ceca and carcasses”. In: *Poultry Science* 87.10, pp. 2152–2159.
- De Jong, I. et al. (2016). “Simplifying the Welfare Quality® assessment protocol for broiler chicken welfare”. In: *Animal* 10.1, pp. 117–127.
- DEFRA (2020). *United Kingdom Poultry and Poultry Meat Statistics – September 2020*. Department for Environment Food & Rural Affairs (DEFRA). URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936520/poultry-statsnotice-22oct20.pdf.
- Dingle, K. et al. (2001). “Multilocus sequence typing system for *Campylobacter jejuni*”. In: *Journal of Clinical Microbiology* 39.1, pp. 14–23.
- Djennad, A. et al. (2019). “Seasonality and the effects of weather on *Campylobacter* infections”. In: *BMC Infectious Diseases* 19.1, pp. 1–10.
- Domingues, A. et al. (2012). “Source attribution of human campylobacteriosis using a meta-analysis of case-control studies of sporadic infections”. In: *Epidemiology & Infection* 140.6, pp. 970–981.
- Donahue, M. J., Holyoak, M., and Feng, C. (2003). “Patterns of dispersal and dynamics among habitat patches varying in quality”. In: *The American Naturalist* 162.3, pp. 302–317.
- Dorazio, R. M. (2016). “Bayesian data analysis in population ecology: motivations, methods, and benefits”. In: *Population Ecology* 58.1, pp. 31–44.
- Doyle, M. P. and Roman, D. (1981). “Growth and survival of *Campylobacter fetus* subsp. *jejuni* as a function of temperature and pH”. In: *Journal of Food Protection* 44.8, pp. 596–601.
- EFSA Panel on Biological Hazards (BIOHAZ) (2011). “Scientific Opinion on *Campylobacter* in broiler meat production: control options and performance objectives and/or targets at different stages of the food chain”. In: *EFSA Journal* 9.4, p. 2105.
- Evans, S. and Sayers, A. (2000). “A longitudinal study of *Campylobacter* infection of broiler flocks in Great Britain”. In: *Preventive Veterinary Medicine* 46.3, pp. 209–223.
- Fan, G. et al. (2015). “Stability and persistence in ODE models for populations with many stages”. In: *Mathematical Biosciences & Engineering* 12.4, p. 661.

- Fayyad, U. and Irani, K. (1993). “Multi-interval discretization of continuous-valued attributes for classification learning”. In:
- Feng, Y. et al. (2020). “Inferring competitive outcomes, ranks and intransitivity from empirical data: A comparison of different methods”. In: *Methods in Ecology and Evolution* 11.1, pp. 117–128.
- Fox, J. and Monette, G. (1992). “Generalized collinearity diagnostics”. In: *Journal of the American Statistical Association* 87.417, pp. 178–183.
- Fraser, R. W. et al. (2010). “Reducing *Campylobacter* and salmonella infection: two studies of the economic cost and attitude to adoption of on-farm biosecurity measures”. In: *Zoonoses and Public Health* 57.7-8, e109–e115.
- Frosth, S. et al. (2020). “Identification of transmission routes of *Campylobacter* and on-farm measures to reduce *Campylobacter* in chicken”. In: *Pathogens* 9.5, p. 363.
- FSA (2010). *The Joint Government and Industry Target to Reduce Campylobacter in UK Produced Chickens by 2015*. Food Standards Agency (FSA). URL: <https://acss.food.gov.uk/sites/default/files/multimedia/pdfs/campytarget.pdf>.
- Gallien, L. et al. (2017). “The effects of intransitive competition on coexistence”. In: *Ecology Letters* 20.7, pp. 791–800.
- Ge, B. et al. (2003). “Antimicrobial-resistant *Campylobacter* species from retail raw meats”. In: *Applied and Environmental Microbiology* 69.5, pp. 3005–3007.
- Gelman, A. (2006). “Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)”. In: *Bayesian Analysis* 1.3, pp. 515–534.
- Ghareeb, K., Awad, W. A., Mohnl, M., Porta, R., et al. (2012). “Evaluating the efficacy of an avian-specific probiotic to reduce the colonization of *Campylobacter jejuni* in broiler chickens”. In: *Poultry Science* 91.8, pp. 1825–1832.
- Ghareeb, K., Awad, W. A., Mohnl, M., Schatzmayr, G., et al. (2013). “Control strategies for *Campylobacter* infection in poultry production”. In: *World’s Poultry Science Journal* 69.1, pp. 57–76.
- Gharib, N. K., Rahimi, S., and Khaki, P. (2012). “Comparison of the effects of probiotic, organic acid and medicinal plant on *Campylobacter jejuni* challenged broiler chickens”. In: *Journal of Agricultural Science and Technology*.
- Gibbens, J. et al. (2001). “A trial of biosecurity as a means to control *Campylobacter* infection of broiler chickens”. In: *Preventive Veterinary Medicine* 48.2, pp. 85–99.
- Gielda, L. M. and DiRita, V. J. (2012). “Zinc competition among the intestinal microbiota”. In: *mBio* 3.4, e00171–12.
- Gölz, G. et al. (2014). “Relevance of *Campylobacter* to public health—the need for a One Health approach”. In: *International Journal of Medical Microbiology* 304.7, pp. 817–823.
- Gomes, A. et al. (2014). “Overcrowding stress decreases macrophage activity and increases *Salmonella enteritidis* invasion in broiler chickens”. In: *Avian Pathology* 43.1, pp. 82–90.
- Gopinath, S., Carden, S., and Monack, D. (2012). “Shedding light on *Salmonella* carriers”. In: *Trends in Microbiology* 20.7, pp. 320–327.
- Grant, A. J. et al. (2005). “Signature-tagged transposon mutagenesis studies demonstrate the dynamic nature of cecal colonization of 2-week-old chickens by *Campylobacter jejuni*”. In: *Applied Environmental Microbiology* 71.12, pp. 8031–8041.
- Grilli, J. et al. (2017). “Higher-order interactions stabilize dynamics in competitive network models”. In: *Nature* 548.7666, pp. 210–213.

- Guardia, S. et al. (2011). “Effects of stocking density on the growth performance and digestive microbiota of broiler chickens”. In: *Poultry Science* 90.9, pp. 1878–1889.
- Hald, B., Rattenborg, E., and Madsen, M. (2001). “Role of batch depletion of broiler houses on the occurrence of *Campylobacter* spp. in chicken flocks”. In: *Letters in Applied Microbiology* 32.4, pp. 253–256.
- Hald, B., Skovgård, H., et al. (2004). “Flies and *Campylobacter* infection of broiler flocks”. In: *Emerging Infectious Diseases* 10.8, p. 1490.
- Hald, B., Sommer, H. M., and Skovgård, H. (2007). “Use of fly screens to reduce *Campylobacter* spp. introduction in broiler houses”. In: *Emerging Infectious Diseases* 13.12, p. 1951.
- Han, Z., Willer, T., Li, L., et al. (2017). “Influence of the gut microbiota composition on *Campylobacter jejuni* colonization in chickens”. In: *Infection and Immunity* 85.11.
- Han, Z., Willer, T., Pielsticker, C., et al. (2016). “Differences in host breed and diet influence colonization by *Campylobacter jejuni* and induction of local immune responses in chicken”. In: *Gut Pathogens* 8.1, pp. 1–14.
- Hannu, T. et al. (2002). “*Campylobacter*-triggered reactive arthritis: a population-based study”. In: *Rheumatology* 41.3, pp. 312–318.
- Hanski, I. and Gyllenberg, M. (1997). “Uniting two general patterns in the distribution of species”. In: *Science* 275.5298, pp. 397–400.
- Hanski, I. and Ovaskainen, O. (2003). “Metapopulation theory for fragmented landscapes”. In: *Theoretical Population Biology* 64.1, pp. 119–127.
- Hartnett, E. et al. (2001). “A quantitative risk assessment for the occurrence of *Campylobacter* in chickens at the point of slaughter”. In: *Epidemiology & Infection* 127.2, pp. 195–206.
- Heckert, R. et al. (2002). “Effects of density and perch availability on the immune status of broilers”. In: *Poultry Science* 81.4, pp. 451–457.
- Herman, L. et al. (2003). “Routes for *Campylobacter* contamination of poultry meat: epidemiological study from hatchery to slaughterhouse”. In: *Epidemiology & Infection* 131.3, pp. 1169–1180.
- Hermans, D. et al. (2011). “*Campylobacter* control in poultry by current intervention measures ineffective: urgent need for intensified fundamental research”. In: *Veterinary Microbiology* 152.3-4, pp. 219–228.
- Heyndrickx, M. et al. (2002). “Routes for *Salmonella* contamination of poultry meat: epidemiological study from hatchery to slaughterhouse”. In: *Epidemiology & Infection* 129.2, pp. 253–265.
- Hibbing, M. E. et al. (2010). “Bacterial competition: surviving and thriving in the microbial jungle”. In: *Nature Reviews Microbiology* 8.1, pp. 15–25.
- Higham, D. J. (2001). “An algorithmic introduction to numerical simulation of stochastic differential equations”. In: *SIAM Review* 43.3, pp. 525–546.
- Hirakawa, R. et al. (2020). “Heat stress causes immune abnormalities via massive damage to effect proliferation and differentiation of lymphocytes in broiler chickens”. In: *Frontiers in Veterinary Science* 7, p. 46.
- Hofreuter, D. et al. (2006). “Unique features of a highly pathogenic *Campylobacter jejuni* strain”. In: *Infection and Immunity* 74.8, pp. 4694–4707.
- Holland, D. and Mahmoudzadeh, N. (2020). *Foodborne Disease Estimates for the United Kingdom in 2018*. Food Standards Agency. URL: https://www.food.gov.uk/sites/default/files/media/document/foodborne-disease-estimates-for-the-united-kingdom-in-2018_0.pdf.

- Höök, H. et al. (2005). "Genotype dynamics of *Campylobacter jejuni* in a broiler flock". In: *Veterinary Microbiology* 106.1-2, pp. 109–117.
- Hu, L. and Kopecko, D. J. (1999). "Campylobacter jejuni 81-176 associates with microtubules and dynein during invasion of human intestinal cells". In: *Infection and Immunity* 67.8, pp. 4171–4182.
- Humphrey, S. et al. (2014). "Campylobacter jejuni is not merely a commensal in commercial broiler chickens and affects bird welfare". In: *mBio* 5.4.
- Humphrey, T., Henley, A., and Lanning, D. (1993). "The colonization of broiler chickens with *Campylobacter jejuni*: some epidemiological investigations". In: *Epidemiology & Infection* 110.3, pp. 601–607.
- Humphrey, T. (2006). "Are happy chickens safer chickens? Poultry welfare and disease susceptibility". In: *British Poultry Science* 47.4, pp. 379–391.
- Huppert, A. and Katriel, G. (2013). "Mathematical modelling and prediction in infectious disease epidemiology". In: *Clinical Microbiology and Infection* 19.11, pp. 999–1005.
- Iannetti, L. et al. (2020). "Animal welfare and microbiological safety of poultry meat: Impact of different at-farm animal welfare levels on at-slaughterhouse *Campylobacter* and *Salmonella* contamination". In: *Food Control* 109, p. 106921.
- Jore, S. et al. (2010). "Trends in *Campylobacter* incidence in broilers and humans in six European countries, 1997–2007". In: *Preventive Veterinary Medicine* 93.1, pp. 33–41.
- Jorgensen, F., Ellis-Iversen, J., et al. (2011). "Influence of season and geography on *Campylobacter jejuni* and *C. coli* subtypes in housed broiler flocks reared in Great Britain". In: *Applied Environmental Microbiology* 77.11, pp. 3741–3748.
- Jorgensen, F., Madden, R. H., et al. (2015). *FSA Project FS241044 - Survey report - A Microbiological survey of Campylobacter contamination in fresh whole UK produced chilled chickens at retail sale (2014-15)*. Public Health England. URL: <https://www.food.gov.uk/sites/default/files/media/document/Final%5C%20Report%5C%20for%5C%20FS241044%5C%20Campylobacter%5C%20Retail%5C%20survey.pdf>.
- Kapperud, G. et al. (1993). "Epidemiological investigation of risk factors for *Campylobacter* colonization in Norwegian broiler flocks". In: *Epidemiology & Infection* 111.2, pp. 245–256.
- Kendall, M. G. and Smith, B. B. (1940). "On the method of paired comparisons". In: *Biometrika* 31.3/4, pp. 324–345.
- Khachatourians, G. G. (1998). "Agricultural use of antibiotics and the evolution and transfer of antibiotic-resistant bacteria". In: *Canadian Medical Association Journal* 159.9, pp. 1129–1136.
- Kohavi, R. and Sahami, M. (1996). "Error-based and entropy-based discretization of continuous features." In: *KDD*, pp. 114–119.
- Kovats, S. et al. (2005). "Climate variability and *Campylobacter* infection: an international study". In: *International Journal of Biometeorology* 49.4, pp. 207–214.
- Kretzschmar, M. (2020). "Disease modeling for public health: added value, challenges, and institutional constraints". In: *Journal of Public Health Policy* 41.1, pp. 39–51.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- Kudirkienė, E. et al. (2010). "Demonstration of persistent strains of *Campylobacter jejuni* within broiler farms over a 1-year period in Lithuania". In: *Journal of Applied Microbiology* 108.3, pp. 868–877.
- Laird, R. A. and Schamp, B. S. (2006). "Competitive intransitivity promotes species coexistence". In: *The American Naturalist* 168.2, pp. 182–193.

- Lawes, J. et al. (2012). "Investigation of prevalence and risk factors for *Campylobacter* in broiler flocks at slaughter: results from a UK survey". In: *Epidemiology & Infection* 140.10, pp. 1725–1737.
- Leeson, S. and Summers, J. D. (2010). *Broiler breeder production*. Nottingham University Press.
- Leibold, M. A. et al. (2004). "The metacommunity concept: a framework for multi-scale community ecology". In: *Ecology Letters* 7.7, pp. 601–613.
- Lenhart, S. and Workman, J. T. (2007). *Optimal control applied to biological models*. CRC press.
- Li, X. et al. (2008). "The paternal effect of *Campylobacter jejuni* colonization in ceca in broilers". In: *Poultry Science* 87.9, pp. 1742–1747.
- Locht, H. and Krogfelt, K. (2002). "Comparison of rheumatological and gastrointestinal symptoms after infection with *Campylobacter jejuni/coli* and enterotoxigenic *Escherichia coli*". In: *Annals of the Rheumatic Diseases* 61.5, pp. 448–452.
- Louis, V. R. et al. (2005). "Temperature-driven *Campylobacter* seasonality in England and Wales". In: *Applied and Environmental Microbiology* 71.1, pp. 85–92.
- Lu, J. et al. (2003). "Diversity and succession of the intestinal bacterial community of the maturing broiler chicken". In: *Applied and Environmental Microbiology* 69.11, pp. 6816–6824.
- Lydekaitienė, V. L. and Kudirkienė, E. (2020). "Prevalence and genetic diversity of *C. jejuni* isolated from broilers and their environment using *flaA*-RFLP typing and MLST analysis". In: *Annals of Animal Science* 1.ahead-of-print.
- Maiden, M. C. et al. (2013). "MLST revisited: the gene-by-gene approach to bacterial genomics". In: *Nature Reviews Microbiology* 11.10, pp. 728–736.
- Mastroeni, P. et al. (2009). "A dynamic view of the spread and intracellular distribution of *Salmonella enterica*". In: *Nature Reviews Microbiology* 7.1, pp. 73–80.
- May, R. M. and Nowak, M. A. (1994). "Superinfection, metapopulation dynamics, and the evolution of diversity". In: *Journal of Theoretical Biology* 170.1, pp. 95–114.
- McCarthy, N. D. and Giesecke, J. (2001). "Incidence of Guillain-Barré syndrome following infection with *Campylobacter jejuni*". In: *American Journal of Epidemiology* 153.6, pp. 610–614.
- Meerburg, B. G. and Kijlstra, A. (2007). "Role of rodents in transmission of *Salmonella* and *Campylobacter*". In: *Journal of the Science of Food and Agriculture* 87.15, pp. 2774–2781.
- Menanteau, P. et al. (2018). "Role of systemic infection, cross contaminations and super-shedders in *Salmonella* carrier state in chicken". In: *Environmental Microbiology* 20.9, pp. 3246–3260.
- Meunier, M. et al. (2017). "Promising new vaccine candidates against *Campylobacter* in broilers". In: *PLoS One* 12.11, e0188472.
- Mohamed, M. A. and Hanson, R. (1980). "Effect of social stress on Newcastle disease virus (LaSota) infection". In: *Avian Diseases*, pp. 908–915.
- Morishita, T. Y. et al. (1997). "Evaluation of an avian-specific probiotic to reduce the colonization and shedding of *Campylobacter jejuni* in broilers". In: *Avian Diseases*, pp. 850–855.
- Mountzouris, K. et al. (2007). "Evaluation of the efficacy of a probiotic containing *Lactobacillus*, *Bifidobacterium*, *Enterococcus*, and *Pediococcus* strains in promoting broiler performance and modulating cecal microflora composition and metabolic activities". In: *Poultry Science* 86.2, pp. 309–317.

- Murphy, C., Carroll, C., and Jordan, K. (2006). "Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni*". In: *Journal of Applied Microbiology* 100.4, pp. 623–632.
- Murray, S. M. and Howard, M. (2019). "Center finding in *E. coli* and the role of mathematical modeling: past, present and future". In: *Journal of Molecular Biology* 431.5, pp. 928–938.
- Nagaraj, M. et al. (2007). "Effect of high-protein and all-vegetable diets on the incidence and severity of pododermatitis in broiler chickens". In: *Journal of Applied Poultry Research* 16.3, pp. 304–312.
- Nagarajan, R., Scutari, M., and Lèbre, S. (2013). *Bayesian Networks in R*. New York: Springer.
- Nauta, M. et al. (2005). "Risk assessment of *Campylobacter* in the Netherlands via broiler meat and other routes". In: *RIVM Rapport 250911006*.
- Newell, D. and Wagenaar, J. (2000). *Poultry infections and their control at the farm level*. *Campylobacter*, 2nd. Washington D.C.: ASM Press.
- Newell, D., Elvers, K., et al. (2011). "Biosecurity-based interventions and strategies to reduce *Campylobacter* spp. on poultry farms". In: *Applied and Environmental Microbiology* 77.24, pp. 8605–8614.
- Newell, D. and Fearnley, C. (2003). "Sources of *Campylobacter* colonization in broiler chickens". In: *Applied and Environmental Microbiology* 69.8, pp. 4343–4351.
- Nothhaft, H. et al. (2016). "Engineering the *Campylobacter jejuni* N-glycan to create an effective chicken vaccine". In: *Scientific Reports* 6.1, pp. 1–12.
- Nylen, G. et al. (2002). "The seasonal distribution of *Campylobacter* infection in nine European countries and New Zealand". In: *Epidemiology & Infection* 128.3, pp. 383–390.
- Ongeng, D. et al. (2014). "Modeling the fate of *Escherichia coli* O157: H7 and *Salmonella enterica* in the agricultural environment: current perspective". In: *Journal of Food Science* 79.4, R421–R427.
- Pan, D. and Yu, Z. (2014). "Intestinal microbiome of poultry and its interaction with host and diet". In: *Gut Microbes* 5.1, pp. 108–119.
- Park, S. F. (2002). "The physiology of *Campylobacter* species and its relevance to their role as foodborne pathogens". In: *International Journal of Food Microbiology* 74.3, pp. 177–188.
- Patrick, M. E. et al. (2004). "Effects of climate on incidence of *Campylobacter* spp. in humans and prevalence in broiler flocks in Denmark". In: *Applied Environmental Microbiology* 70.12, pp. 7474–7480.
- Peh, E. et al. (2020). "Antimicrobial activity of organic acids against *Campylobacter* spp. and development of combinations—A synergistic effect?" In: *PLoS One* 15.9, e0239312.
- Petersen, L., Nielsen, E., and On, S. L. (2001). "Serotype and genotype diversity and hatchery transmission of *Campylobacter jejuni* in commercial poultry flocks". In: *Veterinary Microbiology* 82.2, pp. 141–154.
- Pielsticker, C., Glünder, G., Rautenschlein, S., et al. (2016). "Colonization pattern of *C. jejuni* isolates of human and avian origin and differences in the induction of immune responses in chicken". In: *Veterinary Immunology and Immunopathology* 169, pp. 1–9.
- Pigrau, C. et al. (1997). "Bacteremia due to *Campylobacter* species: clinical findings and antimicrobial susceptibility patterns". In: *Clinical Infectious Diseases* 25.6, pp. 1414–1420.

- Pires, S. M. et al. (2010). “Using outbreak data for source attribution of human salmonellosis and campylobacteriosis in Europe”. In: *Foodborne Pathogens and Disease* 7.11, pp. 1351–1361.
- Plummer, M. (2007). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*. Available at: <http://mcmc-jags.sourceforge.net/>.
- Plummer, M., Stukalov, A., and Denwood, M. (2016). *rjags: Bayesian graphical models using MCMC*. Available at: <http://CRAN.R-project.org/package=rjags>. R package version.
- Poniatowski, D. et al. (2018). “Patch occupancy of grassland specialists: Habitat quality matters more than habitat connectivity”. In: *Biological Conservation* 225, pp. 237–244.
- Powell, L. et al. (2012). “The prevalence of *Campylobacter* spp. in broiler flocks and on broiler carcasses, and the risks associated with highly contaminated carcasses”. In: *Epidemiology & Infection* 140.12, pp. 2233–2246.
- Psifidi, A. et al. (2021). “Quantitative trait loci and transcriptome signatures associated with avian heritable resistance to *Campylobacter*”. In: *Scientific Reports* 11.
- Public Health England (2018a). *Campylobacter data 2007 to 2016*. PHE publications gateway number 2018070. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/796964/campylobacter_data_2007_to_2016_may_2018.pdf.
- Public Health England (2018b). *Zoonoses Report, UK 2017*. PHE publications gateway number GW-65. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/918089/UK_Zoonoses_report_2017.pdf.
- Raines, D. J. et al. (2016). “Bacteria in an intense competition for iron: Key component of the *Campylobacter jejuni* iron uptake system scavenges enterobactin hydrolysis product”. In: *Proceedings of the National Academy of Sciences* 113.21, pp. 5850–5855.
- Rapp, D. et al. (2012). “Differences in the fecal concentrations and genetic diversities of *Campylobacter jejuni* populations among individual cows in two dairy herds”. In: *Applied and Environmental Microbiology* 78.21, pp. 7564–7571.
- Rawson, T., Colles, F. M., Smith, A. L., et al. (2021). “Can good broiler flock welfare prevent colonization by *Campylobacter*?” In: *bioRxiv*. URL: <https://doi.org/10.1101/2021.03.30.437710>.
- Rawson, T., Colles, F. M., Terry, J. C. D., et al. (2021). “Mechanisms of biodiversity between *Campylobacter* sequence types in a flock of broiler-breeder chickens.” In: *bioRxiv*. URL: <https://doi.org/10.1101/2021.04.14.439797>.
- Rawson, T., Dawkins, M. S., and Bonsall, M. B. (2019). “A Mathematical Model of *Campylobacter* Dynamics Within a Broiler Flock”. In: *Frontiers in Microbiology* 10, p. 1940.
- Rawson, T., Paton, R. S., et al. (2020). “A Mathematical Modeling Approach to Uncover Factors Influencing the Spread of *Campylobacter* in a Flock of Broiler-Breeder Chickens”. In: *Frontiers in Microbiology* 11, p. 2481.
- Ridley, A. M. et al. (2008). “Real-time PCR approach for detection of environmental sources of *Campylobacter* strains colonizing broiler flocks”. In: *Applied and Environmental Microbiology* 74.8, pp. 2492–2504.
- Rivoal, K. et al. (1999). “Molecular characterization of the diversity of *Campylobacter* spp. isolates collected from a poultry slaughterhouse: analysis of cross-contamination”. In: *Letters in Applied Microbiology* 29.6, pp. 370–374.

- Rushton, S. et al. (2009). "Campylobacter in housed broiler chickens: a longitudinal study of risk factors". In: *Epidemiology & Infection* 137.8, pp. 1099–1110.
- Sahin, O., Kassem, I. I., et al. (2015). "Campylobacter in poultry: ecology and potential interventions". In: *Avian Diseases* 59.2, pp. 185–200.
- Sahin, O., Zhang, Q., et al. (2001). "Prevalence, antigenic specificity, and bactericidal activity of poultry anti-Campylobacter maternal antibodies". In: *Applied and Environmental Microbiology* 67.9, pp. 3951–3957.
- Saint-Cyr, M. J. et al. (2016). "Recent advances in screening of anti-Campylobacter activity in probiotics for use in poultry". In: *Frontiers in Microbiology* 7, p. 553.
- Saltelli, A. et al. (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons.
- Santini, C. et al. (2010). "Characterization of probiotic strains: an application as feed additives in poultry against Campylobacter jejuni". In: *International Journal of Food Microbiology* 141, S98–S108.
- Scheffer, M. and van Nes, E. H. (2006). "Self-organized similarity, the evolutionary emergence of groups of similar species". In: *Proceedings of the National Academy of Sciences* 103.16, pp. 6230–6235.
- Schoeni, J. L. and Doyle, M. P. (1992). "Reduction of Campylobacter jejuni colonization of chicks by cecum-colonizing bacteria producing anti-C. jejuni metabolites." In: *Applied and Environmental Microbiology* 58.2, pp. 664–670.
- Scutari, M. (2009). "Learning Bayesian networks with the bnlearn R package". In: *arXiv*.
- Scutari, M. and Denis, J.-B. (2014). *Bayesian networks: with examples in R*. CRC press. Chap. More Complex Cases: Hybrid Bayesian Networks, pp. 65–84.
- Shanawany, M. (1984). "Inter-relationship between egg weight, parental age and embryonic development". In: *British Poultry Science* 25.4, pp. 449–455.
- Shanker, S., Lee, A., and Sorrell, T. (1990). "Horizontal transmission of Campylobacter jejuni amongst broiler chicks: experimental studies". In: *Epidemiology & Infection* 104.1, pp. 101–110.
- Sheppard, A. (2004). *The structure and economics of broiler production in England*. Centre for Rural Policy Research, University of Exeter.
- El-Shibiny, A., Connerton, P. L., and Connerton, I. F. (2007). "Campylobacter succession in broiler chickens". In: *Veterinary Microbiology* 125.3-4, pp. 323–332.
- El-Shibiny, A., Scott, A., et al. (2009). "Application of a group II Campylobacter bacteriophage to reduce strains of Campylobacter jejuni and Campylobacter coli colonizing broiler chickens". In: *Journal of Food Protection* 72.4, pp. 733–740.
- Shreeve, J. et al. (2000). "Sequential spread of Campylobacter infection in a multipen broiler house". In: *Avian Diseases*, pp. 983–988.
- Sibanda, N. et al. (2018). "A review of the effect of management practices on Campylobacter prevalence in poultry farms". In: *Frontiers in Microbiology* 9, p. 2002.
- Skirrow, M. (1977). "Campylobacter enteritis: a "new" disease." In: *British Medical Journal* 2.6078, pp. 9–11.
- Soliveres, S. and Allan, E. (2018). "Everything you always wanted to know about intransitive competition but were afraid to ask". In: *Journal of Ecology* 106.3, pp. 807–814.
- Soliveres, S., Lehmann, A., et al. (2018). "Intransitive competition is common across five major taxonomic groups and is driven by productivity, competitive rank and functional traits". In: *Journal of Ecology* 106.3, pp. 852–864.

- Soliveres, S., Maestre, F. T., et al. (2015). "Intransitive competition is widespread in plant communities and maintains their species richness". In: *Ecology Letters* 18.8, pp. 790–798.
- Stahl, M., Butcher, J., and Stintzi, A. (2012). "Nutrient acquisition and metabolism by *Campylobacter jejuni*". In: *Frontiers in Cellular and Infection Microbiology* 2, p. 5.
- Stelling, J. (2004). "Mathematical models in microbial systems biology". In: *Current Opinion in Microbiology* 7.5, pp. 513–518.
- Stern, N. J., Clavero, M., et al. (1995). "Campylobacter spp. in broilers on the farm and after transport". In: *Poultry Science* 74.6, pp. 937–941.
- Stern, N. J., Cox, N. A., Bailey, J., et al. (2001). "Comparison of mucosal competitive exclusion and competitive exclusion treatment to reduce Salmonella and Campylobacter spp. colonization in broiler chickens". In: *Poultry Science* 80.2, pp. 156–160.
- Stern, N. J., Cox, N. A., Musgrove, M. T., et al. (2001). "Incidence and levels of Campylobacter in broilers after exposure to an inoculated seeder bird". In: *Journal of Applied Poultry Research* 10.4, pp. 315–318.
- Stern, N. J., Meinersmann, R. J., et al. (1990). "Influence of host lineage on cecal colonization by Campylobacter jejuni in chickens". In: *Avian Diseases*, pp. 602–606.
- Strachan, N. J. and Forbes, K. J. (2010). "The growing UK epidemic of human campylobacteriosis". In: *The Lancet* 376.9742, pp. 665–667.
- Strachan, N. J., Rotariu, O., et al. (2013). "Operationalising factors that explain the emergence of infectious diseases: a case study of the human campylobacteriosis epidemic". In: *PLOS One* 8.11, e79331.
- Sutherland, C., Elston, D., and Lambin, X. (2014). "A demographic, spatially explicit patch occupancy model of metapopulation dynamics and persistence". In: *Ecology* 95.11, pp. 3149–3160.
- Tam, C. C. and O'Brien, S. J. (2016). "Economic cost of campylobacter, norovirus and rotavirus disease in the United Kingdom". In: *PLOS One* 11.2, e0138526.
- Thibodeau, A. et al. (2015). "Chicken caecal microbiome modifications induced by Campylobacter jejuni colonization and by a non-antibiotic feed additive". In: *PLOS One* 10.7, e0131978.
- Travers, K. and Michael, B. (2002). "Morbidity of infections caused by antimicrobial-resistant bacteria". In: *Clinical Infectious Diseases* 34.Supplement_3, S131–S134.
- Tresse, O., Alvarez-Ordóñez, A., and Connerton, I. F. (2017). "About the Foodborne Pathogen Campylobacter". In: *Frontiers in Microbiology* 8, p. 1908.
- Ulrich, W. et al. (2014). "Matrix models for quantifying competitive intransitivity from species abundance data". In: *Oikos* 123.9, pp. 1057–1070.
- Van Gerwe, T. et al. (2005). "Quantifying transmission of Campylobacter spp. among broilers". In: *Applied and Environmental Microbiology* 71.10, pp. 5765–5770.
- Wales, A. D. et al. (2019). "Field interventions against colonization of broilers by Campylobacter". In: *Comprehensive Reviews in Food Science and Food Safety* 18.1, pp. 167–188.
- Wallace, J. et al. (1997). "Seasonality of thermophilic Campylobacter populations in chickens". In: *Journal of Applied Microbiology* 82.2, pp. 219–224.
- Welkos, S. (1984). "Experimental gastroenteritis in newly-hatched chicks infected with Campylobacter jejuni". In: *Journal of Medical Microbiology* 18.2, pp. 233–248.

- Westrell, T. et al. (2009). "Zoonotic infections in Europe in 2007: a summary of the EFSA-ECDC annual report". In: *Eurosurveillance* 14.3, p. 19100.
- Wieczorek, K., Wołkiewicz, T., and Osek, J. (2020). "MLST-based genetic relatedness of *Campylobacter jejuni* isolated from chickens and humans in Poland". In: *PLOS One* 15.1, e0226238.
- Wigley, P. (2015). "Blurred lines: pathogens, commensals, and the healthy gut". In: *Frontiers in Veterinary Science* 2, p. 40.
- Wiles, T. J. et al. (2016). "Host gut motility promotes competitive exclusion within a model intestinal microbiota". In: *PLOS Biology* 14.7, e1002517.
- Willis, W. L. and Reid, L. (2008). "Investigating the effects of dietary probiotic feeding regimens on broiler chicken production and *Campylobacter jejuni* presence". In: *Poultry Science* 87.4, pp. 606–611.
- Wilson, D. J. et al. (2008). "Tracing the source of campylobacteriosis". In: *PLOS Genetics* 4.9.
- Young, G., Ermentrout, B., and Rubin, J. E. (2015). "A boundary value approach to optimization with an application to salmonella competition". In: *Bulletin of Mathematical Biology* 77.7, pp. 1327–1348.
- Yu, D. W. and Wilson, H. B. (2001). "The competition-colonization trade-off is dead; long live the competition-colonization trade-off". In: *The American Naturalist* 158.1, pp. 49–63.
- Zautner, A. et al. (2014). "Seroprevalence of campylobacteriosis and relevant post-infectious sequelae". In: *European Journal of Clinical Microbiology & Infectious Diseases* 33.6, pp. 1019–1027.
- Zuidhof, M. J. et al. (2014). "Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 2005". In: *Poultry Science* 93.12, pp. 2970–2982.