Elucidating the Aetiology of Human Campylobacter coli Infections

Citation for published version:

Digital Object Identifier (DOI):
10.1371/journal.pone.0064504

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
PLoS One

Publisher Rights Statement:
Copyright: © 2013 Roux et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Elucidating the Aetiology of Human Campylobacter coli Infections

Francois Roux¹, Emma Sproston¹, Ovidiu Rotariu¹, Marion MacRae², Samuel K. Sheppard³, Paul Bessell⁴, Alison Smith-Palmer⁵, John Cowden⁵, Martin C. J. Maiden⁶, Ken J. Forbes², Norval J. C. Strachan¹

¹School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom, ²School of Medicine and Dentistry, University of Aberdeen, Aberdeen, United Kingdom, ³Medical Microbiology and Infectious Diseases, College of Medicine, Swansea University, Swansea, United Kingdom, ⁴The Roslin Institute, University of Edinburgh, Edinburgh, United Kingdom, ⁵Health Protection Scotland, National Services Scotland, Glasgow, United Kingdom, ⁶Department of Zoology, University of Oxford, Oxford, United Kingdom

Abstract

There has been little research on the determinants of Campylobacter coli infection, despite its contributing up to 10% of human Campylobacter infections. A case-control and two case-case study methods explored the aetiology of C. coli over a one year period across Scotland. The case-control multivariate model found an increased risk of C. coli infection in people older than 19 years (O.R. = 3.352), and during the summer months (O.R. = 2.596), while residing in an urban area decreased the risk (O.R. = 0.546). The first case-case study compared C. coli and C. jejuni cases and also showed a higher risk of C. coli during the summer (O.R. = 1.313) and in people older than 19 years (O.R. = 0.791). Living in an urban area was associated with a reduced risk of infection (O.R. = 0.769). Multi-locus sequence typing (MLST) indicated that sheep and chicken C. coli sequence types (STs) were most frequently found in humans whilst those from cattle and pigs were rarer. MLST diversity is high in isolates from pigs and chicken, intermediate in human isolates, and low in ruminant isolates. The second case-case study used MLST data to ascribe putative sources of infection to the cases. The putative source for 40% of cases was chicken, with 60% acquired from other sources (ruminants 54% and pigs 6%). The case-case analysis also showed that female gender was a risk factor (O.R. = 1.940), which may be explained by females being more likely to prepare poultry in the home. These findings indicate differences between the aetiology of C. coli and C. jejuni infections: this should be taken into account by public health professionals when developing strategies to reduce the burden of human campylobacteriosis.

Introduction

Human campylobacteriosis is the most commonly reported bacterial gastrointestinal infectious disease in the world [1,2] with an estimated 572,000 community cases in the UK during 2009 [3] and 845,000 cases in the USA annually [4]. Campylobacter jejuni and Campylobacter coli are the commonest species to cause human infections, with approximately 9% of human infections being caused by C. coli in the USA [5] and approximately 7% in England and Wales [6]. Consequently most research has concentrated on the epidemiology of C. jejuni, and there is a more limited understanding of the aetiology of human C. coli infections [7].

The symptoms of human campylobacteriosis include diarrhoea (which can be bloody), abdominal pain and fever [8]. About 10% of reported cases are hospitalised [9] and, although rare, severe sequelae include Guillain-Barré syndrome, arthritis, or gastrointestinal perforation and occasionally death [8,10]. In England and Wales the symptoms caused by C. jejuni and C. coli appear to be clinically indistinguishable, [6] however in the Netherlands diarrhoea is reported in fewer cases of C. coli than C. jejuni [11].

C. jejuni and C. coli are zoonoses and both species are frequently carried asymptomatically in a wide range of domesticated livestock (cattle, sheep, pigs, chickens, and turkeys) and wildlife (birds, voles, insects etc.) [12]. They can also be found in symptomatic cats and dogs [13]. Pigs usually have a higher prevalence of C. coli than C. jejuni [14,15] whilst most other animals tend to carry a higher proportion of C. jejuni (e.g., >65% for poultry, sheep, cattle and wild birds [15]). Most human Campylobacter infections are sporadic and outbreaks are rare [16]. The vehicles of infection in recognised household and community Campylobacter spp. outbreaks include contaminated water, unpasteurized milk, and chicken liver pâté [17].

Case-control studies have been conducted on sporadic campylobacter cases (C. jejuni and C. coli combined or C. jejuni alone). The main source of infection identified in these studies is fresh chicken, including both the handling of raw and consumption of undercooked chicken [18,19]. Environmental


Editor: Thomas Alter, Free University of Berlin, Germany

Received February 20, 2013; Accepted April 15, 2013; Published May 29, 2013

Copyright: © 2013 Roux et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was partly funded by Food Standards Agency Scotland which involved the following contracts: The Molecular Epidemiology of Scottish Campylobacter Isolates from Human Cases of Infection using Multilocus Sequence Typing (MLST)- Contract S14006 (URL: http://www.food.gov.uk/science/research/devolvedadmins/scotlandresearch/scotlandresearch/ScotlandProjectList/s14006/), Factors Associated with Geographic and Temporal Variation in Campylobacter Isolates from Human Cases of Infection using Multilocus Sequence Typing (MLST)- Contract S14004 (URL: http://www.food.gov.uk/science/ research/devolvedadmins/scotlandresearch/scotlandresearch/ScotlandProjectList/s14004/). No additional external funding was received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: n.strachan@abdn.ac.uk
sources (e.g. contaminated water), contact with domesticated and wild animals and recent travel (particularly foreign) are also important in some settings [2,20–22]. However, at most only half of all cases are explained in the majority of studies, and the only published case-control study of _C. coli_ involved small numbers of cases (121) [11].

A case-case methodology [6] identified differences in risk factors between the two species, where cases of _C. coli_ infection were more likely to drink bottled water, eat pâté, and tended on average to be older than _C. jejuni_ cases. Cases of _C. jejuni_ infection were more likely to have had contact with farm animals, and develop illness during the summer months. The case-case methodology minimizes a number of possible biases inherent in case-control studies that include representativeness of reporting in the health care system. However, it is worth noting that the _C. jejuni_ case controls are not representative of the population as a whole and hence it is not possible to extrapolate the results to the general population [23].

The _Campylobacter_ genome is highly variable and frequent recombination complicates the typing of isolates. The advent of sequence-based typing methods, in particular multi locus sequence typing (MLST) [24], has helped both the characterisation of isolates and provided evidence of host association (i.e. strains that are more commonly found from a particular animal reservoir). MLST has the advantage of being unambiguous, reproducible, and portable allowing rapid exchange of data between laboratories and the creation of reference databases (e.g. PubMLST www.pubmlst.org/campylobacter). Source attribution has employed MLST data to identify the putative origin of combined _C. jejuni_ and _C. coli_ clinical isolates with poultry being identified as the main source for _C. jejuni_. Poultry and sheep were the main source sources for _C. coli_ [25]. MLST-based source attribution has also been combined with risk factor analysis for _C. jejuni_ in a case-case study that compared ruminant and poultry types [26]. It was found that women were at greater risk of infection from poultry types and it was hypothesised that this was because they were involved in preparation of chicken in the home. In the Netherlands [18] a case-control study combined MLST source attribution data with risk factors. These researchers reported that chicken and ruminant associated genotypes only partially explained foodborne transmission and that it was likely that environmental transmission (i.e. following contact with a contaminated environment) was also important. No studies have previously been performed that combine case-case and case control studies solely on _C. coli_ using genotyping data.

Scotland, with a population of 5.25 million, is an appropriate area to conduct investigations into the aetiology of human _C. coli_ infection because of its relatively high disease incidence (approximately 95 cases per 100,000 [13], its spectrum of demographic (e.g. rural and urban) and social (e.g. affluent and deprived) characteristics and the wide range of risk factors to which its population is exposed. The aim of this paper is investigate the aetiology of human _C. coli_ infections using genotyped isolates by conducting and analysing (1) a simulated case-control study where _C. coli_ cases are compared to randomly generated controls from the human population, (2) a case-case study that compares _C. coli_ cases to _C. jejuni_ cases, (3) comparison MLST genotypes from humans and animals to determine their genealogy, source attribution and diversity and (4) a case-case study that compares _C. coli_ cases attributed to chicken with those assigned to other animal reservoirs.

### Materials and Methods

#### Data

A clinical dataset comprising 2,733 _C. jejuni_ and 307 _C. coli_ cases typed by MLST was collected from across Scotland from 1st September 2005 to 31st August 2006. This comprised 52% of the total reported Scottish cases over this period. Case information was anonymous but included the postcode sector of main residence, age, gender, and the date of the laboratory report [13] (See File S1). Human population data stratified by age, postcode sector and gender was obtained from the 2001 Scottish census. The Carstairs index of deprivation was used to describe the socioeconomic status of the human population [27]. Cattle, pig, sheep and poultry numbers in 2 by 2 km tetrads were obtained from the 2004 Scottish agricultural census and these were integrated into postcode sectors using ArcView 3.3 (ESRI, Redlands, California, USA).

#### Risk Factors for Case-control and Case-case Analysis

Six parameters were available as putative risk factors; (1) age (young - 0 - 19 yrs old and adult - >20 yrs old), (2) gender (male or female), (3) season (summer - June to August - or the rest of the year), (4) rural or urban human population density (rural - <200 individuals/km^2, urban - ≥200 individuals/km^2), (5) deprived (Carstairs index≥0) or affluent (Carstairs index<0) and (6) animal population density. The animal population density (cattle, pigs, poultry and sheep) were subdivided into four groups: group 1 (null density), group 2 (low density), group 3 (medium density), group 4 (high density) for each postcode sector (see File S1).

All of the predictive variables were used in three observational analyses employing univariate and multivariate logistic regression employing the epidemiological modelling software package EGRET (EGRET, version 2.0.5, Cytel Software Corporation, Cambridge, MA, USA). Results for each risk factor were considered as statistically significant when P<0.05. Factors from the univariate analyses with a P value of <0.25 were used in the multivariate analysis.

#### Case-control

This analysis compared the 307 _C. coli_ clinical cases with 921 controls generated by randomly sampling the human population as described by the Scottish census (www.scrol.gov.uk).

#### Case-case _C. coli_ versus _C. jejuni_

307 _C. coli_ cases were compared with the 2,733 _C. jejuni_ cases as controls.

#### MLST Analysis

The clonal genealogy of _C. coli_ sequence types (STs) was estimated using a model-based approach for determining bacterial microevolution: ClonalFrame software (version 1.0; http://www2.warwick.ac.uk/fac/sci/statistics/staff/research/dideлот/ clonalframe/[28]). This approach incorporates both point mutation and recombination events. The program was run with 50,000 “burn-in” iterations which are discarded to minimise the effects of initial values followed by 50,000 data collection iterations. The consensus trees represent combined data from three independent runs, with 75% consensus required for inference of relatedness. The probable reservoir origin of _C. coli_ MLST sequence types (STs) was investigated by STRUCTURE genetic population software [29]. Using this method, STs can be probabilistically assigned to ancestral populations based on their frequency. A source dataset of _C. coli_ strains with known origins was used as a source reference population and clinical isolates were attributed to
this based on ST similarities (See File S2). This source dataset comprised 85 cattle, 322 pigs, 439 chicken and 57 sheep isolates (see File S3) obtained from both the PubMLST database and the CaMPS study [13]. The diversity of cases was determined by Simpson's index [30] where a value of 0 indicates homogenous STs and a value of 1 indicates a heterogeneous population with maximum diversity. Confidence intervals were calculated using a bootstrap method from the PopTools add-in for Microsoft Excel (available from http://www.cse.csiro.au/poptools).

**Case-case Chicken Attributed Strains versus Non Chicken Strains**

The *C. coli* STs from cases were assigned to putative source (chicken or non-chicken - cattle, pigs and sheep) when the attribution score was greater than 0.6 (See Files S3 and S4). This analysis then compared 113 *C. coli* cases attributed to chicken with the 181 non chicken cases as controls. Scores from 13 cases were too ambiguous to determine source and were removed from this further analysis.

**Ethics Statement**

The Multi-Centre Research Ethics Committee (MREC) for Scotland granted an ethical approval (REC ref: 06/MRE00/85) for acquisition and use of the dataset; additionally, approval for the research was obtained from the Research and Development Committee in each of the NHS Health Boards.

**Results**

**Case-control Study**

In univariate analysis *C. coli* cases were more common in adults than children, in rural rather than urban environments, in affluent as opposed to deprived areas, in postcode sectors with a high pig density and during the summer compared to the remainder of the year (Table 1). These were the only statistically significant factors used in the multivariate analysis as none of the rest had P values <0.25. The multivariate analysis also found that human campylobacteriosis from *C. coli* was statistically significantly associated with being an adult, living in a rural area, and contracting the disease during the summer months.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Unit</th>
<th>O.R.</th>
<th>C.I. (95%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Univariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>child</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>3.346</td>
<td>2.234–5.013</td>
<td>0.000*</td>
</tr>
<tr>
<td>Gender</td>
<td>male</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>0.878</td>
<td>0.678–1.137</td>
<td>0.323</td>
</tr>
<tr>
<td>Season</td>
<td>rest of year</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>summer</td>
<td>2.531</td>
<td>1.935–3.311</td>
<td>0.000*</td>
</tr>
<tr>
<td>Location</td>
<td>rural</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td>0.573</td>
<td>0.437–0.751</td>
<td>0.000*</td>
</tr>
<tr>
<td>Carstairs</td>
<td>affluent</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>deprived</td>
<td>0.654</td>
<td>0.502–0.851</td>
<td>0.002*</td>
</tr>
<tr>
<td>cattle density*</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>0.985</td>
<td>0.880–1.103</td>
<td>0.796</td>
</tr>
<tr>
<td>pig density*</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>1.167</td>
<td>1.050–1.298</td>
<td>0.004*</td>
</tr>
<tr>
<td>poultry density*</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>1.034</td>
<td>0.924–1.157</td>
<td>0.557</td>
</tr>
<tr>
<td>sheep density*</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>1.023</td>
<td>0.911–1.149</td>
<td>0.701</td>
</tr>
<tr>
<td>(B) Multivariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>child</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>3.552</td>
<td>2.221–5.059</td>
<td>0.000*</td>
</tr>
<tr>
<td>season</td>
<td>rest of year</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>summer</td>
<td>2.596</td>
<td>1.969–3.423</td>
<td>0.000*</td>
</tr>
<tr>
<td>location</td>
<td>rural</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td>0.546</td>
<td>0.411–0.724</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

*Odd ratios and their associated p-value for all the selected cases in the univariate models. Factors with P≤0.05 are considered as significant (*). Factors with a P<0.25 are entered in the multivariate model (†). (A) Factors with a P-value at each step, are not shown. The program used to execute the analysis gave P = 0.0000 for the overall model fit equal to 0.0000. (B) Odd ratios for the final multivariate model. Previous steps, consisting in removing one by one the factors with the highest P-value at each step, are not shown. The program used to execute the analysis gave P = 0.0000 for the overall model fit equal to 0.0000. (A) Odd ratios and their associated p-value for all the selected cases in the univariate models. Factors with P≤0.05 are considered as significant (*). Factors with a P < 0.25 are entered in the multivariate model (†). (B) Odd ratios and P-values for the final multivariate model. Previous steps, consisting in removing one by one the factors with the highest P-value at each step, are not shown. The program used to execute the analysis gave P = 0.0000 for the overall model fit equal to 0.0000. (A) Odd ratios and their associated p-value for all the selected cases in the univariate models. Factors with P≤0.05 are considered as significant (*). Factors with a P<0.25 are entered in the multivariate model (†). (B) Odd ratios and P-values for the final multivariate model. Previous steps, consisting in removing one by one the factors with the highest P-value at each step, are not shown. The program used to execute the analysis gave P = 0.0000 for the overall model fit equal to 0.0000.

**Discussion**

The case-case (*C. colii* C. jejuni) study shows that there are a higher proportion of *C. coli* than *C. jejuni* cases in adults than children. This finding has been reported previously [6,13,31] where it was found that *C. coli* incidence is higher in older than younger people. The reasons for this are unknown although it is likely to be due to behavioural factors, influencing exposure, or physiological factors, influencing susceptibility, or a combination of both. One possibility is differential acid resistance between *C. coli* and *C. jejuni*. This would have the greatest impact in the adult/elderly population, where proton pump inhibitors are more heavily used and have been demonstrated to be associated with

**Case-case Studies and MLST Analysis**

The first case-case analysis comparing *C. coli* cases to those from *C. jejuni* found that *C. coli* cases were more frequent in adults and during the summer months (Table 2). Only one other factor – residence in a rural area - had P<0.25, and was added to the multivariate model. The multivariate analysis showed the same pattern with an increased probability of *C. coli* infection in adults, living in a rural areas and during the summer.

The ClonalFrame phylogeny of *C. coli* sequence types (Fig. 1A) shows that particular clades dominated particular hosts. It was observed that 31% of cattle, 100% of sheep, 17% of pigs and 62% of chicken STs are also found in humans. Attribution by structure (Fig. 1B) assigns 41% of human clinical cases to sheep, 40% to chicken and lower proportions to cattle (14%) and pigs (66%). Simpson’s index (Fig. 1C) shows that pigs and chickens have the greatest diversity of *C. coli* STs, whilst cattle and sheep the least with humans being intermediate.

The univariate case-case analysis comparing chicken attributed STs to non-chicken STs (Table 2) showed females more likely to be infected than males to be infected by chicken strains. Similarly, in the multivariate analysis where only gender and season were used in the analysis (P<0.25), only gender was statistically significant (P = 0.006), supporting the observation that *C. coli* infections involving strains attributed to chicken were more common in females.
increased risk of campylobacteriosis [10]. The seasonality of human campylobacteriosis has been researched extensively, although this has been primarily on all Campylobacter infections [32,33]. The case-control study indicates that there is a higher incidence of C. coli infection in the summer months and this can potentially be explained by the same risk factors associated with increased C. jejuni infection in summer (e.g. travel, greater exposure to environmental sources, greater prevalence in poultry resulting in increased human exposure and therefore infection). Further work is required to establish which of these factors are the most relevant, by attributing seasonal C. coli cases to source. However, the case-case finding that C. coli infection has a higher summer incidence that for C. jejuni which differs from results previously published from England and Wales [6]. The reasons for this are unclear. Most of the poultry consumed within the UK is farmed, processed and distributed within the country, so differences in farming or production between Scotland and the rest of the UK is unlikely to provide an explanation. Travel, particularly abroad, is likely to be more common across the UK during the summer months. England has easier access to the continent than Scotland does, and has a higher proportion of first and second generation immigrants who may be more likely to travel abroad to meet family etc. [6]. Again, further work is required to understand the difference of increased C. coli incidence during the summer months.

Table 2. Results of the logistic regression for the case-case studies.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Reference</th>
<th>O.R.</th>
<th>C.I. (95%)</th>
<th>P-value</th>
<th>O.R.</th>
<th>C.I. (95%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Univariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>child</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>1.696</td>
<td>1.147–2.506</td>
<td>0.008†</td>
<td>0.816</td>
<td>0.371–1.795</td>
<td>0.614</td>
</tr>
<tr>
<td>gender</td>
<td>male</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>1.091</td>
<td>0.862–1.382</td>
<td>0.469</td>
<td>1.940</td>
<td>1.205–3.125</td>
<td>0.006†</td>
</tr>
<tr>
<td>season</td>
<td>rest of year</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>summer</td>
<td>1.285</td>
<td>1.014–1.628</td>
<td>0.038‡</td>
<td>1.362</td>
<td>0.850–2.182</td>
<td>0.200†</td>
</tr>
<tr>
<td>location</td>
<td>rural</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1.143</td>
<td>0.705–1.853</td>
<td>0.589</td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td>0.793</td>
<td>0.622–1.010</td>
<td>0.060†</td>
<td>0.830</td>
<td>0.510–1.350</td>
<td>0.452</td>
</tr>
<tr>
<td>Carstairs</td>
<td>affluent</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>deprived</td>
<td>1.021</td>
<td>0.801–1.301</td>
<td>0.866</td>
<td>0.830</td>
<td>0.510–1.350</td>
<td>0.452</td>
</tr>
<tr>
<td>cattle density</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>0.962</td>
<td>0.867–1.069</td>
<td>0.473</td>
<td>1.056</td>
<td>0.860–1.296</td>
<td>0.604</td>
</tr>
<tr>
<td>pig density</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>0.975</td>
<td>0.888–1.071</td>
<td>0.597</td>
<td>1.107</td>
<td>0.918–1.336</td>
<td>0.287</td>
</tr>
<tr>
<td>poultry density</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>0.969</td>
<td>0.876–1.071</td>
<td>0.533</td>
<td>0.999</td>
<td>0.816–1.222</td>
<td>0.991</td>
</tr>
<tr>
<td>sheep density</td>
<td>low density</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high density</td>
<td>1.026</td>
<td>0.921–1.144</td>
<td>0.643</td>
<td>1.017</td>
<td>0.827–1.251</td>
<td>0.874</td>
</tr>
<tr>
<td>(B) Multivariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gender</td>
<td>child</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>gender male</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>1.791</td>
<td>1.209–2.653</td>
<td>0.004*</td>
<td>female</td>
<td>1.940</td>
<td>1.205–3.125</td>
</tr>
<tr>
<td>season</td>
<td>rest of year</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>summer</td>
<td>1.313</td>
<td>1.035–1.665</td>
<td>0.025*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location</td>
<td>rural</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td>0.769</td>
<td>0.603–0.981</td>
<td>0.034*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(A) Odd ratios and their associated P-value for all the selected cases in the univariate models. Factors with P<0.05 are considered as significant (*). Factors with a P<0.25 are entered in the multivariate model (!).

(B) Odd ratios and p-values for the final multivariate models. Previous steps, consisting in removing one by one the factors with the highest p-value at each step, are not shown. The program used to execute the analysis gave P=0.0060 for the overall model fit for the chicken versus non chicken case-case study, and P=0.0006 for the C. coli versus C. jejuni case-case study. Because gender is the only factor kept at the end of the multivariate model in the chicken versus non chicken study, odd ratio and P-value are the same as in the univariate gender model.

Animals are grouped into four density groups (see File S1) and the odds ratio indicates the relative amount by which the odds of the outcome changes when the value of the predictor value is increased by 1.0 unit.

doi:10.1371/journal.pone.0064504.t002
animals etc.), which may help to explain this finding. Pigs are a potential source of human *C. coli* infection because they have a high prevalence (e.g. up to 99% [14]), especially when compared to other sources. However source attribution suggests that they are relatively unimportant, with only 6% of human *C. coli* cases being attributed to pigs. There is a high diversity of genotypes observed in pigs and only a small proportion of these STs (17%) were associated with human illness in the current study. Prevalence rates on retail pork are low (<0.5%) compared to those found on retail chicken (52–90%) [35,36] which suggests the foodborne route is not a major contributing factor. The univariate case-control study showed a significant association of *C. coli* infection with increasing pig density. However, the odds ratio being close to one (O.R. = 1.167) suggests that pig density can only potentially explain a small proportion of the actual number of cases. The univariate analysis also found that there was a greater risk of *C. coli* infection for cases living in more affluent areas and this has already been reported elsewhere for *Campylobacter* spp. infections [37]. It remains unclear whether this effect is real or associated with a reporting artefact [38]. In the multivariate model, affluence is no longer statistically significant and it is likely that this is correlated to some extent with population density since the poorest areas in Scotland are generally located in postal sectors with the highest population density.

The case-case study utilising source attribution data found that female gender was associated with chicken STs. This agrees with a similar study for *C. jejuni* [26]. It is unknown whether this difference is a behavioural or physiological phenomenon; a possible explanation is that women are more likely to handle and prepare raw chicken in the home and as a consequence are at a greater risk. This is corroborated by research findings in the USA which report that on average women spend three times longer than men preparing food each week [39]. However, it is also possible that increased female susceptibility plays a role [40].

The source attribution analysis indicated that both sheep (41%) and chicken (40%) were the main sources of human *C. coli* infection. This is different to source attribution studies for *C. jejuni* in Scotland, North-West England [41], New Zealand [42] and The Netherlands [18] which all showed that chicken had a higher attribution to human cases (57–80%) than sheep (2.5–24%). This raises the question of why sheep are potentially as important for *C. coli* infection as chicken in Scotland. Further, which pathways (e.g. foodborne or environmental - e.g. direct contact with the environment or waterborne) are most important? Retail surveys of chicken [43,44] show that approximately 30% are contaminated with *C. coli* whereas *C. coli* and *C. jejuni* are rarely isolated from red meats e.g. *C. coli* was absent in 1,056 retail lamb samples from the UK [35]. In addition the consumption of chicken is higher than that of sheep products [45]. This suggests that the foodborne route for human *C. coli* infection is more likely to be associated with chicken, which is counter to the source attribution findings. When considering environmental transmission to humans in rural areas, human contact through animal faeces, either directly or indirectly (e.g. via water), is more likely with sheep faeces than with chicken since almost all chickens are housed, with faeces and litter collected and disposed (a significant proportion in the UK is actually incinerated). The case-control study did find that living in a rural area was associated with an increased likelihood of becoming infected by *C. coli* rather than *C. jejuni*, but there was no association with increasing density of sheep. This was also supported by the results from the case-case study that also failed to find an association with sheep density. These data provide little evidence that a large proportion of human cases are of sheep origin. A possible explanation is that there is a large diversity of

---

Figure 1. A, ClonalFrame tree of *C. coli* by host (brown – cattle, green – sheep, pink – pigs, yellow – chicken and red – human clinical. B, probabilistic assignment of the host of human *C. coli* infections using STRUCTURE attribution model (four equal sized columns would be expected in the absence of any genetic differentiation by host). C, Simpson’s index of diversity by host.

doi:10.1371/journal.pone.0064504.g001
STs in chicken, a number of which may exhibit a lower virulence in humans. All of the C. coli strains originating from sheep in the current study are also found in both chicken and humans. It may be that these strains are more likely to cause disease in humans. Indeed, the sheep-associated strains excreted into the environment could be a source for chicken colonisation which ultimately infects humans. Sheep have a C. coli prevalence of 11%, can shed up to $10^6$ CFU/g in their faeces [15], can be grazing nearby to broiler houses and only a small breakdown of biosecurity (e.g. via insects, rodents, contaminated drinking water, physical transfer via the soles of boots) can lead to contamination of the broiler flock. This explanation highlights one of the main disadvantages of human source attribution, in that it does not take account of the transmission pathway. It only compares the distribution of strains in the source reservoirs with the human host. That being said, the comparison of source attribution reservoirs to risk factors has been demonstrated to be biologically feasible for C. jejuni [18]. Further work is required to improve our understanding of the infection routes, sources and epidemiology of C. coli infections that are evidently different from C. jejuni. A case-control study, with associated genotyping could be conducted to include more proximate risk factors (e.g. contact with sheep, drinking from private water supplies, handling raw chicken etc.).

Conclusions

The aetiology of human C. coli infections is similar in a number of respects to C. jejuni but there are important differences. There is an increased risk of C. coli infection in the older people, in people who live in rural areas and during the summer months. Public health together with national and international food safety agencies should take these differences into account when considering interventions to reduce the incidence of this gastro-intestinal pathogen.

Supporting Information

File S1 C. coli and C. jejuni human clinical data including all 9 raw and transformed (for logistic regression) variables together with MLST sequence type. (XLSX)

File S2 Host attribution scores for each C. coli ST. (XLSX)

File S3 List of all isolates by ST and host. (XLSX)

File S4 Cut-off attribution scores used to classify human clinical C. coli isolates as chicken or non-chicken strains. (XLSX)

Acknowledgments

This study used the Campylobacter Multilocus Sequence Typing Web site (http://pubmlst.org/campylobacter/), developed by Keith Jolley and Man-Suen Chan at the University of Oxford.

Author Contributions

Conceived and designed the experiments: FR OR ASP JC KF ES MCJM NS. Performed the experiments: FR OR PB MM. Analyzed the data: FR OR SS PB ASP JC ES MF NS. Contributed reagents/materials/analysis tools: MM ASP JC KF. Wrote the paper: FR OR MM SS PB ASP JC MCJM ES MF NS.

References


