**A novel sampling method for assessing human-pathogen interactions in the natural environment using boot socks and citizen scientists, with an application to the seasonality of *Campylobacter***

**Authors: Natalia R Jones1#, Caroline Millman2, Mike van der Es3, Miroslava Hukelova2\*, Ken J Forbes4, Catherine Glover5, Sam Haldenby**6**, Paul R Hunter3, Kathryn Jackson6, Sarah J O’Brien7, Dan Rigby2, Norval J C Strachan8, Nicola Williams9, Iain R Lake1, for the Enigma consortium.**

1School of Environmental Sciences, University of East Anglia, Norwich, UK

2Department of Economics, School of Social Sciences, University of Manchester, Manchester, UK

3Norwich Medical School, University of East Anglia, Norwich, UK

4School of Medicine, Medical Sciences & Nutrition, University of Aberdeen, UK

5Institute of Infection and Global Health, University of Liverpool, Liverpool, UK

6Centre for Genomic Research, University of Liverpool, Liverpool, UK

7Institute of Psychology, Health & Society, University of Liverpool, Liverpool, UK

8School of Natural and Computing Sciences, University of Aberdeen, UK

9Institute of Infection and Global Health, University of Liverpool, Liverpool, UK

Running title: Environmental sampling by citizen scientists

# Address correspondence to Natalia R Jones, [n.jones@uea.ac.uk](mailto:n.jones@uea.ac.uk)

\*Present address: Miroslava Hukelova, University of Hull, Hull, UK

**Abstract**

This paper introduces a novel method for sampling pathogens in natural environments. It uses fabric boot socks worn over walkers’ shoes allowing collection of composite samples over large areas. Wide area sampling is better suited to studies focussing upon human exposure to pathogens (e.g. recreational walking). This sampling method is implemented using a Citizen Science approach: groups of three walkers wearing boot socks undertook one of six routes, 40 times over 16 months in the North West (NW) and East Anglian (EA) regions of England. To validate this methodology we report the successful implementation of this Citizen Science approach, the observation that *Campylobacter* was detected on 47% of boot socks, and the observation that multiple boot socks from individual walks produced consistent results. Findings indicate elevated *Campylobacter* presence in the livestock dominated NW in comparison to EA (55.8% vs 38.6%). Seasonal variation in *Campylobacter* presence was found between regions, with indications of winter peaks in both regions, but a spring peak in NW. *Campylobacter* presence on boot socks was negatively associated with ambient temperature (p=0.011) and positively associated with precipitation (p<0.001), results which are consistent with our understanding of *Campylobacter* survival and the probability of material adhering to boot socks. *C. jejuni* was the predominant species found, with *C. coli* largely restricted to the livestock dominated NW. Source attribution analysis indicated that the potential source of *C. jejuni* was predominantly sheep in NW and wild birds in EA but did not vary between peak and non-peak periods of human incidence.

**Importance**

There is debate in the literature on the pathways through which pathogens transfer from the environment to humans. We report on the success of a novel method for sampling human-pathogen interactions using boot socks and citizen science techniques, which enable us to sample human-pathogen interactions that may occur through visits to natural environments. This contrasts with traditional environmental sampling, which is based upon spot sampling techniques and does not sample human-pathogen interactions. Our methods are of practical value to scientists trying to understand transmission of pathogens from the environment to people. Our findings provide insight into the risk of *Campylobacter* from recreational visits and an understanding of how these risks vary seasonally and the factors behind these patterns. We highlight the *Campylobacter* species predominantly encountered and the potential sources of the *C. jejuni*.

**Introduction**

*Campylobacter* is the most common bacterial cause of diarrhoeal disease in the developed world (1). Whilst most cases are self-limiting, longer-term sequelae include irritable bowel syndrome and Guillain-Barré syndrome (2). In the European Union (EU) it is the most frequently reported food-borne illness, estimated to cause over 190,000 human cases annually. Due to under-reporting the true incidence is believed to be around 9 million annually. The cost of *Campylobacter* infections to public health systems and in lost productivity in the EU is estimated at EUR 2.4 billion a year (3). In England and Wales more than one million cases have been recorded officially since its emergence in the 1970s (1), and given under-reporting that is known to occur, this equates to over 10 million cases in total (4).

Campylobacter epidemiology is complex, with one of the most prominent features being a strong seasonal peak in cases. This occurs towards the end of June in the UK (5). The causes of this seasonal peak are unknown, with proposed hypotheses including changes in food preparation (e.g. barbeques), increases in chicken flock colonisation, and seasonal changes in human engagement with the outdoor environment in which *Campylobacter* might be present (6). Human *Campylobacter* infections have also been shown to be positively associated with temperature (5) and precipitation (7) although the precise mechanisms for these associations are unclear. Within England there are distinct regional variations in cases, with greater numbers of cases seen in northern and western regions compared to southern and eastern England (1, 5). Although foreign travel is associated with 20% of *Campylobacter* cases in the UK (8), the peak in travel associated cases occurs later in the summer (9). Conventionally it is assumed that most *Campylobacter* cases are transmitted from food, primarily poultry meat. However, recent studies estimate that food-borne routes of exposure account for around 50% of the disease burden (10, 11). Therefore, non-food pathways of transmission are likely to be important for *Campylobacter* infection.

The predominant human pathogens (*C. jejuni* andto a lesser extent *C. coli*) are found in ruminants, pigs, poultry, and wild birds and animals (12). Furthermore, studies indicate that high levels of *Campylobacter* are found in individual environmental components such as soils, faeces, and water (13). Recreational visits to the countryside have been hypothesised as one pathway through which *Campylobacter* in the environment could be transferred to people (6). The mechanism for this human environment interaction would be via material containing *Campylobacter* being picked up by people as they walk through these environments as part of outdoor recreation via (for example) their shoes, their hands touching fences and the ground, or splashes from mud onto clothes. This material could ultimately be transferred into mouths, facilitating the transfer of *Campylobacter* to humans.

To investigate this we deploy a novel method for sampling human-environment interactions, specifically using fabric boot socks worn over citizen scientists’ boots. This has advantages over traditional methods of environmental sampling, such as small site samples of water or soil, in that it is a composite sample over a large geographical area. However, most importantly, this method samples the human-environment interaction and indicates whether after such an interaction, in this case a countryside walk, the individual’s footwear is positive or negative for *Campylobacter*. It also samples the entire foot unlike some previous studies which have taken footwear swabs (14). Boot sock sampling methods have been used before in small scale farm sampling (e.g. broiler houses; 15), but here we extend their use to examine the method’s utility for outdoor environmental sampling over far greater distances (up to 4km) and durations (c 60 minutes). This study is also novel as we utilise citizen science techniques (16) in which citizens participate in the collection of microbiological data. Specifically we utilised networks of volunteer walkers to undertake the environmental sampling in 40 individual weeks over a 16 month period.

In this paper we present our methodology for using citizen scientists to sample human-environment interactions using boot socks. We then validate this methodology by considering the utility of citizen scientists for environmental sampling and examine the internal consistency of our method by comparing the results from multiple walkers traversing the same route in the environment. We show that our methods are able to detect *Campylobacter* on walkers’ shoes and highlight the *Campylobacter* species found. We also attribute the *C. jejuni* found to potential sources through source attribution using microbial subtyping. We present results on the seasonality of *Campylobacter* positive boot socks in two contrasting regions of England and subsequently generate statistical models to examine the role of a variety of environmental variables (e.g. temperature, rainfall) on the probability of *Campylobacter* positive boot socks.

**Methods**

The study was conducted in two English regions: North West (NW) and East Anglia (EA), which have different *Campylobacter* infection rates. In NW the seasonal peak in human *Campylobacter* infection is more pronounced than in EA. The agricultural characteristics and climate of the two regions are also markedly different with NW being wetter and dominated by livestock in comparison to EA, which is drier and mostly arable.

Within EA and NW candidate walking routes were selected. To ensure that these were frequented by walkers, within each region areas in the top quintile of visits were selected. This was achieved using a gridded density map of visits to natural areas (17).

To ensure that these areas were regionally representative in terms of likely *Campylobacter* sources, candidate sites were identified that had numbers of livestock (sheep, cows, or cows and sheep combined), within the median regional quintile. These data were obtained from the UK Agricultural Census, which provides data on the numbers of livestock in 2km grid squares (18). Areas satisfying both these criteria and 4km2 in area (to provide a large walk area) were retained. Within or nearby to these areas, large scale maps (OS 1:25,000; 19) were used to identify circular 4km walking routes (to keep the time and cost requirements of sampling manageable) along public rights of way. Routes had to have an accessible starting point. Potential walks were inspected by the research team to ensure that the majority was on natural surfaces (e.g. soil or grass), that the route was safe in all weathers and not proximate to a likely *Campylobacter* hotspot (e.g. large poultry farm) in order to avoid skewed data. Routes did not include any long sections (>100m) across arable fields, as during trials excessive amounts of soil from newly ploughed fields caused the boot sock to detach. Three routes were selected in each region (i.e. six in total) which were spatially dispersed, being at least 30km apart, and which ensured the various landscapes in the regions were included. The six routes and a description of each were:

NW Upland, Cumbria, predominantly sheep

Foothills, Lancashire, predominantly cattle

Lowlands, Cheshire, predominantly cattle and sheep

EA Lowland, Norfolk, predominantly arable, cattle and sheep

Lowland, Suffolk, predominantly arable, cattle

Lowland, Norfolk, predominantly arable interspersed with wetlands, sheep

The six routes were walked weekly from April 2013 to July 2013 and from April 2014 to July 2014. In between these two periods walks were conducted every 3 weeks. This frequency ensured that sampling was intensified around the seasonal peak of *Campylobacter*. In total walks were undertaken on 40 separate weeks on each of the six walking routes (i.e. 240 walks in total). Each of the walks was traversed by three individuals walking together, resulting in three samples for each walk and a total of 720 boot socks across the whole study. Three walkers were chosen so that replicates would be produced for each walk (i.e. 3 walkers on each of the 6 walking routes on every week when walks occurred).

A boot sock protocol was developed to ensure their correct use by walkers. At the start of each walk the walkers placed a plastic overshoe over the shoe on their leading foot to prevent contamination between shoe and boot sock. A new boot sock (Tunika 60gm disposable shoe cover) was placed over the plastic overshoe. The walkers then traversed the specified route, and were asked to walk normally. During the walk the participants were asked to check their boot sock for large tears. If any were found then another boot sock was placed over the existing boot sock. At the end of the walk participants removed their boot socks using a clean glove. Each individual walker then placed their boot sock(s) in an individual biohazard specimen bag, which was sealed and labeled using a pre-completed laboratory sample form. The biohazard bag was placed inside a second plastic sealable bag to prevent leakage, and an absorbent sachet was added to this second bag to remove excess moisture. This plastic bag was then placed in a plastic pre-addressed envelope, which specified (to conform to postal regulations) it contained a UN3373 Biological Substances Diagnostic specimen, and posted to the Preston Public Health England Laboratory for the first month of the study and subsequently to the University of Liverpool laboratories to be tested for the presence of *Campylobacter*. Studies have indicated that *Campylobacter* can survive for up to two weeks at ambient temperatures (20), and thus the postal service was deemed an appropriate method for transporting boot socks.

Walkers were recruited from local walking groups and through advertisements placed close to the selected routes. Written consent to take part in the study was obtained (UEA consent form 150113). Walkers were screened using a health questionnaire to identify those able to walk the route safely. As already mentioned, each walk was traversed by three individuals walking together for both replication and safety. Walking with pets was not allowed to prevent possible contamination. All walks were undertaken at a similar time each week (Mondays 10am-2pm). Walkers were offered a small remuneration of supermarket vouchers for each completed walk. A training event, including a pilot walk, was held to explain the boot sock protocol and walkers were provided with written instructions and guidance on how to deal with any eventuality. An emergency phone was supplied.

A subset of walkers were designated as walk leaders and provided with extra training regarding the correct procedure to collect, prepare and post the boot socks to the laboratory. One walk leader was present on each walk. Walkers were sent a reminder (text, phone or email) by the research team in advance of each walk. Members of the research team periodically accompanied the walkers to ensure that the protocol was adhered to. To maintain engagement with such a large group of walkers, a newsletter about the project was produced quarterly and sent to all walkers.

To further characterise each walk, walkers were asked to complete an observational survey on each walk. For this purpose each walk leader was provided with a smartphone that was used to collect and submit observational data. The route was divided into three sections (contiguous parcels of visible land), and for each walk, walkers were asked to categorise the number (0, 0-10, >10) and type of livestock (sheep, cows, horses and pigs) seen in each section. They were also asked to record the weather and the underfoot conditions. Additionally, the route taken during the walk was recorded using a GPS tracking application. This was automatically uploaded to the research team and provided validation that the walk had been successfully completed.

At the laboratory, multiple boot socks from one walker (i.e. where large tears had developed in the boot sock during the walk and a second boot sock had been placed over the first) were pooled to produce one result per walker (hereafter “boot sock”). Therefore, three boot sock samples were analysed for each walk, resulting in 720 boot socks (6 walking routes x 40 weeks x 3 boot socks). On receipt in the laboratory, boot sock details were logged and 100 mL of buffered peptone water (BPW) was added to each boot sock and hand palpated for one minute. Boot socks were left for 10 minutes to allow the sediment to settle and 10 mL of the supernatant was then added to 10 mL of double concentrated Exeter broth (1.1 L nutrient broth, 11 mL lysed defibrinated horse blood, C*ampylobacter* enrichment supplement SV59 (Mast Group Ltd, Bootle, UK) and *Campylobacter* growth supplement SV61 (Mast Group Ltd)). This was incubated under microaerobic conditions (80% N2, 12% CO2, 5% O2 and 3% H2) in a variable atmosphere incubator (Don Whitley Scientific Ltd, Shipley UK) at 41+1oC for 48 hours.

A 47 mm, 0.45 microns cellulose nitrate filter (Sartorius, Epsom UK) was placed on top of a modified blood-free charcoal *Campylobacter* (mCCDA) selective agar plate containing cefoperazone (32 mg/L) and amphotericin B (10 mg/L). Altogether, 100 L of enriched broth was added to the agar plate, then spread over the surface of the filter and allowed to sit at room temperature for 30 minutes. After this time, the filter was removed and the plates incubated as above for up to 72 hours. The addition of the filter aided recovery of *Campylobacter* species from the broth, as, due to their size and mobility, they are able to easily pass through the filter onto the agar plate surface, with other contaminating bacteria being retained on the filter surface. Incubation plates were examined for the presence of *Campylobacter* spp. and up to four suspect *Campylobacter* colonies were sub-cultured onto two Columbia blood agar plates containing 5% defibrinated horse blood. One plate was incubated under microaerobic conditions for 48 hours at 41oC, and the other under aerobic conditions at 30°C for 48 hours. Isolates which grew under microaerobic conditions only were retained for further study. All media and supplements unless stated otherwise were from Lab M (Bury, UK) and blood obtained from Southern Group Labs Ltd (Corby, UK).

Cell lysates were prepared using a chelex method (21) from fresh cultures for PCR identification. Briefly, a loopful of bacteria was suspended in 300 mL of a 20% (w/v) chelex solution (chelex-100 in 10 mM Tris-HCL and 1 mM EDTA, pH 8) (Bio-Rad, UK), which was heated at 95oC for 10 minutes and then subject to centrifugation at 10,000 rpm for two minutes. Fifty L of supernatant added to 450 L of sterile deionised water was used for PCR amplification. A genus specific PCR assay was used to confirm an isolate as *Campylobacter* species as previously described (22). For each sample identified as *Campylobacter* species one isolate was subject to a LPX PCR assay to determine if an isolate was *C. jejuni* or *C. coli,* or, if negative for these species, *C. upsaliensis* or *C. lari* (23). Isolates not assigned to the latter species were further screened using *C. hyointestinalis* and *C. fetus* PCR assays (24). Isolates still not assigned to a species were subject to amplification of the *gro*EL gene followed by Sanger sequencing as previously described (25, 26). All PCR reagents were obtained from Solis BioDyne, Estonia.

One mL of the enriched broth samples also underwent the DNA chelex extraction method as stated above and were subject to the M. Katzav, et al. (22) PCR assay to determine whether they were positive for *Campylobacter* species. Therefore, even if no positive boot socks were found during the culture process, PCR analysis was still undertaken on the boot socks, which meant that boot socks could have been found to be positive by culture only, PCR only or both methods.

Further analysis was undertaken on two thirds of the boot socks positive for *Campylobacter* on culture, with a single pick of *C. jejuni* selected per boot sock walk for whole genome sequencing. TruSeq Nano DNA Illumina libraries were prepared for each DNA sample and subsequently sequenced on the Illumina HiSeq 2500 or 4000 platforms to generate 125bp paired end reads. Reads were trimmed to remove low quality bases and sequencing adaptors using Sickle v1.200 (27) and Cutadapt v1.2.1 (28), respectively. Each set of reads were de novo assembled into scaffolds using SPAdes v3.7.0 (29). These scaffolds were interrogated to determine the multi-locus sequence typing (MLST) alleles and sequence type for each isolate. This was achieved by aligning MLST allele sequences obtained from pubmlst.org against each genome assembly using Bowtie2 (30). For each locus, if an allele sequence perfectly aligned, the sample was assigned this allele. In the case of detection of novel alleles (i.e. the best alignment for a locus was a non-perfect, but full length match), these were submitted to pubmlst.org.

Structure software (31) was used to assign boot sock MLST isolate data to one of five potential infection sources (cattle, sheep, pigs, wild birds (all faecal isolates) and chicken (all retail food isolates)). This was achieved by determining the genotype frequencies among each of the potential infection sources, with the uncertainty regarding this determination of genotype frequency resulting from the probabilistic, rather than the absolute, nature of the attribution scores for each isolate (using the method of S. K. Sheppard, et al. (11)). The average proportion attributed by source was calculated and the corresponding 95% confidence intervals (CIs) (Percentiles) calculated using 10,000 Monte Carlo steps (32).

To characterise further each walk, environmental conditions on the day of the walk and for a period of 28 days previously were obtained. These environmental variables were hypothesised to be related to the *Campylobacter* presence on boot socks. The maximum morning temperature (oC) for the six routes was obtained from the Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Dataset (33). These data were from the nearest (1 - 37km) Automatic Weather Station to the route with complete data for the time period required. Using the same data source daily rainfall (mm) from the nearest (2.5 - 8km) Water Authority station to the route was obtained for the time period. As a measure of soil moisture (which may relate to periods when material is more likely to adhere to the boot sock), river flow data was obtained for the six different walk locations from the National River Flow archive (34). The nearest (1.5 - 9.5km) gauging station for the river catchment within which each walk was located in was determined and the gauged daily flow (cubic metres per second) for the study period was obtained. Measured soil moisture was obtained for two locations from the UK Environmental Change Network. However because these values were not specific to individual routes they could not be used. For similar reasons one measurement of UV, obtained from the Department for Food and Rural Affairs, was excluded.

Temperature, rainfall and river flow for the day previous to the walk and an average for the seven days and the 28 days prior to the walk were calculated. These data on environmental conditions were combined with the data collected from the citizen scientists (livestock observed, weather and underfoot conditions) and used as explanatory variables of positive boot socks. All analysis was undertaken using STATA 11, using a mixed effects logistic regression model with route and week as random effects.

The study received ethical clearance from the Faculty of Medicine and Health Science Research Ethics Committee at the University of East Anglia (2012/2013-40).

Accession number(s). The raw sequence data were deposited in the EBI ENA database (<http://www.ebi.ac.uk/ena>) with the accession no. PRJEB20152.

**Methodological Validation and Results**

Our experience of using ‘citizen scientists’ in the form of networks of walkers is overwhelmingly positive. Altogether 60 people volunteered to participate in the study. They were enthusiastic in their training and willing to take part in the study. Demographically they were a mix of genders and ranged in age from 18 to 80. Training was simple and the walkers engaged fully with the study. This extended to submitting photos, text and even a poem to the research team for use in quarterly newsletters. In terms of science delivery all 240 walks were successfully undertaken, and all 720 boot socks were posted to the laboratory on the day of the walk. For all walks, observations of livestock, underfoot conditions and weather were successfully taken. Variable mobile data signals in several of the locations meant that some walkers recorded these data on paper forms and uploaded them later in the day, as opposed to using the smart phone. This did not affect the GPS recordings for the walks, which were automatically recorded and uploaded.

The livestock seen by walkers on the different walk routes did not always reflect the livestock which was expected from the agricultural survey. In particular, in the Lancashire foothills walk, sheep rather than cows were seen. In Suffolk no livestock were seen. Pigs were seen just once in each region.

The 720 boot socks (6 walking routes x 40 weeks x 3 boot socks) were all received at the laboratory for analysis within 72 hours of the walk taking place. Overall 94% arrived within 24 hours, with 5% arriving between 24 and 48 hours after the walk and just 1% arriving between 48 and 72 hours. After each walk meetings were held between the walker management and the laboratory teams to confirm that each walk had successfully taken place and to identify any sampling issues that could be fed back to the walkers. The most common of these was the failure to place the absorbent sachet in the second bag, and occasionally not placing the boot sock in the biohazard bag.

The method proved capable of detecting *Campylobacter* on boot socks. In this study boot socks were classed as positive if they were found to be so by either PCR or culture. Of the 240 walks undertaken, 156 (65.0%) had at least one boot sock which tested positive for *Campylobacter* either by culture or PCR. Table 1 shows the total number of boot socks which tested positive, and how this varied by region. Overall 340 (47.2%) boot socks were positive for *Campylobacter* spp. by either culture or PCR with 55.8% of NW boot socks positive compared with 38.6% of EA boot socks (NW vs EA, p<0.001, by Fishers Exact Test). Table 1 also highlights the number of boot socks which were found to be positive by the two different test methods (culture and PCR). Almost all the positive NW boot socks were found to be positive by PCR (89.1%) and culture (85.6%). The positive boot socks from EA were mostly found to be positive by PCR (93.5%), with only 54.7% positive using culture (NW vs EA PCR positive not significant; NW vs EA culture positive p<0.001).

To validate the boot sock as a method for environmental sampling we examined the internal consistency of the boot socks from each walk. All 240 walks had three boot socks (from three individual walkers). An internally consistent walk was defined as one where all three boot socks showed the same result, i.e. all three boot socks were either positive or negative for C*ampylobacter* (culture or PCR). The internal consistency of walks is show in Figure 1. From the 240 walks, 70 (29.2%) had three positive boot socks and 84 (35.0%) had three negative boot socks. If the walks were not internally consistent then in terms of probability we would expect a clear majority of walks to be mixtures of positive and negative boot socks. This was not observed, providing evidence of internal consistency.

The species of *Campylobacter* found on the boot socks in the two regions is shown in Table 2. Species was only identified on the culture positive boot socks (n=248). This Table shows that the dominant species in both regions was *C. jejuni* (NW 62.2%; EA 84.2%), while almost all *C. coli* boot socks were found in the NW (NW 27.4%; EA 2.6%). This difference in species between region was statistically significant (p<0.001).

Figure 2a shows the potential source attribution of the *C. jejuni* found in each region. In EA the dominant potential source of *C. jejuni* was wild birds (70%), while in NW the sources were more mixed. Sheep was the largest potential source in NW (36%) with around 25% of isolates being derived from both wild birds and retail chicken. Figure 2b show the potential source attribution of the *C. jejuni* found in each region subdivided between the spring peak (April to July) and other times of the year. Within regions potential source attribution was similar between times of the year.

In terms of variation in positive boot socks over the study period, , Figure 3 shows the percentage of boot socks which were *Campylobacter* positive each week in NW and EA. The line on the graph shows the 5 week (current week and two sampling weeks either side) running median. Care needs to be take in interpreting these data, due to small sample sizes and inter week variability. However, the figure is suggestive of peaks in both spring periods and in winter in the NW region. In EA, the highest levels of positive boot socks were found in the winter months, with little consistent evidence for a peak in spring across the two spring periods.

A mixed effects logistic model was run to determine the associations between positive boot socks and the environmental conditions in the walk area. The results are presented in Table 3. The week and route of the walk were included as random effects in the model. Significant associations with temperature and rainfall were found (Wald χ2 = 24.61; p>χ2 <0.001). The strongest associations were found for the environmental conditions in the seven days prior to the walk date. The average (over seven days) maximum morning air temperature was negatively association with positive boot socks, while the average (over seven days) daily rainfall had a positive but non-linear association, as indicated by the inclusion of the quadratic term. The quadratic term was included rather than a logged predictor as the relationship between positive bootsocks and rainfall was not monotonic. In these final models river flow and the variables collected by the walkers were insignificant. Examination of the model deviance residuals show they were normally distributed and a plot of the deviance residuals against the fitted values produced straight lines, signifying an appropriate model structure (Figure S1 and S2 in Supplementary Material).

**Discussion and Conclusions**

Here we present a novel methodology utilising boot socks and citizen scientists to evaluate human-pathogen interactions. Working with citizens requires significant inputs in terms of delivering appropriate training and providing regular support and feedback. However, if these are in place then citizen science can be used for successful, independent, long term, and systematic environmental sampling. All walks were successfully completed, every boot sock posted to the laboratory, and all additional observations submitted.

Almost half of the boot socks (47.2%) worn on a 4km walk in the countryside were positive for *Campylobacter* either by PCR or culture. Previous studies using boot socks have done so in more controlled environments such as poultry sheds (15) or swab sampled shoes as opposed to whole footwear sampling (14). We demonstrate that using boot socks with citizen scientists, on walks which were substantial in distance, and sending boot socks to the laboratory through the postal system, we were able to extract *Campylobacter* from a high proportion of boot socks. This highlights the potential of spatially distributed, citizen science based, environmental sampling (35) for pathogens or potentially other particles in the environment.

Here we present boot sock positives by either PCR or culture. The ability to culture *Campylobacter* or detect it by PCR is affected by the loading of C*ampylobacter* in the environment. *Campylobacter* can only be detected by these methods if there is enough pathogen present. PCR is the more sensitive method and more of our boot socks were positive by PCR in comparison to culture (90.9% vs 72.9%). When a sample is positive on PCR only, this suggests that the bacteria may be present in low numbers, may be dead, stressed, or in what is known as a viable but non-culturable state (VBNC). In the NW a similar percentage of boot socks were positive by PCR and culture methods. In EA far more boot socks were positive by PCR as opposed to culture. This suggests that the *Campylobacter* found in EA may be present only in low numbers or may be in a VBNC state. There is however some evidence that bacteria in a VBNC state can still be infectious to humans and pose a public health risk (36). Some boot socks were found to be positive on culture, but negative on PCR. This could be due to carry over of potential inhibitors during the chelex extraction which inhibited the PCR assay, which is unusual but can occur when there are low numbers of pathogen.

Our confidence in using boot socks as a method for sampling the environment is enhanced by the observation that there was evidence of consistency among the three boot socks worn on each walk. On 64% of walks all three boot socks produced the same result. Our methodological confidence is enhanced by the adoption of this approach by Public Health England to sample for the presence of *Escherichia coli* (*E. coli*) O55 in the environment. This was during an outbreak in Dorset in Summer 2015 and the research team was contacted based upon a presentation at a National *Campylobacter* conference (37). This approach involved public health officials as opposed to citizen scientists, and the method was adapted to be used for hands as well as feet.

*C. jejuni*, was the predominant species found in the environment (~70%), with *C. coli* found at lower frequencies (~20%). This supports other studies that have shown *C. jejuni* to be the predominant species found in animal faeces (13). The majority of *C. coli* was found in NW region. This may be a reflection of the high numbers of sheep present in this region (18) as observed by our walkers in the NW. *C. coli* is found commonly in sheep faeces (13). It is also present in pigs (38), but these were rarely seen.

The source attribution element of the study showed that the potential source of *C. jejuni in* the environment varies considerably between the regions. However, within each region it does seem to match the likely potential sources of exposure based upon each region’s agricultural characteristics. In EA wild birds were the main potential source of *C. jejuni* (71%), which is likely to reflect the low levels of livestock in EA and hence the likely predominance of wild bird sources. In the NW, however, only 25% of *C. jejuni* was potentially derived from wild birds, with sheep (32%) being the main potential source of *C. jejuni.* This is not unexpected as sheep are abundant in the NW and were common in all three NW walks. There was no evidence that the potential sources of *C. jejuni* in the environment varied between peak and non-peak periods of human incidence.

Interpretation of the variation in our results between the spring and winter periods needs to be cautious as sampling only occurred in two spring and one winter season, sample numbers were not large and there was variability in the numbers of positive boot socks from week to week. However, our findings show some evidence of a spring peak in NW, with generally higher levels of *Campylobacter* in the months of May and June in both years compared to other months. This finding is interesting, given the recognised seasonal peak in human *Campylobacter* cases in the UK (1, 5). However in EA, there was no suggestion of a peak in the spring months. This is also interesting as in this region the spring peak in human *Campylobacter* cases is less distinct than in NW (1, 5).

In EA, and to a lesser degree NW, there was a suggestion of a peak in *Campylobacter* positive boot socks during the winter months. This may indicate that during these cool winter months *Campylobacter* survival is high and the damp conditions are ideal for material adhering to the boot sock. It is unclear why this pattern may be lower in NW but may possibly be associated with lower overall rainfall amounts in EA leading to less dilution. A winter peak is not observed in human *Campylobacter* cases, but at this time of the year visits to the countryside are typically low.

An understanding of these patterns is enhanced by the multivariable analysis of *Campylobacter* positive boot socks. In interpreting these results it is important to recognise that a positive boot socks is reliant on the *Campylobacter* being present in the environment and positive material adhering to the boot sock. Our finding that the presence of *Campylobacter* on boot socks is associated with increased rainfall over the previous seven days is supportive of results from other studies that indicate *Campylobacter* survives better in damper conditions (39-41). However, it may also be indicative of damper conditions enhancing soil moisture and hence the probability of material adhering to the boot sock. The quadratic relationship with rainfall could be indicative of higher water levels flushing *Campylobacter* from the system or a dilution effect (42, 43). The statistical analysis also indicated that *Campylobacter* positive boot socks were negatively associated with ambient temperature. This is consistent with studies indicating enhanced *Campylobacter* survival at low temperatures (44, 45).

There are limitations in using boot socks for environmental sampling, the first being that they produce a composite environmental sample and it is not possible to determine precisely where *Campylobacter* was encountered. It would be possible to use multiple boot socks over smaller sections of each walk, although this may present challenges when using citizen scientists. While this presents issues for environmental sampling, our method is ideal for sampling human-environment interactions. It is unknown how the risk explored in this study using boot socks relates to the risk of a walkers using a shoe or walking boot with features such as treads and laces, although a previous study where shoes were swabbed suggests that deep treads enhanced the adhesion of pathogens to shoes (14). Further work is needed to relate the presence of *Campylobacter* on footwear to disease risk in humans. Another limitation of using boot socks is that it provides results which are presence / absence; there is no measure of quantity or ‘loading’.

To conclude, we present a novel methodology for assessing human-pathogen interactions in the natural environment using boot socks and citizen scientists. Our method of using networks of volunteer walkers wearing boot socks was shown to be a successful data collection method, and produced boot socks from which *Campylobacter* was successfully extracted. Confidence in our results is enhanced through the observation that where multiple boot socks were worn on the same walk there was a high degree of internal consistency.

The boot sock method using citizen scientists can provide important insights into *Campylobacter* in the environment and implications for human health. Boot socks measure potential human-pathogen interactions, and the results indicated that just under half of all boot socks were positive for *Campylobacter* either via culture or PCR after a 4km walk in the countryside. *C. jejuni* was the predominantly found *Campylobacter* in the environment with *C. coli* largely restricted to the livestock dominated NW. Source attribution modelling indicated the *C. jejuni* matched the region livestock characteristics with wild bird sources providing an important component, especially in EA. There was no indication from this modelling that source varied between the spring peak in human cases and other times of the year. We highlighted the seasonal variations in *Campylobacter* presence on boot socks and indicate a spring peak in positivity in the livestock dominated NW but not in EA. This is coincident with the spring peak in human cases. Both regions show a peak in the winter, but at a time when outdoor recreation is limited. *Campylobacter* positive boot socks show significant associations with decreasing temperature and increased rainfall, results which are consistent with our understanding of *Campylobacter* survival and the probability of material adhering to boot socks.

**Acknowledgements**

We would like to thank Mrs Ruth Ryvar and Dr Amy Wedley for assistance with laboratory work. Also our large network of “citizen scientists” who irrespective of the weather spend their Monday morning walking through the countryside with white boot socks over their shoe!

We acknowledge the Medical Research Council, Natural Environment Research Council, Economic and Social Research Council, Biotechnology and Biosciences Research Council, and Food Standards Agency for the funding received for this project (ENIGMA) through the Environmental & Social Ecology of Human Infectious Diseases Initiative, grant reference G1100799/1.

We thank members of the ENIGMA Consortium for helpful discussion. The ENIGMA Consortium investigators comprise Sarah O’Brien (PI), Rob Christley, Christiane Hertz-Fowler, Paul Wigley, Nicola Williams and Craig Winstanley (University of Liverpool), Peter Diggle (Lancaster University), Iain Lake, Kevin Hiscock and Paul Hunter (University of East Anglia), Ken Forbes and Norval Strachan (University of Aberdeen), Rachel Griffith and Dan Rigby (University of Manchester), Paul Cross (Bangor University), Stephen Rushton (Newcastle University), Tom Humphrey (Swansea University), Malcolm Bennett (University of Nottingham), David Howard (Centre for Ecology and Hydrology), and Brendan Wren (London School of Hygiene and Tropical Medicine).

P.H., S.B and I.L. are funded in part by the National Institute for Health Research, Health Protection Research Unit in Gastroenterology at Liverpool University. P.H. and I.L. and also funded in part by the National Institute for Health Research, Health Protection Research Unit in Emergency Preparedness and Response at King’s College London.

**References**

1. **Nichols GL, Richardson JF, Sheppard SK, Lane C, Sarran C.** 2012. Campylobacter epidemiology: A descriptive study reviewing 1 million cases in England and Wales between 1989 and 2011. BMJ Open **2**.

2. **Keithlin J, Sargeant J, Thomas MK, Fazil A.** 2014. Systematic review and meta-analysis of the proportion of Campylobacter cases that develop chronic sequelae. BMC Public Health **14**.

3. **EFSA Panel on Biological Hazards.** 2011. Scientific Opinion on Campylobacter in broiler meat production: control options and performance objectives and/or targets at different stages of the food chain. EFSA Journal **9:**141

4. **Tam CC, Rodrigues LC, Viviani L, Dodds JP, Evans MR, Hunter PR, Gray JJ, Letley LH, Rait G, Tompkins DS, O'Brien SJ.** 2012. Longitudinal study of infectious intestinal disease in the UK (IID2 study): Incidence in the community and presenting to general practice. Gut **61:**69-77.

5. **Louis VR, Gillespie IA, O'Brien SJ, Russek-Cohen E, Pearson AD, Colwell RR.** 2005. Temperature-driven campylobacter seasonality in England and Wales. Applied and Environmental Microbiology **71:**85-92.

6. **Nichols GL.** 2005. Fly transmission of Campylobacter. Emerging Infectious Diseases **11:**361-364.

7. **Nichols G, Lane C, Asgari N, Verlander NQ, Charlett A.** 2009. Rainfall and outbreaks of drinking water related disease and in England and Wales. Journal of Water and Health **7:**1-8.

8. **Gillespie IA, O'Brien SJ, Penman C, Tompkins D, Cowden J, Humphrey TJ.** 2008. Demographic determinants for Campylobacter infection in England and Wales: Implications for future epidemiological studies. Epidemiology and Infection **136:**1717-1725.

9. **Strachan NJC, Rotariu O, Smith-Palmer A, Cowden J, Sheppard SK, O'Brien SJ, Maiden MCJ, MacRae M, Bessell PR, Matthews L, Reid SWJ, Innocent GT, Ogden ID, Forbes KJ.** 2013. Identifying the seasonal origins of human campylobacteriosis. Epidemiology and Infection **141:**1267-1275.

10. **Davis MA, Moore DL, Baker KNK, French NP, Patnode M, Hensley J, MacDonald K, Besser TE.** 2013. Risk factors for campylobacteriosis in two washington state counties with high numbers of dairy farms. Journal of Clinical Microbiology **51:**3921-3927.

11. **Sheppard SK, Dallas JF, Strachan NJC, MacRae M, McCarthy ND, Wilson DJ, Gormley FJ, Falush D, Ogden LD, Maiden MCJ, Forbes KJ.** 2009. Campylobacter genotyping to determine the source of human infection. Clinical Infectious Diseases **48:**1072-1078.

12. **Jones K.** 2001. Campylobacters in water, sewage and the environment. Journal of Applied Microbiology Symposium Supplement **90:**68S-79S.

13. **Brown PE, Christensen OF, Clough HE, Diggle PJ, Hart CA, Hazel S, Kemp R, Leatherbarrow AJH, Moore A, Sutherst J, Turner J, Williams NJ, Wright EJ, French NP.** 2004. Frequency and spatial distribution of environmental Campylobacter spp. Applied and Environmental Microbiology **70:**6501-6511.

14. **Rashid T, VonVille HM, Hasan I, Garey KW.** 2016. Shoe soles as a potential vector for pathogen transmission: a systematic review. Journal of Applied Microbiology **121:**1223-1231.

15. **Berghaus RD, Thayer SG, Law BF, Mild RM, Hofacre CL, Singer RS.** 2013. Enumeration of Salmonella and Campylobacter spp. in environmental farm samples and processing plant carcass rinses from commercial broiler chicken flocks. Applied and Environmental Microbiology **79:**4106-4114.

16. **Kullenberg C, Kasperowski D.** 2016. What is citizen science? - A scientometric meta-analysis. PLoS ONE **11**.

17. **Sen A, Harwood AR, Bateman IJ, Munday P, Crowe A, Brander L, Raychaudhuri J, Lovett AA, Foden J, Provins A.** 2014. Economic Assessment of the Recreational Value of Ecosystems: Methodological Development and National and Local Application. Environmental and Resource Economics **57:**233-249.

18. **DEFRA.** 2013. Livestock numbers in England and the UK. Department for Environment, Food and Rural Affairs.

19. **Ordnance Survey.** 2013. 1:25 000 Raster tiles. EDINA Digimap

20. **Nicholson FA, Groves SJ, Chambers BJ.** 2005. Pathogen survival during livestock manure storage and following land application. Bioresource Technology **96:**135-143.

21. **Walsh PS, Metzger DA, Higuchi R.** 1991. Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques **10:**506-513.

22. **Katzav M, Isohanni P, Lund M, Hakkinen M, Lyhs U.** 2008. PCR assay for the detection of Campylobacter in marinated and non-marinated poultry products. Food Microbiology **25:**908-914.

23. **Klena JD, Parker CT, Knibb K, Claire Ibbitt J, Devane PML, Horn ST, Miller WG, Konkel ME.** 2004. Differentiation of Campylobacter coli, Campylobacter jejuni, Campylobacter lari, and Campylobacter upsaliensis by a multiplex PCR developed from the nucleotide sequence of the lipid A gene lpxA. Journal of Clinical Microbiology **42:**5549-5557.

24. **Linton D, Owen RJ, Stanley J.** 1996. Rapid identification by PCR of the genus Campylobacter and of five Campylobacter species enteropathogenic for man and animals. Research in Microbiology **147:**707-718.

25. **Williams NJ, Jones TR, Leatherbarrow HJ, Birtles RJ, Lahuerta-Marin A, Bennett M, Winstanley C.** 2010. Isolation of a novel campylobacter jejuni clone associated with the bank vole, Myodes glareolus. Applied and Environmental Microbiology **76:**7318-7321.

26. **Kärenlampi RI, Tolvanen TP, Hänninen ML.** 2004. Phylogenetic analysis and PCR-restriction fragment length polymorphism identification of Campylobacter species based on partial groEL gene sequences. Journal of Clinical Microbiology **42:**5731-5738.

27. **Joshi NA, Fass, J. N.** 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33), Available at <https://github.com/najoshi/sickle>.

28. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal **17.1**.

29. **Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA.** 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads, p 158-170. *In* Deng M, Jiang R, Sun F, Zhang X (ed), Research in Computational Molecular Biology: 17th Annual International Conference, RECOMB 2013, Beijing, China, April 7-10, 2013 Proceedings doi:10.1007/978-3-642-37195-0\_13. Springer Berlin Heidelberg, Berlin, Heidelberg.

30. **Langmead B, Salzberg, S. .** 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods **9:**357-359.

31. **Pritchard JK, Stephens M, Donnelly P.** 2000. Inference of population structure using multilocus genotype data. Genetics **155:**945-959.

32. **Strachan NJC, Gormley FJ, Rotariu O, Ogden ID, Miller G, Dunn GM, Sheppard SK, Dallas JF, Reid TMS, Howie H, Maiden MCJ, Forbes KJ.** 2009. Attribution of campylobacter infections in northeast scotland to specific sources by use of multilocus sequence typing. Journal of Infectious Diseases **199:**1205-1208.

33. **Met Office.** 2012. Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current), NCAS British Atmospheric Data Centre. <http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0>

34. **Centre for Ecology and Hydrology.** 2015. National River Flow Archive, Natural Environment Research Council. <http://nrfa.ceh.ac.uk/>.

35. **Bonney R, Shirk JL, Phillips TB, Wiggins A, Ballard HL, Miller-Rushing AJ, Parrish JK.** 2014. Next steps for citizen science. Science **343:**1436-1437.

36. **Oliver JD.** 2005. The viable but nonculturable state in bacteria. Journal of Microbiology **43:**93-100.

37. **Wall T.** 2015. Experts fail to identify cause of mystery E. coli outbreak, p *In* Environmental Health News. <http://www.ehn-online.com/news/article.aspx?id=14914>.

38. **Ogden ID, Dallas JF, MacRae M, Rotariu O, Reay KW, Leitch M, Thomson AP, Sheppard SK, Maiden M, Forbes KJ, Strachan NJC.** 2009. Campylobacter excreted into the environment by animal sources: Prevalence, concentration shed, and host association. Foodborne Pathogens and Disease **6:**1161-1170.

39. **Doyle MP, Roman DJ.** 1982. Sensitivity of Campylobacter jejuni to drying. J Food Prot **45:**507-510.

40. **Fernández H, Vergara M, Tapia F.** 1985. Dessication resistance in thermotolerant campylobacter species. Infection **13:**197.

41. **Oosterom J, De Wilde GJA, De Boer E, De Blaauw LH, Karman H.** 1983. Survival of Campylobacter jejuni during poultry processing and pig slaughtering. J Food Prot **46:**702-706.

42. **Sterk A, Schijven J, de Roda Husman AM, de Nijs T.** 2016. Effect of climate change on runoff of Campylobacter and Cryptosporidium from land to surface water. Water Research **95:**90-102.

43. **Sinton LW, Braithwaite RR, Hall CH, Mackenzie ML.** 2007. Survival of indicator and pathogenic bacteria in bovine feces on pasture. Applied and Environmental Microbiology **73:**7917-7925.

44. **Kelana LC, Griffiths MW.** 2003. Use of an Autobioluminescent Campylobacter jejuni to Monitor Cell Survival as a Function of Temperature, pH, and Sodium Chloride. Journal of Food Protection **66:**2032-2037.

45. **Blaser MJ, Hardesty HL, Powers B, Wang WLL.** 1980. Survival of Campylobacter fetus subsp. jejuni in biological milieus. Journal of Clinical Microbiology **11:**309-313.

**Figure Legends**

Figure 1: Percentage of walks with zero, one, two and three positive boot socks per week as a measure of the internal consistency of walks

Figure 2: Proportion of boot socks positive for *C.* *jejuni* attributed to potential sources of *C. jejuni by* a. region and b. region and season. With 95% bootstrap confidence intervals.

Figure 3: Variation in the percentage of positive boot socks each week across the study period, by region. Bars indicate the percentage of boot socks which were positive for *Campylobacter* each week (n = 9). Line indicates the running median over five sample-weeks.