THE EPIDEMIOLOGY AND MOLECULAR EPIDEMIOLOGY OF GIARDIASIS IN NORTH WEST ENGLAND

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Corrado Minetti

November 2014

AUTHOR'S DECLARATION

Apart from the help and advice acknowledged, this thesis represents the unaided work of the author

.....

Corrado Minetti

November 2014

This research was carried out in the Department of Infection Biology, University of Liverpool

ACKNOWLEDGEMENTS

I want to thank everyone that helped and supported me through my PhD experience, and without which this project wouldn't have been possible. I first would like to thank my supervisors Prof. Jonathan Wastling and Dr Kenneth Lamden for their invaluable scientific input and comments throughout the entire project, and also Public Health England for funding this PhD project. A special thank goes to André Charlett of the Public Health England Statistics Unit for his precious advice and comments on statistics and data analysis, and to Dr Nadine Randle for her help in the molecular laboratory. Many thanks also to Caroline Durband and Dr John Cheesbrough of the Microbiology Department at the Lancashire Teaching Hospitals NHS Trust, as well as to the clinical laboratory personnel at the Royal Blackburn Hospitals and the Manchester Royal Infirmary for the collection and storage of stool samples used in this project. Many thanks also to Prof. Sarah O'Brien, Prof. Andrew Fox, Dr Robert Hogg and Dr Michaela Giles for their precious comments and help during the project.

I also would like to thank all the people that contributed to the collection of the epidemiological data and made the *Giardia* case-control study possible: Catherine Platt, Lea Madgwick, Susan Johnson, Jackie Parker, Clare J. Devine and all the rest of the staff of the Cumbria and Lancashire Health Protection Unit for their help with the case notifications and questionnaires; the staff and the children of the Liverpool Medicines for Children Research Network that helped in making the study questionnaire child-friendly; all the environmental health officers, general practitioners and staff of the Health Protection Units, Lancashire and South Cumbria Agency and NHS Manchester for their help in selecting and contacting the cases and the controls.

Last but not least, I am grateful to my family, all my friends and people of the Department of Infection Biology for their invaluable support during these four years.

ABSTRACT

THE EPIDEMIOLOGY AND MOLECULAR EPIDEMIOLOGY OF GIARDIASIS IN NORTH WEST ENGLAND

Corrado Minetti

Giardiasis, cause by the parasitic protozoan *Giardia duodenalis*, is one of the most common infectious gastrointestinal diseases in humans worldwide. However, its true population burden and epidemiology and in particular its zoonotic transmission potential are still poorly understood. Furthermore, *G. duodenalis* is not a uniform parasite but a complex of seven genetic assemblages or cryptic species (named A to G) that infect humans and a variety of domesticated and wild animals, and that can only be distinguished using molecular genotyping methods. Although there is some evidence that the two *Giardia* assemblages infecting humans (namely A and B) may differ in their virulence and major transmission routes, data are still scarce. In the UK, several studies suggested that giardiasis is considerably under-diagnosed and a few data are available on the genetic diversity of the parasite causing infection and disease in this country. We investigated the burden, clinical outcomes, risk factors and molecular diversity of giardiasis in North West England using both a descriptive and analytical approach.

In Chapter 2, we analysed the self-reported clinical and exposure data collected over four years from clinical cases of giardiasis in Central Lancashire, as part of an enhanced surveillance program on the illness. The resulting average disease rate of 22.5 cases/100,000 population was high when compared to the available national figures. Giardiasis was particularly abundant in adults in their 30s and children under five, and the disease rate in males was significantly higher than in females. Furthermore, the clinical picture of the cases confirmed the high morbidity associated with this infection particularly in terms of the length of illness and severity of symptoms. Only 32% of the cases reported foreign travel during the exposure window. The results suggested the presence of a hidden burden of disease in adults and males, and indicated that local transmission of *Giardia* can be more common than expected.

In Chapter 3, we performed a case-control study to determine the significant risk factors for symptomatic giardiasis in North West England, by recruiting clinical cases of *Giardia* and age and sex matched controls from Central and East Lancashire and Greater Manchester. The multivariable logistic regression analysis done on 118 cases and 226 controls revealed that overall travelling abroad (particularly to developing countries) was an important risk factor for the illness (OR 9.59). Following the exclusion of participants that reported foreign travel, four risk factors were significant for the acquisition of giardiasis: going to a swimming pool (OR 2.67), changing nappies (OR 3.38), suffering irritable bowel syndrome (OR 3.66) and drinking un-boiled water from the tap (OR 8.17). The results indicated the important role of swimming pools and contact with children in nappies for the transmission of the parasite.

In Chapter 4, whole faecal DNA was extracted from the faecal samples of the cases part of the surveillance and case-control studies and the *Giardia* assemblages and sub-assemblages causing infection were determined using PCR amplification and DNA sequencing of up to four parasite genes (*beta-giardin*, *glutamate dehydrogenase*, *triose-phosphate isomerase* and *small-subunit ribosomal RNA*). The majority of infections (64%) were caused by assemblage B, followed by assemblage A (33%), whereas mixed-assemblage infections were rare (3%). The majority of the assemblage A isolates belonged to the sub-assemblage AII and showed completed identity with previously described isolates, and six multi-locus genotypes were identified. The level of genetic sub-structuring as revealed by phylogenetic analysis was significantly higher in assemblage B isolates compared with A isolates: a higher proportion of novel assemblage B sequences was detected compared to what was observed in assemblage A isolates. A high number of assemblage B sequences showed heterogeneous nucleotide positions that prevented the unambiguous assignment to a specific sub-assemblage. Up to 17 different assemblage B multi-locus genotypes were found. The molecular genotyping results showed that *Giardia* assemblage B was responsible for the majority of the clinical infections and confirmed the occurrence of a high diversity of parasite multi-locus genotypes.

In Chapter 5, we integrated the epidemiological and the molecular data generated by the enhanced surveillance and case-control studies and we studied the clinico-epidemiological differences between cases infected with *Giardia* assemblage A or B. Our results showed a difference in the age prevalence between the two assemblages, with assemblage A being more common in older cases. Cases infected with assemblage B reported a series of symptoms more frequently than cases infected with assemblage A, as well as reporting a longer illness. Although the exposure profile of the cases largely overlapped between the two assemblages, two different types of exposures were reported more frequently in the two groups of cases: keeping a dog in assemblage A cases and the presence in the household of children and children at nursery in assemblage B cases. The results suggested that assemblage A could have a major zoonotic reservoir, whereas assemblage B could be transmitted more commonly via the human-to-human route.

TABLE OF CONTENTS

AUTHOR'S DECLARATION ii
ACKNOWLEDGEMENTS iii
ABSTRACT iv
TABLE OF CONTENTS vi
LIST OF FIGURES x
LIST OF TABLES xii
LIST OF ABBREVIATIONS xv
CHAPTER ONE: GENERAL INTRODUCTION 1
1.1 INTRODUCTION TO *GIARDIA* PARASITES 2
1.2 BIOLOGY OF *GIARDIA* 3
1.2.1 Morphology and life cycle 3
1.2.2 Host-parasite interactions in *Giardia* infection and disease 6
1.3 *GIARDIA* IN HUMANS 8
1.3.1 Clinical manifestations 8
1.3.2 Laboratory diagnosis and treatment 10
1.3.3 Epidemiology 11
1.3.3.1 Infection rates and disease burden 11
1.3.3.2 Transmission routes and population at-risk 13
1.3.3.3 Risk factors for giardiasis 14
1.3.3.4 Zoonotic transmission of *Giardia* 17
1.4 *GIARDIA* IN DOMESTICATED ANIMALS 19
1.4.1 *Giardia* in pets 19
1.4.2 *Giardia* in livestock 20
1.5 THE BIOLOGICAL AND MOLECULAR VARIATION OF *GIARDIA DUODENALIS* 21
1.5.1 Phenotypic and genomic variation between *Giardia* assemblages 21
1.5.2 *Giardia* molecular genotyping methods: current status and limitations 23
1.5.3 The molecular epidemiology of *G. duodenalis* 27
1.5.3.1 Humans 27
1.5.3.2 Domesticated animals 29
1.5.3.3 Wildlife and non-mammalian species 30

1.6 AIMS OF THE THESIS 31
CHAPTER TWO: ENHANCED SURVEILLANCE STUDY ON THE BURDEN OF GIARDIASIS IN CENTRAL LANCASHIRE 32
2.1 INTRODUCTION 33
2.2 AIMS OF THE STUDY 34
2.3 MATERIALS AND METHODS 35
2.3.1 Study duration and catchment area 35
2.3.2 Cases identification and inclusion 35
2.3.3 Clinico-epidemiological data collection 36
2.3.4 Data cleaning and analysis 37
2.4 RESULTS 39
2.4.1 Notified *Giardia* cases 39
2.4.2 Population rates and seasonality of giardiasis in Central Lancashire 40
2.4.3 Clinical and exposure profile of the cases 45
2.4.4 Comparison of cases with and without a history of travel abroad 50
2.4.5 Comparison of male and female cases exposure history 54
2.5 DISCUSSION 55
CHAPTER THREE: CASE-CONTROL STUDY OF RISK FACTORS FOR GIARDIASIS IN NORTH WEST ENGLAND 63
3.1 INTRODUCTION 64
3.2 AIMS OF THE STUDY 66
3.3 MATERIALS AND METHODS 67
3.3.1 Study duration and catchment area 67
3.3.2 Ethical approval, study conduct and registration 67
3.3.3 Study sample size 68
3.3.4 Case definition and inclusion 68
3.3.5 Control definition and selection 69
3.3.6 Recruitment and questionnaire data collection 69
3.3.7 Data management and quality control 72
3.3.8 Data analysis 72
3.3.8.1 Univariable analysis 72
3.3.8.2 Multivariable analysis and modelling 73
3.4 RESULTS 74
3.4.1 Notified *Giardia* cases 74

3.4.2 Recruited cases 74
3.4.3 Cases self-reported clinical outcomes 76
3.4.4 Household clusters description 77
3.4.5 Recruited controls and health details 79
3.4.6 Case-control matching 79
3.4.7 Eligibility for risk factor analysis 79
3.4.8 Risk factors for giardiasis 81
3.4.8.1 Univariable analysis 81
3.4.8.2 Multivariable analysis of general risk factors for giardiasis 97
3.4.8.3 Multivariable analysis of risk factors for indigenous giardiasis 98
CHAPTER FOUR: THE MOLECULAR DIVERSITY OF *GIARDIA* IN SYMPTOMATIC PATIENTS IN NORTH WEST ENGLAND 107
4.1 INTRODUCTION 108
4.2 AIMS OF THE STUDY 109
4.3 MATERIALS AND METHODS 110
4.3.1 Collection and storage of faecal specimens 110
4.3.2 Faecal DNA extraction 111
4.3.3 Molecular genotyping of *G. duodenalis* 112
4.3.3.1 General PCR reactions protocol 113
4.3.3.2 Amplification of the beta-giardin (*bg*) gene 113
4.3.3.3 Amplification of the glutamate dehydrogenase (*gdh*) gene 114
4.3.3.4 Amplification of the triose-phosphate isomerase (*tpi*) gene 114
4.3.3.5 Amplification of the small-subunit ribosomal RNA (*ssu-rRNA*) gene 115
4.3.3.6 DNA sequencing 117
4.3.3.7 Genotypic and phylogenetic analysis 117
4.3.4 Relationship between *Giardia* cyst load and PCR success 118
4.3.5 Statistical analysis of overall and locus-specific PCR success 119
4.4 RESULTS 120
4.4.1 PCR amplification success 120
4.4.2 Relationship between *Giardia* cyst load and PCR success 122
4.4.3 Overall *Giardia* assemblages prevalence 123
4.4.5 *Giardia* sub-assemblages and genotypes by locus 124
4.4.5.1 Beta-giardin 124
4.4.5.2 Glutamate dehydrogenase 127

4.4.5.3 Triose phosphate isomerase 130
4.4.5.4 Small subunit ribosomal RNA 133
4.4.6 Multi-locus genotypes 134
4.4.6.1 Assemblage A multi-locus genotypes 134
4.4.6.2 Assemblage B multi-locus genotypes 134
4.5 DISCUSSION 136
CHAPTER FIVE: DIFFERENCES IN THE CLINICAL OUTCOME AND TRANSMISSION ROUTES BETWEEN *G. DUODENALIS* ASSEMBLAGES ... 141
5.1 INTRODUCTION 142
5.2 AIMS OF THE STUDY 143
5.3 MATERIALS AND METHODS 144
5.3.1 General statistical methods 144
5.3.2 Analysis outline 144
5.3.2.1 Gender, age and spatio-temporal distribution of assemblage A and B 144
5.3.2.2 Clinical outcomes of patients infected with assemblage A or B 145
5.3.2.3 Exposure profile of patients with assemblage A or B infection and assemblage-specific risk factors 145
5.4 RESULTS 147
5.4.1 Gender and age distribution of *Giardia* assemblages 147
5.4.2 *Giardia* assemblages distribution in Central Lancashire by month of report and local authority 148
5.4.3 Clinical outcomes of patients infected with assemblage A or B 149
5.4.4 Comparison of reported exposures of *Giardia* assemblage A and B cases from the surveillance dataset 154
5.4.5 Risk factors for *Giardia* assemblage A and B infection 156
5.4.6 *Giardia* assemblages and multi-locus genotypes in household clusters 169
5.5 DISCUSSION 171
CHAPTER SIX: FINAL CONCLUSIONS 177
APPENDIX 1: Surveillance study case-case comparisons 182
APPENDIX 2: Case-control study forms 191
APPENDIX 3: Case-control study univariable analysis (variables with $p>0.2$) .. 205
APPENDIX 4: *Giardia* molecular genotyping results 218
APPENDIX 5: Univariable comparisons between assemblage A and B cases 236
REFERENCES 249

LIST OF FIGURES

FIGURE 1.1: MORPHOLOGICAL CHARACTERISTICS OF THE TROPHOZOITE (A) AND THE CYST (B) OF *GIARDIA*. 3
FIGURE 1.2: LIFE CYCLE OF *GIARDIA*..... 5
FIGURE 1.3: TWO SENSITIVE METHODS FOR THE LABORATORY DIAGNOSIS OF GIARDIASIS 11
FIGURE 2.1: DIAGNOSTIC WORKFLOW IMPLEMENTED AT THE ROYAL PRESTON HOSPITAL FOR THE DIAGNOSIS AND SUBSEQUENT INCLUSION OF *GIARDIA* CASES IN THE ENHANCED SURVEILLANCE STUDY. 36
FIGURE 2.2: AGE DISTRIBUTION OF THE 423 CASES NOTIFIED DURING THE ENHANCED SURVEILLANCE PERIOD 39
FIGURE 2.3: DISTRIBUTION BY LEVEL OF NEIGHBOURHOOD DEPRIVATION OF THE 423 CASES NOTIFIED DURING THE ENHANCED SURVEILLANCE PERIOD. 40
FIGURE 2.4: RATES OF GIARDIASIS IN CENTRAL LANCASHIRE AND ITS RESPECTIVE LOCAL AUTHORITIES, 2008-2012. 41
FIGURE 2.5: RATES OF GIARDIASIS BY AGE AND GENDER, CENTRAL LANCASHIRE 2008-2012. 44
FIGURE 2.7: SELF-REPORTED SYMPTOMS IN 238 SYMPTOMATIC CASES THAT ANSWERED THE SURVEILLANCE QUESTIONNAIRE AND REPORTING AT LEAST ONE SYMPTOM. 46
FIGURE 2.8: MONTHLY (A) AND SEASONAL (B) DISTRIBUTION OF CASES THAT DID NOT TRAVEL ABROAD (N=140) AND THOSE THAT TRAVELLED ABROAD (N=63) IN THE MONTH PRIOR TO ILLNESS, 2008-11. 53
FIGURE 3.1: FLOWCHART ILLUSTRATING THE CASE-CONTROL STUDY DESIGN AND IMPLEMENTATION. 71
FIGURE 3.2: GENDER (A) AND AGE (B) DISTRIBUTION OF THE RECRUITED (N=123) AND THE TOTAL NOTIFIED CASES (N=236). 75
FIGURE 3.3: FREQUENCY OF SELF-REPORTED SYMPTOMS IN THE 118 SYMPTOMATIC CASES. 76
FIGURE 3.4: FLOWCHART ILLUSTRATING THE ELIGIBILITY OF CASE AND CONTROL RECORDS FOR THE INCLUSION IN THE RISK FACTOR ANALYSIS. 80
FIGURE 4.1: EXAMPLES OF A FORMED (A), SEMI-FORMED (B) AND LIQUID (C) STOOL SPECIMEN. 110
FIGURE 4.2: *GIARDIA* MOLECULAR GENOTYPING FLOWCHART. 112
FIGURE 4.3: PREVALENCE OF *GIARDIA* ASSEMBLAGES DETERMINED IN 247 SAMPLES. 123
FIGURE 4.4: OVERALL SUB-ASSEMBLAGE TYPING RESULTS AT THE BG LOCUS IN 47 ASSEMBLAGE A (A) AND 108 ASSEMBLAGE B ISOLATES (B)..... 124
FIGURE 4.5: MAXIMUM LIKELIHOOD TREE OF THE *BG* GENE SEQUENCE OF 118 ISOLATES WITHOUT HETEROGENEOUS POSITIONS..... 125
FIGURE 4.6: OVERALL SUB-ASSEMBLAGE TYPING RESULTS AT THE *GDH* LOCUS IN 51 ASSEMBLAGE A (A) AND 95 ASSEMBLAGE B ISOLATES (B).127
FIGURE 4.7: MAXIMUM LIKELIHOOD TREE OF THE *GDH* GENE SEQUENCE OF 115 ISOLATES WITHOUT HETEROGENEOUS POSITIONS..... 128

FIGURE 4.8: OVERALL SUB-ASSEMBLAGE TYPING RESULTS AT THE *TPI* LOCUS IN 52 ASSEMBLAGE A (A) AND 111 ASSEMBLAGE B ISOLATES (B)..... 130
FIGURE 4.9: MAXIMUM LIKELIHOOD TREE OF THE *TPI* GENE SEQUENCE OF 124 ISOLATES WITHOUT HETEROGENEOUS POSITIONS..... 131
FIGURE 4.10: MAXIMUM LIKELIHOOD TREE OF THE *SSU-RRNA* GENE SEQUENCE OF 32 ISOLATES. 133
FIGURE 5.1: AGE PREVALENCE OF ASSEMBLAGE A AND B CASES (N=296).... 147
FIGURE 5.2: DISTRIBUTION BY MONTH (A) AND SEASON (B) OF REPORT OF ASSEMBLAGE A AND B IN 179 GENOTYPED CASES RESIDENT IN CENTRAL LANCASHIRE, 2008-2012. 148
FIGURE 5.3: DISTRIBUTION BY LOCAL AUTHORITY OF ASSEMBLAGE A AND B IN 179 GENOTYPED CASES RESIDENT IN CENTRAL LANCASHIRE, 2008-2012..... 149
FIGURE A: CASE-CONTROL STUDY INVITATION LETTER. 191
FIGURE B: CASE-CONTROL STUDY PARTICIPANT INFORMATION LEAFLET. 192
FIGURE C: CASE-CONTROL STUDY CONSENT FORM AND CASE QUESTIONNAIRE..... 194
FIGURE D: CASE-CONTROL STUDY CONTROL QUESTIONNAIRE HEALTH DETAILS SECTION. 203
FIGURE E: CASE-CONTROL STUDY CONTROL REMINDER LETTER. 204

LIST OF TABLES

TABLE 1.1: EFFECT OF DIFFERENT ENVIRONMENTAL MATRICES AND PARAMETERS ON THE VIABILITY OF *GIARDIA* CYSTS..... 5
TABLE 1.2: MAJOR ROUTES FOR *GIARDIA* TRANSMISSION IN HUMANS. 13
TABLE 1.3: OVERVIEW AND FINDINGS OF MAJOR *GIARDIA* CASE-CONTROL STUDIES. 14
TABLE 1.4: SUMMARY OF THE MAJOR PHENOTYPIC AND GENOMIC DIFFERENCES OBSERVED BETWEEN *GIARDIA* ASSEMBLAGES..... 22
TABLE 1.5: OVERVIEW OF THE HOST OCCURRENCE OF *G. DUODENALIS* ASSEMBLAGES AND SUB-ASSEMBLAGES AS SUMMARIZED FROM FENG AND XIAO (2011) AND RYAN AND CACCIÒ (2013). 30
TABLE 2.1: RATES OF GIARDIASIS IN CENTRAL LANCASHIRE 2008-12.. 41
TABLE 2.2: RATES OF CRYPTOSPORIDIOSIS AND CAMPYLOBACTERIOSIS IN CENTRAL LANCASHIRE, 2008-2012. 43
TABLE 2.3: RATES OF GIARDIASIS BY GENDER AND MALE TO FEMALE RISK RATIOS IN PRESTON, SOUTH RIBBLE AND CHORLEY, 2008-12. 43
TABLE 2.4: SIGNIFICANT DIFFERENCES (P<0.05) IN THE CLINICO-EPIDEMIOLOGICAL CHARACTERISTICS OF *GIARDIA* CASES THAT DID NOT TRAVEL ABROAD (N=162) AND CASES THAT TRAVELLED ABROAD (N=76) IN THE MONTH PRIOR TO ILLNESS AND THAT RETURNED THE SURVEILLANCE QUESTIONNAIRE..... 51
TABLE 2.5: SIGNIFICANT DIFFERENCES (P<0.05) IN THE EXPOSURE HISTORY REPORTED IN THE MONTH PRIOR TO ILLNESS BY MALE (N=138) AND FEMALE (N=108) *GIARDIA* CASES THAT RETURNED THE SURVEILLANCE QUESTIONNAIRE..... 54
TABLE 3.1: REPORTED DURATION (MEDIAN NUMBER OF DAYS) OF THE SYMPTOMS REPORTED IN THE 118 SYMPTOMATIC CASES. 77
TABLE 3.2: CHARACTERISTICS AND EXPOSURE PROFILE OF THE 13 CASES THAT WERE PART OF THE FIVE HOUSEHOLD CLUSTERS IDENTIFIED IN THE CASE-CONTROL STUDY. 78
TABLE 3.3: VARIABLES SHOWING AN ASSOCIATION WITH GIARDIASIS WITH A P-VALUE≤0.2 IN UNIVARIABLE ANALYSIS. 86
TABLE 3.4: CONSTRUCTED MULTIVARIABLE MODEL FOR THE GENERAL RISK FACTORS FOR GIARDIASIS BASED ON 303 VALID (E.G. WITHOUT MISSING DATA) OBSERVATIONS (99 CASES AND 204 CONTROLS). 97
TABLE 3.5: CONSTRUCTED MULTIVARIABLE MODEL FOR THE RISK FACTORS FOR INDIGENOUS GIARDIASIS BASED ON 241 VALID (E.G. WITHOUT MISSING DATA) OBSERVATIONS (69 CASES AND 172 CONTROLS). 99
TABLE 4.1: NUMBER OF *GIARDIA*-POSITIVE FAECAL SPECIMENS BY STUDY, COLLECTION TIMEFRAME AND PRESERVATION METHOD..... 110
TABLE 4.2: GENES USED FOR THE *GIARDIA* MOLECULAR GENOTYPING AND PCR REACTIONS CONDITIONS..... 116
TABLE 4.3: NUMBER AND PERCENTAGE OF PCR-POSITIVE SPECIMENS AT ANY OF THE FOUR LOCI BY YEAR OF COLLECTION, PRESERVATION STATUS

AND AGE (E.G. THE NUMBER OF MONTHS FROM COLLECTION TO DNA EXTRACTION). 120
TABLE 4.4: MULTIVARIABLE LOGISTIC REGRESSION MODEL OF THE VARIABLES INDEPENDENTLY ASSOCIATED WITH THE OVERALL PCR AMPLIFICATION SUCCESS IN 406 EXTRACTED FAECAL SPECIMENS. 121
TABLE 4.5: MULTIVARIABLE LOGISTIC REGRESSION MODEL OF THE VARIABLES INDEPENDENTLY ASSOCIATED WITH THE PCR AMPLIFICATION SUCCESS AT THE *BG*, *GDH* AND *TPI* LOCI IN 406 EXTRACTED FAECAL SPECIMENS. 121
TABLE 4.6: ASSEMBLAGE TYPING RESULTS IN THE SEVEN SPECIMENS WITH A MIXED ASSEMBLAGE A AND B INFECTION. 123
TABLE 4.7: ASSEMBLAGE A MULTI-LOCUS GENOTYPES IN 31 ISOLATES UNAMBIGUOUSLY GENOTYPED AT THE *BG*, *GDH* AND *TPI* LOCI. 134
TABLE 4.8: ASSEMBLAGE B MULTI-LOCUS GENOTYPES IN 45 ISOLATES UNAMBIGUOUSLY GENOTYPED AT THE *BG*, *GDH* AND *TPI* LOCI. 135
TABLE 5.1: COMPARISON OF THE SELF-REPORTED CLINICAL OUTCOMES BETWEEN SYMPTOMATIC PATIENTS INFECTED WITH ASSEMBLAGE A (N=36) AND WITH ASSEMBLAGE B (N=67) FROM THE ENHANCED SURVEILLANCE STUDY. 151
TABLE 5.2: COMPARISON OF THE SELF-REPORTED CLINICAL OUTCOMES BETWEEN SYMPTOMATIC PATIENTS WITH ASSEMBLAGE A (N=28) AND WITH ASSEMBLAGE B (N=57) IN THE CASE-CONTROL STUDY. 152
TABLE 5.3: EXPOSURE VARIABLES DIFFERENT AT $P \leq 0.2$ BETWEEN CASES INFECTED WITH ASSEMBLAGE A (N=39) AND B (N=69) THAT RETURNED THE SURVEILLANCE STUDY QUESTIONNAIRE. 155
TABLE 5.4: VARIABLES SHOWING AN ASSOCIATION WITH A $P \leq 0.2$ WITH INFECTION WITH EITHER ASSEMBLAGE A (N=28 CASES) OR ASSEMBLAGE B (N=57 CASES) COMPARED WITH HEALTHY CONTROLS (N=226) THAT RETURNED THE CASE-CONTROL STUDY QUESTIONNAIRE. 158
TABLE 5.5: CONSTRUCTED MULTIVARIABLE MODEL FOR THE RISK FACTORS FOR *GIARDIA* ASSEMBLAGE A INFECTION BASED ON 243 VALID (E.G. WITHOUT MISSING DATA) OBSERVATIONS (23 ASSEMBLAGE A CASES AND 220 CONTROLS). 168
TABLE 5.6: CONSTRUCTED MULTIVARIABLE MODEL FOR THE RISK FACTORS FOR *GIARDIA* ASSEMBLAGE B BASED ON 257 VALID (E.G. WITHOUT MISSING DATA) OBSERVATIONS (44 ASSEMBLAGE B CASES AND 213 CONTROLS). 169
TABLE 5.7: *GIARDIA* ASSEMBLAGES AND MULTI-LOCUS GENOTYPES FOUND IN 13 CASES PART OF FIVE HOUSEHOLD CLUSTERS. 170
TABLE 2.4.1: COMPARISONS ($P \geq 0.05$) OF THE CLINICAL AND EXPOSURE CHARACTERISTICS BETWEEN THE *GIARDIA* CASES THAT DID NOT TRAVEL ABROAD (N=162) AND THOSE THAT TRAVELLED ABROAD (N=76) IN THE MONTH PRIOR TO ILLNESS AND THAT RETURNED THE SURVEILLANCE QUESTIONNAIRE. 182
TABLE 2.5.1: COMPARISONS ($P \geq 0.05$) OF THE EXPOSURE HISTORY REPORTED IN THE MONTH PRIOR TO ILLNESS BETWEEN MALE (N=138) AND FEMALE (N=108) *GIARDIA* CASES THAN RETURNED THE SURVEILLANCE QUESTIONNAIRE. 186

TABLE 3.3.1: LIST OF VARIABLES THAT RETURNED A *P*-VALUE>0.2 IN UNIVARIABLE ANALYSIS. 205
TABLE 4A: *GIARDIA* REFERENCE SEQUENCES USED FOR SUB-ASSEMBLAGE GENOTYPING. 218
TABLE 4B: LIST OF THE 157 ISOLATES SUCCESSFULLY SEQUENCED AT THE *BG* GENE ORDERED BY SUB-ASSEMBLAGE AND GENOTYPE. 219
TABLE 4C: NUCLEOTIDE POLYMORPHISMS AT THE *BG* LOCUS IN THE ASSEMBLAGE A ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS. 221
TABLE 4D: HETEROGENEOUS POSITIONS IN THE ISOLATE 12/C71 *BG* SEQUENCE. 221
TABLE 4E: NUCLEOTIDE POLYMORPHISMS AT THE *BG* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS AND WITHOUT HETEROGENEOUS POSITIONS. 222
TABLE 4F: NUCLEOTIDE POLYMORPHISMS AT THE *BG* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS AND SHOWING HETEROGENEOUS POSITIONS. 223
TABLE 4G: LIST OF THE 146 ISOLATES SUCCESSFULLY SEQUENCED AT THE *GDH* GENE ORDERED BY SUB-ASSEMBLAGE AND GENOTYPE. 225
TABLE 4H: NUCLEOTIDE POLYMORPHISMS AT THE *GDH* LOCUS IN THE ASSEMBLAGE A ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS. 227
TABLE 4I: NUCLEOTIDE POLYMORPHISMS AT THE *GDH* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS AND WITHOUT HETEROGENEOUS POSITIONS. 228
TABLE 4L: NUCLEOTIDE POLYMORPHISMS AT THE *GDH* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS AND SHOWING HETEROGENEOUS POSITIONS. 229
TABLE 4M: LIST OF THE 163 ISOLATES SUCCESSFULLY SEQUENCED AT THE *TPI* GENE ORDERED BY SUB-ASSEMBLAGE AND GENOTYPE 231
TABLE 4N: NUCLEOTIDE POLYMORPHISMS AT THE *TPI* LOCUS IN THE ASSEMBLAGE A ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS. 233
TABLE 4O: NUCLEOTIDE POLYMORPHISMS AT THE *TPI* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS WITHOUT HETEROGENEOUS POSITIONS. 233
TABLE 4P: NUCLEOTIDE POLYMORPHISMS AT THE *TPI* LOCUS IN ASSEMBLAGE B ISOLATES SUCCESSFULLY SEQUENCED IN BOTH DIRECTIONS AND SHOWING HETEROGENEOUS POSITIONS. 234
TABLE 5A: LIST OF EXPOSURE VARIABLES THAT RETURNED A *P*>0.2 IN THE COMPARISON BETWEEN CASES INFECTED WITH ASSEMBLAGE A (N=39) AND B (N=69) THAT RETURNED THE ENHANCED SURVEILLANCE STUDY QUESTIONNAIRE. 236
TABLE 5B: LIST OF VARIABLES THAT RETURNED A *P*-VALUE>0.2 IN UNIVARIABLE ANALYSIS OF *GIARDIA* ASSEMBLAGE A CASES (N=28), ASSEMBLAGE B CASES (N=57) AND HEALTHY CONTROLS (N=226) THAT RETURNED THE CASE-CONTROL STUDY QUESTIONNAIRE. 239

LIST OF ABBREVIATIONS

%	Percent
$(NH_4)_2SO_4$	Ammonium sulphate
°C	Degree Celsius
µl	Microlitre
µM	Micromolar
µm	Micrometre
ASH	Allele sequence heterozygosity
bg	Beta-giardin
bp	Base pair
CH	Chorley
CI	Confidence interval
CLHPU	Central Lancashire Health Protection Unit
df	Degrees of freedom
DFA	Direct fluorescence antibody
DNA	Deoxyribonucleic acid
EIA	Enzyme immunoassay
ELISA	Enzyme-linked immunosorbent assay
EtOH	Ethanol
g	Gram
gdh	Glutamate dehydrogenase
GMHPU	Greater Manchester Health Protection Unit
GP	General practitioner
HPU	Health Protection Unit
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
IMD	Index of multiple deprivation
KCl	Potassium chloride
kg	Kilogram
KH_2PO_4	Potassium dihydrogen phosphate
LA	Local authority
LaSCA	Lancashire and South Cumbria Agency
mg	Milligram

$MgCl_2$	Magnesium chloride
min	Minute
ml	Millilitre
MLG	Multi-locus genotype
MLST	Multi-locus sequence typing
mM	Millimolar
Na_2HPO_4	Disodium phosphate
NaCl	Sodium chloride
nM	Nanomolar
OR	Odds ratio
p	*p*-value
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pH	Percent hydrogen
PR	Preston
RFLP	Restriction fragment length polymorphism
RR	Risk ratio
s	Second
sd	Standard deviation
spp.	Several species
SR	South Ribble
ssu-rRNA	Small-subunit ribosomal RNA
tpi	Triose phosphate isomerase
UK	United Kingdom
V	Volt
VSP	Variant surface protein
x g	Centrifugal force

CHAPTER ONE: GENERAL INTRODUCTION

1.1 INTRODUCTION TO *GIARDIA* PARASITES

The genus *Giardia* (order Giardiida, subclass Diplozoa, class Trepomonadea, superclass Eopharyngia, subphylum Trichozoa, phylum Metamonada) (Cavalier-Smith, 2003) is a group of flagellated protozoa inhabiting the intestine of numerous vertebrate animals (Roberts *et al.*, 2009). These organisms exist in two distinct forms. The motile trophozoite stage reproduces in the host intestinal tract, whereas the dormant and environmentally resistant cyst stage is disseminated in the environment through the host faeces and is responsible for transmission. Trophozoites possess two identical nuclei, a ventral adhesive disk used to attach to the intestinal epithelial cells and two dark-staining median bodies of unknown function, but they completely lack mitochondria, peroxisomes and the Golgi apparatus. Due to these peculiar characteristics, *Giardia* parasites have been long considered primitive organisms that diverged from the ancestral eukaryotes prior to their acquisition of mitochondria. However molecular data recently confirmed that these parasites are highly derived organisms, and that the loss or alternative evolution of organelles occurred multiple times probably as an adaptation to the parasitic lifestyle (Thompson & Monis, 2012).

The nomenclature and taxonomy of the genus *Giardia* underwent several changes throughout the first half of the 20th century (Thompson & Monis, 2004). Six species are currently recognized based on both morphological and molecular evidences: *G. agilis* is found in amphibians, *G. muris* in mice and other rodents, *G. ardeae* and *psittaci* in birds, *G. microti* in voles and muskrats and *G. duodenalis* in a wide range of mammalian hosts including humans, dogs, cats, livestock and wildlife (Feng & Xiao, 2011). A potentially new *Giardia* species was found in an Australian marsupial, the quenda (*Isoodon obesulus*) (Adams *et al.*, 2004), but its taxonomic status has not been confirmed yet.

The first species to be described and also the most known due to its both medical and veterinary importance is *G. duodenalis* (synonyms *lamblia/intestinalis*). Unless otherwise stated, the term *Giardia* used from this point onwards will refer to *G. duodenalis*.

1.2 BIOLOGY OF *GIARDIA*

1.2.1 Morphology and life cycle

The parasite occurs in two developmental stages: the trophozoite and the cyst (**Figure 1.1**). Trophozoites are usually 12-15 μm long and 5-15 μm wide, with two identical diploid nuclei with a deeply-stained karyosome. The cell has a convex dorsal side and an enlarged anterior and a narrower posterior, giving the trophozoite a peculiar pear-like shape. Cytoskeletal structures include eight posteriorly-oriented flagella arranged in four pairs (anterior, ventral, posterior-lateral and caudal), a pair of claw-shaped median bodies (microtubular structures of unknown function) and a rigid bilobed adhesive disk located at the anterior part of the ventral side of the cell. The trophozoite lacks a typical Golgi apparatus, peroxysomes and mitochondria. A variable number of double-membraned structures called mitosomes are located in the centre and the periphery of the cell, as well as lysosome-like vesicles scattered beneath the cell plasma membrane. The parasite cyst is transparent and ovoid, smooth-walled, 8-12 μm by 7-10 μm in size. When mature and excreted in the host faeces, the cyst contains four tetraploid nuclei and the disassembled components of the flagella and adhesive disk.

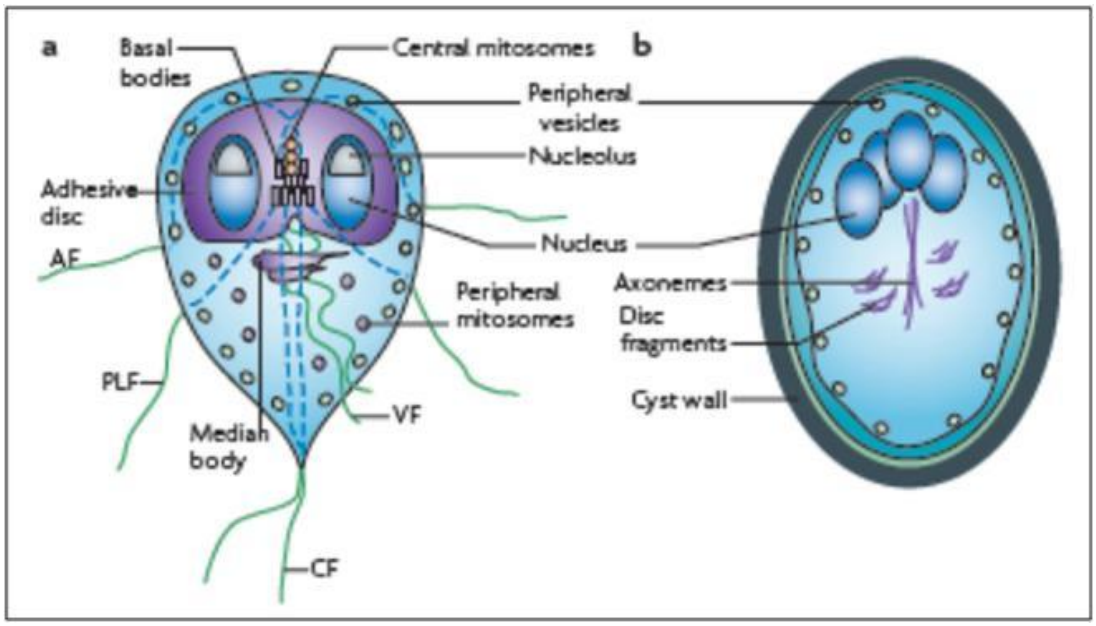


Figure 1.1: Morphological characteristics of the trophozoite (A) and the cyst (B) of *Giardia*. AF=anterior flagella; VF=ventral flagella; PLF: posterior/lateral flagella; CF=caudal flagella (taken from Ankarklev *et al.*, 2010).

Giardia has a direct and simple life cycle (**Figure 1.2**). Infection in the host is caused by the ingestion of the cystic stage. Ingestion of the cyst can occur directly via faeco-oral contact, or indirectly through contaminated drinking water or food. The infectious dose of *Giardia* is low (25-100 cysts), but even ten cysts are sufficient to establish the infection in a human (Rendtorff, 1954). The swallowed cysts pass through the stomach and hatch in the host duodenum. The excystation process is triggered by the host stomach acidic pH (Ankarklev *et al.*, 2010) and gives rise to a short-lived stage of the parasite called the excyzoite, which then divides twice generating four trophozoites. Trophozoites live in the duodenum, jejunum and upper ileum, where they attach to the epithelium and reproduce by binary fission. *Giardia* is an aerotolerant anaerobe, and it uses glucose as the primary substrate for respiration (Adam, 2001). In the host colon when passing faeces start to dehydrate, trophozoites differentiate to form the cystic stage. The encystation process seems to be induced by an alkaline pH and by the presence of high levels of bile and low levels of cholesterol (Lauwaet *et al.*, 2007). The cysts are immediately infectious upon excretion, and up to 300 million cysts can be shed in the stools from an infected person (Roberts *et al.*, 2009). Trophozoites can also be excreted in loose stools (Roberts & Zeibig, 2013), but they cannot survive long outside the host.

While in the environment, *Giardia* cysts are resistant to various stresses and their persistence seems to be affected primarily by temperature, a factor that appeared consistently across different studies and environmental matrices (**Table 1.1**). Most researchers have used *G. muris* in their inactivation studies, and the cysts' viability has been usually assayed either by mice infectivity tests or success of excystation (% cysts hatched) *in vitro*. Cysts tend to retain infectivity for longer in damp and cool environments, particularly if they are in freshwater. On the other side, in matrices that are subject to desiccation (like soil or faeces) the viability can be drastically reduced (Olson *et al.*, 1999). The same study also confirmed that extremely low (i.e. freezing) or warm (25^{o}C) temperatures have a strongly negative effect on cysts infectivity. No data are available about the viability of *Giardia* cysts on dry fomites (like toilet surfaces).

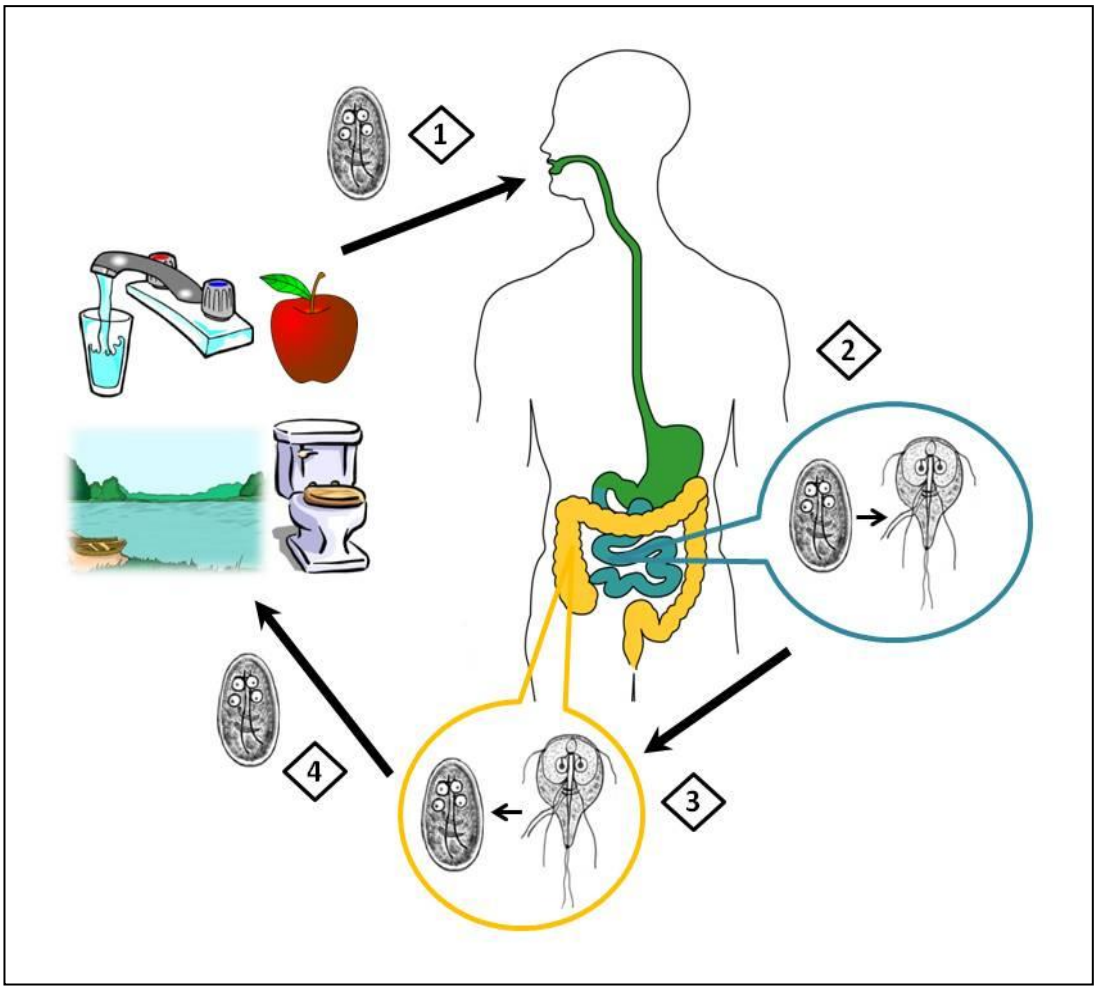


Figure 1.2: Life cycle of *Giardia*. (1) The cyst is ingested by the host; (2) In the small intestine the trophozoite hatches from the cyst and multiplies; (3) In the large intestine the trophozoite encysts; (4) The cysts (and occasionally trophozoites) are shed in faeces and contaminate the environment. Adapted from: http://www.cdc.gov/dpdx/giardiasis.

Table 1.1: Effect of different environmental matrices and parameters on the viability of *Giardia* cysts.

Matrix(ices)	Tested parameter(s)	*Giardia* cyst viability	Reference
Lake/river/tap water	Temperature and water turbidity, hardness, dissolved oxygen	Prolonged (up to 84 days) if temperature <10^{o}C (*)	(DeRegnier *et al.*, 1989)
Seawater	Water salinity, sunlight	Up to 60 days, reduced by increasing salinity and exposure to light (*)	(Johnson *et al.*, 1997)
Distilled water, soil, cattle faeces	Temperature	Reduced in soil and faeces compared to water, longest at 4^{o}C (*)	(Olson *et al.*, 1999)
Cattle slurry	Storage at constant temperature, no light	Reduced after 45 days, nearly complete inactivation at 90 days (#)	(Grit *et al.*, 2012)

*species tested: *G. muris;* #species tested: *G. duodenalis*

The combination of the high numbers of parasites excreted in stools and the resilience of the cysts in the environment leads to high levels of environmental contamination, particularly in fresh waters where *Giardia* cysts can survive for several weeks. The recovery of *Giardia* is common in fresh waters, where this parasite is usually observed more frequently than *Cryptosporidium* (Smith *et al.*, 2006a). The contamination of both surface or ground waters is commonly reported (Castro-Hermida *et al.*, 2014) and it can be the result of either direct faecal pollution from humans and animals, or the result of the input of untreated wastewaters or agricultural runoff from manure-fertilized fields (Plutzer *et al.*, 2010). *Giardia* cysts are commonly recovered in large numbers in untreated wastewaters worldwide (Nasser *et al.*, 2012). In water treatment plants the parasite cysts can be removed from water through flocculation/coagulation and sedimentation followed by filtration, but disinfection by chlorination is also necessary (Betancourt & Rose, 2004). It has been shown that *Giardia* cysts can still be recovered in treated waters (Lobo *et al.*, 2009; Castro-Hermida *et al.*, 2014). Other effective water treatment methods include the use of monochloramine, chlorine dioxide, ozonation and UV light (World Health Organisation, 2008). Boiling water effectively kills the parasite, and portable kits using iodine or chlorine are available to disinfect water for personal use (Hill, 2001).

1.2.2 Host-parasite interactions in *Giardia* infection and disease

The biology of *Giardia* infection in the intestine and the interactions between the parasite and the host have been characterised mostly in rodent models and *in vitro* cell culture systems (Thompson & Monis, 2012)

The pathophysiology of giardiasis is influenced by the interaction between both parasite and host factors (Cotton *et al.*, 2011). These include the parasite genetic makeup, the host age, gender, nutritional and immune status, and the presence of other infections or chronic conditions. The different mechanisms by which *Giardia* causes disease still remain poorly understood. Differently from other gut parasites like *Entamoeba* spp., *Giardia* is not invasive and apparently does not secrete any known toxins (Ankarklev *et al.*, 2010). During infection, the intestinal permeability is increased via both parasite-induced apoptosis of enterocytes (Panaro *et al.*, 2007)

and/or disruption of the junctions between these cells (Maia-Brigagão *et al.*, 2012). Also the shortening of the epithelial brush border microvilli has been observed (Cotton *et al.*, 2011). All these morphological alterations of the gut epithelium modify the mechanisms of water and nutrients absorption, resulting in increased intestinal motility rates, anionic hypersecretion, malabsorption and maldigestion. The malabsorptive syndrome associated with *Giardia* infection has been shown to determine deficiencies of various vitamins (Robertson *et al.*, 2010), as well as of iron and zinc (Abou-Shady *et al.*, 2011). The trophozoite adhesive disk, flagella and variant surface proteins (VSPs) have all been hypothesized as potential virulence factors (Ankarklev *et al.*, 2010). VSPs are glycoproteins that constitute an important mechanism of evasion of the host immune responses (Prucca *et al.*, 2011): by changing their expression on the cell surface some parasites are able to escape from complete clearance by immune cells. When incubated with intestinal cells, *Giardia* trophozoites have been shown to secrete several metabolic enzymes (Ringqvist *et al.*, 2008). Some of these enzymes either exert an anti-inflammatory effect or down-regulate the production of nitric oxide in the epithelium (Ankarklev *et al.*, 2010).

During *Giardia* infection both innate (in the intestinal mucosa) and acquired (in mucosal secretions and serum) immune responses are activated (Roxström-Lindquist *et al.*, 2006). Cellular-based mechanisms play a prominent role in fighting infection, with the involvement of both T-cells and mast cells (Solaymani-Mohammadi & Singer, 2010). The host immune defences play a major role in the severity of disease as well. For example, chronic and prolonged giardiasis is more frequently observed in immunosuppressed rather than immunocompetent patients (John & Petri, 2006; Robertson *et al.*, 2010). Data on the presence of long-term immunity to *Giardia* following infection are scarce. During a waterborne outbreak of *Giardia* at a ski resort in Colorado, people that lived in the area for more than two years had a lower incidence rate of disease compared with short-term residents (Istre *et al.*, 1984). This finding may suggest that repeated exposure to the parasite (for example through drinking water) can lead to some kind of protection against re-infection. In another study, peripheral blood mononuclear cells were isolated from patients that got *Giardia* during a waterborne outbreak five years before and they were exposed to a variety of parasite antigens: the observed proliferation and activation of CD4 T-cells

were significantly higher in the patients group compared to a group of unexposed people (Hanevik *et al.*, 2011).

1.3 *GIARDIA* IN HUMANS

Giardia duodenalis is the most frequently reported enteric parasite in humans in many countries (Olson *et al.*, 2000). Giardiasis is one of the most common infectious gastrointestinal diseases worldwide and it has also been included in the Neglected Disease Initiative by the World Health Organisation since 2004 (Savioli *et al.*, 2006).

1.3.1 Clinical manifestations

Giardiasis typically shows a variable symptomatology, and completely asymptomatic infections are a common occurrence. According to Hill (2001) between 25 and 50% of the exposed individuals develop a clinical diarrhoeal syndrome, whereas the percentage of those expected to become asymptomatic cyst passers is estimated between five and 15%.

The incubation period for clinical giardiasis is usually between six and 15 days following infection (Roxström-Lindquist *et al.*, 2006), and the usual duration of illness (seven to ten days) is relatively long if compared to other etiologic agents of diarrhoea (Hill, 2001). The most commonly reported symptom is always diarrhoea (in nearly 90% of cases) associated with foul-smelling and greasy stools. In non-outbreak cases of giardiasis, intermittent rather than continuous diarrhoea was reported (Cantey *et al.*, 2012). Other frequent symptoms include flatulence, abdominal pain with cramps and anorexia with weight loss (up to four kilograms on average) (John & Petri, 2006). Vomiting and fever can also be associated with giardiasis but less frequently. Furthermore, intestinal malabsorption syndrome is frequently observed in chronic infections (Robertson *et al.*, 2010). Although they are thought to be rare, extra-intestinal manifestations (including eye, skin and joints symptoms) were also reported in nearly 34% of the cases in a study from the United States (Cantey *et al.*, 2012).

Chronic giardiasis produces a symptomatology that is also consistent with irritable bowel syndrome (IBS) (Stark *et al.*, 2007). Such overlap in symptomatology may lead to the misdiagnosis of cases of giardiasis as cases of IBS. A study from Italy showed that 6% of the 137 patients in a clinic and presenting with IBS and dyspepsia were positive for *Giardia* in their stools (Grazioli *et al.*, 2006), with the authors suggesting that some patients diagnosed with IBS may represent actual cases of giardiasis (and so the presence of *Giardia* should be assessed in these patients). However, this study did not include appropriate control patients without IBS to verify whether the prevalence of infection was actually higher in patients with IBS compared to the general population.

The appearance of IBS-like symptoms following infection has been reported in several intestinal pathogens (Stark *et al.*, 2007). A recent study from Norway seemed to suggest an association between acute giardiasis and the development of IBS (and chronic fatigue) following the clearing of infection (Wensaas *et al.*, 2012). In this study the relative risk of developing IBS and chronic fatigue was determined over three years in a cohort of patients that suffered from acute giardiasis during an outbreak, in comparison with a group of unexposed control patients. Results showed that the risk of both IBS and chronic fatigue was significantly higher in the exposed (46.1% frequency of IBS and fatigue) compared to the control group (14% and 12%, respectively) (adjusted relative risk 3.4 for IBS and 4 for chronic fatigue) (Wensaas *et al.*, 2012). Although the study included an appropriately selected control group it is possible that people with a history of giardiasis were more aware of the symptoms and more prone to find them abnormal, resulting in an overestimation of reported IBS and fatigue in this group. Nevertheless, the prevalence of reported IBS in the control group overlapped the frequency of report of this condition in the general population reported in other studies.

Despite the suggestive evidence from the above study, more data are then needed to confirm whether acute giardiasis is a significant risk factor for the development of post-infectious IBS. The pathogenesis of post-infectious IBS in relation to the exposure to the parasite should be evaluated in future studies, as well as the potential role played by co-infections and by the patients' psychological elements.

Furthermore, a strong association of previous *Giardia* infection with both IBS and chronic fatigue was shown in a cohort of patients that were infected during an outbreak in Norway (Wensaas *et al.*, 2012).

1.3.2 Laboratory diagnosis and treatment

The diagnosis of *Giardia* in both humans and animals is traditionally made following the microscopic examination of the stools for ova and parasites (O&P) (Hill, 2001), and is based on the identification of the parasites trophozoites or cysts in wet or stained faecal smears. Since cysts are often excreted sporadically, examination of at least three stool samples from the same patient collected at regular time intervals is recommended (John & Petri, 2006; Manser *et al.*, 2014). When diagnosis is based on microscopy, the examination of multiple stool samples can return a parasite recovery yield of up to 90% compared to 50-70% observed when only one sample is checked (Hill, 2001). Alternatively, trophozoites can be found in the duodenal fluid collected from the patient through intubation or via the string test (Enterotest) (Roberts & Zeibig, 2013). Examination of faeces by direct fluorescence antibody (DFA) microscopy (**Figure 1.3a**) is a good alternative to wet microscopy and can increase the chances of detection of low burdens of infection. This approach has been shown to be more sensitive than traditional microscopy (Garcia *et al.*, 1992).

Various enzyme immunoassays (EIA, ELISA) and non-enzymatic immunochromatographic assays that detect the parasite antigens in the faeces are available (Aldeen *et al.*, 1998) (**Figure 1.3b**). These approaches are more sensitive than microscopy (Mank *et al.*, 1997), and their use is recommended either for routine screening of large number of samples or if low levels of infection or asymptomatic carriage are suspected. In recent years, several real-time PCR assays for the amplification of parasite DNA directly from stools have been developed (Amar *et al.*, 2002; Verweij *et al.*, 2003; Almeida *et al.*, 2010; Calderaro *et al.*, 2010). Real-time PCR assays have been shown to be at least as sensitive as the antigen-detection methods (Verweij *et al.*, 2003) and significantly more sensitive than microscopy (sensitivity 100% compared with 86.7% in microscopy) (Calderaro *et al.*, 2010). These molecular assays recently started to be implemented in clinical laboratories and they have been incorporated into multiplex PCR assays capable of detecting a wide array of intestinal pathogens (not exclusively parasites) in stools (Verweij, 2014).

Treatment of giardiasis relies on the use of nitroimidazolic compounds. Metronidazole is usually considered the drug of choice (John & Petri, 2006), but also other 5-nitroimidazole compounds such as tinidazole, ornidazole or secnidazole are effective (Robertson *et al.*, 2010). Other drugs including benzimidazoles (albendazole, mebendazole), quinacrine, paromomycin or nitazoxanide can be used in the treatment of refractory cases to nitroimidazolic compounds (Robertson *et al.*, 2010).

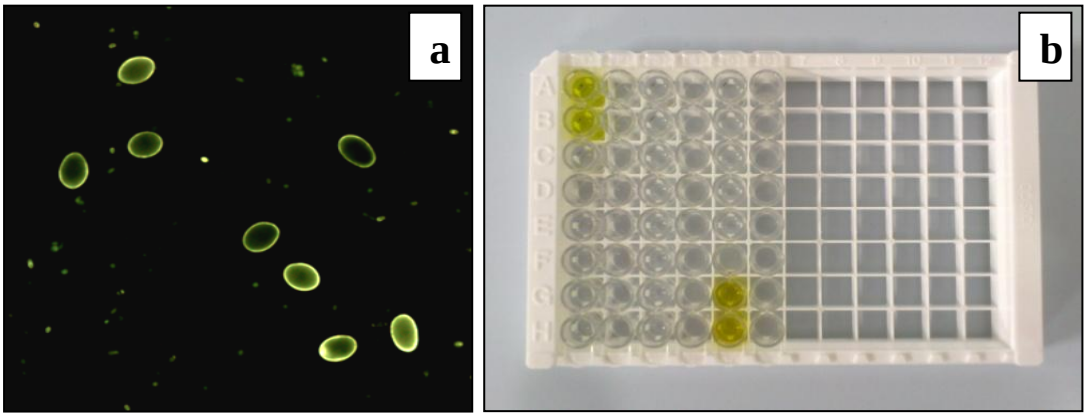


Figure 1.3: Two sensitive methods for the laboratory diagnosis of giardiasis. The direct fluorescence antibody (DFA) microscopy (MERIFLUOR® *Cryptosporidium/Giardia*, Meridian Diagnostics Inc.) (A) and the faecal antigen enzyme immunoassay (GIARDIA/CRYPTOSPORIDIUM CHEK®, Techlab®) (B) (pictures by Corrado Minetti).

1.3.3 Epidemiology

1.3.3.1 Infection rates and disease burden

The burden of *Giardia* in human populations is still poorly understood, and any direct comparison between the results of different studies must take into account whether the presence of the parasite is determined in people with or without symptoms of intestinal disease and also the sensitivity of the diagnostic method(s) used.

The estimated overall prevalence of *Giardia* in human populations is ~20% (4-43%) and 5% (3-7%) in developing and developed countries, respectively (Roxström-Lindquist *et al.*, 2006), and in terms of clinical disease *Giardia* contributes to an

estimated 280 million new symptomatic infections per year worldwide (Ankarklev *et al.*, 2010). In developing countries subclinical infections are frequent. In 274 asymptomatic children aged 5-15 years in Bangkok the prevalence of *Giardia* by microscopy was 12.41% (Popruk *et al.*, 2011), also with peaks of 33% (Laishram *et al.*, 2012). Using real-time PCR, 40.7% of the 108 inhabitants sampled in three villages in Uganda tested positive for *Giardia* (Johnston *et al.*, 2010).

A few studies have investigated the rate of asymptomatic carriage of *Giardia* in developed countries. In 857 symptomatic and 574 asymptomatic patients from the Netherlands the prevalence of *Giardia* determined by microscopy was 5.4% and 3.3%, respectively (de Wit *et al.*, 2001). Another study done in Portugal reported a microscopy prevalence of 3.9% in 177 asymptomatic children below 12 years of age (Almeida *et al.*, 2006). In Melbourne, *Giardia* was found as the most prevalent pathogen (1.6% of 1091 faecal specimens by microscopy) in asymptomatic people randomly selected from the community (Hellard *et al.*, 2000). In the United Kingdom, in the infectious intestinal disease (IID) study the faecal samples from symptomatic cases and asymptomatic controls were screened for various pathogens, and the prevalence of *Giardia* as determined by microscopy and PCR was 1.8% and 1.4% in symptomatic and asymptomatic people, respectively (Amar *et al.*, 2007). The aforementioned studies and the non significant difference in the prevalence of the parasite between symptomatic and asymptomatic patients revealed in the UK IID study suggest that sub-clinical infections may be more common than previously thought. Furthermore the presence of *Giardia* in some cases of symptomatic intestinal disease may not be the direct cause of the symptoms, with the parasite acting only as a transient colonizer of the intestine.

Cases of giardiasis can be significantly underreported. The second infectious intestinal disease (IID2) study estimated a total of 52,434 cases of giardiasis in the UK community only in 2008-2009 (Tam *et al.*, 2012): as a comparison, the estimated number of cases of cryptosporidiosis in the same period was 43,834. Using a capture-recapture technique it was estimated that only half of the actual cases were notified in Auckland (Hoque *et al.*, 2005). Using a meta-analytic approach, it was estimated that in Nordic countries for each *Giardia* case registered in the national registers there are 254-867 undetected cases (Hörman *et al.*, 2004). Several factors can

contribute to the underestimation of the true burden of disease, including the severity of illness, the educational level and behaviour of the patients and the choices of general practitioners (GPs) (Tam *et al.*, 2003). Cases presenting with a mild illness or with a lower educational level are less likely to refer to their GP, and the general practitioner may require *Giardia* testing mostly for cases with a history of foreign travel.

1.3.3.2 Transmission routes and population at-risk

In humans *Giardia* can be transmitted through both direct and indirect routes (**Table 1.2**). These routes are not mutually exclusive, since the spreading of the disease can follow different paths within the same setting or during an outbreak (Katz *et al.*, 2006). As for other enteric infections, factors like poor sanitation, overcrowding and poor water quality are all known to favour the transmission of *Giardia*, particularly in developing countries (Younas *et al.*, 2008).

Table 1.2: Major routes for *Giardia* transmission in humans.

Direct, person-to-person	poor fecal-oral hygiene settings (developing countries, child-care centres, custodial institutions, anal-oral sexual practices)
Indirect, waterborne	untreated drinking water, exposure to recreational waters (lakes, rivers, swimming and wading pools)
Indirect, foodborne	fresh products contaminated by run-off or irrigation water, infected food handler

At temperate latitudes the occurrence of giardiasis usually shows seasonal variability, with the number of cases increasing during summer and early autumn (Gray *et al.*, 1994; Rodríguez-Hernández *et al.*, 1996; Yoder *et al.*, 2010). In some circumstances seasonality is probably due to an increased frequency of outdoor activities during the warm season, with the consequent exposure to contaminated waters. In a study conducted in the United States, a significantly higher proportion of cases were reported from the northern states compared to the south of the country (Yoder *et al.*, 2010). It is possible that the increased cyst viability in colder climates may partially explain the above findings, but differences in the efficacy of case reporting among states cannot be excluded.

Giardiasis also shows a bimodal age-related distribution, with most cases occurring in children under ten (and in particularly in those under four) and adults between 25 and 44 years of age (Gray *et al.*, 1994; Hoque *et al.*, 2002a; Stuart *et al.*, 2003; Ellam *et al.*, 2008; Yoder *et al.*, 2010). This trend may reflect differences in the immune status or in the frequency of exposure to the parasite among different age groups.
Giardia infections are known to be a very common occurrence in day care centres (Steketee *et al.*, 1989; Rauch *et al.*, 1990). In these settings two high-risk groups have been identified, namely children under five years of age and childbearing women (John & Petri, 2006).

Person-to-person transmission has been shown to be favoured also in custodial institutions (Naiman *et al.*, 1980). Sexual transmission of several intestinal parasites is also possible (Abdolrasouli *et al.*, 2009), and *Giardia* can be commonly transmitted in this way particularly among people used to anal-oral sexual practices like homosexual men (Keystone *et al.*, 1980). Interestingly, in several studies the majority of reported cases of giardiasis were males rather than females (Espelage *et al.*, 2010; Yoder *et al.*, 2010).

1.3.3.3 Risk factors for giardiasis

In epidemiology, factors that are associated with an increased risk for a particular disease are usually determined through the implementation of case-control (or cohort) studies. The findings of some major case-control (or cohort) studies done on giardiasis so far are summarized in **Table 1.3.** These studies mostly included only cases of clinical disease, and controls were selected without being checked for the presence of asymptomatic infection. Asymptomatic carriers of the parasite are then either missed or miss-classified as controls. As stated by Espelage *et al.* (2010) regarding their own study, risk factor analysis should be then more appropriately referred to symptomatic giardiasis and not necessarily to parasite transmission. Furthermore, both the significance and the effect size of exposure factors can vary among different studies. Although this may indicate that different populations and communities are exposed to the parasite in different ways, a lack of consistency in published results can also be due to the presence of bias and differences in the study design.

Travel abroad, and in particular to developing countries, is usually one of the most consistent risk factors found (Gray *et al.*, 1994; Hoque *et al.*, 2002b; Faustini *et al.*, 2006; Gagnon *et al.*, 2006). *Giardia* is well-known as one of the major causes of travellers' diarrhoea, being the most common enteric pathogen isolated from tourists showing signs of gastrointestinal disorders at their return (Swaminathan *et al.*, 2009). Almost 85% of the cases of giardiasis reported between 1997 and 2003 in Sweden were not Swedish and they were mainly immigrants, refugees and internationally adopted children (Ekdahl & Andersson, 2005). Geographical regions considered to be at high-risk for *Giardia* are in particular Asia and Africa (Hoque *et al.*, 2002b) and Central and South America (Espelage *et al.*, 2010). In a comprehensive analysis of more than 25,000 cases of travel-related GI infection, the highest rate of giardiasis was found in people returning from South Asia, Middle East and South America (Swaminathan *et al.*, 2009).

Table 1.3: Overview and findings of major *Giardia* case-control studies.

Country and study type	Case selection (no. cases)	Control selection (no. controls)	Significant risk factors by multivariate analysis (OR, 95% CI)*	Reference
United Kingdom, case-control	All ages, with diarrhoea and *Giardia* by microscopy, primary cases only, no travel abroad during exposure window (N=33)	Age/sex matched, from same GP register and living within five miles of the case, no diarrhoea or travel abroad during exposure window (five controls/case, N=112)	Having contact with pets (14.55, 4.18-50.62) and with farm animals (4.77, 1.31-17.38)	(Warburton *et al.*, 1994)
United Kingdom, case-control	All ages, with diarrhoea and *Giardia* by microscopy, primary cases only (N=74)	Age/sex matched, from same GP register of the case (two controls/case, N=108)	Swimming (2.4, 1-6.1), travelling to developing countries (7.6, 0.8-70.1) camping or caravanning (8.4, 0.8-70.1)	(Gray *et al.*, 1994)
United Kingdom, case-control	All ages, with diarrhoea and *Giardia* by microscopy, primary cases and residents only, no travel abroad during exposure window (N=192)	Age/sex matched, from same GP register of the case, no diarrhoea or travel abroad during exposure window (three controls/case, N=492)	Swallowing water while swimming (6.2, 2.3-16.6), having contact with fresh waters (5.5, 1.9-15.9), drinking any additional glass of tap water/day (1.3, 1.1-1.5), eating lettuce (2.2, 1.2-4.3)	(Stuart *et al.*, 2003)
New Zealand (NZ), case-control	Only 15-64 years of age, with diarrhoea and *Giardia* by microscopy, only residents in Auckland (N=183)	Age matched, randomly selected from phone book (two controls/case, N=336)	Drinking water outside NZ (8.78, 3.82-20.16) or from non-mains in NZ (2.1, 1.26-3.49), exposure to child's nappy (7.03, 4.31-11.48), occupational contact with human wastes (5.26, 2.27-12.20), non occupational contact with human wastes (3.33, 1.14-9.74)	(Hoque *et al.*, 2002b)
Canada, case-control	Only ≥1 year of age, with diarrhoea and *Giardia* by microscopy, residents in Quebec only (N=139)	Matched by administrative region and exposure window (plus age matched but only with cases ≤13 years), no diarrhoea and not living in healthcare facilities during study period	In cases 1-13 years: drinking filtered or unfiltered surface or well water (6.13, 1.61-23.32). In cases 14-64 years: drinking untreated water from natural environment (2.46, 1.21-5.02), camping (2.10, 1.06-4.15), travelling to an at-risk country (other than Canada, USA or	(Gagnon *et al.*, 2006)

			Western Europe) (28.68, 9.05-90.85), changing diapers (3.19, 1.61-6.30)	
Germany, case-control	All ages, with diarrhoea/abdominal cramps/bloating and *Giardia* by microscopy/ELISA, only primary cases and no travel abroad (N=120)	Age and county of residence matched, no diarrhoea or travel abroad (two controls/case, N=240)	Male sex (2.5, 1.4-4.4), impaired immunity (15.3, 1.8-127) and eating green salad daily (2.9, 1.2-7.2)	(Espelage *et al.*, 2010)
Cuba, case-control	Only <17 years of age, admitted to hospital and presenting or not diarrhoea and *Giardia* by microscopy (N=94)	Same hospital and admission window matching, without *Giardia* (N=257)	Biting nails (3.2, 1.8-5.7) and eating unwashed vegetables raw (2.9, 1.4-6.2)	(Bello *et al.*, 2011)
Israel, cohort	Children weighing >2500 grams at birth, without conditions affecting normal feeding/growth, enrolled at three months of age and followed up to 18 months (N= 247)	Not applicable	Spring or summer season (5.43, 2.77-10.70), household exposure to cattle/chickens (4.36, 1.62-11.70), having a prior infection with entero-aggregative *E. coli* (1.11), each unit of increase in weight at six months of age	(Coles *et al.*, 2009)

*odds ratio with 95% confidence intervals

Risk factors analysis demonstrated that exposure to human wastes (both occupational and not) and related factors play a major role in the person-to-person and household transmission of the disease. An increased risk of giardiasis has been associated in particular to the presence of children in nappies at home (Hoque *et al.*, 2003) as well as to changing nappies and being a nursing mother (Hoque *et al.*, 2001; Hoque *et al.*, 2002b; Gagnon *et al.*, 2006). These data are indicative that children in nappies are likely to be the major vector for transmission among family members. Occupational exposure to human wastes too has been shown to be a risk factor (Hoque *et al.*, 2002b).

Giardiasis is a typical waterborne infection. This is confirmed by the fact that both drinking or having recreational contact (swimming above all) with potentially contaminated water consistently appear as significant risk factors (**Table 1.3**). *Giardia*, along with *Cryptosporidium* spp., accounts for the majority of parasitic waterborne outbreaks worldwide (Karanis *et al.*, 2007). In developed countries waterborne transmission mainly occurs because of disruptions in the water supply systems, contamination of aquifers with sewage spillages or insufficient water treatment. Contaminated drinking water has been incriminated as the cause of some large community outbreaks in North America (Daly *et al.*, 2010) and Europe (Rimhanen-Finne *et al.*, 2010). For example, during a severe outbreak in Norway a leakage from the sewage system into the lake serving the city centre water supply was incriminated as the possible contaminating source (Nygård *et al.*, 2006). The occurrence of *Giardia* outbreaks related to contaminated municipal waters varies among different countries, probably reflecting differences in the procedures of water treatment and distribution. An increased risk of contamination is also expected to be associated with private rather than public water supplies (Smith *et al.*, 2006b).

Outbreaks of giardiasis due to contaminated water in swimming pools have been frequently reported in the United States (Harter *et al.*, 1984; Porter *et al.*, 1988; Levy *et al.*, 1998) and the UK (Smith *et al.*, 2006b). Transmission in swimming or wading pools is caused by accidental faecal releases coupled with insufficient filtration or disinfection of the pool water. Swimming or bathing in surface waters (like rivers or lakes) also showed to be an important exposure factor (Hoque *et al.*, 2002b; Faustini *et al.*, 2006). Surface waters can be easily contaminated by animal or human faeces. Indeed, the unintentional ingestion of water in these settings has been correlated to

small outbreaks in the United States (Kramer *et al.*, 1996). The risk of *Giardia* transmission in marine settings is still an open question, mainly due to a lack of information (Fayer *et al.*, 2004), although significant amounts of *Giardia* cysts have been sometimes reported in beach waters (Johnson *et al.*, 1995).

Transmission of *Giardia* through contaminated food has been reported too. More precisely, eating raw vegetables (and in particular salad) appeared sometimes as a risk factor (Stuart *et al.*, 2003; Faustini *et al.*, 2006; Mohammed Mahdy *et al.*, 2008; Espelage *et al.*, 2010). Crops can be easily contaminated with *Giardia* cysts through both water runoff from contaminated soils or by the use of unsafe irrigation water (Smith *et al.*, 2007). The presence of parasites has been confirmed in a variety of fresh products (Cook *et al.*, 2007; Robertson *et al.*, 2014). Also shellfish has been hypothesized as a possible transmission vehicle for foodborne protozoa including *Giardia*, especially if the molluscs are eaten raw (Robertson, 2007b). Bivalve molluscs are able to accumulate *Giardia* parasites from water and viable cysts were found in oysters up to two weeks post-exposure (Graczyk *et al.*, 2006). Foodborne outbreaks have also been reported (Smith *et al.*, 2007), and the contamination of food is usually caused by an infected food handler (Mintz *et al.*, 1993). Cases of foodborne giardiasis have been associated to either rodent faeces contaminating a pudding (Conroy, 1960) or to the offal from an infected sheep used to prepare a soup (Karabiber & Aktas, 1991).

Other risk factors associated to giardiasis can be found in the published literature, although less consistently. Outdoor activities like camping or caravanning have been associated to giardiasis (Gray *et al.*, 1994; Gagnon *et al.*, 2006), but they are probably indirectly related to exposure to contaminated water or food.

1.3.3.4 Zoonotic transmission of *Giardia*

The fact that *G. duodenalis* circulates in both domestic and wild animals populations is crucial in respect of the possibility of zoonotic transmission. This parasite has been considered to have zoonotic potential for 30 years (World Health Organisation, 1979). However, the occurrence and frequency of animal-to-human transmission have been largely debated.

The ability of this parasite to infect different hosts has been tested through cross-transmission experiments. These experiments consisted in feeding different animals cysts isolated from patients, to test if they were able to establish an infection. Human parasites have shown the ability of infecting a wide array of mammals in the laboratory (Thompson & Monis, 2004). Furthermore, in the only one experiment of its kind a human volunteer developed infection after swallowing cysts from a rat (Majewska, 1994). Although sometimes the results are not consistent among laboratories and the parasites were not genetically characterized (Thompson & Monis, 2004), these results show that *G. duodenalis* (or at least certain genotypes) can circulate between humans and animals. For example, the possible implication of infected beavers in the spread of waterborne giardiasis has been often hypothesized (Dykes *et al.*, 1980). However, most of the epidemiological reports involved only the finding of infected beavers in source watersheds without any concrete evidence of their involvement in human infections (Xiao & Fayer, 2008).

Some epidemiological evidence of zoonotic transmission is available, but the results are not consistent among different studies or areas. In New Zealand, no correlation between the incidence rate of giardiasis and farm animals density was found (Snel *et al.*, 2009b). However, in a study from Quebec the incidence rate of giardiasis in children below five years of age showed a significant increase with increasing cattle density (Kaboré *et al.*, 2010). Children in a rural village in Thailand were at nearly five time greater risk of getting giardiasis if they had cats at home (Boontanom *et al.*, 2011). A study in a rural area of England found that having contact with pets and farm animals was a significant risk factor for the disease (Warburton *et al.*, 1994), and the presence of cattle and chickens in or nearby the household was associated with higher rates of giardiasis in infants from Israel (Coles *et al.*, 2009). Another study done in temple communities in Bangkok reported the presence of genetically related parasites in humans and dogs (Traub *et al.*, 2009). So far, the only study reporting a more consistent evidence of zoonotic transmission supported by both molecular and epidemiological data is the one by Traub *et al.* (2004) from a tea-growing community in India. In this setting, genetically similar (although not identical) parasites were recovered in dogs and humans living in the same household. Furthermore a highly significant association was found between dog ownership, the presence of an infected dog in the household and the prevalence of *Giardia* in humans.

1.4 *GIARDIA* IN DOMESTICATED ANIMALS

G. duodenalis is a parasite found in a large variety of domestic and wild animals and is usually considered one of the most common intestinal parasites of pets and livestock (Thompson *et al.*, 2008). However, compared to our knowledge of human disease, data about the epidemiology and health impact of *Giardia* infections in animals are relatively scarce. Furthermore, most of the available information comes only from domesticated mammals.

1.4.1 *Giardia* in pets

Giardia infections are usually asymptomatic in dogs and cats (Thompson *et al.*, 2008). Clinical disease is mainly manifested in young animals and in those living in overcrowded kennel or cattery situations (Robertson *et al.*, 2000). The reported disease pre-patent period in dogs and cats varies between 5-16 days, and the most frequently observed symptom is small bowel diarrhoea (Thompson *et al.*, 2008).

Giardia is one of the most commonly reported enteric protozoa in dogs. The prevalence varies greatly among different populations, but most studies showed that it is usually higher in puppies and in kennel or stray dogs (Batchelor *et al.*, 2008; Scaramozzino *et al.*, 2009; Ballweber *et al.*, 2010). In a study from Belgium *Giardia* was found in only 9.3% of household dogs, but in 43.9% of kennel dogs (Claerebout *et al.*, 2009). Results from an extensive study in Europe showed that the prevalence of *Giardia* among owned symptomatic dogs can be up to 24.8%, with significant association between being positive for the parasite and presenting diarrhoea (Epe *et al.*, 2010). In another study a higher prevalence of *Giardia* was observed in diarrhoeic dogs compared to asymptomatic animals, but approximately 12% of screened animals were revealed to be asymptomatic carriers (Scaramozzino *et al.*, 2009).

The prevalence of *Giardia* in cats is usually lower than that observed in dogs, as has been reported in some large surveys (Carlin *et al.*, 2006; Palmer *et al.*, 2008; Tzannes *et al.*, 2008). In a Europe-wide study Epe *et al.* (2010) reported the presence of *Giardia* in approximately 20% of symptomatic owned cats. As observed in dogs, a

significantly higher prevalence of *Giardia* in cats is observed in younger animals (Ballweber *et al.*, 2010).

1.4.2 *Giardia* in livestock

Giardia in cattle is usually considered pathogenic (Taylor *et al.*, 2007), although asymptomatic infected animals are commonly reported. The disease pre-patent period in calves is normally 7-8 days (Thompson *et al.*, 2008). It has been observed that infections in cattle occur more commonly at the end of the neonatal period, and they are often chronic (O'Handley & Olson, 2006). The association of *Giardia* infection with diarrhoea in cattle has been under debate for a long time, mainly due to the largely multifactorial aetiology of this symptom (O'Handley & Olson, 2006). However, the pathogenic power of *Giardia* in cattle has been recently demonstrated since morphological alterations have been recently described in the intestinal epithelium of scouring calves with giardiasis (Barigye *et al.*, 2008). Furthermore, a significant difference in weight gain has been reported in calves treated for giardiasis compared to untreated infected animals (Geurden *et al.*, 2010). The last finding suggests that *Giardia* infection may have important detrimental effects on cattle productivity. *Giardia* is a very common intestinal protozoan in cattle worldwide, with reported prevalences up to 100% in both dairy and beef calves (Thompson *et al.*, 2008). Several studies showed that the average infection prevalence tends to be higher in both pre- (40%) and post-weaned (52%) calves (Trout *et al.*, 2004; Trout *et al.*, 2005) compared to adult cows (~27%) (Trout *et al.*, 2007; Castro-Hermida *et al.*, 2007a).

In sheep and goats *Giardia* is commonly considered non-pathogenic (Taylor *et al.*, 2007), but clinical cases of disease in these animals are reported. The pre-patent period seems to be 6-10 days and 10-21 days for goats and sheep, respectively (Thompson *et al.*, 2008). In goat kids the disease has been associated with apathy, decrease in voluntary food intake and softening of the faeces (Castro-Hermida *et al.*, 2005). Also severe weight loss was recorded in infected lambs during a confirmed *Giardia* outbreak (Aloisio *et al.*, 2006). A study from Spain showed that the prevalence of *Giardia* in sheep and goats (33%) was comparable between the two ruminants (Castro-Hermida *et al.*, 2007b).

The impact of giardiasis in pigs is poorly understood, and affected animals are usually asymptomatic. Few studies have been carried out to evaluate the prevalence of *Giardia* in these animals. It has been shown that weaner pigs tend to be more commonly infected (Maddox-Hyttel *et al.*, 2006) and that prevalence varies among farms with different management systems (Xiao *et al.*, 1994).
The presence of *Giardia* has been reported in horses too (Xiao & Herd, 1994; Olson *et al.*, 1997; Atwill *et al.*, 2000).

1.5 THE BIOLOGICAL AND MOLECULAR VARIATION OF *GIARDIA DUODENALIS*

In the past 20 years it has become apparent that *G. duodenalis* is not a uniform parasite, and it is now recognized that it actually a group of at least seven so called genetic assemblages (Feng & Xiao, 2011). These assemblages (named A to G) are morphologically indistinguishable at both the trophozoite and cyst level, but they differ in terms of host occurrence and for a series of phenotypic and molecular traits. There has been increasing consensus between authors that the assemblages are in fact different (cryptic) species, and the traditional assemblage nomenclature should now be abandoned in favour of a more appropriate classification at the species level (Monis *et al.*, 2009; Thompson & Monis, 2012). However, since the species status of the *G. duodenalis* assemblages has not yet been recognised officially the assemblage nomenclature will still be used throughout this thesis for clarity.

1.5.1 Phenotypic and genomic variation between *Giardia* assemblages

The phenotypic variability present among the parasites classified as *G. duodenalis* was already evident more than 50 years ago, with preliminary experiments showing that different isolates of the parasite greatly differed in both their infectivity and host specificity (Thompson & Monis, 2004). The existence of genetically distinct groups within *G. duodenalis* was then officially demonstrated by using allozyme electrophoresis in the 80s (Andrews *et al.*, 1989), and subsequent studies showed that human isolates belonged to two groups, that were named assemblage A and B (Mayrhofer *et al.*, 1995). Following the analysis of protein polymorphisms at 23

different loci in *Giardia* isolates from humans and animals, the seven assemblages (A-G) were confirmed as well as the genetic sub-structuring of assemblage A and B (both representing a group of four sub-assemblages, named AI-IV and BI-IV) and E (with three sub-assemblages) (Monis *et al.*, 2003 Throughout the years, a large amount of information has been collected about the phenotypic characteristics of *Giardia* assemblages *in vitro* and *in vivo*. However, the studies have been limited so far to a few isolates and assemblages (A, B and E). The major findings from some of these studies are summarised in **Table 1.4**. Also other differences have been reported throughout the years between the assemblages in the trophozoites metabolism, drug sensitivity and susceptibility to a double-stranded RNA virus (Thompson & Monis, 2004).

Table 1.4: Summary of the major phenotypic and genomic differences observed between *Giardia* assemblages

Trait(s) investigated	Finding(s)	Reference
Infectivity of human volunteers	Infection successfully established only by the GS/M (assemblage B) isolate and not by the Isr (assemblage A) isolate	(Nash *et al.*, 1987)
Trophozoites multiplication and virulence in the gerbil intestine	Higher parasite load and pathogenicity (e.g. increased intestinal mucosa damage and inflammation) caused by AI and B compared to AII and E	(Bénéré *et al.*, 2012a)
In vitro susceptibility of trophozoites to nitric oxide (NO)	Higher susceptibility showed by B and E compared to AI and AII	(Bénéré *et al.*, 2012b)
Genomic organisation and gene content	<80% nucleotidic and amino-acidic identity between A and B; differences in *Giardia*-specific gene families (VSPs, NEK kinases and high-cysteine rich membrane proteins) between A, B and E; higher frequency of allelic sequence heterozygosity in B compared to A and E	(Jerlström-Hultqvist *et al.*, 2010a; Jerlström-Hultqvist *et al.*, 2010b)

The genomes of assemblage A (Morrison *et al.*, 2007), B (Franzén *et al.*, 2009) and E (Jerlström-Hultqvist *et al.*, 2010a) have now been fully sequenced. There is significant variation between these three genomes, with important differences observed in terms of chromosomal arrangements, gene content and polymorphisms. As a comparison with other protozoa, the genomic divergence between the assemblages A and B is similar to what is observed between *Theileria parva* and *annulata* and the difference between A and E is very close to what has been detected between *Leishmania major* and *Leishmania infantum* (Jerlström-Hultqvist *et al.*,

2010a). The differences observed in certain gene repertoires, particularly those coding for the trophozoite variant surface proteins (VSPs), are particularly interesting since these molecules are considered as potential virulence factors (Ankarklev *et al.*, 2010) and they are involved in the host immune system evasion (Prucca *et al.*, 2011). It has also been proposed that differences in VSPs expression could expand the infectivity towards new species and so the parasite's host range (Singer *et al.*, 2001). Overall, these findings suggest that the assemblages may interact differently with their hosts and these differences may be the basis of the variation in host occurrence or clinical outcome of infection.

1.5.2 *Giardia* molecular genotyping methods: current status and limitations

In recent years, the use of PCR and DNA sequencing showed that the *Giardia* assemblages show extensive and consistent genetic variation across multiple genes.
PCR-based methods exploiting this sequence variation now allow the determination of the assemblage and sub-assemblage of *Giardia* isolates directly from faeces or environmental samples (Ryan & Cacciò, 2013). Several molecular assays based on PCR amplification and DNA sequencing targeting a wide array of genetic loci have been developed. The molecular markers most commonly used by different authors include the small subunit (18S) ribosomal DNA (*ssu-rDNA*) (Hopkins *et al.*, 1997), the beta giardin (*bg*) (Cacciò *et al.*, 2002), the triose phosphate isomerase (*tpi*) (Sulaiman *et al.*, 2003) and the glutamate dehydrogenase (*gdh*) (Read *et al.*, 2004) genes. For the same markers, also restriction fragment length polymorphism (RFLP) assays have been developed, particularly for the *bg* locus (Cacciò *et al.*, 2002). More recently, assays based on the *tpi* gene that allow a more sensitive detection of mixed-assemblage infections have been developed and successfully applied to both human (Geurden *et al.*, 2009) and animal (Geurden *et al.*, 2008) samples.
Overall, the differentiation of the assemblages using sequence data can be readily achieved across all the aforementioned markers due to the relatively large amount of nucleotidic differences between them, whereas the univocal identification of the sub-assemblages and the subtypes is more difficult due to the low number of variant nucleotides.

There are some important issues that must be considered related to our incomplete knowledge of the genetics of the parasite, the methods employed for genotyping and the nomenclature used to identify the genetic variants found in different studies.

The reliable genetic discrimination between strains or genotypes of a pathogen is a fundamental requisite in order to elucidate its actual host occurrence and transmission routes. As stated by Monis and Andrews (1998), the reliability of a diagnostic marker can be undermined by the inaccuracy of the systematics of the pathogen and by a lack of information on the parasite genetic sub-structuring at the population level. The extent of variability (and rate of evolution) of the chosen marker(s) has a profound influence on the level of genetic heterogeneity that can be detected (Traub *et al.*, 2005a). Although the segregation of the *Giardia* assemblages is clear by comparing the sequence of different genes as shown by phylogenetic analysis, different loci can be informative or not depending on the level of genetic sub-structuring that has to be studied. For example, the *ssu-rDNA* is a highly conserved and slow-evolving locus suitable for studying the phylogenetic relationships between *Giardia* species and assemblages and its multi-copy nature favours its applicability for the sensitive detection of the parasite DNA from faeces or environmental samples (Wielinga & Thompson, 2007). However, it is not applicable as a sub-typing marker due to the absence of sequence variation at this locus between the sub-assemblages (Sprong *et al.*, 2009). Conversely, more variable loci like the *bg*, *gdh* and *tpi* have a higher discriminatory power and they allow the genotyping at the within-assemblage level (Wielinga & Thompson, 2007). However, the level of genetic variability and sub-structuring differ among the assemblages and different loci show a different ability in resolving these differences (Feng & Xiao, 2011): for example, the three major assemblage A sub-assemblages (AI-III) are easily distinguished based on the *gdh* sequence, whereas the AI and AII subgroups do not show the same robust clustering using the *bg* or *tpi* gene sequence. As consistently shown across different loci, the level of genetic sub-structuring is significantly higher within assemblage B than within A. However, there are no DNA sequence polymorphisms that are either specific for the subgroups BIII/IV or consistent across the *bg*, *tpi* and *gdh* loci (Feng & Xiao, 2011). As a result sub-typing assemblage B isolates is usually more complicated and the assignment of an isolate to a particular sub-assemblage is inconsistent in most cases.

Also the use of an imprecise terminology in identifying the parasite isolates identified from an individual host is a common occurrence in the *Giardia* literature (Ryan & Cacciò, 2013), with the term sub-assemblage used sometimes to indicate sequence variants that should be indicated more appropriately as subtypes within a particular sub-assemblage.

Another issue is the reliability of assigning a parasite isolate to a specific assemblage. An unambiguous classification is absolutely essential if the molecular evidence has to be used to support epidemiological data, especially in the discrimination between potentially host-specific and zoonotic genotypes. It has been shown sometimes that different markers assign the same isolate to different assemblages, a phenomenon known as 'assemblage swapping' (Wielinga & Thompson, 2007). Relying on a single locus approach can then lead to different conclusions depending on the gene that has been targeted, and the concomitant presence of multiple parasite assemblages in the sample can be partly responsible for the assemblage swapping phenomenon. The actual prevalence of mixed-assemblage infections can be highly underestimated by the conventional PCR assays that tend to amplify the most abundant parasite population (Ryan & Cacciò, 2013). The use of either assemblage-specific primers (Geurden *et al.*, 2008; Geurden *et al.*, 2009; Levecke *et al.*, 2011) or a sensitive quantitative real-time PCR assay (Almeida *et al.*, 2010) has shown that mixed assemblage infections are more common than previously expected.

It is evident that a multi-locus sequence typing (MLST) approach is more appropriate for *Giardia* molecular epidemiology. The loci that are commonly used for *Giardia* genotyping are all unlinked, making them suitable for a MLST approach (Cacciò *et al.*, 2008). The development of a standardized multi-locus protocol is essential to permit the comparison of sequence data obtained from different studies around the world (Sprong *et al.*, 2009). For this purpose, whenever a MLST protocol for assemblage A isolates is implemented the *Giardia* isolates should be always named using the nomenclature proposed by Cacciò *et al.* (2008): capital letter for the assemblage (A), roman number for the sub-assemblage (AI), and number for the specific sequence type (genotype) (AI-1). In recent years, an increasing number of studies adopting a MLST approach have been published (Geurden *et al.*, 2009;

Lebbad *et al.*, 2010; Beck *et al.*, 2012). However, a consistent nomenclature for assemblage B multi-locus genotypes is still lacking.

The occurrence of heterogeneous templates is commonly reported in *Giardia* genotyping, particularly in assemblage B parasites. Heterogeneous templates consist in sequences containing ambiguous nucleotide positions in their chromatograms (sometimes referred to as sequences with double peaks in the *Giardia* literature).
The detection of mixed templates can be due to the presence of either mixed-assemblage infections (with DNA amplification from genetically distinct cysts in the same sample) or allele sequence heterozygosity (ASH). Allele sequence heterozygosity (i.e. the difference in sequence between alleles of the same gene) is expected to be high in *Giardia*. This is due to the presence of two distinct diploid nuclei accumulating mutations and evolving separately (Adam, 2000; Cacciò & Sprong, 2010). Curiously, ASH at the sequence level has been found to be far lower than expected in the genome of assemblage A (Morrison *et al.*, 2007). On the contrary, a significantly higher level of ASH has been estimated in the assemblage B (Franzén *et al.*, 2009). Allele sequence heterozygosity commonly occurs at the single cell level in assemblage B parasites, as recently demonstrated through genotyping of individual cysts (Ankarklev *et al.*, 2012a).

The occurrence of genetic recombination in *Giardia* has always been largely debated, since the asexual multiplication of *Giardia* should suggest a clonal population structure. However, there are some indirect evidences suggesting the possibility of some sort of sexual recombination in this parasite (Birky, 2010). It is clear that the proven occurrence of recombination would directly impact the current taxonomy of *G. duodenalis*, and more importantly our understanding of its molecular epidemiology (Cacciò & Sprong, 2010). Currently there are no genotyping methods able to consistently discriminate between true recombinants and mixed-assemblage infections, despite the fact that a recent population-based analysis based on an extensive collection of isolates showed that recombination between assemblages is either very rare or absent (Takumi *et al.*, 2012).

To sum up, the results emerging coming from molecular epidemiological studies should be critically evaluated following the limitations described above. The

development of new and more powerful molecular markers to better understand the population genetics of *Giardia* is nevertheless needed. The use of highly variable markers as microsatellites has been increasingly applied to different parasites to understand their population genetic structure (Traub *et al.*, 2005a). For example, this approach has been crucial in understanding the population genetics structure of *Cryptosporidium* spp. (Mallon *et al.*, 2003). A microsatellite assay for *Giardia* genotyping has never been developed so far, probably because these repeated sequences are rarely found in the genome of this parasite (Ryan & Cacciò, 2013).

1.5.3 The molecular epidemiology of *G. duodenalis*

In recent years, more insights in the genetic diversity and host occurrence of *G. duodenalis* assemblages in humans and animals from all around the world have been obtained. These data have been recently and comprehensively reviewed by Feng and Xiao (2011) and Ryan and Cacciò (2013). The host occurrence of the assemblages and major sub-assemblages are reported in **Table 1.5**. Overall, it is clear that some assemblages (e.g. A and B) have a wide host range (encompassing both humans and a variety of animal species) and show zoonotic transmission potential. Conversely, other assemblages (e.g. C to G) seem to be predominantly associated to particular animal taxa although occasional spill-overs to other species may occur.

1.5.3.1 Humans

Worldwide, humans are infected by both assemblage A and B. Overall assemblage B seems to be more prevalent than A worldwide (Sprong *et al.*, 2009; Feng & Xiao, 2011), particularly in symptomatic patients from European countries such as Belgium (Geurden *et al.*, 2009), Germany (Broglia *et al.*, 2013), Sweden (Lebbad *et al.*, 2011) and the UK (Amar *et al.*, 2002; Breathnach *et al.*, 2010; Elwin *et al.*, 2013). However, some studies from both developing and developed countries showed a different picture. For example, assemblage A was found in the majority of the samples of symptomatic children in Saudi Arabia (Al-Mohammed, 2011), asymptomatic people of various ages from both urban and rural areas in Ethiopia (Gelanew *et al.*, 2007) and symptomatic patients from Italy (Cacciò *et al.*, 2002) and Portugal (Sousa *et al.*, 2006). Mixed infections with both assemblage A and B have

been reported but the prevalence greatly varies between studies, ranging from 32.4% in a study from Belgium (Geurden *et al.*, 2009) to only 3% in a study from the UK (Elwin *et al.*, 2013). As highlighted in the previous section, these differences likely reflect the heterogeneity in the molecular typing methods used between studies so the actual prevalence of mixed assemblage infections in humans is likely to be highly underestimated (Ryan & Cacciò, 2013).

Within assemblage A the sub-assemblage AII is the most frequently observed in humans followed by sub-assemblage AI (Feng & Xiao, 2011). The occurrence of these two sub-assemblages has been confirmed also by studies using a multi-locus genotyping approach: it was shown that humans can be infected by a variety of assemblage A multi-locus genotypes but these genotypes mostly belong to the sub-assemblage AII, whereas the occurrence of AI-type genotypes seems limited (Cacciò *et al.*, 2008; Lebbad *et al.*, 2011). Sub-assemblage AIII has not been reported in human samples so far.

Sub-typing of assemblage B isolates revealed a much more complicated picture. The two major sub-assemblages (BIII and BIV) have both been consistently reported amongst humans worldwide (Feng & Xiao, 2011). However, the high level of genetic sub-structuring of assemblage B, the lack of resolution of the current markers and the absence of a clear nomenclature for the multi-locus genotypes of this assemblage do not allow to draw consistent conclusions about the distribution of B sub-assemblages amongst human populations.

Supposedly animal-specific assemblages have been sporadically reported in humans. Assemblage C was found in a human from Egypt by RFLP and sequencing of the *bg* gene (Soliman *et al.*, 2011). Seven people from Ethiopia were found harbouring a mixed assemblage A and F infection by RFLP of the *bg* locus (Gelanew *et al.*, 2007). Also, assemblage E as found in three people from Egypt following sequencing of the *tpi* gene (Foronda *et al.*, 2008). These results may suggest that assemblages other than A or B can occasionally infect humans in certain circumstances (e.g. in areas with either high levels of contact between people and animals or environmental contamination with animal faeces). However in the aforementioned studies the presence of these unusual assemblages was never confirmed following the analysis

of multiple loci, or the genotyping was based on the use of markers showing poor resolution in distinguishing related assemblages such as A and F (Foronda *et al.*, 2008).

1.5.3.2 Domesticated animals

Molecular epidemiological studies on companion animals mostly involved dogs and cats, and in a lesser extent pet rodents and horses. Assemblage C and D are the only or the most commonly found in dogs (Ballweber *et al.*, 2010) and they are often found in combination (Claerebout *et al.*, 2009; Lebbad *et al.*, 2010). Dogs are also infected with assemblage A parasites, and most notably by sub-assemblage AI: in dog specimens from Europe that were diagnosed with assemblage A, AI was found in 73% of the samples compared to AII (27%) (Sprong *et al.*, 2009). Also assemblage B has been occasionally reported in dogs (Ryan & Cacciò, 2013).
Assemblage F is the most frequently isolated assemblage from cats (Ballweber *et al.*, 2010), and cats are also commonly infected with both sub-assemblage AI and AII (Cacciò *et al.*, 2008; Sprong *et al.*, 2009; Lebbad *et al.*, 2011).

Assemblage B appear to be common in pet rodents such as guinea pigs (Lebbad *et al.*, 2010), rabbits (Sulaiman *et al.*, 2003; Lebbad *et al.*, 2010) and chinchillas (Levecke *et al.*, 2011). Horses have been found to be infected with sub-assemblages AI and AII and assemblage B (Traub *et al.*, 2005b; Santín *et al.*, 2013) and also with assemblage E (Veronesi *et al.*, 2010).

Assemblage E is the most prevalent assemblage found in livestock, including cattle, sheep and goats (Trout *et al.*, 2007; Castro-Hermida *et al.*, 2007a; Geurden *et al.*, 2008) and pigs (Armson *et al.*, 2009). The same animals are also commonly infected with assemblage A but predominantly with sub-assemblage AI, and also assemblage B infections have been reported in livestock worldwide (Ryan & Cacciò, 2013). Mixed infections with both assemblage E and A appear to be common in cattle (Geurden *et al.*, 2008), and unusual assemblages including C, D and F have been recently reported in UK livestock (Minetti *et al.*, 2014).

1.5.3.3 Wildlife and non-mammalian species

Although *Giardia* is a common parasite in wildlife (Thompson *et al.*, 2010), a complete picture of the actual host range of *Giardia* assemblages in wild animals is still lacking. Data so far indicate that both assemblage A and B occur in a variety of wild ungulates, wild canids and non-human primates, as well as in marine mammals, rodents and marsupials (Ryan & Cacciò, 2013). Despite the repeated findings of *G. duodenalis* assemblage A, B and also and D in non-mammal vertebrates like aquatic birds or freshwater and fish (Lasek-Nesselquist *et al.*, 2010; Yang *et al.*, 2010; Soares *et al.*, 2011; Ghoneim *et al.*, 2012) there is still disagreement whether these species are natural hosts for the parasite or they simply act as mechanical vectors for the cyst. A new assemblage (named H) was apparently identified in a seal and a gull (Lasek-Nesselquist *et al.*, 2010), but the identification was based on a single locus approach and so it awaits confirmation by sequencing at multiple loci.

Table 1.5: Overview of the host occurrence of *G. duodenalis* assemblages and sub-assemblages as summarized from Feng and Xiao (2011) and Ryan and Cacciò (2013).

Assemblage name	Species name	Major sub-assemblages (if available) and host occurrence	Other host occurrences (sporadic or controversial findings/isolates not typed at the sub-assemblage level)
A	*G. duodenalis*	**AI**: frequent in cattle and other hoofed livestock (pigs, sheep, goats, buffalos), dogs, cats, wildlife; less common in humans	Rodents (mouse, beaver, chinchilla), horses, non-human primates, wild canids (wolf, coyote, red fox), alpacas, ferrets, marsupials, marine mammals (dolphin, seal), birds (gull, eider), fish
		AII: the most frequent sub-assemblage in humans – also reported in cats, livestock, dogs	
		AIII: almost exclusively in wild ungulates (deer, moose); sporadic in cats and cattle	
B	*G. enterica*	The most frequent in humans and non-human primates. Also reported in dogs and both wild (beaver) and pet rodents (guinea pig, rabbit, chinchilla)	Cats, horses, aquatic rodents (nutria, muskrat), hoofed livestock (cattle, pig), wild canids, marine mammals, birds (ostrich), fish (sharks, various freshwater species)
C, D	*G. canis*	The most frequent in dogs and other canids (wolf, coyote)	Mice, cats, marsupials, chinchillas, humans
E	*G. bovis*	The most frequent in cattle and other hoofed livestock (water buffalo, yak) and wild ruminants	Cats, dogs, chinchillas, marsupials, freshwater fish
F	*G. cati*	The most frequent in cats	Pigs, bush rats, humans
G	*G. simondi*	Rat	-

1.6 AIMS OF THE THESIS

Compared to other communicable intestinal diseases, giardiasis has received relatively little public health attention in the United Kingdom. This is largely due to the fact that it is considered a predominantly travel-acquired infection and associated with a low socio-economic burden. More data are needed in order to determine the actual population burden of this infection, particularly in terms of the infections not acquired through travel abroad. Also the assemblages and sub-assemblages of *G. duodenalis* circulating in the UK population have not been characterized extensively: in particular, no data are available on the parasite multi-locus genotypes in the UK population. Furthermore, the existence of preferential transmission routes for the different *Giardia* assemblages has not been tested before in developed countries like the UK.

For these reasons, we decided to investigate the burden, epidemiology and molecular diversity of giardiasis in North West England using both a descriptive and analytical approach.

In this project, we aimed to investigate in particular three aspects. First, we determined the population burden, clinical outcomes and risk factors for giardiasis in North West England through the analysis of surveillance data and by implementing a case-control study. Second, we studied the prevalence and diversity of the *G. duodenalis* assemblages in symptomatic patients and the multi-locus genotypes of the parasite. Third, we explored the presence of potential differences between the parasite assemblages in terms of clinical outcome of infection and transmission routes.

CHAPTER TWO: ENHANCED SURVEILLANCE STUDY ON THE BURDEN OF GIARDIASIS IN CENTRAL LANCASHIRE

2.1 INTRODUCTION

The burden of giardiasis in the United Kingdom is not well understood. Between 2000 and 2010 a total of 37,121 *Giardia* laboratory identifications were reported to the Health Protection Agency in England and Wales (Health Protection Agency, 2011a): most cases were reported in the month of September and were children under the age of five. The incidence of disease seems also to vary geographically within the UK, with the south of England showing the highest rate (Health Protection Agency, 2011a). However, these numbers likely represent only a portion of the actual cases of disease (Tam *et al.*, 2012).

In the UK, faecal specimens are usually tested for *Giardia* on the basis of selective criteria, like age and reported travel, and these criteria may vary between laboratories (Ellam *et al.*, 2008). The presence of travel history in particular is a very common criterion for *Giardia* testing (Health Protection Agency, 2013). As a result not all faecal samples from cases of gastroenteritis in the community are normally tested for *Giardia*, leading to an underestimation of the actual burden of infection in certain age groups (particularly adults) and of infections acquired in the UK and not through foreign travel. Furthermore, diagnosis of *Giardia* in stool samples still relies on the use of relatively insensitive microscopy methods: a recent survey done on 18 diagnostic laboratories the UK and Europe showed that the majority (78%) of them still use light microscopy for the identification of this parasite in stool specimens (Manser *et al.*, 2014).

The impact of using selective criteria for specimen testing and insensitive diagnostic methods on the estimation of the burden of *Giardia* is well depicted by the recent case of the Royal Preston Hospital microbiology laboratory serving the Central Lancashire area, in North West England. In 2002, the hospital replaced conventional microscopy with an enzyme immunoassay (EIA) test for assaying for *Giardia* any faecal specimen submitted by diarrhoeic patients following request of their general practitioner. Following this change, the reported incidence of giardiasis in Central Lancashire showed a significant increase, going from 10.1 cases in 2002 to 33.6 cases/100,000 population in 2006 (Ellam *et al.*, 2008). The increase in the number of *Giardia* cases could not be clearly explained by factors other than improved

detection, since during the same period no variation in the number of reports was observed in North Lancashire and Cumbria where specimens were examined by light microscopy following selective criteria for age and history of travel.

Following these findings, in 2007 an enhanced *Giardia* surveillance program was established in Central Lancashire with the purpose to collect more information on the actual burden and potential transmission routes for giardiasis in the area. The program was established by the Cumbria and Lancashire Health Protection Unit (CLHPU), and undertaken with the collaboration of the three Environmental Health Departments in Central Lancashire (CLEHDs) and the microbiology laboratory of the Royal Preston Hospital. It involved the collection via questionnaire of self-reported clinical and exposure information from laboratory-confirmed clinical cases of giardiasis notified in Central Lancashire and surrounding areas.

2.2 AIMS OF THE STUDY

The aim of this chapter was to analyse the socio-demographic, self-reported clinical and exposure data collected during the enhanced surveillance study to describe the population burden, epidemiological trends and clinical outcome of giardiasis from Central Lancashire, and to generate hypotheses about the major transmission routes for the parasite in the area.

2.3 MATERIALS AND METHODS

2.3.1 Study duration and catchment area

The study started in mid-September 2007 and finished in mid-February 2012 spanning a period of slightly over four years, and it included patients with diarrhoea referring to the Royal Preston Hospital for diagnosis and with a confirmed *Giardia* infection. Patients were mostly resident in Central Lancashire (e.g. local authorities of Preston, South Ribble and Chorley), but occasionally also those coming from other Lancashire LAs and referrals to the Royal Preston hospital for diagnosis were included.

2.3.2 Cases identification and inclusion

For the entire duration of the study, all the diarrhoeal samples submitted by family doctors to the Royal Preston Hospital were screened for *Giardia*. Faecal specimens were first assayed using a monoclonal EIA antigen detection method for the simultaneous detection of *Giardia* and *Cryptosporidium* (*GIARDIA/CRYPTOSPORIDIUM CHEK*®, Techlab). Positive samples (indicating the presence of either of the two parasites) were then confirmed with a *Giardia*-specific immunochromatographic assay (RIDA®QUICK *Giardia*, R-Biopharm). Faecal specimens were tested also for the presence of other intestinal pathogens including *Salmonella*, *Shigella*, *Campylobacter*, *Escherichia coli* 0157 and *Vibrio* spp. (the latter only in case of reported foreign travel). The diagnostic workflow is summarised in **Figure 2.1**. Patients that were found co-infected with either *Cryptosporidium* or at least one of the aforementioned bacterial pathogens were excluded from the study. All the laboratory-confirmed cases of giardiasis were then reported by the hospital microbiology department to both the Central Lancashire Health Protection Unit (CLHPU) and their local Environmental Health Department, as part of routine communicable disease surveillance.

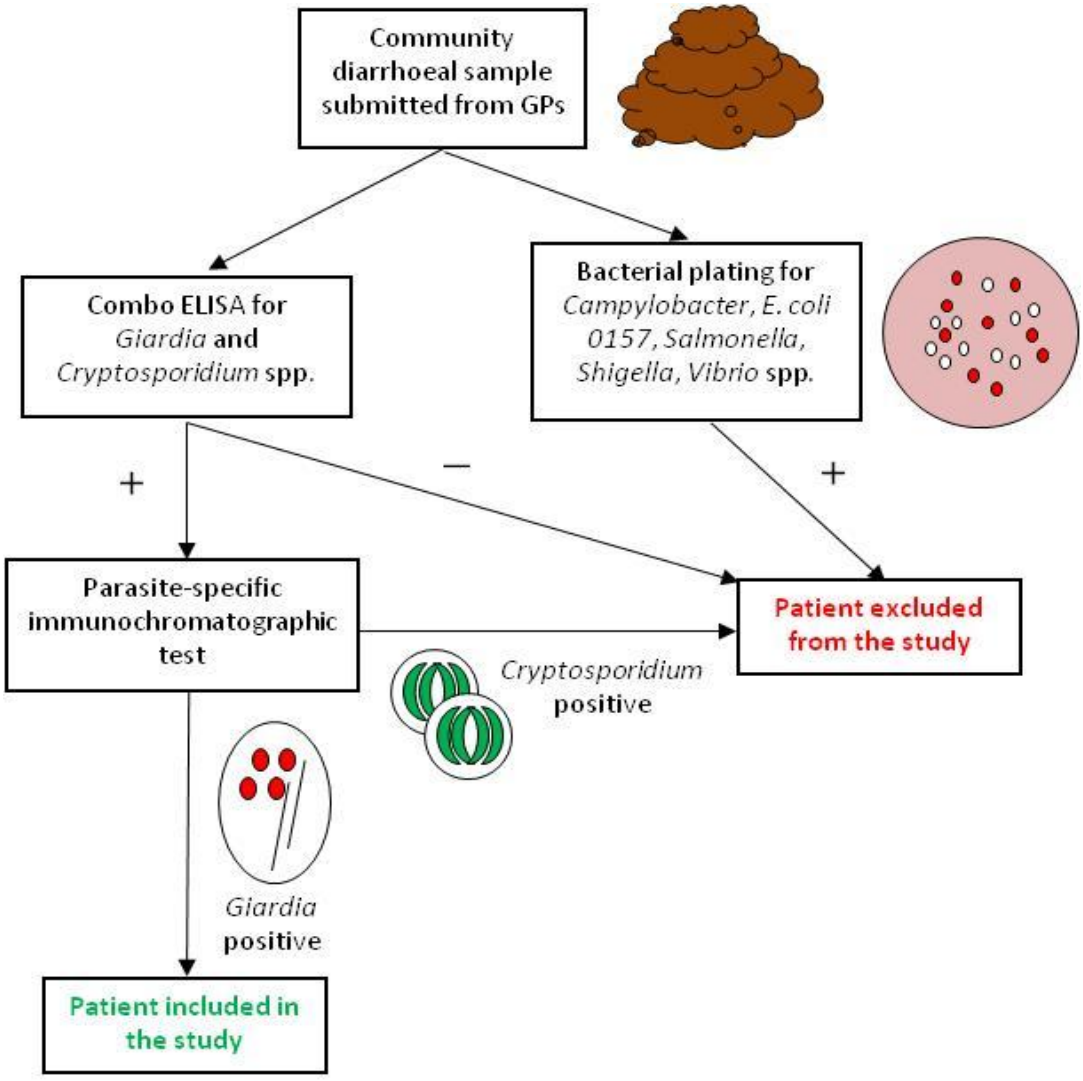


Figure 2.1: Diagnostic workflow implemented at the Royal Preston Hospital for the diagnosis and subsequent inclusion of *Giardia* cases in the enhanced surveillance study.

2.3.3 Clinico-epidemiological data collection

Information about age, gender and postcode was recorded for all the cases. Additionally, an epidemiological questionnaire was administered to the Central Lancashire cases to collect information about experienced symptoms and potential exposure factors. The questionnaire included questions about socio-demographic details (e.g. ethnicity, household composition), self-reported clinical outcomes (e.g. experienced symptoms, illness duration, days of normal activity prevented due to illness etc.) and potential exposures experienced in the month prior to illness

(exposure window) (e.g. travel abroad, water and food consumption, outdoor recreational activities, pet ownership, visits to animal premises etc.). Cases were either sent the questionnaire through post by the CLHPU, or directly contacted by the Environmental Health Department and interviewed face-to-face. Reminders were not sent to the cases that didn't reply. Completed questionnaires were then returned to the CLHPU, and the data were entered by the Unit staff onto the study database on a regular basis. At the end of the study period an Excel copy of the completed study database was sent to the Department of Infection Biology at University of Liverpool for data analysis.

2.3.4 Data cleaning and analysis

Before analysis, data held in the surveillance database were checked for duplicate records (which were deleted) and entry errors and discrepancies (that were identified by checking the questionnaire fields for unusual values and/or comparing related fields in the same section). Wrong values or missing fields were then corrected whenever possible by accessing the original questionnaire (done by the CLHPU staff following request) or by using the information from a related question in the same section. No multiple imputation methods were used.

Statistical analyses were done using both Epi Info™ version 7 (Centre for Disease Control and Prevention, USA) or IBM® SPSS® Statistics 20 (IBM, USA). For disease rates calculation, regional population (mid-2011 resident population estimates) and local authority data were obtained by accessing the website of the Office for National Statistics (Office for National Statistics, 2013a). To check for a potential association between giardiasis and area deprivation, the cases' postcode information was entered onto the Office for National Statistics website (Office for National Statistics, 2013b) to obtain the lower super output area (LSOA) Index of Multiple Deprivation (IMD) 2010. This index is calculated considering 37 different indicators related to seven different dimensions (income, employment, health deprivation and disability, education skills and training, barriers to housing and services, crime and living environment) and it is used to generate an IMD rank. Each of the 32,482 LSOAs within England is classified by level of deprivation according to the IMD rank, spanning from 1 (for the most deprived area) to 32,482 (for the

least deprived). Neighbourhoods were then classified into four categories of deprivation based on the quartiles of distribution: high (IMD rank ≤ 8,120), moderate (8,121≤ IMD rank ≤ 16,241), low (16,242 ≤ IMD rank ≤ 24,362) and very low (24,363 ≤ IMD rank ≤ 32,482).

Descriptive statistics were calculated for all the socio-demographic, clinical and exposure variables. Correlations between continuous or ordinal variables were tested with the Spearman's rank (rho) correlation coefficient, whereas cross-tabulations were used to explore associations between categorical variables. Case-case analysis was performed to compare the characteristics and exposures between different categories of cases. Tests of significance used included the Pearson's Chi-Square (χ^2) (or Fisher's Exact when data were sparse) and the Mann-Whitney's U test, for categorical and continuous variables respectively. Trend analysis was performed using the Chi-Square test for trend. In the analysis some variables were treated both as continuous and categorical, and the results from both were compared and reported. All tests were two-sided and *p*-values of less than 5% were considered to be statistically significant.

2.4 RESULTS

2.4.1 Notified *Giardia* cases

From September 2007 to February 2012, a total of 423 *Giardia* cases were diagnosed at the Royal Preston Hospital and notified to the CLHPU. Thirty-six cases were notified from September to December 2007, 120 in 2008, 91 in 2009, 101 in 2010, 67 in 2011 and eight in the first two months of 2012. Of the 423 cases notified, 358 (84.6 %) were resident in Central Lancashire, with 163 cases from Preston, 120 from South Ribble and 75 from Chorley. The remaining 65 cases (15.4 %) came from other local authorities in Lancashire including Wyre (32 cases), Ribble Valley (21), Fylde (seven), Blackburn with Darwen (two) and Pendle, Lancaster and West Lancashire (one each). No outbreaks were detected during the study period (Ken Lamden, personal communication).

Of the 423 cases, males were significantly more frequent than females (255 males, 60% and 168 females, 40%) (Pearson's χ^2, p<0.001). The age distribution of the cases was bimodal (**Figure 2.2**), with the higher number in adults 30 to 40 years of age. Children below five represented around 10% of the cases. Young people from five to 19 years of age had the lowest proportion of cases.

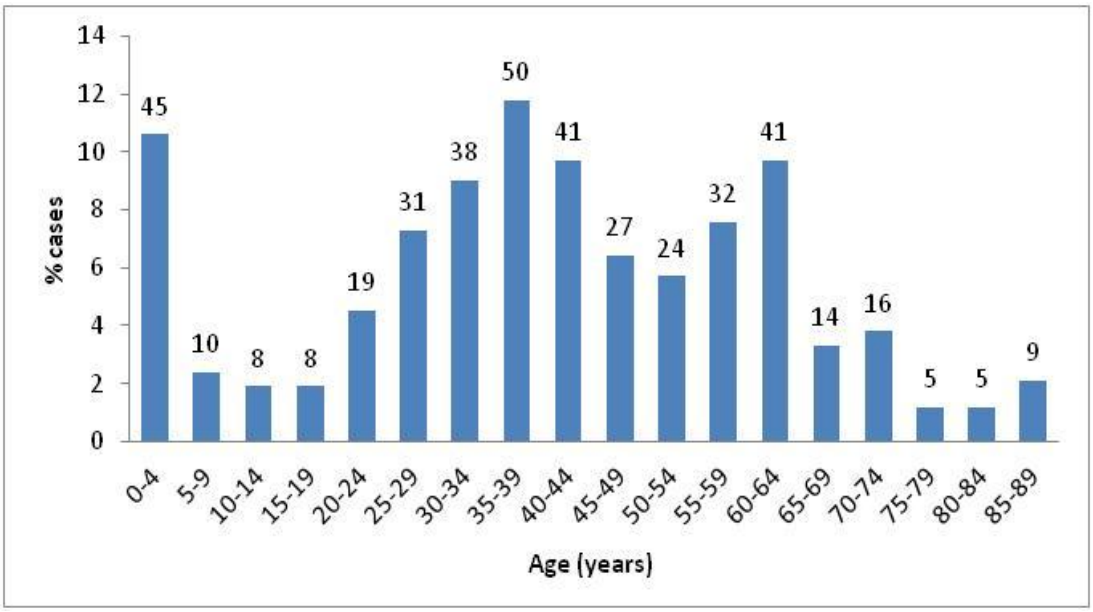


Figure 2.2: Age distribution of the 423 cases notified during the enhanced surveillance period. The number of cases is reported above the bars.

By considering the neighbourhood deprivation, the median neighbourhood IMD rank of the 423 cases was 19,819 (range 79 to 31,828). The cases were not equally distributed in terms of their neighbourhood deprivation level (Pearson's χ^2, p<0.001), with cases living in the least deprived areas being more frequent (143 over 423, 33.8%) (**Figure 2.3**).

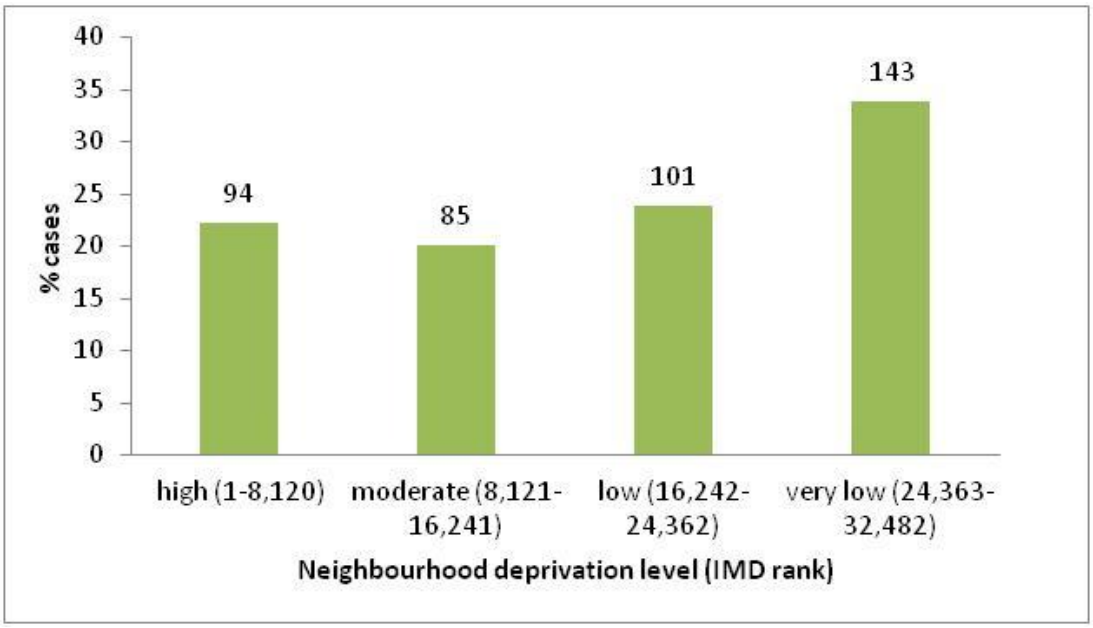


Figure 2.3: Distribution by level of neighbourhood deprivation of the 423 cases notified during the enhanced surveillance period. The number of cases is reported above the bars.

2.4.2 Population rates and seasonality of giardiasis in Central Lancashire

The population rates and seasonality of giardiasis were determined only for Central Lancashire since faecal specimens from this area were exclusively referred to the Royal Preston Hospital, whereas those from other areas were not and so they did not represent the total number of cases in their respective LAs. Data were also available on 86 *Giardia* cases resident in Central Lancashire and notified during 2012. These cases were included in the disease rates analysis to cover a time period of five full years (2008-2012). The estimated rates of giardiasis in Central Lancashire for the period 2008-2012 are shown in **Table 2.1**.

Table 2.1: Rates of giardiasis in Central Lancashire 2008-12. Rates were calculated using as denominator the estimated resident population (in thousands) at mid-2011 (Preston: 139,800; South Ribble: 109,200; Chorley: 107,800; Central Lancashire: 356,800) (Office for National Statistics, 2013a).

	Rate/100,000 population (no. cases)			
Year	**Preston**	**South Ribble**	**Chorley**	**Central Lancashire**
2008	35.8 (50)	32.1 (35)	16.7 (18)	**28.9 (103)**
2009	30 (42)	19.2 (21)	13 (14)	**21.6 (77)**
2010	26.5 (37)	31.1 (34)	13 (14)	**23.8 (85)**
2011	17.9 (25)	16.5 (18)	12.1 (13)	**15.7 (56)**
2012	18.6 (26)	30.2 (33)	19.5 (21)	**22.4 (80)**
Cumulative rate	**128.8 (180)**	**129.1 (141)**	**74.2 (80)**	**112.4 (401)**
Mean annual rate (±sd)	***25.8 (±7.6)***	***25.8 (±7.4)***	***14.8 (±3.1)***	***22.5 (±4.7)***

Overall the rate of giardiasis in Central Lancashire varied between years and generally decreased over time, although not reaching levels of conventional statistical significance (χ^2 for trend 3.79, p=0.052) (**Figure 2.4**). Considering each local authority separately, the disease rate significantly decreased over time only in Preston (χ^2 for trend 10.39, p=0.001) but in neither South Ribble (χ^2 for trend 0.20, p=0.655) nor Chorley (χ^2 for trend 0.93, p=0.335).

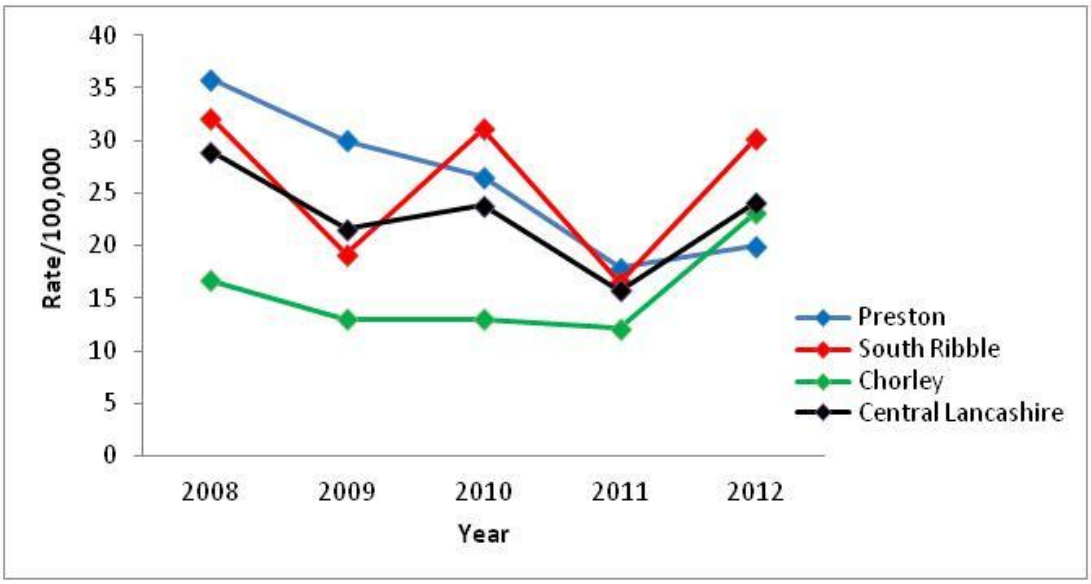


Figure 2.4: Rates of giardiasis in Central Lancashire and its respective local authorities, 2008-2012.

In order to verify whether the observed variation in the rates of giardiasis followed a pattern similar to other intestinal pathogens diagnosed in the area within the same diagnostic workflow, available data on the number of cases of cryptosporidiosis and campylobacteriosis in Central Lancashire residents diagnosed by the Royal Preston Hospital in the years 2008-2012 were analysed. The rate of cryptosporidiosis showed a significant increase over time in the whole Central Lancashire (χ^2 for trend 6.36, p=0.012). Considering the three local authorities separately, the increasing trend in the rate from 2008 to 2012 was significant in Chorley (χ^2 for trend 14.97, p<0.001) and marginally significant in South Ribble (χ^2 for trend 3.66, p=0.056), but there was no significant trend in Preston (χ^2 for trend 1.93, p=0.164). The rate of campylobacteriosis significantly increased in the whole Central Lancashire from 2008 to 2012 (χ^2 for trend 10.15, p=0.001), but the increasing trend was significant in South Ribble (χ^2 for trend 5.40, p=0.020), marginally significant in Chorley (χ^2 for trend 3.43, p=0.064) and there was no significant trend in Preston (χ^2 for trend 1.94, p=0.164).

The rate of giardiasis varied between the three local authorities within Central Lancashire, with the highest and the lowest disease rates observed in Preston and Chorley respectively (**Table 2.1**). The rate of disease was significantly higher in Preston compared to Chorley (Risk ratio RR 1.53, 95% CI 1.19-1.97, Mantel-Haenszel χ^2 corrected for study year: p=0.001) and in South Ribble compared to Chorley (RR 1.93, 95% CI 1.35-2.75, Mantel-Haenszel χ^2 corrected for study year: p<0.001), whereas it was not significantly higher in Preston than South Ribble (RR 1.28, 95% CI 0.92-1.79, Mantel-Haenszel χ^2 corrected for study year: p=0.164). For comparison, the rates of cryptosporidiosis and campylobacteriosis in the three local authorities of Central Lancashire for the years 2008-2012 were also calculated. The rate of cryptosporidiosis was significantly higher in Chorley than in Preston, whereas the opposite was observed for the rate of campylobacteriosis (**Table 2.2**).

Table 2.2: Rates of cryptosporidiosis and campylobacteriosis in Central Lancashire, 2008-2012. Rates were calculated using as denominator the estimated resident population (in thousands) at mid-2011 (Preston: 139,800; South Ribble: 109,200; Chorley: 107,800; Central Lancashire: 356,800) (Office for National Statistics, 2013a).

	Cumulative rate/100,000 population (no. cases)			
Pathogen	**Preston PR**	**South Ribble SR**	**Chorley CH**	**LA vs LA risk ratio (95%CI, *p*-value)***
Cryptosporidium	100.9 (141)	101.6 (111)	128.9 (139)	PR vs SR: 0.922 (0.77-1.27, p=1.000) SR vs CH: 0.79 (0.61-1.01, p=0.070) PR vs CH: 0.78 (0.62-0.99, p=0.045)
Campylobacter	595.1 (832)	584.2 (638)	554.7 (598)	PR vs SR: 1.06 (0.94-1.20, p=0.359) SR vs CH: 1.09 (0.92-1.29, p=0.361) PR vs CH: 1.18 (1.01-1.38, p=0.041)

*Mantel-Haenszel χ^2 corrected for study year

The rate of giardiasis in Central Lancashire was significantly higher in the male population (141.8 cases/100,000 males) compared to the female population (86.5 cases/100,000 females) (RR 1.64, 95% CI 1.34-2.00, Mantel-Haenszel χ^2 corrected for local authority: p<0.001). Rates were consistently and significantly higher in males in all the three local authorities (**Table 2.3**).

Table 2.3: Rates of giardiasis by gender and male to female risk ratios in Preston, South Ribble and Chorley, 2008-12. Rates were calculated using as denominator the estimated resident male and female populations (in thousands) at mid-2011 (Preston: 70,200 males, 69,600 females; South Ribble: 53,400 males, 55,800 females; Chorley: 54,100 males, 53,700 females) (Office for National Statistics, 2013a).

Local authority	**Rate/100,000 male population (no. cases)**	**Rate/100,000 female population (no. cases)**	**Male to female risk ratio (95% CI, *p*-value)**
Preston	152.4 (107)	107.8 (75)	1.41 (1.05-1.90, p=0.021)
South Ribble	170.4 (91)	89.6 (50)	1.90 (1.35-2.68, p<0.001)
Chorley	99.8 (54)	55.9 (30)	1.79 (1.14-2.79, p=0.012)

The rate of giardiasis in Central Lancashire showed a bimodal distribution in relation to age, with the largest rate observed in children below five years and another one in people in their 30s (**Figure 2.4**). The lowest rates were observed from five to 19 years of age. In adults 30 years of age or older the rate decreased with increasing age. Although case males were more frequent than females across all age groups, the rate of giardiasis was significantly higher in males in those aged 10-19y (RR 4.16, 95% CI 1.18-14.58, p=0.022) and in adults in their 40s (RR 2.43, 95% CI 1.48-3.98, p<0.001) and 50s (RR 2.32, 95% CI 1.24-4.34, p=0.008) (**Figure 2.5**). There was

only weak evidence to suggest that the rate in males was higher than the rate in females in children below five years (RR 1.89, 95% CI 1.02-3.52, p=0.051).

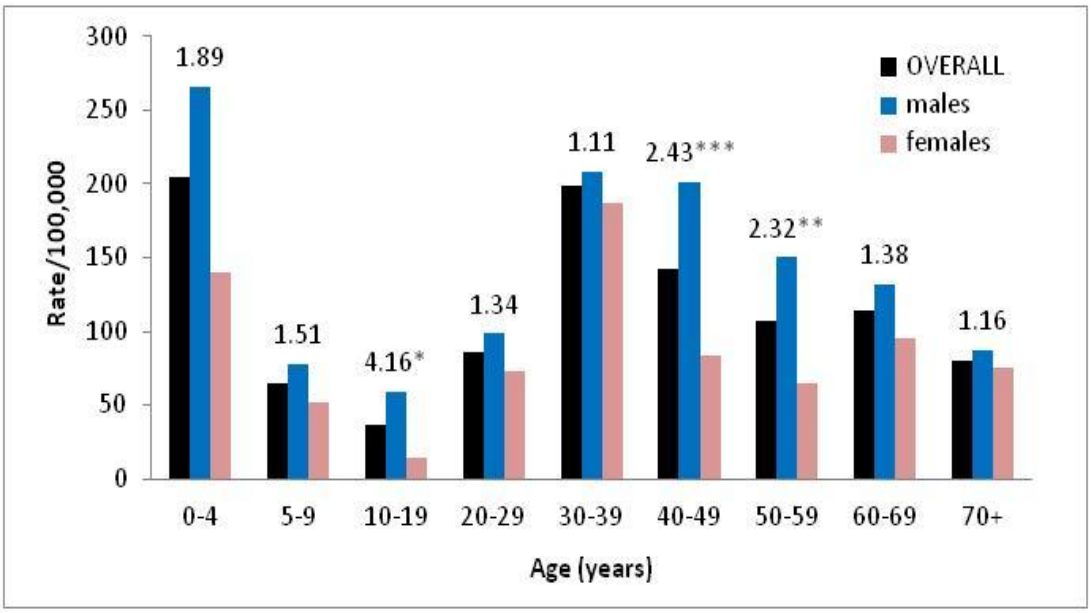


Figure 2.5: Rates of giardiasis by age and gender, Central Lancashire 2008-2012. The male to female risk ratio (in italics) is reported above the bars representing the rates by age group, with the asterisks indicating the presence of a significant difference in the rates between genders (*p<0.05; **p<0.01; ***p<0.001).

The seasonality of giardiasis was explored by using both the month of report of the cases and the estimated month of exposure, which was calculated as the month of report minus one month (e.g. taking into account the disease incubation period, the time passed from the start of illness to the consultation of the GP and the subsequent specimen testing). By considering the total number of cases reported in Central Lancashire from 2008 to 2012, the proportion of cases did not significantly differ between months (Pearson's χ^2, df=11, p=0.338) and no trend in the number of reported cases was observed from January to December (correlation between number of cases and month of report: Spearman's rho=0.415, p=0.180). The number of reported cases peaked in the month of September (corresponding to August month of exposure) and also in November (October month of exposure), and a smaller peak was observed in April (March month of exposure) (**Figure 2.6a**).

By grouping the cases in four seasons, a significant trend with an increasing number of cases reported from Winter to Autumn was observed (correlation between number of cases and season of report: Spearman's rho=0.994, p=0.006). The trend was less

apparent when the season of exposure was considered (Spearman's rho=0.602, *p*=0.398), with the number of cases peaking during the Summer (**Figure 2.6b**).

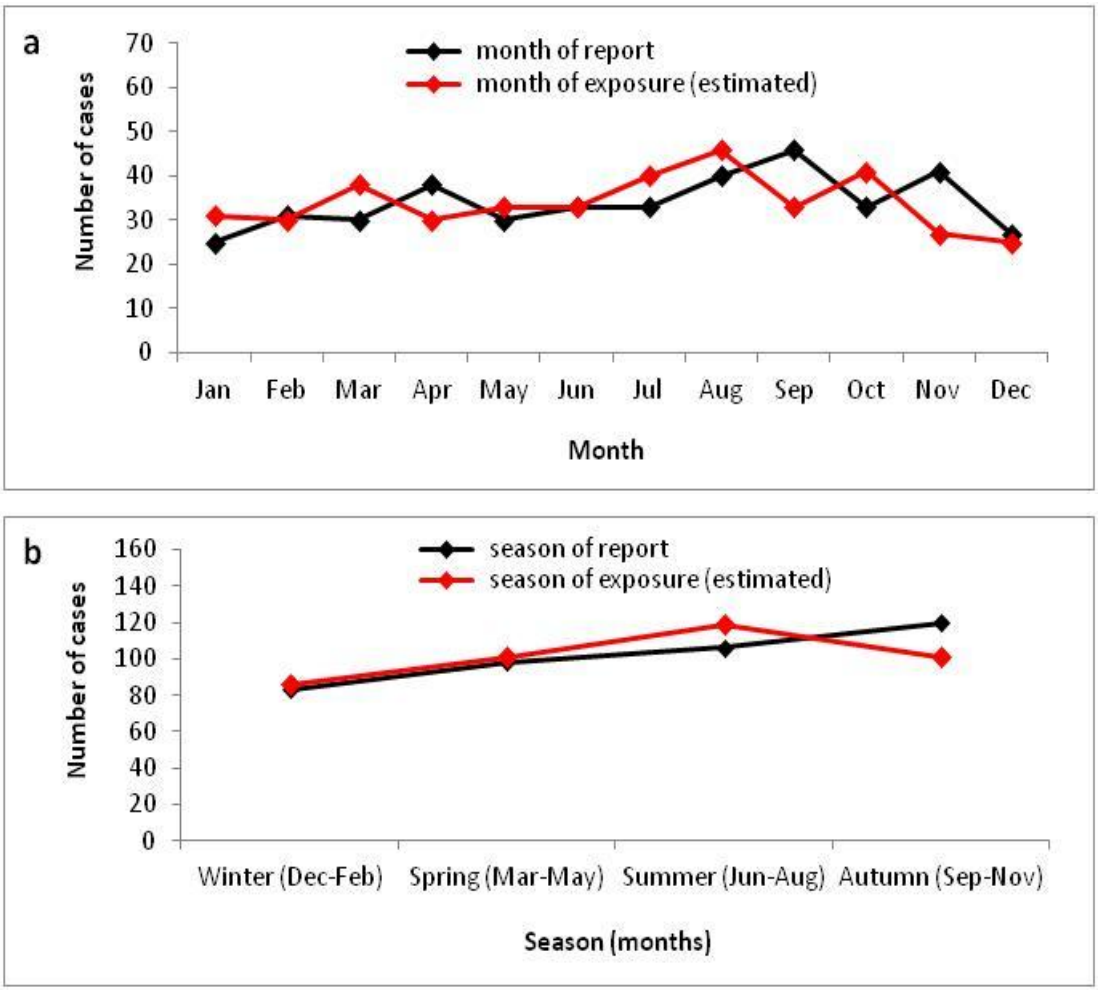


Figure 2.6: Cumulative number of cases of giardiasis reported per month of report (and estimated month of exposure) (a) and by season of report (and estimated season of exposure) (b) in Central Lancashire, 2008-12 (n=407).

2.4.3 Clinical and exposure profile of the cases

Of the 423 cases that were notified and contacted during the surveillance study period (September 2007-February 2012), 246 (58.1% response rate) returned the study questionnaire. Response rates varied according to both gender and age, with males having a significantly lower response rate (138 returned over 255 male cases, 54.1%) compared to females (108 returned over 168 female cases, 64.3%) (Pearson's χ^2, p=0.038). Response rates significantly varied between age groups (Fisher's Exact χ^2, p=0.003): the highest response rate was observed in cases over 60 years of age

(69 out of 90, 76.7%) and the lowest in young adults from 20 to 39 years (65 out of 138, 47.1%), with similar response rates observed in the 40-59 years age group (72 out of 124, 58.1%), 5 to 19 age group (15 out of 26, 57.7%) and in children below five (25 out of 45, 55.6%). The reported clinical and exposure profile of the cases was described. The number of responders to each question varied and only those responding have been used in the analysis. The denominator of the percentages presented provides the number of responders.

Self-reported clinical outcomes

Diarrhoea was the most commonly reported symptom, followed by abdominal pain, vomiting and fever (**Figure 2.7**). The presence of blood in the stools was reported by only 10.5% of the cases.

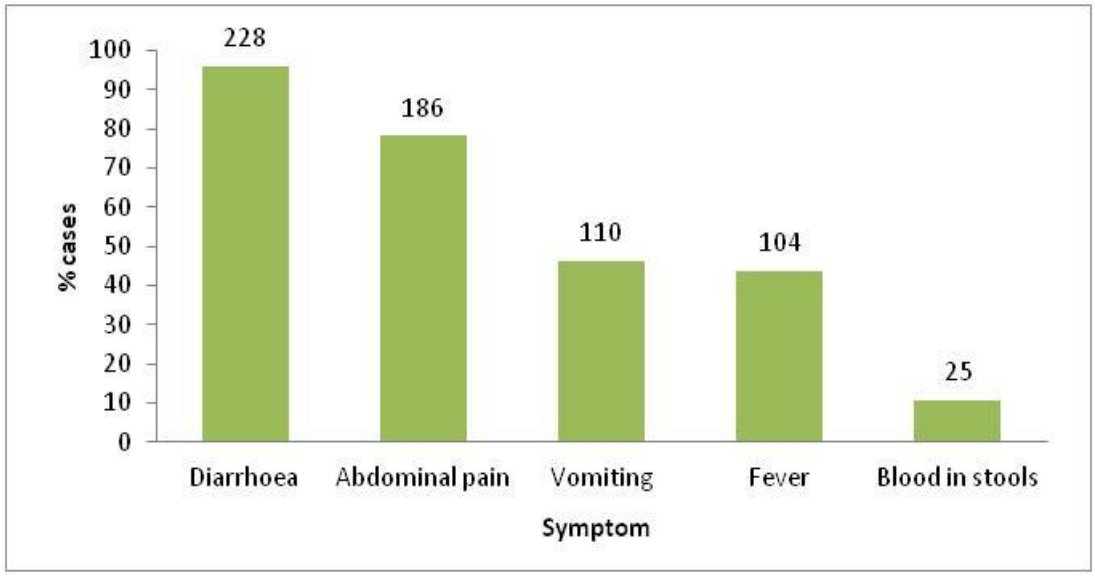


Figure 2.7: Self-reported symptoms in 238 symptomatic cases that answered the surveillance questionnaire and reporting at least one symptom. The number of cases is reported over the bars.

There were no significant differences in the reported symptoms between males and females. The median age of cases that reported experiencing vomiting or fever was significantly lower than the age of those that did not (38 vs 46 years for vomiting, Mann-Whitney U, p=0.026; 40 vs 44.5 years for fever, Mann-Whitney U, p=0.032).

The median number of different symptoms experienced at the same time was three (range one to five). By looking at symptoms that were reported at the same time,

reporting fever was found to be significantly and positively associated with both reporting diarrhoea (Fisher's Exact, p=0.046) and vomiting (Pearson's χ^2, p=0.009). The number of reported symptoms was not correlated to the age of the cases (Spearman's rho=-0.110, p=0.091), and there was no significant difference in the number of reported symptoms between males and females (Mann-Whitney U, p=0.473).

Of the 234 cases that answered the question, 69 (29.5%) reported that they still felt ill when they filled the questionnaire. Of the 148 cases that reported they were no longer ill, the median length of illness was 17.5 days (range one to 180). There was no correlation between the length of illness and the age of the cases (Spearman's Rho = -0.140, p=0.074), but the median length of illness was higher in females compared to males (21 days in females and 14 days in males, Mann-Whitney U, p=0.003).

Information about the number of days of normal activity prevented due to illness was available from 117 cases, which reported a median of seven days (range none to 90). There was no correlation between the number of days of normal activity prevented and the age of the cases (Spearman's rho = -0.011, p=0.907), but the median number of days of activity prevented was higher in females compared to males (10 days in females and five days in males, Mann-Whitney U, p=0.015).

Out of 219 cases that answered the question, 23 (10.5%) reported being admitted to hospital as a result of their illness. Their median age was 63 years (range one to 89) and 18 of them (78.3%) were males. In 15 cases that specified also the duration of hospitalisation, the median number of days was five (range one to 47). Both reporting fever and blood in stools were significantly and positively associated with being hospitalised due to illness (Fisher's Exact, p=0.014 and 0.006 for fever and blood in stools respectively).

Eight cases were reported more than once during the study period, due to either recurrent or relapsing infection. Six cases suffered a second episode of giardiasis less than one year from the first one. Seven cases were males and the median age was 38 years (range one to 63).

<u>Ethnicity</u>

Out of 205 cases that answered the question about their ethnicity, 188 (91.7%) were Whites, followed by 15 (7.3%) that were Asians, one White/Indian and one Black-Caribbean.

Travel history

Of the 238 cases answering the question on their travel history in the month prior to illness, 138 (58%) reported that they did not travel. Of the 100 cases that travelled, 76 travelled outside the UK and 74 reported the countries they travelled to: of these, 46 (62.2%) travelled to one or more potentially at-risk countries for giardiasis such as Middle East and Asia (including Turkey) (32 cases, 69.6%), Africa (including Canary Islands) (ten cases, 21.7%) or Central and South America (two cases). The remaining 28 cases that travelled abroad reported travelling to Europe (25 cases) or North America (three cases).

Water and food consumption

A total of 237 cases reported the source(s) from which they drank un-boiled water in the month prior to illness: 85.2% (202/237) reported drinking from the mains water supply, 1.3% (3/237) drank water from a private supply, 17.7% (42/237) from a water filter, 60.8% (144/237) drank bottled water and 1.3% (3/237) potentially unsafe water from a river, a spring or a pond. The median number of glasses of un-boiled water (e.g. taken straight from the tap or mixed with cordial or squash) consumed daily by the cases was three (range zero to 20).

Out of 225 cases answering the question 207 (92%) reporting eating fresh fruit in the month prior to illness and 144 (64%) ate it at least three times a week, with most cases purchasing it at a supermarket (178 cases out of 207 reporting the provenience, 86%). A total of 227 cases answered the question on salad items (e.g. lettuce, tomatoes, ready-made or homemade salads, or salad items in sandwiches), of which 86.3% (199/227) ate any salad item in the month prior to illness, with 99 cases (43.6%) reporting eating them at least three times a week. Salad items were purchased mainly at a supermarket (167/197 reporting the provenience, 84.8%). Uncooked vegetables other than salad items were eaten at least once only by 101 (46.3%) of the 218 responding cases, most being purchased at a supermarket (88/143 reporting the provenience, 85.4%). A total of 166 (75.1%) of the 221 responding cases drank fruit or vegetable juice in the month prior to illness, with 39.8% (88/221) reporting drinking juice at least three times a week and juice mostly purchased at a supermarket (146/166, 88%). A total of 31 cases also reported eating any of the aforementioned food items from a place different from those specified (supermarket,

greengrocers, market or home-grown), and these mostly including hotel restaurants and farm shops.

Outdoor recreational activities

A total of 246 cases answered the questions regarding outdoor recreational activities practised in the month prior to illness: 35.4% (87/246) went to the swimming pool at least once, for a median of three times (range one to 30) in the 60 cases that reported the information, 7.3% (18/246) swam in a lake, pond or river, 10.9% (27/246) practised water sports in freshwater or at sea, 32.9% (81/246) walked in the countryside and 10.2% (25/246) went picnicking, 5.7% (14/246) practised fieldsports, 5.7% (14/246) went caravanning, 4.9% (12/246) went camping and 3.3% (8/246) went fishing.

Animal contact

Out of 232 cases answering the question about pet ownership, 113 (48.7%) reported either owning a pet (110) or having pets at the workplace (three cases, of which two working at a kennel and a cattery). For the pet owners, dogs were most frequently reported (66 cases, 60%), followed by cats (44 cases, 40%), birds (including chickens) (16 cases, 14.5%), rabbits (12 cases, 10.9%) and horses (two cases, 1.8%). Other reported pets included fish (seven cases) and hamsters or guinea pigs (five cases). Of the 91 pet owners that responded, only 6 (6.6%) reported any of their pets having diarrhoea in the exposure window. A total of 215 cases responded to the question of visiting a premise with animals in the exposure window, of which 61 (28.4%) reported having do so including a farm, stable or a horse riding school (28 cases) and a wildlife park or a zoo (nine). Of 54 responders, 38 (70.4%) reported touching at least one animal during the visit. Most commonly touched animals were sheep and goats (20 cases, 52.6%), horses (12, 31.6%) and cattle (ten, 26.3%), followed by chickens (nine, 23.7%) and pigs (four, 10.5%). Other types of animal touched included birds (parrot, starling, owlet) (four cases), dogs and cats (three), elephants (three), rabbits and guinea pigs (two), reptiles (tortoise, snake, lizard) (two), deer and camels (two).

Household composition and human contact

In 235 cases reporting this information the median number of adults (i.e. 16 years of age or more) in the household was two (range one to 10), and 112 (47.7%) reported at least one child below 16 years of age. The median number of children was two (range one to five). The median number of people living in the same household was three (range one to ten). Of 106 cases with children in the household and reporting this information, 59 (55.7%) reported at least one child being in nappies. Of the 226 cases answering the question, 41 (18.1%) reported changing nappies (not only of their own children) at least once in the exposure window. A total of 30 (13%) of the 230 that responded reported the presence in the household of at least another person ill with symptoms similar to those experienced the case. A total of 24 (11.1%) of 216 responders reported also they had contact with a person ill with similar symptoms outside the household.

2.4.4 Comparison of cases with and without a history of travel abroad

The socio-demographics, clinical outcome and exposure history were compared between and the cases that either did not report any travel or that only travelled within the UK (defined as indigenous) and those reporting travelling abroad in the month prior to illness. The two groups of cases did not differ in terms of their gender distribution (Pearson's χ^2, p=0.895), but significant differences between travel-related and indigenous cases were found in relation to age (Fisher's Exact, p=0.020): travelling abroad was reported significantly more frequently in adults in their 50s, whereas the significant majority of people aged 70+ years did not report any travel abroad. Indigenous and travel-related cases did not differ in terms of their neighbourhood level of deprivation (Fisher's Exact, p=0.117). The significant differences (p<0.05) found in the clinical and exposure profile of the two groups of cases are reported in **Table 2.4**, whereas non-significant comparisons ($p\geq$0.05) are reported in **Table 2.4.1**, **Appendix 1**.

Table 2.4: Significant differences (p<0.05) in the clinico-epidemiological characteristics of *Giardia* cases that did not travel abroad (n=162) and cases that travelled abroad (n=76) in the month prior to illness and that returned the surveillance questionnaire.

			Travel abroad (outside the UK)		
Variable	**No. valid (% missing)***	**Category**	**No n (%)****	**Yes n (%)****	***p*-value**
Admitted to hospital due to the illness	219 (7.8)	No	130 (86.1)	65 (95.6)	0.037
		Yes	21 (13.9)	3 (4.4)	
Drinking un-boiled tap water from a mains water supply	231 (2.9)	No	16 (10.1)	18 (24.7)	0.004
		Yes	142 (89.9)	55 (75.3)	
Drinking bottled water	231 (2.9)	No	73 (46.2)	16 (21.9)	<0.001
		Yes	85 (53.8)	57 (78.1)	
Going to the swimming pool	238	No	124 (76.5)	30 (39.5)	<0.001
		Yes	38 (23.5)	46 (60.5)	
Frequency of going to the swimming pool (no. times)	212 (10.9)	0	124 (81)	30 (50.8)	<0.001[M]
		1-2	19 (12.4)	6 (10.2)	
		3-4	5 (3.3)	10 (16.9)	
		5-6	5 (3.3)	2 (3.4)	
		7+	0 (0)	11 (18.6)	
Swimming in a lake, pond or river	238	No	155 (95.7)	65 (85.5)	0.006
		Yes	7 (4.3)	11 (14.5)	
Practising watersports in freshwater	238	No	161 (99.4)	71 (93.4)	0.014
		Yes	1 (0.6)	5 (6.6)	
Practising watersports at sea	238	No	157 (96.9)	60 (78.9)	<0.001
		Yes	5 (3.1)	16 (21.1)	
Went fishing	238	No	160 (98.8)	70 (92.1)	0.014
		Yes	2 (1.2)	6 (7.9)	
Eating fresh fruit purchased from a market	221 (7.1)	No	145 (95.4)	59 (85.5)	0.011
		Yes	7 (4.6)	10 (14.5)	
Eating fresh fruit from a place other than market supermarket, greengrocers or homegrown	221 (7.1)	No	147 (96.7)	57 (82.6)	<0.001
		Yes	5 (3.3)	12 (17.4)	
Eating salad items purchased from a market	223 (6.3)	No	151 (98.1)	63 (91.3)	0.027
		Yes	3 (1.9)	6 (8.7)	
Eating salad items from a place other than market supermarket,	223 (6.3)	No	148 (96.1)	58 (84.1)	0.002

greengrocers or homegrown		Yes	6 (3.9)	11 (15.9)	
Eating uncooked vegetables from a place other than market, supermarket, greengrocers or homegrown	214 (10.1)	No	145 (98.6)	61 (91)	0.013
		Yes	2 (1.4)	6 (9)	
Drinking juice purchased from a market	216 (9.2)	No	147 (99.3)	63 (92.6)	0.013
		Yes	1 (0.7)	5 (7.4)	
Drinking juice from a place other than market supermarket, greengrocers or homegrown	216 (9.2)	No	142 (95.9)	60 (88.2)	0.041
		Yes	6 (4.1)	8 (11.8)	

*percentages refer to the proportion of participants with missing information for the variable; **percentages refer to the proportion among cases that answered the question;
[M]Mann-Whitney U test

The seasonality of the cases reported in the period 2008-2011 in relation to their reported travel history was also compared. There was no significant association between the cases travel history and their distribution by either month (Fisher's Exact, p=0.188) or season (Pearson's χ^2, p=0.107). Cases reporting travel abroad exhibited a larger peak in August, September and an upswing in December (**Figure 2.8**), their numbers significantly increasing in the second half of the year (e.g. Summer and Autumn) (correlation between number of cases and month of report: Spearman's rho=0.685, p=0.014 by month and Spearman's rho=0.655, p=0.345 by season). On the other hand, cases that did not travel abroad showed a major peak during September and a minor peak in May-June, but their distribution did not show any significant trend throughout the year (Spearman's rho=0.00, p=1.000 by month and Spearman's rho=0.880, p=0.120 by season).

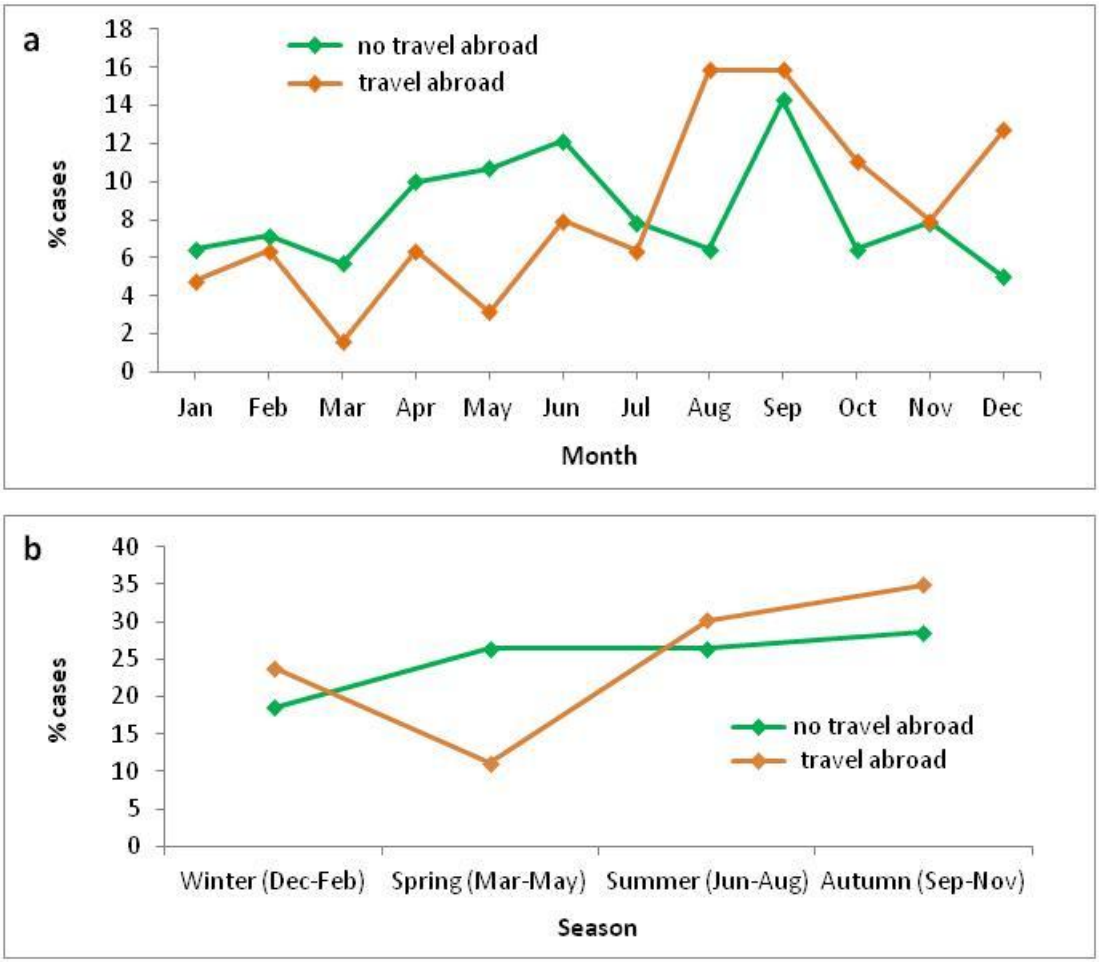


Figure 2.8: Monthly (a) and seasonal (b) distribution of cases that did not travel abroad (n=140) and those that travelled abroad (n=63) in the month prior to illness, 2008-11.

2.4.5 Comparison of male and female cases exposure history

The exposure history reported in the month prior to illness was also compared between genders. The significant differences ($p<0.05$) found in the exposure profile of male and female cases are reported in **Table 2.5**, whereas non-significant comparisons ($p\geq0.05$) are reported in **Table 2.5.1** in **Appendix 1**.

Table 2.5: Significant differences ($p<0.05$) in the exposure history reported in the month prior to illness by male (n=138) and female (n=108) *Giardia* cases that returned the surveillance questionnaire.

Variable	No. valid (% missing)	Category	Males n (%)*	Females n (%)*	*p*-value
Eating salad items purchased from a market	228 (7.3)	No	118 (93.7)	101 (99)	0.045
		Yes	8 (6.3)	1 (1)	
Keeping a cat	229 (6.9)	No	109 (86.5)	76 (73.8)	0.015
		Yes	17 (13.5)	27 (26.2)	
Changing nappies	226 (8.1)	No	112 (87.5)	73 (74.5)	0.012
		Yes	16 (12.5)	25 (25.5)	

*percentages refer to the proportion among cases that answered the question

2.5 DISCUSSION

We explored the burden of giardiasis and its major clinico-epidemiological features in Central Lancashire, using data collected during an enhanced surveillance program in the area.

Results showed that giardiasis commonly occurs in Central Lancashire, with an average rate of 22.5 cases/100,000 population in the period 2008-2012. This rate was four times higher than the average rate reported in either North West England (5.5 cases/100,000) or in England and Wales (6.2/100,000) in the period 2008-2009 (Health Protection Agency, 2011a) and also higher than the one reported in Europe in 2009 (5.6/100,000) by the European Centre for Disease Prevention and Control (European Center for Disease Prevention and Control, 2011).

It is not possible to tell if the incidence of disease was actually higher in Central Lancashire compared to other regions of England or Europe. Incidence rates cannot be directly compared across countries (or regions within the same country) if different surveillance systems or diagnostic procedures are being used (Yoder *et al.*, 2010). The procedure currently in place at the Royal Preston Hospital, based on testing all diarrhoeal samples submitted from the community with a highly sensitive diagnostic method, allowed the detection of infections that would normally be missed in other contexts: under this light, the rates of giardiasis observed in Central Lancashire can be considered reliable. It is important to note that in this study mostly symptomatic infections were taken into account, whereas asymptomatic infections were occasionally found only when the household members of the index case were tested for *Giardia* too. Not many data on the prevalence of asymptomatic *Giardia* infection in the UK population are available. In a study on preschool children at day care facilities, the parasite prevalence in this group of subjects was found to be only 1.3% (Davies *et al.*, 2009). The application of a PCR assay on faecal samples collected from asymptomatic controls during the English case-control Infectious Intestinal Disease Study revealed the presence of *Giardia* DNA in 1.4% of the specimens (Amar *et al.*, 2007). It is evident that more data on the asymptomatic carriage of this parasite in the general population are needed in order to understand its true burden.

Rates of giardiasis varied considerably over time and between the three local authorities of Central Lancashire. Spatio-temporal variation was also observed in the rates of cryptosporidiosis and campylobacteriosis. Neither the way faecal samples were tested by the hospital lab nor the way cases were notified to the CLHPU changed during the study period, and the sample submission rate (i.e the number of faecal samples submitted per local resident population) did not vary between the local authorities (John Cheesbrough, personal communication). Variations in disease rates observed at such a small scale must not be over-interpreted since they could be the result of variations in the population exposure levels to the specific pathogen (i.e. changes in travel habits or likelihood of outdoor activities) or they may be the result of small outbreaks going unnoticed.

Giardiasis did not show a pronounced seasonality in the area under study, but the case reports increased in the warm season and in particular they peaked in the month of September. Results confirm the data available from England and Wales, where the majority of cases of disease are usually reported in August and September (Gray *et al.*, 1994; Health Protection Agency, 2011a). A peak in the number of cases of giardiasis is usually observed in the warm season in other developed countries like the United States (Yoder *et al.*, 2010) and Finland (Rimhanen-Finne *et al.*, 2011). The increase in the number of cases at the end of the summer is commonly observed in cryptosporidiosis as well (Chalmers *et al.*, 2011; European Center for Disease Prevention and Control, 2011). Travelling increases over the Summer, and hot weather is associated with increased water consumption and recreational water use. It has also been proposed that parasite transmission can also be amplified by person-to-person transmission due to the close quarters and low levels of hygiene of activities like camping and swimming (Naumova *et al.*, 2007). A different tendency in monthly reporting was observed when cases were compared according to their history of foreign travel, with most cases reporting travelling abroad being reported at the end of Summer and beginning of Autumn. The late-summer peak (and partially also in December) in cases reporting travel abroad supports the role of exposure by increased foreign travel as observed also in cryptosporidiosis (Chalmers *et al.*, 2011). The tendency of an increased number of reports from April to June observed in cases that did not travel abroad could be an indication of increased local environmental exposure during that period of the year.

Children under five and adults in their 30s were the most affected by giardiasis. Available nationwide data show that young children (below ten years) are usually the most represented age group for giardiasis (Yoder *et al.*, 2010; HPA, 2011a). Nevertheless, our finding of high rates of disease in adults is in accordance with some previous studies from England (Stuart *et al.*, 2003; Ellam *et al.*, 2008), but not from other studies (Rimhanen-Finne *et al.*, 2011). Interestingly, by considering other enteric pathogens a bimodal age distribution (i.e a major peak in infancy, followed by a decrease and a second increase in rates in young adults) has also been observed for campylobacteriosis in the UK (Gillespie *et al.*, 2008) (Health Protection Agency, 2011b). Secondary transmission from the infected child to the other members of the household (through nappy changing, shared toilet etc.) is thought to play a major role in this (Gillespie *et al.*, 2008), and this scenario is applicable to giardiasis as well.

The most striking result was the overall excess of male compared to female cases. A disproportion between genders in the number of samples submitted for diagnosis may have biased our result. However, the majority (54%) of specimens received by the Royal Preston Hospital lab during 2011 actually came from females, and only 46% from males (John Cheesbrough, personal communication). It has also been shown that males over a wide range of ages (in particular from 15 to 60 years) tend to have lower GP consultation rates compared to females (Hippisley-Cox & Vinogradova, 2009). Both the aforementioned situations should have then caused males to be underrepresented. Although previous studies done in England did not show it (Gray *et al.*, 1994; Stuart *et al.*, 2003; Breathnach *et al.*, 2010), a slight preponderance of male cases can be nevertheless observed in the giardiasis surveillance data from Europe (European Center for Disease Prevention and Control, 2011), United States (Yoder *et al.*, 2010) and New Zealand (Snel, 2009a). Male sex is a highly significant risk factor for symptomatic giardiasis in Germany (Espelage *et al.*, 2010), and men were found to be at a higher risk for infection in other two studies (Ekdahl & Andersson, 2005; Laupland & Church, 2005). It is important to consider that in our study, as in those mentioned above, only symptomatic cases were included. Whether the prevalence of *Giardia* infection in the asymptomatic general population tends also to be higher in males or is equally distributed between genders has to be verified.

Interestingly, male-biased infection rates and higher infection intensities in males have been shown in 84% of the 58 parasite species in which gender differences are reported (Klein, 2004), and male mice experimentally infected with *Giardia muris* tend to suffer from higher trophozoite burdens compared to female mice (Daniels & Belosevic, 1995). Gender-specific differences in the immune responses play a major role in host-parasite interactions and disease severity in several bacterial and fungal diseases (McClelland & Smith, 2011). Males and females differ in both their innate and acquired immune responses and sex hormones are known to influence the inflammation process, with testosterone having depressive effects on the immune system (Klein, 2004). If the progression of giardiasis differs between genders in humans as well, then men may tend to develop symptomatic infection or more severe disease more easily than females resulting in higher GP consultation rates. Contrasting evidence came from the fact that in cases from Central Lancashire females reported a longer illness than males, but most of the hospitalised cases were males. However, the severity of disease is by itself difficult to define and express quantitatively. Furthermore, differences in illness progression between genders may be subtle and the self-reporting of symptoms may not be the most appropriate method to ascertain them.

In addition to immunological differences, also gender specific behaviours that may favour the exposure to the parasite must be considered (Bundy, 1988). However, no available study on giardiasis has ever shown any gender- specific risk factors. Our results seem to suggest that males and females may be exposed to the parasite differently. Overall females reported a higher frequency of nappy changing, which is a well-known risk factor for giardiasis (Hoque *et al.*, 2001; Hoque *et al.*, 2002a). The high frequency of going swimming we observed in women in their 30s and over 60 years may also be related to this aspect, potentially by accompanying their children or grandchildren to the pool. A study from Ivory Coast reported girls being significantly more infected with *Giardia* than boys and infection was positively associated with the use of pond water (Ouattara *et al.*, 2010). Higher frequencies of foreign travel (and consequently eating food in places such as restaurants) were noticed in males between 40 and 60 years of age. It is possible that a more adventurous behaviour while travelling abroad may expose men to a higher risk of transmission from contaminated water or food. Other behavioural factors may play a role in the excess of male cases. For example, anal-oral practices are common

amongst gay men and can favour the transmission of various intestinal parasites like *Giardia* (Keystone *et al.*, 1980; Abdolrasouli *et al.*, 2009). Also the presence of differences in the level of personal hygiene between males and females cannot be excluded.

The majority (57.7%) of the cases of disease lived in areas of moderate to low deprivation. Our results match what has been reported in New Zealand, where higher notification rates for both giardiasis and cryptosporidiosis were observed in least deprived areas (Snel *et al.*, 2009b). It is important to stress that due to the complexity of the IMD rank (that takes into account more than 30 different indicators) it may be difficult to observe a direct relationship between the overall level of deprivation and risk factors for giardiasis, like poor levels of hygiene or overcrowding.

The analysis of self-reported clinical data confirmed that giardiasis is associated to high morbidity in humans. Illness duration well exceeded two weeks on average (possibly even more considering that nearly 30% of the cases declared feeling still ill when they filled the questionnaire), and it interfered with the normal daily activities for several days. More severe and usually rarer outcomes like fever and vomiting were quite commonly reported. Since patients with more severe intestinal illness show higher rates of GP presentation (Tam *et al.*, 2003), it is possible that the morbidity has been slightly overestimated in our study since cases with mild symptoms may have been missed. The presence of blood in faeces was reported by a number of cases. This finding is surprising since *Giardia* is not an invasive pathogen like *Entamoeba* spp., and blood shouldn't be present in the stools (Hill, 2001). However, this feature was previously reported by a few patients with giardiasis in England (Breathnach *et al.*, 2010). It cannot be excluded that the presence of blood in stools could have been the result of other underlying medical conditions: in our study the cases reporting blood in faeces also showed a higher rate of hospitalisation. So the relevance of blood in stools must be considered cautiously. Females seemed to suffer from a comparatively longer illness than males, but the latter were more frequently hospitalised as a result of giardiasis. Since the clinical outcomes were reported by the cases themselves, the presence of gender specific bias cannot be excluded completely. The perception of illness varies between genders and females tend to report more physical symptoms than males (Barsky *et al.*, 2001), although no

difference was found between the genders in the number or type of symptoms experienced.

The analysis of the exposures reported by the cases in the month prior to illness led to some interesting findings.

Cases reporting foreign travel during the exposure window were only a relatively small proportion (32%), suggesting that transmission of *Giardia* may occur locally (or at least within the UK) more often than expected. *Giardia* is one of the most common etiologic agents of travellers' diarrhoea, and travelling abroad consistently emerged as a highly significant risk factor in several studies (Gray *et al.*, 1994; Hoque *et al.*, 2002b; Gagnon *et al.*, 2006). However, the actual number of cases of disease acquired locally is likely to be underestimated if faecal samples mostly coming from patients with recent travel are selectively tested for *Giardia*. In our study, the non-selective testing of community diarrhoeal specimens allowed the inclusion of a significantly higher number of cases without a foreign travel history. Both people behaviours (like water and food consumption) and so potentially the exposure levels to the parasite (for example due to different environments and levels of hygiene) are likely to change when travelling to another country. Our results confirmed this hypothesis and suggested that during travel abroad the transmission of giardiasis may have an important waterborne and foodborne component: water-related recreational activities (including swimming, water sports etc.) and the consumption of fresh products from markets or restaurants where reported significantly more frequently in cases that travelled abroad. Giardiasis can be associated with swimming (Gray *et al.*, 1994; Stuart *et al.*, 2003) or other activities involving contact with water (Karanis *et al.*, 2007). Also transmission of *Giardia* through contaminated fresh products is possible (Smith *et al.*, 2007), and in several studies their consumption stood up as significant risk factors for the disease (Stuart *et al.*, 2003; Espelage *et al.*, 2010; Bello *et al.*, 2011). The provenience of the products is probably an important aspect, with the risk of infection from products sold in supermarkets likely to be reduced due to more strict food safety measures, if compared to food coming from markets, farmers or eaten abroad. Cases that did not travel abroad prior to illness were hospitalised more frequently than those that did. The same finding was reported in a study from Germany (Espelage *et al.*, 2010).

There is no clear explanation for this result. Cases that did not travel abroad appeared to be slightly older than the others. Older age and weakened immune responses have been proposed as a potential explanation for increased hospitalisation (Espelage *et al.*, 2010). Furthermore, people experiencing recent travel abroad are more prone to consult their GP in the case of diarrhoea (Tam *et al.*, 2003). It is possible that indigenous cases may tend to delay visiting their GP, worsening their condition and ending up requiring hospitalisation. However, it was not possible to draw ultimate conclusions due to the very small number of hospitalised cases.

Only a very small proportion of cases reported drinking water from a private supply source. The use of private water supplies is extremely limited in England and Wales, with nearly all households (99.5%) receiving their drinking water from public supplies (Smith *et al.*, 2006b). It seems then unlikely that transmission through drinking water is of any importance in the area. Contact with animals during the exposure window was commonly reported by the cases, with nearly half of them reporting owning at least one pet. Pet ownership is common in the UK, involving at least 45% of households in 2011 (Pet Food Manufacturers Association, 2013). Most reported pets were dogs and cats, but also pet rodents were well represented. All these animals are commonly infected with *Giardia*, and they can harbour potentially zoonotic genotypes of the parasite (Sprong *et al.*, 2009). A quarter of the cases also reported visiting animal premises (mostly farms) and touching various animals on site, in particular horses. We found that indigenous cases touched horses significantly more often than travel-related cases. This difference could be due to the fact that horses usually have closer contact with people on farm settings compared to other species, and our analysis was done on a very small number of valid cases. Curiously, contact with these particular animals on farms was more frequently reported by cases of giardiasis compared to other enteric cases in Canada according to the Public Health Agency of Canada (Public Health Agency of Canada, 2009). The role of farm animals in the zoonotic transmission of *Giardia* is debated, although some indication of it was found in the UK (Warburton *et al.*, 1994) and molecular epidemiological studies seem to suggest it (Winkworth *et al.*, 2008; Khan *et al.*, 2011).

The presence of children in the household was reported by nearly a half (48%) of the cases, and the majority (56%) reported at least one of their children to be in nappies. Transmission of *Giardia* within the household through person-to-person contact is highly facilitated by the low infectious dose of this parasite (Rendtorff, 1954), and both the presence of children in nappies and the act of changing nappies has been reported as significant risk factors for giardiasis (Hoque *et al.*, 2001; Hoque *et al.*, 2002b; Gagnon *et al.*, 2006). The presence of household members with diarrhoea is also commonly reported by people infected with *Cryptosporidium hominis* (Chalmers *et al.*, 2011), and the same has been observed in patients with campylobacteriosis (Gillespie *et al.*, 2008). Furthermore, the majority of outbreaks of giardiasis reported in a study in South West London were in fact small household clusters (Breathnach *et al.*, 2010).

To sum up our results confirmed that using a non-selective approach to community stool specimens testing allows the detection of a larger number of cases of giardiasis, while revealing at the same time a hidden burden of disease in males, in adults and people without a history of foreign travel. The self-reported clinical outcomes also indicated a high morbidity associated with giardiasis. This study was only descriptive and due to the lack of a control population no inferences about the risk factors for infection in the area could be made. Nevertheless, the analysis of the cases' exposure history prior to illness and the comparison between cases with and without a history of travel abroad highlighted the possible role played by water-related recreational activities and household contacts for the transmission of the parasite, and at the same time the need for more data to assess the presence and importance of zoonotic transmission. Our results provided then a starting point for a more thorough investigation of the risk factors for giardiasis in the area, and in particular for infections not related to foreign travel.

CHAPTER THREE: CASE-CONTROL STUDY OF RISK FACTORS FOR GIARDIASIS IN NORTH WEST ENGLAND

3.1 INTRODUCTION

As the results of the surveillance study presented in the previous chapter had shown, the burden of giardiasis in Central Lancashire was considerable compared with the official figures available for the rest of the UK. In particular, the high proportion of adult cases and cases without a history of foreign travel highlighted the need for more data on the determinants of infection acquired within the UK. Various exposures potentially linked to the acquisition of *Giardia* in the area (including swimming and contact with animals or other ill people) were reported in high frequencies in the cases, but their assessment as actual risk factors for the disease was prevented by the lack of a proper control population. The implementation of case-control (or cohort) studies is the most appropriate strategy to determine which factors are significantly associated with an increased risk for the acquisition of particular disease in a population.

So far, the risk factors for giardiasis in the UK have been investigated in only three case-control studies (Gray *et al.*, 1994; Warburton *et al.*, 1994; Stuart *et al.*, 2003). Gray *et al.* (1994) investigated the general risk factors for giardiasis in patients from Avon and Somerset (South West England),. Major hypotheses included the role of foreign travel, drinking potential contaminated water and recreational water use. When exposures were considered separately, four of them were significantly associated with giardiasis: travelling to a developing country, camping, caravanning, or use of holiday chalets, swimming (in swimming pools, in freshwater or in the sea) and drinking potentially contaminated water (from rivers, streams or wells, ice cubes or tap water when outside the UK). Multivariable analysis confirmed that both travelling to developing countries (Odds ratio OR 7.6, $p<0.001$) and camping, caravanning or use of holiday chalets (OR 8.4, $p=0.017$) were strong independent risk factors for giardiasis, whereas swimming was only marginally associated with it (OR 2.4, $p=0.05$).

The study by Stuart *et al.* (2003) analysed the risk factors for giardiasis specifically in people from South West England who had not recently travelled outside the UK. Contact with recreational waters (both chlorinated and untreated, including swallowing water during the activities), drinking water and food consumption (salad, fruit and dairy products), and contact with farms, animals and day nurseries were all

tested as potential risk factors. In univariable analysis, giardiasis was found to be associated with the consumption of tap water and the number of glasses of tap water consumed per day, eating lettuce, visiting a farm, having contact with a nursery, and various recreational water exposures (including swimming in chlorinated or fresh water, frequency of swimming and head immersion or swallowing water during swimming). Recreational contact with fresh water (OR 5.5, p=0.001), swallowing water while swimming (in either chlorinated or fresh water) (OR 6.2, p<0.001), eating lettuce (OR 2.2, p=0.01) and each additional glass of tap water consumed daily (OR 1.3, p<0.001) were all confirmed as independent risk factors for the disease following multivariable analysis.

The study of Warburton *et al.* (1994) focused specifically on the local risk factors for disease in a rural district of East Anglia, by excluding all cases with a history of travel either abroad or outside the study area. The major hypothesis of zoonotic transmission was tested by analysing the contacts of the cases with pets (dogs and cats) and livestock (such as pigs). Also water consumption and swimming habits were tested as potential risk fators. Only exposures related to contact with animals were associated with an increased risk of giardiasis in univariable analysis: in particular, multivariable analysis confirmed that both having contact with pets (OR 14.55, p<0.001) and having contact with farm animals (OR 4.77, p=0.01) were strongly associated with the disease.

To sum up, the aforementioned case-control studies highlighted three major components in the acquisition of giardiasis in the UK. First, travelling abroad to a developed country is an important route for infection, confirming what has been reported by other studies from developed countries. Second, waterborne transmission (through contact with either fresh or chlorinated recreational waters either, or via drinking water) seems to be important in people without a foreign travel history. Third, it is not clear whether either zoonotic or human-to-human transmission (e.g. exposure to human wastes, changing nappies etc.) play any role for *Giardia* transmission in the country.

More data are then needed from other parts of the country in order to obtain a clearer picture of the major risk factors for giardiasis in the UK, focusing in particular on the transmission of the disease within the country.

3.2 AIMS OF THE STUDY

The aim of the study was to determine the risk factors for giardiasis in people from North West England using a case-control study design.

We first tested the hypothesis of foreign travel (and particularly travelling to developing countries, considered more at-risk for giardiasis) as a prominent risk factor for the disease. Secondly, the studied the risk factors for giardiasis not acquired through foreign travel (e.g. indigenous disease acquired within the UK). Our major hypotheses revolved around the role of zoonotic transmission (through contact with various pet species, livestock and also wildlife) and person-to-person transmission (through contact with young children and their nappies or with other sick people in the household). In order to get a comprehensive picture of the epidemiology of the disease in this area, we also tested the role of potential occupational exposures (through contact with animal or human wastes or freshwaters), swimming (either in swimming pools, lakes or rivers, or at sea) and other outdoor recreational activities (including picnicking, camping or caravanning), drinking water consumption (including also the occurrence of unusual changes in the characteristics of tap water) and food consumption (focusing in particular on fresh products).

3.3 MATERIALS AND METHODS

3.3.1 Study duration and catchment area

The study involved data collection within the North West of England. The total duration of the study was 17 months, with recruitment starting on the 21/02/2012 and finishing on the 27/08/13. Participants were recruited from three selected study areas based on the catchment population of three hospital microbiology departments with participants defined by their local authority (LA) of residence, namely Central Lancashire (LAs of Preston, South Ribble and Chorley), East Lancashire (LAs of Blackburn with Darwen, Burnley, Hyndburn, Pendle, Ribble Valley and Rossendale) and Central Manchester. The inclusion of these three areas was motivated by the need of having a reasonable sample size within the planned study time frame (e.g. one and a half years at maximum) and because all community diarrhoeal specimens in these areas were tested for *Giardia* using the same enzyme immunoassay. Originally, at the start of the study only the areas of Central Lancashire and Central Manchester were included. However, due to a low case response rate observed in the first few months of recruitment it was decided to include East Lancashire as a third catchment area. East Lancashire had become eligible for inclusion after having introduced the enzyme immunoassay test after the start of the study. As a result the timing of recruitment of participants differed between the three areas. Participants from Central Lancashire were included for the whole duration of the study. Recruitment from East Lancashire started in November 2012 and continued until the end of the study. Due to the addition of East Lancashire and the very low recruitment rates from Central Manchester, recruitment from this area was prematurely stopped in March 2013.

3.3.2 Ethical approval, study conduct and registration

Ethical approval for the study was sought through the National Research Ethics Service of England (NRES) and it was obtained from the NRES Committee North West in Lancaster. The study was conducted in accordance with the principles of Good Clinical Practice as laid out in the Committee for Proprietary Medicinal Products guidelines (CPMP, 1990). The study was registered in the UK Clinical

Research Network under the following name: *Study 11259 - Case-control Study of Risk Factors for Giardiasis in Northwest England* (a brief outline of the study is available online at the following address: http://england.ukcrn.org.uk/StudyDetail.aspx?StudyID=11259)

3.3.3 Study sample size

The minimum sample size required for the study was estimated using the Sampsize program in GLIM 4.0 (Generalized Linear Interactive Modelling) (Royal Statistical Society). Since transmission of giardiasis can be associated with a range of exposure factors and appropriate estimates of exposures prevalence in the general population were lacking, a range of potential sample sizes was explored. A decision was made to recruit between 90 (minimum) and 140 (maximum) cases to detect an odds ratio of at least two or more for a range of variables, with 80% power and 95% confidence, considering a five to 75% exposure range in the control group and a case/control ratio of 1:3. In order to take into account potential drop outs, it was decided to recruit five controls for each case.

3.3.4 Case definition and inclusion

To be eligible for the study, a case was defined as any person resident within the study areas suffering from gastroenteritis and with an antigen-confirmed *Giardia* infection. Hospitals involved in the study were the Royal Preston Hospital (Central Lancashire), the Royal Blackburn Hospitals (East Lancashire) and the Manchester Royal Infirmary (Central Manchester). The presence of *Giardia* infection in the cases was confirmed at all three hospitals with the same procedure used in the enhanced surveillance study presented in Chapter 2. As for the aforementioned study, cases found to be co-infected with any other major gastrointestinal pathogen for which faecal samples were screened were excluded. The laboratory-confirmed cases were reported by the hospitals microbiology departments either to the Central Lancashire Health Protection Unit (CLHPU) (Central and East Lancashire) or the Greater Manchester Health Protection Unit (GMHPU) (Central Manchester). Each case was allocated a unique six digit identity number (e.g. 123456). Faecal specimens from the cases were sent by the hospital laboratories to the Department of Infection Biology at

the University of Liverpool for parasite DNA extraction and molecular typing (see Chapter 3).

3.3.5 Control definition and selection

Controls were defined as subjects that were resident within the study areas and who had not had diarrhoea in the previous two weeks. Whenever a case of *Giardia* was reported, the HPUs transmitted the details about the case ID number, gender and age to either the Lancashire and South Cumbria Agency (LaSCA) or NHS Manchester (for Central/East Lancashire and Central Manchester areas, respectively). Controls were then selected at random using the "randomise" function in Microsoft Access (Microsoft Corporation, Redmond, WA, USA) from the respective areas' General Practitioner lists by the Information Departments holding the GP registration database. Controls were selected using a frequency matching design: five controls were selected for each case and they were matched to them for both gender and age (within the age ranges of 0-4, 5-14, 15-44, 45-64 and 65 years and above). Each control was also allocated a unique ID number linking it to the respective case (e.g.123456/L1 to L5 sequentially).

3.3.6 Recruitment and questionnaire data collection

All eligible participants were contacted by post from the respective area HPUs (cases) or LaSCA and NHS Manchester (controls) through a reply paid envelope containing: an invitation letter to invite them to take part in the study, a participant information leaflet providing a detailed overview of the background, purpose and implications of the study, an informed consent form and a copy of the study questionnaire that was created specifically for the study. All the forms are reported in **Figures A** to **D** in **Appendix 2**.

The questionnaire included questions about various socio-demographic, occupational, self-reported clinical details and a range of exposures experienced in the three weeks prior to illness onset (including history of travel, outdoor recreational activities, water and food consumption, pet ownership and contact with animals, household composition). The control questionnaire differed from the case questionnaire in only two aspects: the questions were referred to the three weeks

prior to completion of the questionnaire, and the health details section enquired about experiencing diarrhoea in the two weeks prior to compiling the questionnaire in order to identify ineligible controls. It also included questions about a history of previous or either long-lasting diarrhoeal illness or *Giardia* infection. Both the study questionnaire and the information leaflet were reviewed by the Liverpool Medicines for Children Research Network in order to make it comprehensible and child-friendly.

Participation in the study was voluntary and participants were asked to indicate their willingness to take part by signing the informed consent form to return along with the compiled questionnaire. For participants under 16 years of age the consent to take part was sought from a parent or guardian. The consent for the use of faecal specimens for *Giardia* molecular typing was sought from cases. Cases who did not respond to the initial invitation within one week were approached via telephone or via personal contact by either an environmental health officer (Central and East Lancashire) or a HPU health protection practitioner (Central Manchester). Cases that did not respond after two weeks were not contacted again. Controls who did not respond to the initial invitation within two weeks were sent a reminder letter (**Figure E**, **Appendix 2**) and a new copy of the questionnaire, information leaflet and consent form.

The questionnaires and consent forms were returned to either their respective area HPUs (for cases) or to either LaSCA or NHS Manchester (for controls), and they were then forwarded to the CLHPU where all the personal identifiable information (e.g. name, surname and address) was removed and retained along with the consent forms. The study questionnaires minus the personal identifiable information were then posted in batches on a monthly basis to the Department of Infection Biology at the University of Liverpool for data entry.

The whole case-control study design and implementation is summarised in **Figure 3.1**.

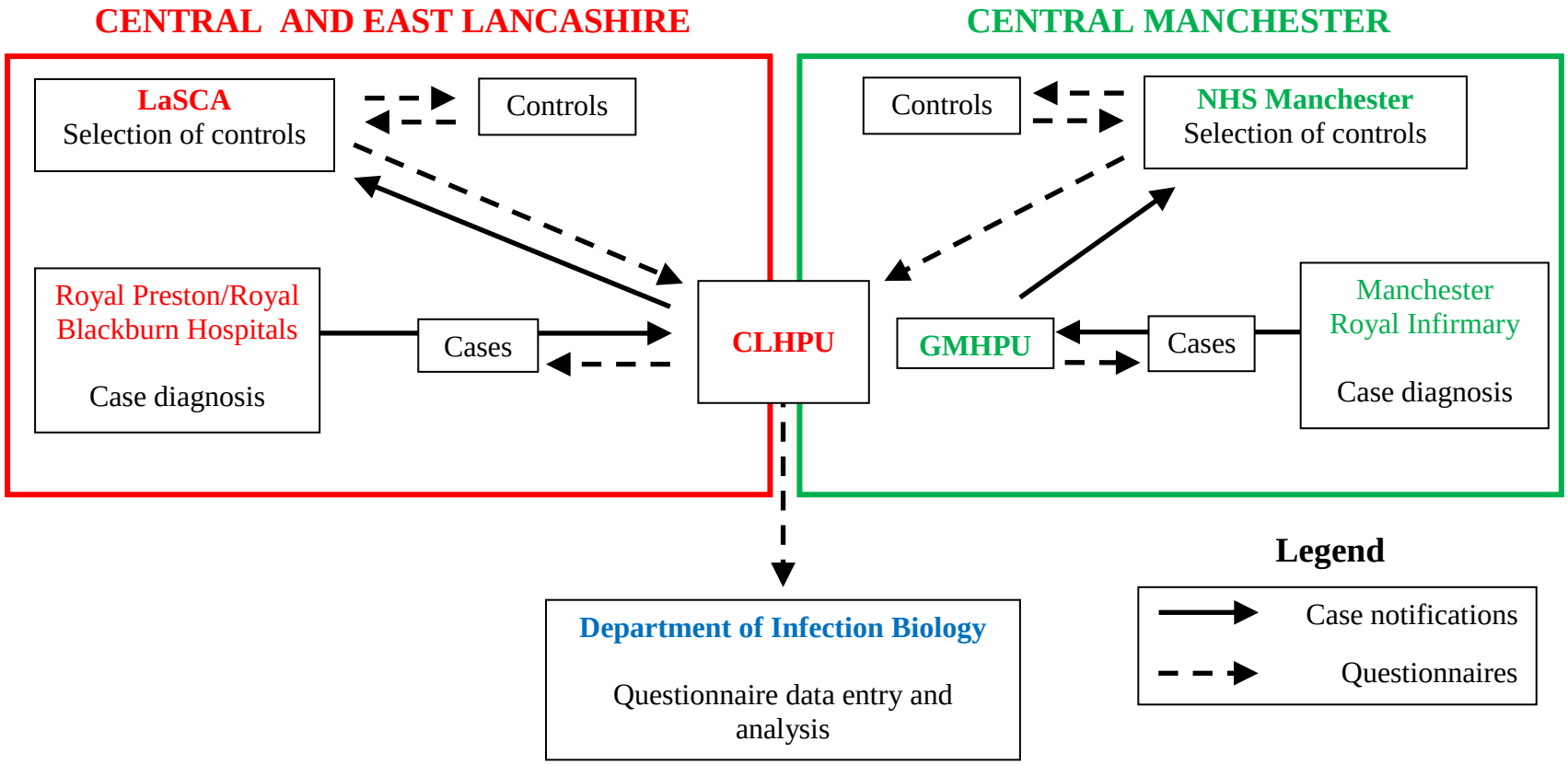


Figure 3.1: Flowchart illustrating the case-control study design and implementation.

3.3.7 Data management and quality control

Questionnaire information was manually entered onto the study database created with EpiData 3.1 (EpiData Association, Odense, Denmark). In order to minimise the occurrence of errors during data entry, data were double-entered by both the research student at Liverpool University and by two administrative staff at the CLHPU. The two copies of the database were then compared using the validate function in EpiData to identify discrepancies between the two datasets. Discrepancies were used to identify data entry errors that were then manually corrected in the dataset used for analysis by accessing the original paper questionnaires. Duplicated records were identified and deleted. All questionnaire fields were checked for unusual (i.e. out of range) values and related fields in the same section were compared. Incorrect values or missing fields, whenever it was possible, were corrected or inferred from the related questions in the same section.

3.3.8 Data analysis

All analyses were done using either EpiInfo™7 (Centre for Disease Control and Prevention, USA) or IBM® SPSS® Statistics 20 (IBM, USA). In order to study the risk factors for giardiasis in the area both in general and specifically in people without a history of foreign travel, two separate analyses were performed and compared: one including all participants and another one including only the participants that did not report travelling abroad in the exposure window (e.g. indigenous).

3.3.8.1 Univariable analysis

Each variable was first assessed singly as a potential risk factor by calculating its odds ratio (OR) estimate and 95% confidence interval (CI). Cross tabulations were produced and two-sided Pearson's Chi-Square test (or Fisher's Exact test when data were sparse) was used to test the null hypothesis of no association. Categorical variables with more than two categories were analysed using univariable logistic regression. Dose-response variables were analysed in their both categorical form (using the Chi-Square test for trend) and continuous form (using binary logistic

regression), and the results from both approaches were compared and reported. A *p*-value of less than 5% was considered to be statistically significant.

3.3.8.2 Multivariable analysis and modelling

Multivariable binary logistic regression modelling was used to determine which variables were independently associated with being a case with symptomatic giardiasis. Variables that had returned a *p*-value of ≤ 0.2 in the previous univariable analysis were selected for inclusion in the model. Variables were entered manually one-by-one, starting from those showing the highest significance by univariable analysis, and the importance of each factor was assessed by its effect on the overall model fit using likelihood ratio tests. Factors that had no significant effect on the model fit were dropped. The only variables that were retained in the final model regardless of their significance were gender and age, and interaction terms of these two variables were included to allow for potential confounding. Potential confounding or effect-measure modification of ethnicity, type of area of living (city, town or village), study area (Central Lancashire, East Lancashire, Greater Manchester) and season on the variables retained in the final model were also tested. Models were fitted to the data and after multiple imputation of missing values the results compared to evaluate the robustness of the model results to data quality. Furthermore the models were fitted following the exclusion of either all the un-matched participants or the suspected secondary household cases.

3.4 RESULTS

3.4.1 Notified *Giardia* cases

A total of 236 *Giardia* cases were reported during the study period. Out of 114 cases from Central Lancashire, 42 (36.8%) were from South Ribble, 41 (36%) from Preston and 31 (27.2%) from Chorley. Out of 79 cases from East Lancashire, 21 (26.6%) were from Blackburn with Darwen, 17 (21.5%) from Hyndburn, 14 (17.7%) from Pendle, 11 (13.9%) from Ribble Valley, nine (11.4%) from Rossendale and seven (8.9%) from Burnley. Of the 34 cases from the Greater Manchester area, 33 (97.1%) were from Manchester and one (2.9%) from Rochdale. The remaining nine cases came from Wyre (six cases), Fylde (two) and Lancaster (one). Although these nine cases were not resident within the study areas, since they suffered from gastroenteritis and were diagnosed with giardiasis at the Royal Preston Hospital they were included within the study.

Overall, 144 cases (61%) were males and 92 (39%) females. People from 15 to 44 years of age were the most frequent (94 cases, 39.8%) followed by adults from 45 to 64 years (56, 23.7%), adults with 65 or more years (34, 14.4%), children below five (30, 12.7%) and young people (5-14y) (22, 9.3%).

3.4.2 Recruited cases

Out of 236 cases originally contacted for the study, 123 returned the questionnaire (52.1% response rate) and were recruited as potentially eligible. The response rate was lower in males (68 questionnaires returned over 144, 47.2%) than in females (55 returned over 92, 59.8%) (Pearson's χ^2, p=0.060). Variation in response rates were observed in the different age groups (Pearson's χ^2, p=0.052): the highest response rate was observed in the elderly (65+y) (25 questionnaires returned over 34, 73.5%), followed by children below five (17 over 30, 56.7%), people from 15 to 44 years of age (46 over 94, 48.9%), adults from 45 to 64 years (27 over 56, 48.2%) and young people (5-14y) (eight over 22, 36.4%). Overall, recruited and notified cases did not differ significantly in either their gender (Pearson's χ^2, p=0.294) or age distribution (Pearson's χ^2, p=0.289) (**Figure 3.2**).

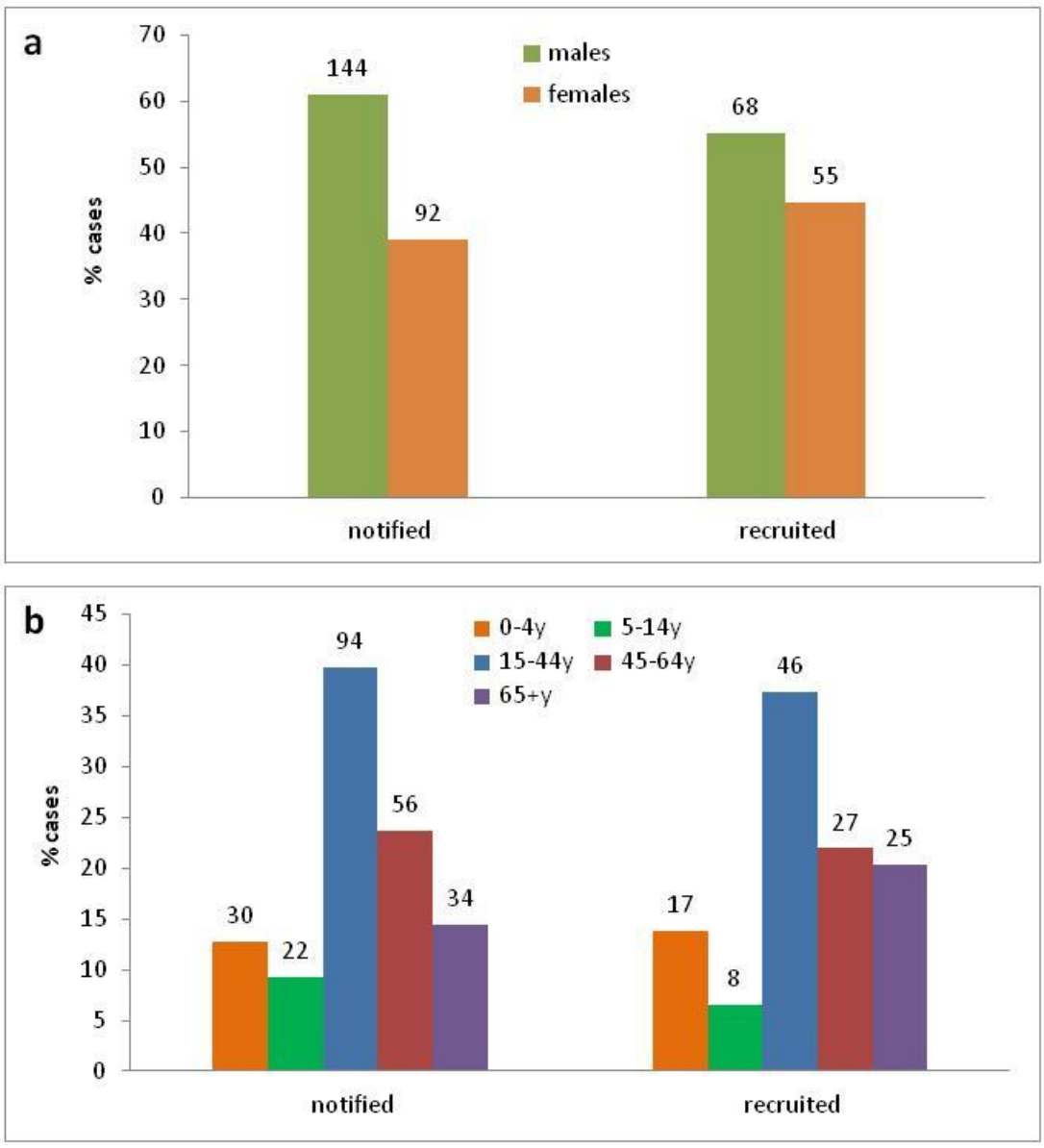


Figure 3.2: Gender (a) and age (b) distribution of the recruited (n=123) and the total notified cases (n=236). The number of cases in each category is reported above the bars.

Excluding the nine cases resident outside the study catchment areas, a higher case response rate was observed from Central (63 responders over 114 notified, 55.3%) and East (41 over 79, 51.9%) Lancashire compared to Greater Manchester (12 over 34, 35.3%), although overall the differences were not significant (Pearson's χ^2, p=0.122).

Three cases were mistakenly sent a wrong questionnaire and they did not return the correct one after it was sent a second time.

3.4.3 Cases self-reported clinical outcomes

Self-reported clinical information was analysed in the 120 cases that were sent and returned the correct questionnaire. Two cases were asymptomatic and they were tested for *Giardia* because at least another person in the same house was a case. The frequency of the self-reported symptoms in the 118 symptomatic cases is shown in **Figure 3.3**.

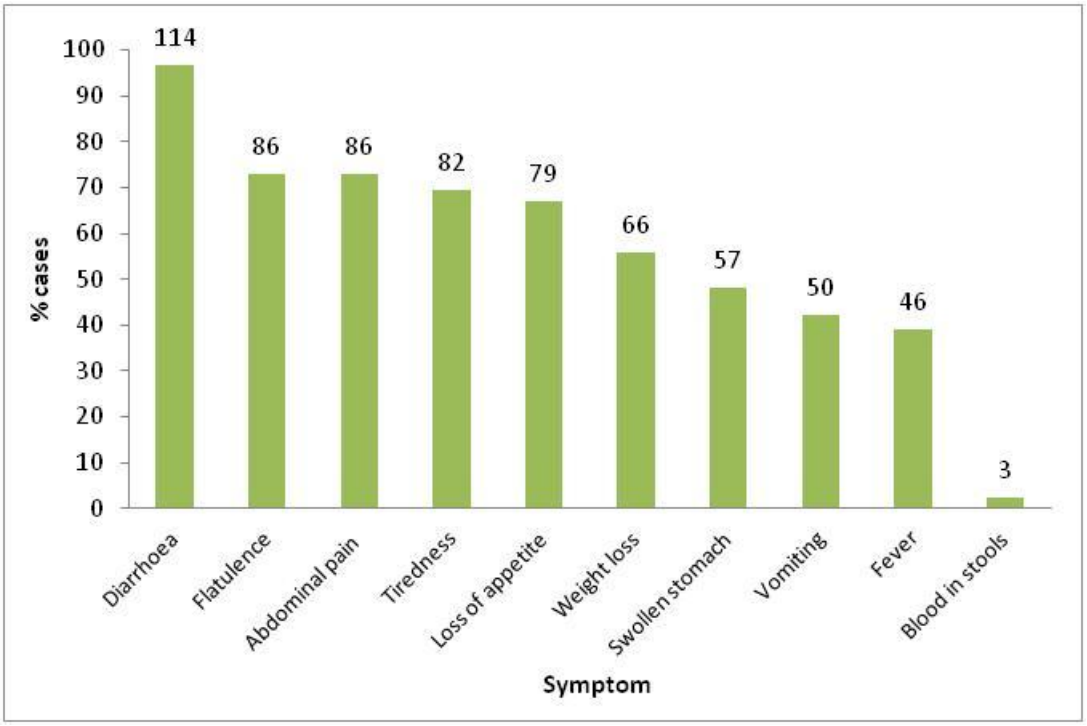


Figure 3.3: Frequency of self-reported symptoms in the 118 symptomatic cases. The number of cases in each category is reported above the bars.

Diarrhoea was reported in nearly all cases. In 106 diarrhoeic cases specifying also the type of diarrhoea, 60 (56.6%) reported it was persistent whereas 46 (43.4%) stated it was intermittent. In 91 cases also reporting the daily number of visits to the toilet due to diarrhoea, the median was four visits (range one to 50). Frequently reported symptoms included flatulence, abdominal pain, tiredness and loss of appetite. Also weight loss was reported by more than half of the cases and in 56 cases that experienced it and reported the information the loss consisted in a median of 3.2 kg (range 0.4-10). Overall, the cases experienced a median of six symptoms (range one to nine). The median duration (in days) of the symptoms in cases that reported this information is shown in **Table 3.1**.

Table 3.1: Reported duration (median number of days) of the symptoms reported in the 118 symptomatic cases. Blood in stools was excluded due to the low number of cases reporting it (n=3).

Symptom	No. valid (% missing)*	Median no. days (range)
diarrhoea	55 (51.7)	9 (1-73)
flatulence	41 (52.3)	14 (1-73)
abdominal pain	39	12 (2-73)
tiredness	34	13 (2-73)
loss of appetite	32	12.5 (5-30)
weight loss	18	14 (2-73)
swollen stomach	22	11 (2-73)
vomiting	25	4 (1-16)
fever	23	6 (1-28)

*percentage of cases that did not report the duration over the number of cases reporting the symptom

Out of 113 cases answering the question, 37 (32.7%) felt still ill at the moment of filling the questionnaire. In 62 cases that were no longer ill and that reported the length of illness, the median was 16 days (range 4-73). In 30 cases that reported to still feel ill when they filled the questionnaire and in which the length of illness could be estimated as the number of days from the date of reported illness onset to the date the questionnaire was signed, the median was 27 days (range seven to 75). Hospitalisation due to the illness was reported by 11 cases out of 115 answering the question (9.6%), and in nine cases reporting also this information hospitalisation lasted for a median of four days (range 0-16). Out of 107 cases answering the question, 11 (10.3%) reported suffering from irritable bowel syndrome (IBS) and one (0.9%) from inflammatory bowel disease (IBD). Six cases declared having suffered from giardiasis in the past: one case was a one year old male, whereas five were females (ranging in age from two to 75 years). A 35 years old female reported she suffered from giardiasis three times.

3.4.4 Household clusters description

In total, five distinct household clusters were identified involving 13 cases. By considering the date of reported illness (whenever available), nine out of 13 cases were identified to be either primary or co-primary cases. The household clusters are described in detail in **Table 3.2**.

Table 3.2: Characteristics and exposure profile of the 13 cases that were part of the five household clusters identified in the case-control study. The suspected primary or co-primary cases are highlighted in bold.

Cluster (no. cases involved)	**Age, gender (family role)**	**Reported date of illness onset**	**Suspected reported exposure(s)**	**Other suspicious exposures or information reported by the case(s)**
1 (3)	**1y, male (son)**	**08/03/2012**	Went to the same swimming pool	Other two family members (mother and the sister) were tested for *Giardia* but were negative; the 1y old son was reported suffering from *Giardia* first
	6y, male (son)	asymptomatic		
	33y, male (father)	missing information	Changed nappies to the youngest son	
2 (2)	**31y, female (mother)**	**24/07/12**	Went to swimming pool together; the mother changed nappies to the son	The mother reported she suffered from *Giardia* first
	3y, male (son)	27/08/12		
3 (4)	**1y, male (son)**	**15/10/2012**	Went to the same swimming pool	The two sons and the father reported touching the brother-in-law dog, that was also reported being sick at the time; tap water reported having an unusual taste of chlorine two weeks before illness
	3y, male (son)	**16/10/2012**		
	33y, male (father)	**14/10/2012**	Went to swimming pool with the sons; changed nappies to the son	
	34y, female (mother)	8/11/2012	Changed nappies to the son	
4 (2)	**1y, male (son)**	**08/01/2013**	Went to the swimming pool and used a Jacuzzi; the household tap water was reported going off and was cloudy when it came back.	Another group of people was at swimming pool but nobody got ill (unclear how the mother got this information)
	30y, female (mother)	**7/01/2013**		
5 (2)	**42y, female (wife)**	**31/12/2012**	Travelled to India from 27/12 to 7/01; drank tap water and ate in restaurants while in India.	Symptoms reported to have started while the cases were still in India
	64y, male (husband)	**30/12/2012**		

3.4.5 Recruited controls and health details

Out of a total of 1180 controls that were contacted, 253 returned the questionnaire (21.4% control response rate) and were recruited as potentially eligible. Of these, 115 (45.4%) were recruited from Central Lancashire, 77 (30.4%) from East Lancashire and 19 (7.5%) from Greater Manchester. Two other controls (0.8%) were from West Lancashire. Health details were available from 250 controls. A total of 24 (9.6%) reported suffering from diarrhoea in the two weeks before filling the questionnaire. Blood in stools, vomiting and abdominal pain were reported by four (1.6%), eight (3.2%) and 34 (13.6%) controls respectively. Furthermore 21 controls (8.4%) reported suffering from a diarrhoeal illness lasting more than three days in the previous year, and four (1.6%) stated they suffered from a *Giardia* infection previously. Of these four controls two reported that it happened more than six months before, whereas one reported it was between three and six months.

3.4.6 Case-control matching

Out of a total 376 recruited participants, 139 (37%) were not matched (e.g. they lacked an associated participant matched by gender and age). The lack of matched participants involved 30.9% of the cases (38 out of 123) and 39.9% of the controls (101 out of 253).

3.4.7 Eligibility for risk factor analysis

A total of 32 records (five cases and 27 controls) were considered not eligible and excluded from the analysis (**Figure 3.4**). Five cases were excluded from the risk factor analysis because they mistakenly sent and returned the wrong questionnaire (three cases) or because they were asymptomatic (two cases). Twenty four controls reporting suffering from diarrhoea in the previous two weeks and three controls with the entire health details section of the questionnaire missing were excluded.

Six cases (four from Wyre, one from Fylde and one from Lancaster) were resident outside the study areas. Of the four cases from Wyre, one had three matched controls from South Ribble, one had two matched controls from South Ribble and Preston respectively and two had no matched controls. The case from Fylde had two matched

controls from Chorley and Burnley respectively, whereas the case from Lancaster had three matched controls from Chorley, South Ribble and Preston respectively. The six cases resident outside the study areas were considered eligible for the analysis since they were symptomatic, they were correctly diagnosed at the Royal Preston Hospital and they were matched with controls that were correctly recruited from Central and East Lancashire. Two controls were recruited from outside the study areas (both from West Lancashire). They were considered eligible for the analysis because they were matched with two cases from South Ribble and Preston, respectively.

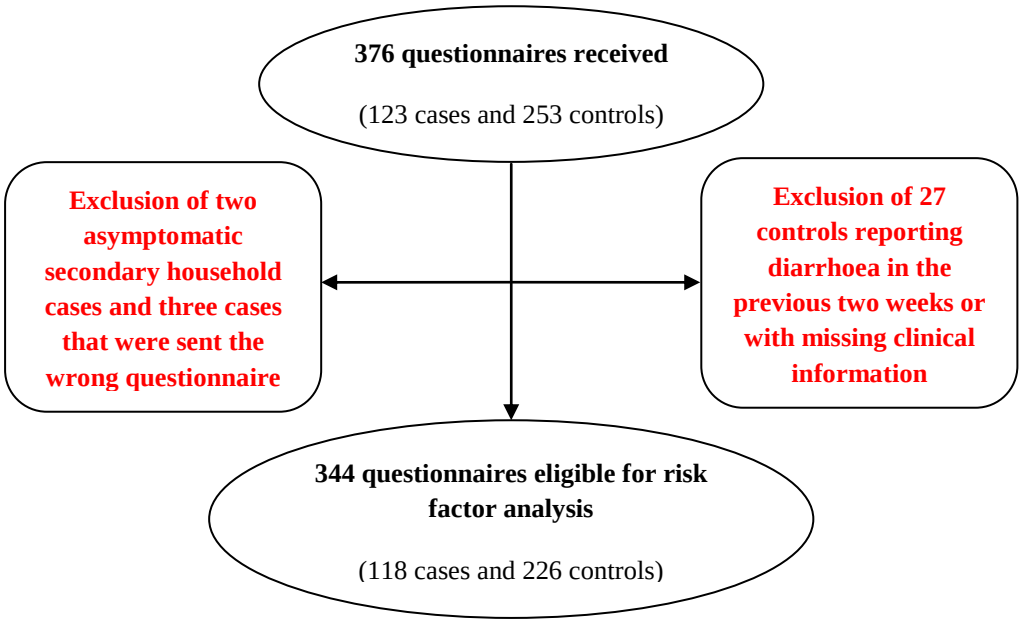


Figure 3.4: Flowchart illustrating the eligibility of case and control records for the inclusion in the risk factor analysis.

A total of 344 questionnaires were considered valid for the risk factors analysis, corresponding to a case-control ratio of 1 case per 1.9 controls. Due to the high number of un-matched questionnaires and the potential negative effect on the study statistical power associated with the exclusion of such records, it was decided to include all of them and to use statistical methods suitable for an unmatched analysis. The multivariable logistic regression models were tested both including and excluding the un-matched records to evaluate the effect of the exclusion on the model statistical power and significance of the predictor variables.

3.4.8 Risk factors for giardiasis

The risk factors for giardiasis were investigated at two levels: in all study participants and in the subgroup of those that did not travel abroad in the three weeks prior to illness (e.g. indigenous).

3.4.8.1 Univariable analysis

Socio-demographic, clinical and exposure variables that showed an association with giardiasis with a *p*-value≤0.2 in either all participants or the indigenous subgroup following univariable analysis are reported in **Table 3.3**. The results of the remaining variables (*p*>0.2) are shown in **Table 3.3.1**, **Appendix 3**.

Socio-demographics, area variables and season

The socio-demographic characteristics of cases and controls, as well as their distribution in relation to the study catchment areas and season, were comparable (**Table 3.3.1, Appendix 3**). No association was found between the risk of illness and either age (age as a categorical variable with 0-4y as reference: *p*=0.230 and 0.330 for all participants and indigenous only, respectively) or gender (*p*=0.508 and 0.191 for all participants and indigenous only, respectively), although a slightly higher proportion of male cases was observed in the group that did not travel abroad. No association was found between the risk of giardiasis and being of white ethnicity (*p*=0.231 and 0.993 for all participants and indigenous only respectively, with whites being the most represented ethnic group), occupation (*p*=0.908 and 0.965 for all participants and indigenous only, respectively), educational status (*p*=0.358 and 0.595 for all participants and indigenous only, respectively) or the level of neighbourhood deprivation (deprivation as categorical variable with Very low deprivation as reference: *p*=0.876 and 0.817 for all participants and indigenous only, respectively). Furthermore, giardiasis was not found to be associated with either the reported type of area of living (City as reference category: *p*=0.373 and 0.737 for all participants and indigenous only, respectively) or with any of the three study areas (Central Lancashire as reference category: *p*=0.738 and 0.162 for all participants and indigenous only, respectively) or with season (Winter as reference category: *p*=0.753 and 0.786 for all participants and indigenous only, respectively).

Clinical variables

A higher frequency of using medicines for indigestion in the exposure window was observed in the cases (31-32%) compared to controls (22-23%) (OR 1.64, *p*=0.052 and OR 1.52, *p*=0.148 for all participants and indigenous only, respectively). Although not significantly, irritable bowel syndrome (IBS) was reported slightly more frequently by the cases (10-11% compared with around 6% in controls) (OR 1.64, *p*=0.244 and OR 1.99, *p*=0.131 for all participants and indigenous only, respectively). Further analysis done amongst the cases showed that reporting irritable bowel syndrome was significantly and positively associated with reporting taking medicines for indigestion (all participants: OR 3.13, 95% CI 1.36-7.18, *p*=0.005; indigenous only: OR 7.86, 95% CI 1.74-35.39, *p*=0.007).

Exposure variables

None of the occupations potentially at-risk for giardiasis was reported more frequently in either the cases or controls, but a detailed analysis was prevented by the fact that a negligible number of participants reported any of the occupational exposures investigated (**Table 3.3.1, Appendix 3**).

Although overall cases reported owning a pet slightly less frequently (38%) than controls (47%) (OR 0.69, *p*=0.113), cases and controls did not differ in the number or type of pets they owned. No difference between cases and controls was found in the frequency of touching or cleaning the pets' faeces or in reporting pets with diarrhoea in the household (**Table 3.3.1, Appendix 3**). Overall, cases reported less frequently touching pets (49%) than controls (63%) (OR 0.56, *p*=0.016) and visiting a premise with animals other than a farm, wildlife park or zoo (reported by 1.2-1.8% of cases and by more than 6% of controls) (OR 0.27, *p*=0.067 and OR 0.16, *p*=0.075 for all participants and indigenous only, respectively). No cases reporting touching any animal while visiting these premises. In all participants, visiting or working at a wildlife park or zoo was more frequently reported by the cases (8.6% compared to 3.6% of the controls) (OR 2.56, *p*=0.048). No difference between cases and controls was found in the frequency of visiting or working at a farm (and touching animals

while at this type of premise), or in the frequency of touching animals in the wild (**Table 3.3.1, Appendix 3**).

The frequency of travelling abroad was significantly higher in the cases (25% compared to 8% of the controls) (OR 3.67, $p<0.001$), and the positive association with giardiasis was even stronger when considering travel to an at-risk destination (e.g. Middle East and Asia including Turkey, Africa including Canary Islands, Central and South America) (OR 22.87, $p<0.001$).

A significant positive association was found between giardiasis and reporting swimming at a swimming pool (reported by the 35-33% of the cases compared with 19-22% of the controls) (OR 1.93, $p=0.009$ and OR 2.08, $p=0.011$ for all participants and indigenous only, respectively). However, neither a dose-response effect in relation to the frequency of swimming nor a significant effect of immersing the head while swimming *per se* was found. Using a Jacuzzi or a hot tub too was associated with an increased risk of giardiasis (reported by the 13-16% of the cases compared with 6% of the controls) (OR 2.09, $p=0.064$ and OR 2.91, $p=0.011$ for all participants and indigenous only, respectively). Further analysis done amongst the cases showed that reporting using a Jacuzzi or a hot tub was significantly and positively associated with reporting going to a swimming pool (OR 24.76, 95% CI 8.05-76.16, $p<0.001$). Although the frequency of swimming in a lake, river or pond or swimming at sea were both not significantly different between cases and controls *per se* (**Table 3.3.1**, **Appendix 3**), all the cases and none of the controls reported immersing their head during swimming. Practising watersports in fresh water or at sea were reported by none and only one of the cases, respectively.

Several outdoor activities were reported less frequently in the cases than in controls, including going picnicking in indigenous cases (4% compared with 10% of the indigenous controls) (OR 0.35, $p=0.084$), practising fieldsports (5-6% of cases and 12-13% of controls) (OR 0.50, $p=0.113$ and OR 0.35, $p=0.052$ for all participants and indigenous only, respectively) and particularly doing gardening (26-31% of cases and 48-47% of controls) (OR 0.38, $p<0.001$ and OR 0.51, $p=0.015$ for all participants and indigenous only, respectively). No difference was found between the cases and the controls in the frequency of reporting going camping or caravanning.

No cases and only two controls reported getting their drinking water from a private supply, with the vast majority having a mains water supply. Although not significantly, indigenous cases reported drinking slightly more frequently un-boiled water from the tap (95%) than indigenous controls (90%) (OR 2.27, p=0.135 compared with OR 1.06, p=0.870 when all participants were analysed). There was no evidence of a dose-response relationship between the number of glasses of tap water and the risk of giardiasis. A nearly significant positive association was found between the risk of illness and drinking bottled water but only when all participants were included (OR 1.54, p=0.066 and OR 1.09, p=0.754 for all participants and indigenous only, respectively). Further analysis done amongst the cases showed that drinking bottled water was significantly and positively associated with travelling abroad (OR 3.82, 95% CI 1.86-7.83, p<0.001). A tendency towards a positive association between giardiasis and reporting the discoloured water coming from the tap was found but it was nearly significant only when all participants were included (OR 3.01, p=0.095 and OR 2.51, p=0.237 for all participants and indigenous only, respectively). A similar tendency was observed also for reporting the water from the tap having an unusual taste (OR 3.33, p=0.127 and OR 3.36, p=0.199 for all participants and indigenous only, respectively). Further analysis done amongst the cases showed that reporting the water from the tap having an unusual taste was significantly and positively associated with reporting the tap water being discoloured (OR 27.43, 95% CI 5.45-137.96, p=0.001). Drinking un-boiled water from a lake, river or stream was not associated with an increased risk of giardiasis (**Table 3.3.1**, **Appendix 3**).

The potential for foodborne transmission was investigated for the consumption of fresh products (e.g. food items more at risk as vehicles of parasite cysts) at two levels. Participants were asked about both their food consumption habits and the food eaten specifically during the exposure window in order to test separately the effect of continuous (habitual consumption) and point exposure (consumption specifically in the exposure window). The risk of giardiasis was significantly and negatively associated with the weekly frequency of consumption of either salad or raw vegetables (OR 0.90, p=0.038 and OR 0.87, p=0.026 for all participants and indigenous only, respectively) and raw fruit (OR 0.91, p=0.008 and OR 0.87, p=0.002 for all participants and indigenous only, respectively). Conversely, no

association was found in the weekly frequency of consumption of cooked vegetables. Cases and controls did not differ significantly in their habit of either peeling or washing raw fruit before consumption. Amongst the food items consumed in the three weeks prior to illness, cases showed a significantly lower frequency of consumption of salad or raw vegetables (81%) compared to controls (91%) (OR 0.41, p=0.008 and OR 0.41, p=0.015 for all participants and indigenous only, respectively). Also the consumption of shellfish was significantly and negatively associated with illness (OR 0.51, p=0.035 and OR 0.32, p=0.007 for all participants and indigenous only, respectively). Cases and controls did not differ in the consumption of cooked vegetables, fruit juice, chicken or fish. Consumption of meat items as beef, lamb and pork seemed to be slightly more frequently reported by cases but none of these differences was significant. Differences in the provenience of the food items consumed in the three weeks prior to illness were observed between cases and controls. Overall, cases reported less frequently than controls eating salad, raw or cooked vegetables and fruit juice purchased from a supermarket, and conversely they reported more frequently than controls eating the same food items and some meat items (lamb, pork and chicken) from other places (nearly all restaurants while being abroad) (**Table 3.3**). In indigenous participants, a tendency towards a negative association between illness and the number of times eating at a pub or a restaurant (OR 0.84, 95% CI 0.70-1.01, p=0.066) or at a barbecue (OR 0.40, 95% CI 0.14-1.12, p=0.081). Eating at a barbecue *per se* was marginally negatively associated with illness in indigenous participants (OR 0.32, 95% CI 0.09-1.12, p=0.062). No differences were found between cases and controls in their habits of eating at a canteen, take-away of fast food.

In the cases' households the number of children was significantly higher than in controls' households (OR 1.26, 95% CI 1.05-1.51, p=0.013 and OR 1.28, 95% CI 1.05-1.56, p=0.014 for all participants and indigenous only, respectively). An association was consistently found between the risk of illness and the presence in the household of children below five years of age and associated exposures, consistently in both all participants and in indigenous only although differences were more pronounced in the latter group. Cases reported more frequently than controls having children attending a nursery or a playgroup in the household, having children in nappies and changing nappies (**Table 3.3**).

Table 3.3: Variables showing an association with giardiasis with a *p*-value ≤0.2 in univariable analysis. Results of both the general risk factor analysis (e.g. including all the eligible118 cases and 226 controls) and the indigenous risk factor analysis (e.g. including only the 86 cases and 207 controls that did not report travelling abroad in the exposure window) are shown for comparison.

Variable	Data subset	No. valid (% missing)*	Category	Cases n (%)**	Controls n (%)**	OR (95% CI)***	*p*-value
HEALTH DETAILS							
Taking any medicine for indigestion	All	341 (0.9)	No	79 (68.1)	175 (77.8)	Ref.	0.052
			Yes	37 (31.9)	50 (22.2)	1.64 (0.99-2.71)	
	Indigenous	290 (1)	No	58 (69)	159 (77.2)	Ref.	0.148
			Yes	26 (31)	47 (22.8)	1.52 (0.86-2.67)	
Suffering from irritable bowel syndrome (IBS)	All	320 (7)	No	96 (89.7)	199 (93.4)	Ref.	0.244
			Yes	11 (10.3)	14 (6.6)	1.63 (0.71-3.72)	
	Indigenous	273 (6.8)	No	69 (88.5)	183 (93.8)	1.00	0.131
			Yes	9 (11.5)	12 (6.2)	1.99 (0.80-4.93)	
ANIMAL CONTACT							
Keeping a pet (any type)	All	342 (0.6)	No	72 (62.1)	120 (53.1)	Ref.	0.113
			Yes	44 (37.9)	106 (46.9)	0.69 (0.44-1.09)	
	Indigenous	291 (0.7)	No	50 (59.5)	108 (52.2)	Ref.	0.254
			Yes	34 (40.5)	99 (47.8)	0.74 (0.44-1.24)	
No. of cats kept	All	339 (1.5)	0	103 (88.7)	182 (81.2)	Ref.	$0.106^{T}/0.135^{C}$
			1	9 (7.8)	30 (13.4)	0.53 (0.24-1.16)	
			2	3 (2.6)	10 (4.5)	0.53 (0.14-1.97)	
			3+	1 (0.9)	2 (0.9)	0.88 (0.08-9.86)	
	Indigenous	289 (1.4)	0	73 (86.9)	167 (81.5)	1.00	$0.325^{T}/0.389^{C}$
			1	8 (9.5)	28 (13.7)	0.65 (0.28-1.50)	
			2	2 (2.4)	9 (4.4)	0.51 (0.11-2.41)	
			3+	1 (1.2)	1 (0.5)	2.28 (0.14-37.07)	
Touching any pet (either own or other people's)	All	314 (8.7)	No	55 (51.4)	77 (37.2)	Ref.	0.016
			Yes	52 (48.6)	130 (62.8)	0.56 (0.35-0.90)	
	Indigenous	269 (8.2)	No	36 (45.6)	74 (38.9)	Ref.	0.314
			Yes	43 (54.4)	116 (61.1)	0.76 (0.45-1.29)	
Visiting or working at a wildlife park or	All	341 (0.9)	No	106 (91.4)	217 (96.4)	Ref.	0.048
			Yes	10 (8.6)	8 (3.6)	2.56 (0.98-6.67)	

zoo	Indigenous	291 (0.7)	No	79 (92.9)	198 (96.1)	Ref.	0.245
			Yes	6 (7.1)	8 (3.9)	1.88 (0.63-5.59)	
Visiting any other premise(s) with animals	All	339 (1.5)	No	112 (98.2)	211 (93.8)	Ref.	0.067
			Yes	2 (1.8)	14 (6.2)	0.27 (0.06-1.20)	
	Indigenous	290 (1)	No	83 (98.8)	192 (93.2)	Ref.	0.075
			Yes	1 (1.2)	14 (6.8)	0.16 (0.02-1.28)	
Touching animals while on any other premise(s)	All	339 (1.5)	Not visiting any other premise(s)	112 (98.2)	212 (94.2)	Ref.	0.034
			Visiting but not touching any animal	2 (1.8)	5 (2.2)	0.76 (0.14-3.96)	
			Visiting and touching an animal	0 (0)	8 (3.6)	n/a****	
	Indigenous	290 (1)	Not visiting any other premise(s)	83 (98.8)	193 (93.7)	Ref.	0.047
			Visiting but not touching any animal	1 (1.2)	5 (2.4)	0.46 (0.05-4.04)	
			Visiting and touching an animal	0 (0)	8 (3.9)	n/a****	
TRAVEL DETAILS							
Travelling abroad (outside the UK)	All	341 (0.9)	No	86 (74.8)	207 (91.6)	Ref.	<0.001
			Yes	29 (25.2)	19 (8.4)	3.67 (1.95-6.90)	
Travelling abroad to an at-risk destination	All	341 (0.9)	Not travelling abroad	86 (74.8)	207 (91.6)	Ref.	<0.001
			Travelling abroad to a not at-risk destination	10 (8.7)	17 (7.5)	1.41 (0.62-3.22)	
			Travelling abroad to any at-risk *destination (e.g. Middle East and Asia including Turkey, Africa including Canary Islands, Central and South America)*	19 (16.5)	2 (0.9)	22.87 (5.21-100.31)	
Travelling in the UK (England, Wales, Scotland)	All	337 (2)	No	90 (79.6)	155 (69.2)	Ref.	0.042
			Yes	23 (20.4)	69 (30.8)	0.57 (0.33-0.98)	
Travelling only in the UK, only abroad or both	All	335 (2.9)	Not travelling at all	64 (57.7)	146 (65.2)	Ref.	<0.001
			Only travelling in the UK	20 (18)	60 (26.8)	0.76 (0.42-1.36)	
			Only travelling abroad	24 (21.6)	9 (4)	6.08 (2.68-13.82)	
			Travelling both in the UK and abroad	3 (2.7)	9 (4)	0.76 (0.20-2.90)	
RECREATIONAL ACTIVITIES							
Swimming or paddling in a swimming pool	All	341 (0.9)	No	75 (65.2)	177 (78.3)	Ref.	0.009
			Yes	40 (34.8)	49 (21.7)	1.93 (1.17-3.17)	
	Indigenous	293	No	58 (67.4)	168 (81.2)	Ref.	0.011
			Yes	28 (32.6)	39 (18.8)	2.08 (1.18-3.68)	
Immersing the head	All	320 (7)	Not going to swimming pool	75 (72.1)	177 (81.9)	Ref.	0.025

underwater while swimming or paddling in a swimming pool			Swimming or paddling without immersing the head	8 (7.7)	4 (1.9)	4.72 (1.38-16.15)	
			Swimming or paddling immersing the head	21 (20.2)	35 (16.2)	1.42 (0.77-2.59)	
	Indigenous	276 (5.8)	Not going to swimming pool	58 (75.3)	168 (84.4)	Ref.	0.033
			Swimming or paddling without immersing the head	7 (9.1)	4 (2)	5.07 (1.43-17.95)	
			Swimming or paddling immersing the head	12 (15.6)	27 (13.6)	1.29 (0.61-2.70)	
Using a Jacuzzi or a hot tub	All	320 (7)	No	90 (87.4)	203 (93.5)	Ref.	0.064
			Yes	13 (12.6)	14 (6.5)	2.09 (0.95-4.64)	
	Indigenous	274 (6.5)	No	64 (84.2)	186 (93.9)	Ref.	0.011
			Yes	12 (15.8)	12 (6.1)	2.91 (1.24-6.79)	
Immersing the head underwater while swimming in a lake, pond or river	All	314 (8.7)	Not swimming in a lake, pond or river	99 (97.1)	209 (98.6)	Ref.	0.010
			Swimming without immersing the head	0 (0)	3 (1.4)	n/a****	
			Swimming immersing the head	3 (2.9)	0 (0)	n/a****	
	Indigenous	271 (7.5)	Not swimming in a lake, pond or river	75 (98.7)	192 (98.5)	Ref.	0.104
			Swimming without immersing the head	0 (0)	3 (1.5)	n/a****	
			Swimming immersing the head	1 (1.3)	0 (0)	n/a****	
Immersing the head underwater while swimming in the sea	All	316 (8.1)	Not swimming in the sea	95 (95)	206 (95.4)	Ref.	0.034
			Swimming without immersing the head	0 (0)	6 (2.8)	n/a****	
			Swimming immersing the head	5 (5)	4 (1.9)	2.71 (0.71-10.32)	
	Indigenous	275 (6.1)	Not swimming in the sea	74 (98.7)	197 (98.5)	1.00	0.420
			Swimming without immersing the head	0 (0)	2 (1)	n/a****	
			Swimming immersing the head	1 (1.3)	1 (0.5)	2.66 (0.16-43.11)	
Going picnicking	All	318 (7.6)	No	102 (93.6)	188 (90)	Ref.	0.279
			Yes	7 (6.4)	21 (10)	0.61 (0.25-1.49)	
	Indigenous	272 (7.2)	No	78 (96.3)	172 (90.1)	Ref.	0.084
			Yes	3 (3.7)	19 (9.9)	0.35 (0.10-1.21)	
Frequency of going picnicking (no. times)	All	312 (9.3)	0	102 (96.2)	188 (91.3)	Ref.	$0.060^{T}/0.145^{C}$
			1-2	4 (3.8)	17 (8.3)	0.43 (0.14-1.32)	
			3+	0 (0)	1 (0.5)	n/a****	
	Indigenous	268 (8.5)	0	78 (97.5)	172 (91.5)	1.00	$0.040^{T}/0.152^{C}$
			1-2	2 (2.5)	15 (8)	0.29 (0.07-1.32)	
			3+	0 (0)	1 (0.5)	n/a****	
Practising fieldsports	All	319 (7.3)	No	100 (93.5)	186 (87.7)	Ref.	

(e.g. football, golf etc)			Yes	7 (6.5)	26 (12.3)	0.50 (0.21-1.19)	0.113
	Indigenous	273 (6.8)	No	76 (95)	168 (87)	Ref.	0.052
			Yes	4 (5)	25 (13)	0.35 (0.12-1.05)	
Frequency of practising fieldsports (no. times)	All	316 (8.1)	0	100 (94.3)	186 (88.6)	Ref.	$0.102^{T}/0.117^{C}$
			1-2	1 (0.9)	8 (3.8)	0.23 (0.03-1.88)	
			3-4	3 (2.8)	6 (2.9)	0.93 (0.23-3.80)	
			5-6	2 (1.9)	6 (2.9)	0.62 (0.12-3.13)	
			7+	0 (0)	4 (1.9)	n/a****	
	Indigenous	270 (7.8)	0	76 (96.2)	168 (88)	1.00	$0.080^{T}/0.131^{C}$
			1-2	0 (0)	8 (4.2)	n/a****	
			3-4	1 (1.3)	6 (3.1)	0.37 (0.04-3.11)	
			5-6	2 (2.5)	5 (2.6)	0.88 (0.17-4.66)	
			7+	0 (0)	4 (2.1)	n/a****	
Frequency of walking in the countryside (no. times)	All	293 (14.8)	0	69 (77.5)	128 (62.7)	Ref.	$0.076^{T}/0.081^{C}$
			1-2	7 (7.9)	38 (18.6)	0.34 (0.14-0.81)	
			3-4	7 (7.9)	19 (9.3)	0.68 (0.27-1.71)	
			5-6	3 (3.4)	6 (2.9)	0.93 (0.22-3.82)	
			7+	3 (3.4)	13 (6.4)	0.43 (0.12-1.55)	
	Indigenous	250 (14.7)	0	50 (78.1)	120 (64.5)	Ref.	$0.088^{T}/0.102^{C}$
			1-2	4 (6.3)	30 (16.1)	0.32 (0.107-0.96)	
			3-4	7 (10.9)	19 (10.2)	0.88 (0.35-2.23)	
			5-6	2 (3.1)	6 (3.2)	0.80 (0.16-4.10)	
			7+	1 (1.6)	11 (5.9)	0.22 (0.03-1.73)	
Doing gardening	All	324 (5.8)	No	79 (73.8)	112 (51.6)	Ref.	<0.001
			Yes	28 (26.2)	105 (48.4)	0.38 (0.23-0.63)	
	Indigenous	280 (4.4)	No	56 (69.1)	106 (53.3)	Ref.	0.015
			Yes	25 (30.9)	93 (46.7)	0.51 (0.29-0.88)	
Frequency of doing gardening (no.times)	All	283 (17.7)	0	79 (87.8)	112 (58)	Ref.	$<0.001^{T}/0.032^{C}$
			1-2	5 (5.6)	31 (16.1)	0.23 (0.08-0.61)	
			3-4	1 (1.1)	27 (14)	0.05 (0.01-0.39)	
			5-6	2 (2.2)	11 (5.7)	0.26 (0.06-1.19)	
			7+	3 (3.3)	12 (6.2)	0.35 (0.10-1.30)	
	Indigenous	240 (18.1)	0	56 (87.5)	106 (60.2)	1.00	$0.001^{T}/0.163^{C}$
			1-2	4 (6.3)	26 (14.8)	0.29 (0.10-0.88)	

			3-4	1 (1.6)	26 (14.8)	0.07 (0.01-0.55)	
			5-6	0 (0)	9 (5.1)	n/a****	
			7+	3 (4.7)	9 (5.1)	0.63 (0.16-2.42)	
WATER CONSUMPTION							
Household water supply	All	342 (0.6)	Mains water supply	115 (98.3)	223 (99.1)	Ref.	0.051
			Private supply (e.g. spring or well)	0 (0)	2 (0.9)	n/a****	
			Mixed	2 (1.7)	0 (0)	n/a****	
Drinking un-boiled water straight from the tap	All	339 (1.5)	No	12 (10.5)	25 (11.1)	Ref.	0.870
			Yes	102 (89.5)	200 (88.9)	1.06 (0.51-2.20)	
	Indigenous	290 (1)	No	4 (4.8)	21 (10.2)	Ref.	0.135
			Yes	80 (95.2)	185 (89.8)	2.27 (0.75-6.83)	
Drinking bottled water	All	324 (5.8)	No	48 (43.2)	115 (54)	Ref.	0.066
			Yes	63 (56.8)	98 (46)	1.54 (0.97-2.44)	
	Indigenous	275 (6.1)	No	43 (53.1)	107 (55.2)	Ref.	0.754
			Yes	38 (46.9)	87 (44.8)	1.09 (0.65-1.83)	
Water from the tap reported having an unusual taste	All	335 (2.6)	No	109 (95.6)	218 (98.6)	Ref.	0.127
			Yes	5 (4.4)	3 (1.4)	3.33 (0.78-14.21)	
	Indigenous	285 (2.7)	No	79 (95.2)	199 (98.5)	Ref.	0.199
			Yes	4 (4.8)	3 (1.5)	3.36 (0.73-15.35)	
Water from the tap reported being discoloured	All	335 (2.6)	No	108 (94.7)	217 (98.2)	Ref.	0.095
			Yes	6 (5.3)	4 (1.8)	3.01 (0.83-10.90)	
	Indigenous	285 (2.7)	No	79 (95.2)	198 (98)	Ref.	0.237
			Yes	4 (4.8)	4 (2)	2.51 (0.61-10.27)	
FOOD CONSUMPTION							
FOOD CONSUMPTION HABITS							
No. times per week eating salads or raw vegetables	All	329 (4.4)	0	22 (19.8)	34 (15.6)	Ref.	$0.062^{T}/0.038^{C}$
			1-2	41 (36.9)	71 (32.6)	0.89 (0.46-1.73)	
			3-4	28 (25.2)	54 (24.8)	0.80 (0.39-1.62)	
			5+	20 (18)	59 (27.1)	0.52 (0.25-1.10)	
	Indigenous	280 (4.4)	0	18 (22.2)	32 (16.1)	Ref.	$0.060^{T}/0.026^{C}$
			1-2	31 (38.3)	68 (34.2)	0.81 (0.40-1.66)	
			3-4	19 (23.5)	51 (25.6)	0.66 (0.30-1.45)	
			5+	13 (16)	48 (24.1)	0.48 (0.21-1.12)	
No. times per week	All	328 (4.7)	0	14 (12.6)	15 (6.9)	Ref.	$0.011^{T}/0.008^{C}$

eating raw fruit			1-2	16 (14.4)	27 (12.4)	0.63 (0.24-1.65)	
			3-4	24 (21.6)	29 (13.4)	0.89 (0.36-2.20)	
			5+	57 (51.4)	146 (67.3)	0.42 (0.19-0.92)	
	Indigenous	280 (4.4)	0	13 (15.9)	15 (7.6)	Ref.	$0.007^{T}/0.002^{C}$
			1-2	12 (14.6)	27 (13.6)	0.51 (0.19-1.40)	
			3-4	19 (23.2)	28 (14.1)	0.78 (0.30-2.01)	
			5+	38 (46.3)	128 (64.6)	0.34 (0.15-0.78)	
FOOD CONSUMPTION DURING THE EXPOSURE WINDOW							
Eating salads or raw vegetables	All	333 (3.2)	No	21 (18.6)	19 (8.6)	Ref.	0.008
			Yes	92 (81.4)	201 (91.4)	0.41 (0.21-0.81)	
	Indigenous	284 (3.1)	No	16 (19.3)	18 (9)	Ref.	0.015
			Yes	67 (80.7)	183 (91)	0.41 (0.20-0.85)	
Provenience of salads or raw vegetables	All	327 (4.9)	Did not eat salad or raw vegetables	21 (19.3)	19 (8.7)	Ref.	<0.001
			Market	7 (6.4)	8 (3.7)	0.79 (0.24-2.60)	
			Greengrocers	4 (3.7)	7 (3.2)	0.52 (0.13-2.05)	
			Supermarket	63 (57.8)	141 (64.7)	0.40 (0.20-0.80)	
			Homegrown	0 (0)	2 (0.9)	n/a****	
			Other place (e.g. restaurant etc.)	7 (6.4)	1 (0.5)	6.33 (0.71-56.32)	
			Multiple proveniences (at least two different)	7 (6.4)	40 (18.3)	0.16 (0.06-0.44)	
	Indigenous	280 (4.4)	Did not eat salad or raw vegetables	16 (19.8)	18 (9)	Ref.	0.044
			Market	2 (2.5)	8 (4)	0.28 (0.05-1.52)	
			Greengrocers	4 (4.9)	7 (3.5)	0.64 (0.16-2.61)	
			Supermarket	50 (61.7)	135 (67.8)	0.42 (0.20-0.88)	
			Homegrown	0 (0)	1 (0.5)	n/a****	
			Other place (e.g. restaurant etc.)	3 (3.7)	1 (0.5)	3.37 (0.32-35.79)	
			Multiple proveniences (at least two different)	6 (7.4)	29 (14.6)	0.23 (0.08-0.70)	
Provenience of cooked vegetables	All	320 (7)	Did not eat cooked vegetables	8 (7.6)	14 (6.5)	Ref.	0.030
			Market	6 (5.7)	11 (5.1)	0.95 (0.25-3.57)	
			Greengrocers	6 (5.7)	8 (3.7)	1.31 (0.33-5.16)	
			Supermarket	65 (61.9)	142 (66)	0.80 (0.32-2.00)	
			Homegrown	0 (0)	3 (1.4)	n/a****	
			Other place (e.g. restaurant etc.)	8 (7.6)	2 (0.9)	7.00 (1.18-41.36)	
			Multiple proveniences (at least two different)	12 (11.4)	35 (16.3)	0.60 (0.20-1.78)	
	Indigenous	274 (6.5)	Did not eat cooked vegetables	5 (6.4)	13 (6.6)	Ref.	0.247

			Market	2 (2.6)	10 (5.1)	0.52 (0.08-3.26)	
			Greengrocers	6 (7.7)	8 (4.1)	1.95 (0.44-8.55)	
			Supermarket	54 (69.2)	137 (69.9)	1.02 (0.35-3.01)	
			Homegrown	0 (0)	2 (1)	n/a****	
			Other place (e.g. restaurant etc.)	4 (5.1)	2 (1)	5.20 (0.71-37.89)	
			Multiple proveniences (at least two different)	7 (9)	24 (12.2)	0.76 (0.20-2.87)	
Provenience of fruit juice	All	323 (6.1)	Did not drink fruit juice	33 (30.3)	58 (27.1)	Ref.	0.088
			Market	1 (0.9)	1 (0.5)	1.76 (0.11-29.03)	
			Greengrocers	1 (0.9)	0 (0)	n/a****	
			Supermarket	68 (62.4)	145 (67.8)	0.82 (0.49-1.38)	
			Other place (e.g. restaurant etc.)	5 (4.6)	2 (0.9)	4.39 (0.81-23.9)	
			Multiple proveniences (at least two different)	1 (0.9)	8 (3.7)	0.22 (0.03-1.83)	
	Indigenous	276 (5.8)	Did not drink fruit juice	26 (32.1)	54 (27.7)	Ref.	0.552
			Market	1 (1.2)	1 (0.5)	2.08 (0.12-34.53)	
			Greengrocers	1 (1.2)	0 (0)	n/a****	
			Supermarket	51 (63)	133 (68.2)	0.80 (0.45-1.41)	
			Other place (e.g. restaurant etc.)	1 (1.2)	2 (1)	1.04 (0.09-11.98)	
			Multiple proveniences (at least two different)	1 (1.2)	5 (2.6)	0.41 (0.05-3.74)	
Eating beef	All	327 (4.9)	No	21 (19.6)	56 (25.5)	Ref.	0.244
			Yes	86 (80.4)	164 (74.5)	1.40 (0.79-2.46)	
	Indigenous	281 (4.1)	No	14 (17.7)	54 (26.7)	Ref.	0.113
			Yes	65 (82.3)	148 (73.3)	1.69 (0.88-3.26)	
Provenience of beef	All	326 (5.2)	Did not eat beef	21 (19.8)	56 (25.5)	Ref.	0.184
			Market stall	0 (0)	4 (1.8)	n/a****	
			Butcher	14 (13.2)	28 (12.7)	1.33 (0.59-3.01)	
			Supermarket	54 (50.9)	104 (47.3)	1.38 (0.76-2.52)	
			Other place (e.g. restaurant etc.)	6 (5.7)	4 (1.8)	4.00 (1.03-15.60)	
			Multiple proveniences (at least two different)	11 (10.4)	24 (10.9)	1.22 (0.51-2.92)	
	Indigenous	280 (4.4)	Did not eat beef	14 (17.9)	54 (26.7)	Ref.	0.075
			Market stall	0 (0)	4 (2)	n/a****	
			Butcher	9 (11.5)	26 (12.9)	1.33 (0.51-3.48)	
			Supermarket	49 (62.8)	95 (47)	1.99 (1.01-3.93)	
			Other place (e.g. restaurant etc.)	0 (0)	4 (2)	n/a****	
			Multiple proveniences (at least two different)	6 (7.7)	19 (9.4)	1.22 (0.41-3.62)	

Eating lamb	All	305 (11.3)	No	43 (42.6)	100 (49)	Ref.	0.288
			Yes	58 (57.4)	104 (51)	1.30 (0.80-2.10)	
	Indigenous	264 (9.9)	No	32 (41.6)	94 (50.3)	1.00	0.198
			Yes	45 (58.4)	93 (49.7)	1.42 (0.83-2.43)	
Provenience of lamb	All	302 (12.2)	Did not eat lamb	43 (43.9)	100 (49)	Ref.	0.185
			Market stall	0 (0)	4 (2)	n/a****	
			Butcher	14 (14.3)	22 (10.8)	1.48 (0.69-3.16)	
			Supermarket	31 (31.6)	67 (32.8)	1.07 (0.62-1.88)	
			Other place (e.g. restaurant etc.)	4 (4.1)	2 (1)	4.65 (0.82-26.36)	
			Multiple proveniences (at least two different)	6 (6.1)	9 (4.4)	1.55 (0.52-4.62)	
	Indigenous	261 (10.9)	Did not eat lamb	32 (43.2)	94 (50.3)	Ref.	0.453
			Market stall	0 (0)	3 (1.6)	n/a****	
			Butcher	11 (14.9)	20 (10.7)	1.62 (0.70-3.73)	
			Supermarket	27 (36.5)	63 (33.7)	1.26 (0.69-2.30)	
			Other place (e.g. restaurant etc.)	0 (0)	1 (0.5)	n/a****	
			Multiple proveniences (at least two different)	4 (5.4)	6 (3.2)	1.96 (0.52-7.38)	
Eating pork	All	311 (9.6)	No	38 (36.9)	84 (40.4)	Ref.	0.553
			Yes	65 (63.1)	124 (59.6)	1.16 (0.71-1.88)	
	Indigenous	264 (9.9)	No	32 (41.6)	94 (50.3)	Ref.	0.198
			Yes	45 (58.4)	93 (49.7)	1.42 (0.83-2.43)	
Provenience of pork	All	308 (10.5)	Did not eat pork	38 (38)	84 (40.4)	Ref.	0.807
			Market stall	1 (1)	4 (1.9)	0.55 (0.06-5.11)	
			Butcher	7 (7)	15 (7.2)	1.03 (0.39-2.74)	
			Supermarket	45 (45)	92 (44.2)	1.08 (0.64-1.82)	
			Other place (e.g. restaurant etc.)	4 (4)	3 (1.4)	2.94 (0.63-13.8)	
			Multiple proveniences (at least two different)	5 (5)	10 (4.8)	1.10 (0.35-3.45)	
	Indigenous	264 (9.9)	Did not eat pork	27 (36)	80 (42.3)	Ref.	0.195
			Market stall	0 (0)	4 (2.1)	n/a****	
			Butcher	4 (5.3)	13 (6.9)	0.91 (0.27-3.03)	
			Supermarket	42 (56)	82 (43.4)	1.52 (0.85-2.69)	
			Other place (e.g. restaurant etc.)	0 (0)	3 (1.6)	n/a****	
			Multiple proveniences (at least two different)	2 (2.7)	7 (3.7)	0.85 (0.17-4.32)	
Provenience of chicken	All	326 (5.2)	Did not eat chicken	5 (4.8)	18 (8.1)	Ref.	0.076
			Market stall	2 (1.9)	4 (1.8)	1.80 (0.25-12.85)	

			Butcher	14 (13.5)	29 (13.1)	1.74 (0.53-5.65)	
			Supermarket	64 (61.5)	149 (67.1)	1.55 (0.55-4.34)	
			Other place (e.g. restaurant etc.)	7 (6.7)	2 (0.9)	12.6 (1.97-80.76)	
			Multiple proveniences (at least two different)	12 (11.5)	20 (9)	2.16 (0.64-7.33)	
	Indigenous	279 (4.8)	Did not eat chicken	4 (5.3)	17 (8.4)	Ref.	0.704
			Market stall	0 (0)	3 (1.5)	n/a****	
			Butcher	11 (14.5)	27 (13.3)	1.73 (0.47-6.32)	
			Supermarket	54 (71.1)	136 (67)	1.69 (0.54-5.24)	
			Other place (e.g. restaurant etc.)	1 (1.3)	2 (1)	2.12 (0.15-29.66)	
			Multiple proveniences (at least two different)	6 (7.9)	18 (8.9)	1.42 (0.34-5.91)	
Eating shellfish	All	285 (17.2)	No	79 (84)	139 (72.8)	Ref.	0.035
			Yes	15 (16)	52 (27.2)	0.51 (0.27-0.96)	
	Indigenous	245 (16.4)	No	64 (90.1)	130 (74.7)	Ref.	0.007
			Yes	7 (9.9)	44 (25.3)	0.32 (0.14-0.76)	
Provenience of shellfish	All	285 (17.2)	Did not eat shellfish	79 (84)	139 (72.8)	Ref.	0.010
			Fishmonger	1 (1.1)	3 (1.6)	0.59 (0.06-5.73)	
			Supermarket	11 (11.7)	41 (21.5)	0.47 (0.23-0.97)	
			Other place (e.g. restaurant etc.)	3 (3.2)	1 (0.5)	5.28 (0.54-51.60)	
			Multiple proveniences (at least two different)	0 (0)	7 (2.5)	n/a****	
	Indigenous	245 (16.4)	Did not eat shellfish	64 (90.1)	130 (74.7)	Ref.	0.013
			Fishmonger	0 (0)	3 (1.7)	n/a****	
			Supermarket	7 (9.9)	36 (20.7)	0.39 (0.17-0.94)	
			Multiple proveniences (at least two different)	0 (0)	5 (2.9)	n/a****	
No. times eating at a pub or a restaurant	All	326 (5.2)	0	46 (43.4)	81 (36.8)	Ref.	$0.143^{T}/0.437^{C}$
			1-2	44 (41.5)	83 (37.7)	0.93 (0.56-1.56)	
			3-4	5 (4.7)	36 (16.4)	0.24 (0.09-0.67)	
			5+	11 (10.4)	20 (9.1)	0.97 (0.43-2.20)	
	Indigenous	283 (3.4)	0	37 (45.7)	80 (39.6)	1.00	$0.043^{T}/0.066^{C}$
			1-2	38 (46.9)	79 (39.1)	1.04 (0.60-1.80)	
			3-4	3 (3.7)	32 (15.8)	0.20 (0.06-0.70)	
			5+	3 (3.7)	11 (5.4)	0.59 (0.15-2.24)	
Eating at a barbecue	All	305 (11.3)	No	91 (90.1)	180 (88.2)	Ref.	0.626
			Yes	10 (9.9)	24 (11.8)	0.82 (0.38-1.80)	
	Indigenous	262 (10.6)	No	73 (96.1)	165 (88.7)	Ref.	

			Yes	3 (3.9)	21 (11.3)	0.32 (0.09-1.12)	0.062
No. times eating at a barbecue	All	302 (12.2)	0	91 (91.9)	180 (88.7)	Ref.	$0.157^{T}/0.198^{C}$
			1-2	8 (8.1)	18 (8.9)	0.88 (0.37-2.10)	
			3-4	0 (0)	5 (2.5)	n/a****	
	Indigenous	261 (10.9)	0	73 (96.1)	165 (89.2)	1.00	$0.036^{T}/0.081^{C}$
			1-2	3 (3.9)	15 (8.1)	0.45 (0.13-1.61)	
			3-4	0 (0)	5 (2.7)	n/a****	
GENERAL HOUSEHOLD DETAILS							
No. of children (<16y) in the house	All	342 (0.6)	0	60 (51.7)	134 (59.3)	Ref.	$0.015^{T}/0.013^{C}$
			1	19 (16.4)	43 (19)	0.99 (0.53-1.83)	
			2	19 (16.4)	34 (15)	1.25 (0.66-2.36)	
			3	10 (8.6)	11 (4.9)	2.03 (0.82-5.04)	
			4+	8 (6.9)	4 (1.8)	4.47 (1.29-15.41)	
	Indigenous	293	0	42 (48.8)	121 (58.5)	Ref.	$0.017^{T}/0.014^{C}$
			1	16 (18.6)	40 (19.3)	1.15 (0.58-2.27)	
			2	14 (16.3)	31 (15)	1.30 (0.63-2.68)	
			3	7 (8.1)	11 (5.3)	1.83 (0.67-5.04)	
			4+	7 (8.1)	4 (1.9)	5.04 (1.40-18.09)	
Any child attending a childcare nursery or a playgroup	All	339 (1.5)	No children in the house	60 (53.1)	134 (59.3)	Ref.	0.054
			Children not attending nursery or playgroup	19 (16.8)	50 (22.1)	0.85 (0.46-1.56)	
			At least one child attending nursery or playgroup	34 (30.1)	42 (18.6)	1.81 (1.05-3.12)	
	Indigenous	292 (0.3)	No children in the house	42 (49.4)	121 (58.5)	Ref.	0.038
			Children not attending nursery or playgroup	15 (17.6)	47 (22.7)	0.92 (0.47-1.81)	
			At least one child attending nursery or playgroup	28 (32.9)	39 (18.8)	2.07 (1.14-3.76)	
Any child in nappies	All	338 (1.7)	No children in the house	60 (52.6)	134 (59.8)	Ref.	0.095
			At least one child but not in nappies	23 (20.2)	52 (23.2)	0.99 (0.55-1.76)	
			At least one child in nappies	31 (27.2)	38 (17)	1.82 (1.04-3.20)	
	Indigenous	291 (0.7)	No children in the house	42 (49.4)	121 (58.7)	Ref.	0.123
			At least one child but not in nappies	19 (22.4)	49 (23.8)	1.12 (0.59-2.11)	
			At least one child in nappies	24 (28.2)	36 (17.5)	1.92 (1.03-3.59)	
Changing nappies	All	335 (2.6)	No children in the house or children not in nappies	83 (74.1)	186 (83.4)	Ref.	0.064

			Child in nappies but not changing nappies	9 (8)	17 (7.6)	1.19 (0.51-2.77)	
			Child in nappies and changing nappies	20 (17.9)	20 (9)	2.24 (1.14-4.39)	
	Indigenous	290 (1)	No children in the house or children not in nappies	61 (71.8)	170 (82.9)	Ref.	0.029
			Child in nappies but not changing nappies	6 (7.1)	16 (7.8)	1.04 (0.39-2.79)	
			Child in nappies and changing nappies	18 (21.2)	19 (9.3)	2.64 (1.30-5.36)	
Any other person with diarrhoea in the house	All	321 (6.7)	No	92 (83.6)	188 (89.1)	Ref.	0.164
			Yes	18 (16.4)	23 (10.9)	1.60 (0.82-3.11)	
	Indigenous	276 (5.8)	No	66 (81.5)	175 (89.7)	Ref.	0.060
			Yes	15 (18.5)	20 (10.3)	1.99 (0.96-4.11)	

*percentages refer to the proportion of participants with missing information for the variable; **percentages refer to the proportion among cases and controls that answered the question; ***Odds ratio with 95% confidence interval; **** Odds ratio not calculated because no variation present in the data; [T] Chi-Square for trend; [C] logistic regression on the continuous form of the variable

3.4.8.2 Multivariable analysis of general risk factors for giardiasis

The constructed multivariable model for the general risk factors for giardiasis is reported in **Table 3.4**. A significantly increased risk for giardiasis was associated with travelling abroad and swimming in a swimming pool, whereas doing gardening, eating salad or raw vegetables in the exposure window and the weekly frequency of raw fruit consumption were all independently associated with a decreased risk. A significant interaction between gender and both travelling abroad and gardening was found. Travelling abroad was more significantly associated with an increased risk of giardiasis in females than in males. Similarly, gardening was significantly associated with a decreased risk of giardiasis in females only.

Table 3.4: Constructed multivariable model for the general risk factors for giardiasis based on 303 valid (e.g. without missing data) observations (99 cases and 204 controls). Estimates following missing data multiple imputation (100 imputed datasets) are reported in parentheses.

Variables included		**B***	**Adjusted OR****	**95% CI*****	***p*-value**
Travelling abroad (outside the UK)		2.26	9.59 (10.53)	2.75-33.51 (3.28-33.80)	0.001 (<0.001)
Swimming or paddling in a swimming pool		1.03	2.79 (2.52)	1.40-5.54 (1.33-4.79)	0.002 (0.005)
Doing gardening		-1.54	0.21 (0.25)	0.09-0.54 (0.10-0.61)	0.001 (0.002)
Weekly frequency of raw fruit consumption (per event)		-0.16	0.85 (0.87)	0.78-0.93 (0.81-0.94)	0.001 (0.001)
Eating salad or raw vegetables		-1.04	0.35 (0.45)	0.16-0.79 (0.21-0.97)	0.009 (0.041)
Male gender		-0.25	0.78 (0.87)	0.38-1.58 (0.44-1.72)	0.539 (0.695)
Age	0-4y		Reference		0.196
	5-14y	-0.30	0.74 (1.03)	0.18-3.03 (0.27-3.83)	0.689 (0.965)
	15-44y	0.84	2.32 (2.38)	0.88-6.12 (0.97-5.85)	0.121 (0.059)
	45-64y	0.20	1.22 (1.25)	0.45-3.32 (0.49-3.17)	0.723 (0.640)
	65+y	0.58	1.79 (2.02)	0.59-5.42 (0.75-5.41)	0.346 (0.164)
Interaction Gender x Travelling abroad#		*-1.67*	*Males* OR 2.06, 95% CI 0.89-4.77, p=0.087 *Females* OR 7.70, 95% CI 2.80-21.19, p<0.001		
Interaction Gender x Gardening##		*1.26*	*Males* OR 0.59, 95% CI 0.30-1.15, p=0.120 *Females* OR 0.22, 95% CI 0.10-0.49, p<0.001		
Constant		*0.52*			

*logistic regression coefficient; ** Odds ratio; ***95% confidence interval; #p=0.028 (0.044); ##p=0.029 (0.090). Model statistics: model χ^2 (12) = 60.49, p<0.001; Nagelkerke R^2 = 0.25; Hosmer and Lemeshow χ^2 (8) = 13.24, p=0.104.

After the exclusion of 32 cases and 83 controls that were un-matched, the odds ratio of swimming in a swimming pool decreased from 2.79 to 2.13 (*p*-value from 0.002 to 0.097). No noticeable change in the ORs or significance levels was observed for the other variables. After the exclusion of the suspected secondary household cases, no loss of significance was observed for any of the predictors.

3.4.8.3 Multivariable analysis of risk factors for indigenous giardiasis

The constructed multivariable model for the risk factors for indigenous giardiasis is reported in **Table 3.5**. A significantly increased risk for giardiasis was associated with swimming in a swimming pool and changing nappies. Also drinking un-boiled water straight from the tap and reporting irritable bowel syndrome were associated with an increased risk of illness. The weekly frequency of consumption of raw fruit, visiting a premise with animals other than a farm, wildlife park or zoo and practising fieldsports were independently associated with a decreased risk of giardiasis. After the exclusion of 16 cases and 66 controls that were un-matched, only two predictors held their significance: changing nappies (OR 4.28, 95% CI 1.32-13.96, p=0.016) and the weekly frequency of consumption of raw fruit (OR 0.82, 95% CI 0.70-0.95, p=0.007). After the exclusion of the suspected secondary household cases the odds ratio of swimming in a swimming pool slightly decreased from 2.67 to 2.38 (*p*-value from 0.023 to 0.047), the odds ratio of changing nappies from 3.38 to 2.83 (*p*-value from 0.016 to 0.046), whereas visiting an animal premise other than a farm, wildlife park or zoo lost entirely its significance (OR from 0.77 to 0.00, p-value from 0.040 to 0.998).

Table 3.5: Constructed multivariable model for the risk factors for indigenous giardiasis based on 241 valid (e.g. without missing data) observations (69 cases and 172 controls). Estimates following missing data multiple imputation (100 imputed datasets) are reported in parentheses.

Variables included		B*	Adjusted OR**	95% CI***	*p*-value
Swimming or paddling in a swimming pool		0.98	2.67 (2.63)	1.14-6.25 (1.16-5.93)	0.023 (0.020)
Changing nappies	No children or children not in nappies		Reference		0.051
	Children in nappies but not changing nappies	0.14	1.14 (1.14)	0.22-6.07 (0.27-4.77)	0.874 (0.859)
	Changing nappies	1.22	3.38 (2.80)	1.25-9.16 (1.04-7.52)	0.016 (0.041)
Reporting irritable bowel syndrome (IBS)		1.30	3.66 (5.06)	1.18-11.37 (1.29-19.80)	0.025 (0.020)
Drinking un-boiled water straight from the tap		2.10	8.17 (3.15)	1.45-46.03 (0.79-12.59)	0.017 (0.104)
Weekly frequency of raw fruit consumption (per event)		-0.26	0.77 (0.82)	0.68-0.87 (0.73-0.91)	<0.00 (<0.001)
Practising fieldsports		-1.55	0.21 (0.25)	0.05-0.89 (0.06-1.01)	0.035 (0.052)
Visiting a premise with animals (other than a farm, wildlife park or zoo)		-2.34	0.10 (0.11)	0.01-0.89 (0.01-1.05)	0.040 (0.055)
Male gender		0.64	1.90 (1.71)	0.98-3.69 (0.93-3.14)	0.057 (0.085)
Age	0-4y		Reference		0.813
	5-14y	-0.51	0.60 (1.85)	0.08-4.23 (0.32-10.64)	0.606 (0.492)
	15-44y	-0.17	0.84 (2.33)	0.21-3.37 (0.66-8.13)	0.810 (0.185)
	45-64y	-0.61	0.54 (0.74)	0.13-2.33 (0.20-2.75)	0.412 (0.650)
	65+y	-0.51	0.60 (1.40)	0.13-2.82 (0.36-5.49)	0.518 (0.630)
Constant		*-1.99*			

*logistic regression coefficient; ** Odds ratio; ***95% confidence interval; Model statistics: model χ^2 (13) = 55.18, p<0.001; Nagelkerke R^2 = 0.293; Hosmer and Lemeshow χ^2 (8) = 5.66, p=0.686.

3.5 DISCUSSION

The risk factors for symptomatic giardiasis in people from North West England were investigated using a typical case-control study design, with cases and controls coming from the same area and matched by gender and age.

Some limitations of this study should be considered.

This study included only clinical cases of disease. The risk factors identified in this study can then be considered related to the acquisition of symptomatic *Giardia* infection, and not necessarily to the transmission of the parasite *per se*: this would have required the inclusion of asymptomatic carriers as well (Espelage *et al.*, 2010). Furthermore, no information was available on the infection status of the cases before their inclusion in the study following the onset of symptoms. If some cases became infected before the study defined exposure period, the significance of certain exposures could have been underestimated. Although the cases were also checked for the presence of other major GI pathogens, we were not able to completely exclude other potential aetiological causes of diarrhoea (including chronic conditions resembling giardiasis, such as IBS). Due to the common occurrence of asymptomatic carriage of *Giardia*, the parasite itself may have not been directly responsible for the illness onset. As a consequence following our case eligibility criteria some non-*Giardia* cases may have been included in the study, resulting in a reduction of the strength of certain associations.

Similarly, it is also important to note that controls were included in the analysis on the basis of the absence of diarrhoea in the two weeks prior to recruitment, but they were not specifically tested for *Giardia*. Misclassification may have then happened, with asymptomatic carriers of the parasite being included as controls. By assuming a prevalence of asymptomatic carriage of 1.4% as the one reported by Amar *et al.* (2007) in our sample of 226 controls, it could be estimated that three controls could have had a subclinical *Giardia* infection. It is unlikely that the inclusion of this minimal amount of asymptomatic controls would have had any significant effect on the risk factors estimates. Response rates to recruitment were low: nearly half of the original notified and contacted cases were lost and only one fifth of the total contacted controls returned the study questionnaire. The recruited cases were

nevertheless representative of the case population in terms of both age and gender distribution.

The consistent loss of participants prevented reaching the original case-control ratio. This could have resulted in a lack of statistical power, particularly in the under-estimation of the importance of exposures occurring at low frequencies in the population. The defined exposure period from which information was collected retrospectively did not completely correspond between cases and controls. In controls the exposure period dated back from the moment the questionnaire was filled, resulting in a delay in time compared to the cases. This could have resulted in bias in assessing the importance of exposures that are likely to be dependent on the month or season. However, no effect of season was observed on the risk factors estimates following multivariable analysis.

The results of the analysis of the clinical information collected from the cases overlapped the findings presented in the previous chapter in terms of reported symptoms, length of illness and hospitalisation rates, and confirmed the high morbidity associated with giardiasis. The additional data collected with the case-control study questionnaire showed that flatulence, tiredness and loss of appetite were also frequently associated with the illness. Weight loss was also common.

Despite the presence of potential bias in terms of comparability of exposure periods between cases and controls and the reduced statistical power due to a consistent drop-out of participants, risk factor analysis successfully identified a series of significant risk factors for giardiasis in North West England.

As previously reported by Gray *et al.* (1994) in the UK and by other case-control studies in developed countries (Hoque *et al.*, 2002b; Gagnon *et al.*, 2006), travelling abroad in the exposure period was overall the most important risk factor for illness. This result confirms *Giardia* as one of the major pathogens causing traveller's diarrhoea (Swaminathan *et al.*, 2009). The type of destination had an important role too, since the strongest association with giardiasis was found for travelling to countries considered at high risk for the illness, including Africa and Asia (Hoque *et al.*, 2002b) and Central and South America (Espelage *et al.*, 2010). Transmission of *Giardia* is likely to occur through drinking contaminated water or eating

contaminated food during travel, since conditions favouring the spread on intestinal pathogens such as poor sanitation and water quality are highly prevalent in developing countries (Younas *et al.*, 2008). Although participants were not asked specifically about exposures they experienced while travelling abroad, the information collected from a couple in one of the household clusters seemed to support such scenario. Both the wife and the husband travelled together to India, and reported drinking tap water and eating at restaurants repeatedly while there. The absence of other potential exposures (including owning pets or touching animals, going swimming or having children in nappies in the household) and the same date of start of illness reported were both consistent with a common source of infection related to contaminated tap water or food.

Swimming or paddling in a swimming pool as an independent risk factor for giardiasis was confirmed both overall (regardless of travel abroad) and in the indigenous subgroup. The same association was observed with the use of a Jacuzzi or a hot tub. Swimming in fresh water bodies as lakes, ponds or rivers did not appear as a risk factor for the disease: the importance of this exposure may have been underestimated though due to the small sample size, since a very small number of participants reported this exposure. Swimming (and associated exposures like swallowing water) in either chlorinated or fresh water has been reported as a risk factor for giardiasis in the UK (Gray *et al.*, 1994; Stuart *et al.*, 2003) and New Zealand (Hoque *et al.*, 2002b), but neither in Germany (Espelage *et al.*, 2010) or Canada (Gagnon *et al.*, 2006). In the study by Gray *et al.* (1994), there was no distinction between swimming in a swimming pool or in fresh water so the relative importance of the former compared to the latter was not properly assessed. Stuart *et al.* (2003) confirmed that exposure to recreational fresh water was more important than exposure to chlorinated water for the acquisition of indigenous giardiasis in the UK. Our results suggest that transmission of *Giardia* in swimming pool settings can occur of North West England. Furthermore, the significance of using a Jacuzzi could be the indication that transmission could be favoured by close contact between people too. A Few outbreaks of giardiasis in swimming pools have been described previously in the UK (Smith *et al.*, 2006b) and United States (Harter *et al.*, 1984; Porter *et al.*, 1988). Outbreaks are usually associated with the accidental faecal contamination of the pool water coupled with insufficient water disinfection. No

large swimming pool-related outbreaks were notified during the study period or were apparent from the study dataset. However such outbreaks may remain unnoticed since cases can be asymptomatic, as it has been shown in an outbreak involving the participants of swim class in the United States (Harter *et al.*, 1984). Information collected from the cases part of two household clusters seemed to suggest that transmission may have occurred at a swimming pool. In one cluster, all the three primary or co-primary cases (the father and the two sons) reported going to the swimming pool, and the same exposure was reported by two children in another cluster too.

Drinking un-boiled water from the tap appeared as a significant risk factor for indigenous giardiasis, but no dose-response effect of the number of glasses drank per day was found. The association between *Giardia* infection and drinking chlorinated water from the tap has not been reported consistently in the literature, with the exception of the study by Stuart et al. (2003) from South West England. Waterborne transmission of *Giardia* in developed countries through drinking water normally occurs following a disruption in the mains water system, and is usually accompanied by large community outbreaks as observed in the United States (O'Reilly *et al.*, 2007; Daly *et al.*, 2010) and Europe (Nygård *et al.*, 2006; Rimhanen-Finne *et al.*, 2010). Very few outbreaks of *Giardia* related to a drinking water source have been reported in the UK: in one case a private source was involved (Smith *et al.*, 2006b), whereas in another occasion the contamination was likely to be associated with the mains water source (Jephcott *et al.*, 1986). There was no evidence of any waterborne outbreak in our study, and the household of nearly all cases (with the exclusion of two reporting a mixed water source) were supplied by the mains water system. Although the difference did not reach statistical significance, univariable analysis revealed that cases also reported more frequently than controls that they observed changes in their tap water (particularly an unusual chlorine taste, or water was reported cloudy). Whether these changes were the results of a disruption in the water system leading to infection could not be concluded. Drinking unsafe surface water directly from the environment (e.g. water taken from a lake, pond or a river) is a well-known risk factor for giardiasis (Hoque *et al.*, 2002b; Gagnon *et al.*, 2006), but no evidence of this route of transmission was found in our study.

The important role of young children for the parasite transmission in the household was confirmed by the fact that cases reported more frequently than controls having children, and particularly children going to a nursery or playgroup or being in nappies. However only changing nappies was independently associated with an increased risk of indigenous giardiasis, highlighting the importance of having actual contact with infected children faeces for the acquisition of *Giardia*. Although this particular risk factor did not appear to be important in the acquisition of indigenous giardiasis in a previous study in the UK (Stuart *et al.*, 2003), the significance of changing nappies (and associated exposures such as the presence in the household of children attending childcare or being a nursing mother) has been reported in other studies (Hoque *et al.*, 2001; Hoque *et al.*, 2002b; Hoque *et al.*, 2003; Gagnon *et al.*, 2006). In two distinct household clusters identified in our study, the two secondary cases (a father and a mother, respectively) reported changing nappies to the youngest primary case. Overall our results confirm that person-to-person transmission of *Giardia* between family members can be common, and it probably explains a significant proportion of the observed household clusters and secondary cases of infection.

In our sample we found that reporting irritable bowel syndrome (IBS) was significantly associated with giardiasis in indigenous cases. No previous case-control study on giardiasis has investigated this aspect. Chronic *Giardia* infections can closely resemble IBS, presenting with abdominal pain that can be continuous and accompanied with intermittent episodes of diarrhoea or loose stools (Stark *et al.*, 2007). The presence of an association between IBS and giardiasis in our study may then be spurious and due to the misclassification of patients with IBS, in which *Giardia* was diagnosed by chance due to the similarity of the symptoms between the two conditions. Furthermore, in our study the presence of IBS was self-reported by the cases, and we couldn't verify whether this condition was officially confirmed by their GP following the appropriate diagnostic criteria. The possibility that some cases in our study may had actually represented chronic infections misdiagnosed as cases of irritable bowel syndrome cannot be ruled out, since in Italy 6% of the 137 patients with IBS and dyspepsia attending a clinic was found to be infected with *Giardia* (Grazioli *et al.*, 2006). In our study there was a nearly significant association between giardiasis and reported using medicines for indigestion (antacids) in the

exposure period, and the use of these medicines was in turn associated with reporting IBS in the cases. An association between giardiasis and being on antacid therapy has been reported previously from Germany, but only when the variable was tested on its own (Espelage *et al.*, 2010). Recent studies also showed that infection with *Helicobacter pylori* was significantly associated with *Giardia* infection (Grazioli *et al.*, 2006; Ankarklev *et al.*, 2012b; Júlio *et al.*, 2012). Whether being on an antacid therapy is associated with the presence of a *Helicobacter* infection and plays a role in causing a major predisposition of patients to *Giardia* infection could not be concluded and requires further investigation.

No evidence of zoonotic transmission appeared from our results. The significance of visiting a wildlife park or a zoo appearing in a higher frequency in cases when analysed on its own was not confirmed by multivariable analysis. There is controversy about the relevance of animals as reservoirs for human *Giardia* infection and the extent of zoonotic transmission risk, since results from different studies are inconsistent. While in some instances having contact with pets and farm animals (Warburton *et al.*, 1994) or having cats at home (Boontanom *et al.*, 2011) appeared to be significantly associated with *Giardia* infection, most case-control studies in developed countries did not find any association between an increased risk of disease and any type of potential animal contact (Hoque *et al.*, 2002b; Stuart *et al.*, 2003; Gagnon *et al.*, 2006). It is possible that appreciable levels of zoonotic transmission are more commonly present in rural contexts as those specifically studied by Warburton *et al.* (1994) and Boontanom *et al.* (2011). Also, the presence of different species and genotypes of *Giardia* which are infective for humans may differ in different populations accounting for the inconsistencies observed between studies. It is important to note that in our study, although not significantly, the frequency of exposures related to pet animals was overall lower in cases compared to controls. A similar finding was reported from Portugal, where the presence of pets and particularly dogs was associated to a decreased risk of *Giardia* infection (Júlio *et al.*, 2012). A similar finding was described in sporadic cryptosporidiosis in Australia, where having animal contact at home appeared protective for the illness (Robertson *et al.*, 2002). The authors of this study proposed that if companion animals are commonly infected with the pathogen and they shed it frequently, then the owners would be constantly exposed to it eventually developing immunity to infection.

Although this scenario can be applied to *Giardia* as well due to its high prevalence in dogs and cats (Ballweber *et al.*, 2010), the presence of bias leading to a higher frequency of pet ownership and animal contact in controls cannot be excluded and further studies are needed to determine the carriage of *Giardia* in owned pet animals.

The apparent protective effect given by certain exposures (e.g. practising fieldsports, visiting premises with animals and doing gardening) could not be easily explained, and it may reflect the presence of bias resulting in a higher frequency of these exposures in controls by chance. Similarly, the consumption of salad and raw vegetables (along with the consumption of shellfish) and the frequency of consumption of raw fruit were significantly associated with a decreased risk of giardiasis. This finding was unexpected, since the consumption of lettuce or other salad items has been reported as a significant risk factor for giardiasis previously (Stuart *et al.*, 2003; Faustini *et al.*, 2006; Espelage *et al.*, 2010; Bello *et al.*, 2011). However, similarly to our study a negative association between the consumption of raw vegetables and the risk of sporadic cryptosporidiosis was observed (Robertson *et al.*, 2002; Hunter *et al.*, 2004). As proposed by Hunter *et al.* (2004), it is possible that repeated exposure to the parasite via contaminated fresh products could lead to immunity. This protective effect would be particularly evident in people consuming fresh products more frequently.

To sum up, with our case-control study we managed to find significant risk factors for giardiasis in North West England. While confirming the important role played by travelling to foreign countries, our results highlighted the importance of transmission of this parasite within the household between family members and in swimming pools (although no outbreaks associated with this type of venue were evident). The role of tap water in the transmission of *Giardia* was not completely clear, and seemed to be associated more with sudden changes in the quality of the water rather than to continuous exposure. Animal contact did not appear to be associated with giardiasis in our study, but more data are needed to assess the zoonotic risk taking into account the genotypic diversity within *Giardia* parasites.

CHAPTER FOUR: THE MOLECULAR DIVERSITY OF *GIARDIA* IN SYMPTOMATIC PATIENTS IN NORTH WEST ENGLAND

4.1 INTRODUCTION

In the past decade, the molecular diversity of *Giardia* parasites circulating in human populations has been investigated using PCR-based molecular typing methods. Data have shown that the majority of human infections (either asymptomatic or clinical) are caused by assemblage B and in a lower proportion by assemblage A or both assemblages (Sprong *et al.*, 2009; Laishram *et al.*, 2012). However, there is a lack of information on the distribution and diversity of multi-locus genotypes of *Giardia* in humans from developed countries, since only a few studies implemented a multi-locus genotyping approach based on the sequencing of the *bg*, *gdh* and *tpi* loci (Cacciò *et al.*, 2008; Geurden *et al.*, 2009; Lebbad *et al.*, 2011).

Little information is available on the molecular diversity of *Giardia* in the human population of the United Kingdom, particularly in terms of the sub-assemblages and multi-locus genotypes of the parasite. Previous studies involved only a relatively small number of samples and mostly relied on the use of a single locus for genotyping. In 33 microscopically-confirmed sporadic cases of giardiasis from London analysed using the *tpi* RFLP assay assemblage B was found in 63.6% of them, followed by assemblage A (27.3%) and mixed-assemblage infections (9.1%) (Amar *et al.*, 2002). The same study found that all the 21 cases involved in a *Giardia* outbreak taking place in a nursery were infected with assemblage B. Another study based on the sequencing of the *tpi* and *ssu-rRNA* loci and involving 199 symptomatic cases from South-West London found a prevalence of 24%, 73% and 3% for assemblage A, B and mixed-assemblage infections respectively (Breathnach *et al.*, 2010). In 78 symptomatic patients from England and Wales genotyped with a newly developed real-time *tpi* PCR assay assemblage B accounted for 72% of the infections the prevalence of mixed-assemblage infections determined was very low (3%) (Elwin *et al.*, 2013). A different picture has been recently reported from Scotland, where the majority (72%) of the 29 cases genotyped at the *bg* locus were infected with assemblage A and the prevalence of mixed-assemblage infections was 10% (Alexander *et al.*, 2014). The authors also reported the presence of an unusual assemblage in one patient that visited Ghana before illness onset, but no information regarding the type of assemblage or sub-assemblage was available. Information on the *Giardia* sub-assemblages infecting the patients was available from two studies

and was based on the sequencing of the *tpi* locus only (Amar *et al.*, 2002; Breathnach *et al.*, 2010): all assemblage A parasites belonged to the sub-assemblage AII, and the BIII and BIV sub-assemblages were identified in assemblage B isolates. No study from the UK so far has used a multi-locus sequence typing method to investigate the molecular diversity of *Giardia* at multiple loci in a large number of patients.

4.2 AIMS OF THE STUDY

The aim of this study was to determine the prevalence and diversity of *Giardia* assemblages, sub-assemblages and multi-locus genotypes in the faecal specimens of symptomatic patients diagnosed with giardiasis in North West England. The standard multi-locus sequence typing approach based on the sequencing of the *bg*, *gdh* and *tpi* loci (including an assemblage-specific PCR assay to improve the detection of mixed-assemblage infections) was used to generate data comparable with other studies.

4.3 MATERIALS AND METHODS

4.3.1 Collection and storage of faecal specimens

Aliquots of the faecal specimens from 406 *Giardia*-positive patients included in both the enhanced surveillance (Chapter 2) and the case-control (Chapter 3) studies were included (**Table 4.1**). Aliquots collected up to November 2010 were first stored unpreserved at 4°C at the Royal Preston Hospital microbiology laboratory; after November 2010, 70% ethanol was added to them in an attempt to stop the degradation of parasite cysts and DNA. Ethanol was also added shortly after collection to 28 specimens collected in the first half of 2011, whereas specimens collected from July 2011 onwards were stored unpreserved.

Table 4.1: Number of *Giardia*-positive faecal specimens by study, collection timeframe and preservation method.

Collection timeframe	Study	No.	No. in 70% EtOH
January 2008 - February 2012	Enhanced surveillance	214	178
March 2012 - August 2013	Case-control	192	0

Specimens were sent by the Royal Preston Hospital and the Manchester Royal Infirmary microbiology laboratories to the Department of Infection Biology at University of Liverpool on a monthly basis. On arrival, unpreserved specimens were classified as formed, semi-formed or liquid based on their physical appearance (**Figure 4.1**). Out of 221 unpreserved specimens for which it was possible to assess the consistency, 81 (36.7%) were formed, 110 (49.8%) semi-formed and 30 (13.6%) liquid. Specimens were stored at 4°C in the dark before DNA extraction.

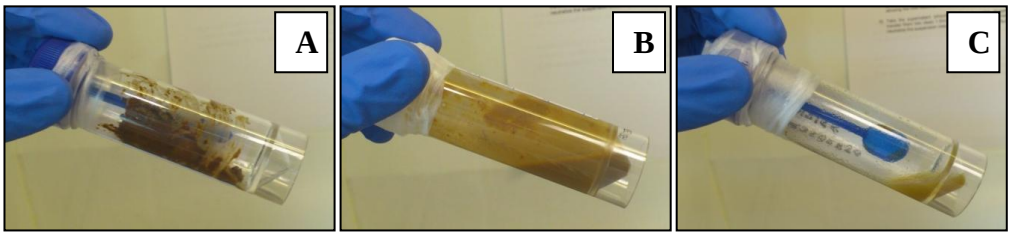


Figure 4.1: Examples of a formed (A), semi-formed (B) and liquid (C) stool specimen.

4.3.2 Faecal DNA extraction

Total genomic DNA was extracted directly from the specimens using the QIAamp® DNA Stool Mini Kit (QIAGEN®) following the manufacturer's instructions with only minor modifications. Ethanol-preserved aliquots were prepared for DNA extraction following a user-developed protocol for the isolation of DNA from formalin-preserved stool samples (available on the QIAGEN website at the following address: http://www.qiagen.com/resources/resourcedetail?id=7b341dd2-8e24-4e3a-9b35-4d2e28e46656&lang=en), with only minor modifications. Briefly, 1.6 ml of stool-ethanol suspension was transferred into a 2 ml microcentrifuge tube and centrifuged at 20,000 x g for 5 min to pellet the stools. After discarding the supernatant, 0.2 ml of the stool sediment was transferred into a new 2 ml tube with 1.6 ml of 1X PBS (containing 37 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 and 2 mM KH_2PO_4) (Ambion®) and vortexed for 15 s, followed by centrifugation at 20,000 x g for 5 min. After discarding the supernatant, 1.6 ml of 1X PBS was added again and the whole process repeated two more times. Unpreserved aliquots were first mixed thoroughly, then 200 mg of formed (or 200 µl of semi-formed or liquid) stools were transferred into a 2 ml microcentrifuge tube. After adding 1.4 ml of stool lysis buffer (buffer ASL), the suspension was vortexed for 1 min and heated for 10 min at 95°C in a heating block (instead of 5 min at 70°C as per the manufacturer's instructions). After incubation the sample was vortexed and centrifuged at 20,000 x g for 1 min, and 1.2 ml of the supernatant was transferred into a new 2 ml tube with an InhibitEX tablet. The sample was vortexed until the tablet was completely suspended, incubated for 3 min at room temperature and centrifuged at 20,000 x g for 3 min. The whole supernatant was transferred into a new 1.5 ml tube and centrifuged again at 20,000 x g for 3 min. Then 200 µl of supernatant was transferred into another 1.5 ml tube containing 15 µl of proteinase K enzyme and 200 µl of buffer AL, and mixed thoroughly. After incubation for 10 min at 70°C in a water bath, 200 µl of 96-100% ethanol was added to the lysate and mixed by vortexing. The lysate was then loaded onto a QIAamp spin column and centrifuged at 20,000 x g for 1 min. The spin column was then washed by adding 500 µl of washing buffer 1 (AW1) and centrifuging at 20,000 x g for 1 min, followed by adding 500 µl of washing buffer 2 (AW2) to the column and centrifuging it at 20,000 x g for 3 min. To ensure the elimination of remaining traces of the washing buffer, the column was

placed into another 2 ml collection tube and centrifuged for an additional 1 min. The column was then placed into a 1.5 ml tube and 100 µl of elution buffer (buffer AE) added directly onto the column membrane. The column was left at room temperature for 5 min and centrifuged for 1 min to elute the DNA. The eluted DNA was stored at -20^{o}C.

4.3.3 Molecular genotyping of *G. duodenalis*

The extracted DNA was used in a series of nested PCR reactions targeting the beta-giardin (*bg*), glutamate dehydrogenase (*gdh*) and triose-phosphate isomerase (*tpi*) genes. For the *tpi* gene, assemblage A and B-specific protocols were first used to improve the detection of mixed-assemblage infections. This was followed by re-amplification of the generic fragment of the same gene. Either the assemblage-specific or the generic *tpi* gene products were sequenced, depending on their amplification success to ensure genotyping at this locus. The small-subunit ribosomal RNA (*ssu-rRNA*) was further amplified from specimens that failed to amplify at the other three loci to ensure assemblage typing of most of the isolates. The molecular genotyping process is represented in **Figure 4.2**.

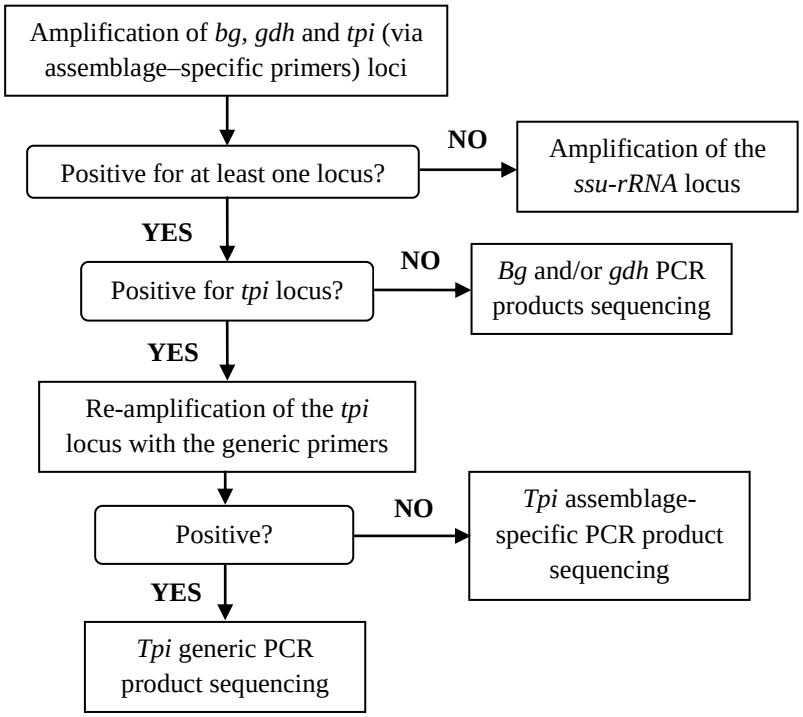


Figure 4.2: *Giardia* molecular genotyping flowchart.

4.3.3.1 General PCR reactions protocol

PCR reactions were prepared in a final volume of 25 µl, each containing 2 µl of genomic DNA, 1X PCR buffer (with Tris-Cl, KCl, $(NH_4)_2SO_4$ and 15 mM $MgCl_2$, pH 8.7) (QIAGEN®), 200 µM of each deoxynucleotide (Sigma-Aldrich®), 250 nM of each primer (Eurofins MWG Operon), 2.5 units of *Taq* DNA polymerase (QIAGEN®) and nuclease-free water (QIAGEN®) up to the final reaction volume. Positive (genomic DNA from *G. duodenalis* assemblage A isolate WBC6 trophozoites and from an assemblage B confirmed clinical isolate) and negative (nuclease-free water) control samples were included in each PCR reaction. The optimal annealing temperature and $MgCl_2$ concentration were determined separately for each set of primers and stage of reaction. All reactions were performed in a DNA Engine Dyad® Peltier Thermal Cycler (MJ Research Inc.). To confirm the successful amplification and the size of the products, 5 µl of PCR products were electrophoresed (30 min, 110 V) onto a 1.5% agarose gel stained with SYBR® Safe DNA gel stain (Invitrogen™) along with 5 µl of 100 bp DNA ladder (Thermo Fisher Scientific Inc.).

A list of all the primer sequences and conditions used for the amplification of each locus are summarised in **Table 4.2**.

4.3.3.2 Amplification of the beta-giardin (*bg*) gene

The beta-giardin gene was amplified using a nested PCR protocol. First, a 753 bp fragment of the gene was amplified using the primers G7 and G759 (Cacciò *et al.*, 2002). Then 1 µl of PCR product from the first reaction was used in a nested PCR reaction that amplified a 511 bp fragment using the primers BGf and BGr (Lalle *et al.*, 2005). PCR conditions were as follows: an initial denaturation step (94°C, 3 min), followed by a set of 35 cycles each consisting of denaturation (94°C, 20 s), annealing for 30 s (at 65°C for the primary reaction and 64°C for the nested reaction) and extension (72°C, 1 min), and concluded with a final extension step (72°C, 10 min). In the primary reaction the concentration of $MgCl_2$ was increased to 2.5 mM.

4.3.3.3 Amplification of the glutamate dehydrogenase (*gdh*) gene

A 754 bp fragment of the glutamate dehydrogenase gene was first amplified using the GDH1 (forward) and GDH2 (reverse) primers followed by the amplification, from 1 µl of PCR product from the first reaction, of a 530 bp fragment in the nested reaction using the GDH3 (forward) and GDH4 (reverse) primers (Cacciò *et al.*, 2008). PCR conditions were as follows: an initial denaturation step (94^{o}C, 3 min), followed by a set of 35 cycles each consisting of denaturation (94^{o}C, 20 s), annealing for 30 s (at 58^{o}C for the primary reaction and 68^{o}C for the nested reaction) and extension (72^{o}C, 1 min), and concluded with a final extension step (72^{o}C, 10 min).
In both the primary and secondary reaction the concentration of $MgCl_2$ was increased to 2.5 mM.

4.3.3.4 Amplification of the triose-phosphate isomerase (*tpi*) gene

The triose-phosphate isomerase gene was amplified using two different PCR assays. In the assemblage-specific PCR assay, samples first underwent the amplification of a 605 bp fragment of the gene using the AL3543 (forward) and AL3546 (reverse) primers (Sulaiman *et al.*, 2003). Then 1 µl of PCR product from the first reaction was used in two nested reactions. One involved the amplification of a 332 bp fragment of the gene from *G. duodenalis* assemblage A using the Af (forward) and Ar (reverse) primers (Geurden *et al.*, 2008). The other consisted of amplifying a 400 bp fragment of the gene from *G. duodenalis* assemblage B with the Bf (forward) and Br (reverse) primers (Geurden *et al.*, 2009). PCR conditions included an initial denaturation step (94^{o}C, 3 min), followed by a set of 35 cycles each consisting of denaturation (94^{o}C, 45 s), annealing for 45 s (at 58^{o}C for the primary reaction and 68 and 62^{o}C for the assemblage A- and B-specific secondary reactions, respectively) and extension (72^{o}C, 1 min), followed by a final extension step (72^{o}C, 10 min). The concentration of $MgCl_2$ was increased to 2.5 mM in both the primary and the assemblage A-specific secondary reactions.
Samples that successfully amplified a product with the assemblage-specific assay were re-amplified for the 605 bp fragment of the gene (using the AL3543 and AL3546 primers as above), and then 1 µl of PCR product from the first reaction was used to amplify a 530 bp fragment of the *tpi* gene using the AL3544 (forward) and

AL3545 (reverse) primers (Sulaiman *et al.*, 2003). PCR conditions were the same as those described above for the *tpi* primary reaction.

The sequencing of the 530 bp product took priority over the assemblage-specific 332-400 bp fragments. Whenever the sequencing of the former failed, the latter was sequenced instead.

4.3.3.5 Amplification of the small-subunit ribosomal RNA (*ssu-rRNA*) gene

Samples that failed to amplify at any of the above loci were additionally amplified for the small-subunit ribosomal RNA gene. A 292 bp fragment of the gene was amplified with the primers RH11 (forward) and RH4 (reverse) (Hopkins *et al.*, 1997). PCR conditions were as follows: started with a denaturation step (94^{o}C, 3 min) followed by 35 cycles each consisting of denaturation (94^{o}C, 20 s), annealing (59^{o}C, 20 s) and extension (72^{o}C, 30 s), and concluded with a final extension step (72^{o}C, 10 min).

Table 4.2: Genes used for the *Giardia* molecular genotyping and PCR reactions conditions.

Gene	PCR type	Primer sequence (5'- 3') (f/r)*	Annealing temperature and time	$MgCl_2$ (mM)	Reference
Beta-giardin *(bg)*	nested	G7 (AAGCCCGACGACCTCACCCGCAGTGC) (f) G759 (GAGGCCGCCCTGGATCTTCGAGACGAC) (r)	65°C for 30 s	2.5	(Cacciò *et al.*, 2002)
		BGf (GAACGAACGAGATCGAGGTCCG) (f) BGr (CTCGACGAGCTTCGTGTT) (r)	64°C for 30 s	1.5	(Lalle *et al.*, 2005)
Glutamate dehydrogenase *(gdh)*	nested	GDH1 (TTCCGTRTYCAGTACAACTC) (f) GDH2 (ACCTCGTTCTGRGTGGCGCA) (r)	58°C for 30 s	2.5	(Cacciò *et al.*, 2008)
		GDH3 (ATGACYGAGCTYCAGAGGCACGT) (f) GDH4 (GTGGCGCARGGCATGATGCA) (r)	68°C for 30 s		
Triose-phosphate isomerase *(tpi)*	nested	AL3543 (AAATIATGCCTGCTCGTCG) (f) AL3546 (CAAACCTTITCCGCAAACC) (r)	58°C for 45 s	2.5	(Sulaiman *et al.*, 2003)
		AL3544 (CCCTTCATCGGIGGTAACTT) (f) AL3545 (GTGGCCACCACICCCGTGCC) (r)	58°C for 45 s	2.5	
		Af (CGCCGTACACCTGTCA) (f) Ar (AGCAATGACAACCTCCTTCC) (r)	68°C for 45 s	2.5	(Geurden *et al.*, 2008)
		Bf (GTTGTTGTTGCTCCCTCCTTT) (f) Br (CCGGCTCATAGGCAATTACA) (r)	62°C for 45 s	1.5	(Geurden *et al.*, 2009)
Small-subunit ribosomal RNA (*ssu-rRNA*)	single-step	RH11 (CATCCGGTCGATCCTGCC) (f) RH4 (AGTCGAACCCTGATTCTCCGCCAGG) (r)	59°C for 20 s	1.5	(Hopkins *et al.*, 1997)

*f = forward primer; r = reverse primer

4.3.3.6 DNA sequencing

PCR products were purified using the QIAquick® PCR Purification kit (QIAGEN®) as per the manufacturer's instructions, and were sent for sequencing to the Core Genomic Facility, Medical School, University of Sheffield, UK. Products were sequenced in both directions (forward and reverse) using an Applied Biosystems® 3730 DNA Analyser. The chromatograms were edited and assembled in sequence contigs using BioEdit ver. 7.2.0 (http://www.mbio.ncsu.edu/bioedit/bioedit.html).
Contigs were first checked in GenBank to confirm they belonged to *G. duodenalis* using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

4.3.3.7 Genotypic and phylogenetic analysis

Sequences were imported into the Molecular Evolutionary Genetics Analysis (MEGA) program ver. 6.06 (Tamura *et al.*, 2013) and aligned using ClustalW, and were made the same length by trimming the excess nucleotides at the extremities.
Sequences successfully sequenced in both directions and without overlapping nucleotides at any position were identified by aligning them with representative sequences from the major *G. duodenalis* sub-assemblages AI-III and BIII-IV downloaded from GenBank. Also representative sequences from the animal assemblages (C-G) and other species of *Giardia* (either *G. muris* or *G. ardeae* depending on the availability of a sequence of the gene examined) were included. The full list of the reference sequences used is reported in **Table 4A**, **Appendix 4**. Maximum likelihood (ML) trees were built using MEGA ver. 6.06 after determining the optimal nucleotide substitution model (i.e. the best combination of evolutionary model and rate variation among sites). Sequences showing at least one difference compared with the references were checked with BLAST in GenBank to determine whether they matched other deposited sequences. The best matching sequences where then retrieved and included in the final ML tree, which was built including one representative sequence from each different genotype found and using 500 replications. Sequences that were only successfully sequenced in one direction were aligned and identified in the same way but separately from the sequences confirmed

in both directions. For clarity and consistency with the other previous *Giardia* studies, the following terminology was used for the description of gene sequences.

Single nucleotide polymorphisms (SNPs) observed in the sequences were distinguished between novel polymorphisms (e.g. an unambiguously determined nucleotide position not reported in previously described sequences) and heterogeneous positions (e.g. a nucleotide position that could not be ambiguously determined due to the presence of two overlapping nucleotides). Sequences with heterogeneous positions were aligned directly with the representative sequences of the genotypes confirmed by sequencing in both directions, and the heterogeneous positions in the sequence were indicated using the IUPAC nucleotide ambiguity codes. Locus–specific subtypes were named after their closest reference sequence or isolate found in GenBank (e.g. AI, A2, Sweh198), followed by a lower case letter (e.g. AIa) if the sequence showed at least one novel polymorphism. Assemblage A multi-locus genotypes (MLGs) were named following the nomenclature proposed by Cacciò *et al.* (2008): capital letter to indicate the assemblage followed by a Roman numeral for the sub-assemblage and an Arabic numeral for the MLG (e.g. AI-1, AI-2).

4.3.4 Relationship between *Giardia* cyst load and PCR success

In order to check whether the parasite cyst load in the stools had any effect on the PCR amplification success, a subset of 25 specimens were checked for the presence of *Giardia* cysts using a commercial direct fluorescence antibody (DFA) assay (MERIFLUOR® *Cryptosporidium/Giardia*, Meridian Bioscience® Europe, London, UK). A drop of faecal material from each sample was transferred and spread on the slide wells, and left to dry at room temperature for 30 min. A positive (formalinized stool with *Giardia* and *Cryptosporidium* cysts/oocysts) and a negative control (parasite-free formalinized stool) were included. A drop of both detection reagent and counterstain were added to the wells and mixed with an applicator stick, and slides were incubated in a humidified chamber at room temperature in the dark for 30 minutes. Reagents were then gently washed away with 1X washing buffer, removing the excess of buffer by tapping the slides edge on a clean paper towel. A drop of mounting medium was added to the wells and the slides covered with a 22x64 coverslip. Slides were then scanned at 40X with a microscope equipped with a QBC®

ParaLens LED apparatus (wavelength 410-511 nm) (QBC® Diagnostics Inc., Port Matilda, Pennsylvania). Parasite cysts were counted in 10 randomly selected fields, and the average number of cysts per field calculated. Since only a very low or very high number of cysts were observed in the slides, specimens were classified in only two categories: low (average ≤10 cysts/field) and high (average ≥10 cysts/field) parasite load. All specimens were analysed within two months of collection, and DNA was extracted as described in Section 4.2.2 at the same time of microscopic examination to avoid DNA degradation. Extracted DNA was visualised onto gel as described in Section 4.3.3.1. DNA bands were assessed visually from the gel and their intensity coded as weak, moderate or strong. All specimens were amplified at the four loci as described in Sections 4.3.3.2-5.

4.3.5 Statistical analysis of overall and locus-specific PCR success

The overall and locus-specific PCR success were analysed in relation to the specimens' year of collection, preservation status and age (e.g. the number of months from collection to DNA extraction). Descriptive statistics were calculated using cross-tabulations, and the Fisher's Exact test was used to test for differences in the PCR success between groups of specimens of different age and preservation status. Multivariable logistic regression analysis was also performed to determine the factors that were independently associated with the overall and locus-specific PCR success. All tests were two-sided and a *p*-value of less than 5% was considered to be statistically significant.

4.4 RESULTS

4.4.1 PCR amplification success

Of the 406 extracted specimens tested for the *bg*, *gdh* or *tpi* loci, 218 (53.7%) amplified at least one: the *tpi* assemblage A/B-specific product amplified in 86.2% (188/218) of the specimens, followed by the *bg* product in 76.1% (166/218) and the *gdh* product in 70.2% (153/218). Of the remaining 188 specimens that were negative at all the three loci and that were further amplified for the *ssu-rRNA* locus, 77 (41%) amplified for this locus and 111 (59%) were negative.

To sum up, in total 72.7% (295/406) of the extracted specimens successfully amplified at least one of the four loci. Of these, 42% (124/295) amplified only one locus whereas 40.7% (120/295) and 17.3% (51/295) amplified three and two loci, respectively.

Differences in the overall PCR success rate were observed in relation to the specimens' year of collection, preservation status and age (**Table 4.3**). Following multivariable analysis (excluding year of collection because it was redundant due to being indistinguishable from specimen age) only ethanol preservation was significantly associated with the overall PCR success (**Table 4.4**): preservation reduced the odds of successful amplification from the specimens by 73%, and there was no significant interaction between the specimen age and preservation.

Table 4.3: Number and percentage of PCR-positive specimens at any of the four loci by year of collection, preservation status and age (e.g. the number of months from collection to DNA extraction).

Year	EtOH preserved	No. samples	No. PCR positive	% PCR positive	Median specimen age in months (range)
2008	yes	45	25	55.6	42.7 (27.4-49)
2009	yes	22	11	50	27.9 (24.6-36.3)
2010	yes	83	49	59	13.5 (4.3-25.7)
2011	yes	28	16	39	8.2 (3-17.1)
2011	no	29	25	61	2.3 (1.2-5.4)
2012	no	120	101	84.2	1.10 (0.1-6.2)
2013	no	79	68	86.1	1.10 (0.2-2.3)
	total	**406**	**295**	**72.7**	**2.3 (0.1-49)**

Table 4.4: Multivariable logistic regression model of the variables independently associated with the overall PCR amplification success in 406 extracted faecal specimens.

Variables included	B*	OR (95% CI)**	*p*-value
Specimen age	-0.01	0.99 (0.97-1.01)	0.510
Ethanol preservation	-1.32	0.27 (0.14-0.50)	<0.001
Constant	*2.11*		

*regression coefficient; **adjusted odds ratio with 95% confidence intervals

The effect of stool consistency on the PCR success was also tested in the 228 unpreserved specimens, but it was not significant (Fisher's Exact, p=1.000).

When the locus-specific PCR amplification success was considered, both the ethanol preservation and age of the specimens were significantly associated with the PCR amplification success of all the three MLST loci but the effect was different for non-preserved and preserved specimens (**Table 4.5**). In unpreserved specimens the odds of amplification success was significantly reduced with the increasing specimen age whereas in preserved specimens the amplification success did not change with sample age, but the ethanol preservation reduced the odds of successful amplification by 88%. Stool consistency did not affect any of the MLST loci amplification success when tested in unpreserved specimens (Fisher's Exact, p>0.05 for all three loci).

Table 4.5: Multivariable logistic regression model of the variables independently associated with the PCR amplification success at the *bg*, *gdh* and *tpi* loci in 406 extracted faecal specimens.

Variables included	B*	OR (95% CI)**	*p*-value
bg			
Specimen age (months)	-0.64	0.53 (0.38-0.74)	<0.001
EtOH preservation	-2.08	0.12 (0.05-0.29)	<0.001
Interaction age and preservation	0.61	1.83 (1.30-2.57)	<0.001
Constant	*1.23*		
gdh			
Specimen age (months)	-0.54	0.58 (0.42-0.81)	0.001
EtOH preservation	-2.53	0.08 (0.03-0.20)	<0.001
Interaction age and preservation	0.52	1.69 (1.22-2.34)	0.002
Constant	*0.99*		
tpi assemblage A/B-specific			
Specimen age (months)	-0.33	0.72 (0.53-0.97)	0.030
EtOH preservation	-1.67	0.19 (0.08-0.42)	<0.001
Interaction age and preservation	0.30	1.36 (1.00-1.83)	0.046
Constant	*1.01*		

*regression coefficient; **adjusted odds ratio with 95% confidence intervals

4.4.2 Relationship between *Giardia* cyst load and PCR success

The majority of the microscopically-examined specimens (17 out of 25, 68%) showed a high cyst load (average ≥10 cysts/field). Stool consistency was intermediate (semi-formed) in the majority of the specimens (14 out of 25, 56%) and formed or liquid in six (24%) and five (20%) specimens, respectively. There was high variability in the extracted whole faecal DNA yield as visualised by agarose gel: half of the samples (13 out of 25, 52%) gave a band of moderate intensity, followed by specimens returning a band of strong (7 out of 25, 28%) or weak to very weak (5 out of 25, 20%) intensity.

Consistency of the stool was not significantly associated with the observed cyst load (Fisher's Exact, p=0.182), although all formed specimens showed a high cyst load. Consistency was not associated with the observed extracted DNA band intensity (Fisher's Exact, p=0.129), but the majority of formed specimens (four out of six, 66.7%) and none of the liquid specimens showed a band of strong intensity on gel. The cyst load was only weakly associated with the DNA band intensity (Fisher's Exact, p=0.095): bands of strong intensity were never observed from specimens with a low cyst load, whereas the majority of those with a high cyst load (14 out of 17, 82.3%) returned bands of moderate to high intensity.

A total of 23 out of 25 (92%) specimens amplified at least one locus. The *ssu-rRNA* gene successfully amplified in 21 (84%) samples, followed by the *tpi* in 19 (76%), the *bg* in 15 (60%) and *gdh* in 14 (56%). Of the two samples that did not amplify at any of the loci tested one was of formed consistency and had a very high cyst load, whereas the other was semi-formed and had a low cyst load. No association was found between stool consistency, cyst load or extracted DNA band intensity and the number of loci successfully amplified (Fisher's Exact, p=0.215, 0.336 and 0.411 for consistency, cyst load and DNA band respectively). The cyst load was marginally significantly associated with the amplification of the *ssu-rRNA* locus (Fisher's Exact, p=0.081): 94.1% of specimens with a high load amplified this locus against 62.5% of the samples with a low cyst load. This trend for association was not observed in the other markers (Fisher's Exact, p=0.194, 0.389 and 0.344 for *bg*, *gdh* and *tpi* respectively).

4.4.3 Overall *Giardia* assemblages prevalence

Overall, the *G. duodenalis* assemblage was determined in 247 of the 295 specimens (83.7%) either by successful sequencing of any of the loci amplified (239 specimens) or by on-gel visualization of the *tpi* assemblage-specific products (eight specimens for which DNA sequencing failed). Assemblage B was found in the majority of the specimens, and only less than five percent showed a mixed assemblage infection (defined as a specimen showing the sequences of two different assemblages at different loci or amplifying the products of both assemblages following the *tpi* assemblage-specific PCRs) (**Figure 4.3**). Of the seven specimens showing a mixed infection (**Table 4.6**), three were diagnosed by the *tpi* assemblage A and B-specific PCR (whereas they were typed as B at both the *bg* and *gdh* loci), three were typed as B at the *tpi* locus but as A at the *gdh* locus (two specimens) or at the *bg* locus (one specimen), and one was typed as A at the *bg* locus but as B at the *gdh* locus.

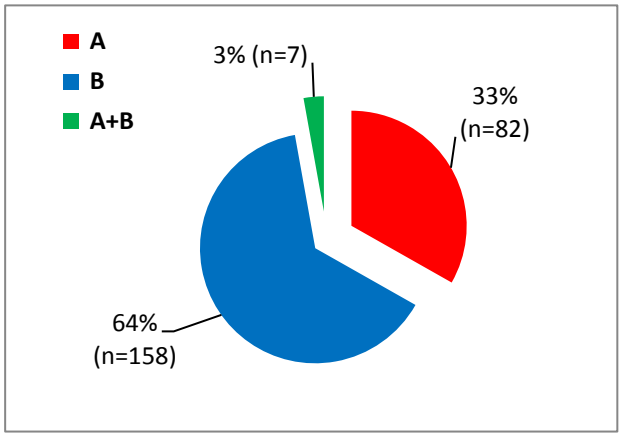


Figure 4.3: Prevalence of *Giardia* assemblages determined in 247 samples.

Table 4.6: Assemblage typing results in the seven specimens with a mixed assemblage A and B infection.

Isolate ID	*Bg*	*Gdh*	*Tpi* A/B specific	*Tpi* generic
13/C155, 13/CF, 12/C11	B	B	A+B	B
10/24	A	B	*negative*	*negative*
11/34	*negative*	A	B	B
10/3	A	*negative*	B	*negative*
8/11	*negative*	A	B	*negative*

4.4.5 *Giardia* sub-assemblages and genotypes by locus

4.4.5.1 Beta-giardin

Out of 166 samples amplifying at the *bg* locus, 157 (94.6%) were successfully sequenced. One isolate sequenced only in one direction showed complete identity with two references sequences (AI and AII), and the sequence of one isolate (12/C71) showed 17 heterogeneous positions possibly compatible with a mixed assemblage A and B template. The full list of isolates genotyped at the *bg* locus is reported in **Table 4B** and the detailed description of the SNPs found in the *bg* sequences is reported in **Tables 4C-F**, **Appendix 4**. Of 155 isolates that could be assigned to a specific assemblage, 131 (of which 118 were confirmed by sequencing in both directions and 13 in only one direction) had sequences without heterogeneous positions and were unambiguously typed at the sub-assemblage level. The overall typing results obtained at the *bg* locus are shown in **Figure 4.4**, and the diversity of the *bg* sequences of the 118 isolates sequenced in both directions and without heterogeneous positions as determined by phylogenetic analysis is reported in **Figure 4.5**.

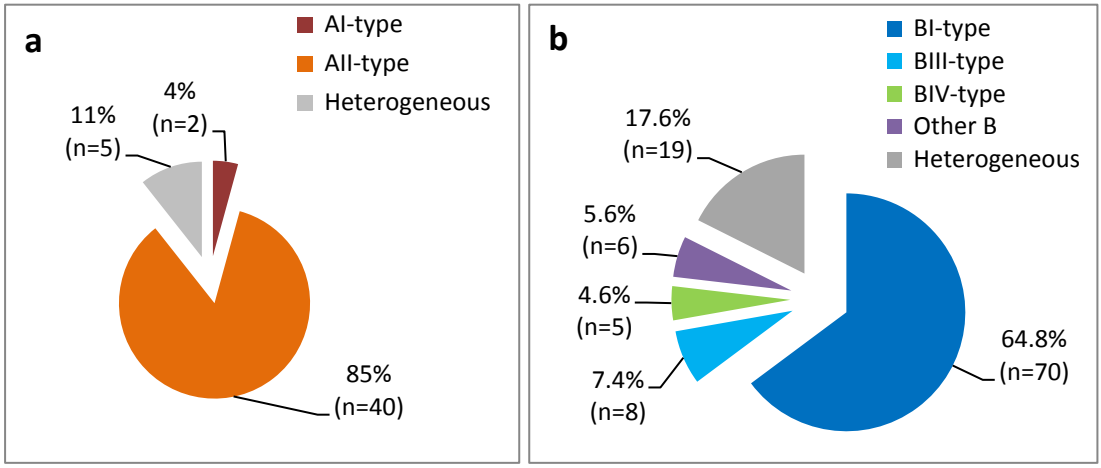


Figure 4.4: Overall sub-assemblage typing results at the bg locus in 47 assemblage A (a) and 108 assemblage B isolates (b). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity or whether their sequence showed heterogeneous positions.

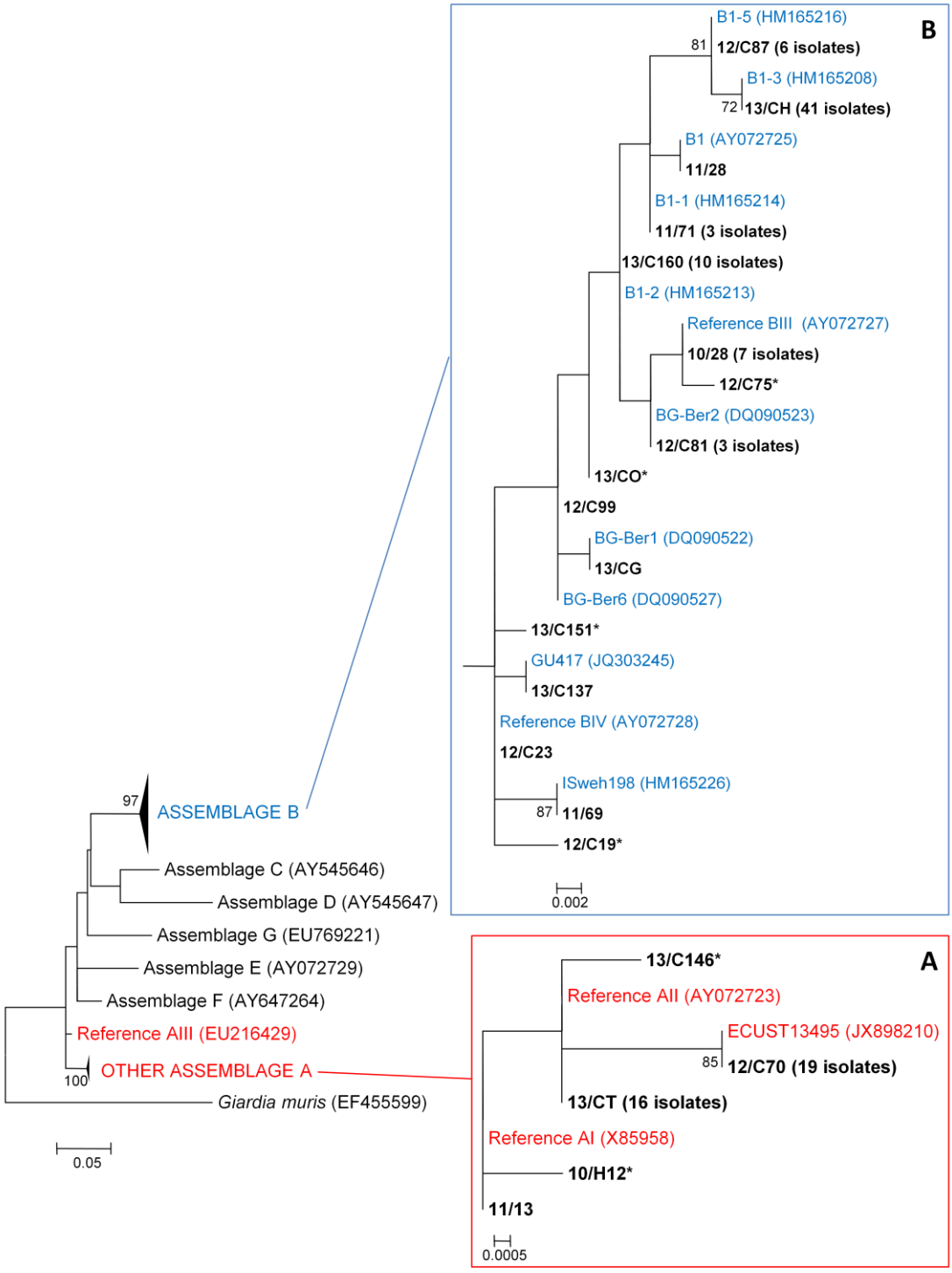


Figure 4.5: Maximum likelihood tree of the *bg* gene sequence of 118 isolates without heterogeneous positions. One representative sequence from each identified genotype (in bold) is shown and it is indicated by its assigned study ID (with the total number of isolates sharing the same sequence reported in parentheses). The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates (in red for assemblage A, in blue for assemblage B and in black for the other assemblages) are indicated by their GenBank accession number. Optimal nucleotide substitution model: Tamura-Nei with Gamma distribution. Only bootstrap values ≥70% for bipartitions are reported.

Out of the 47 assemblage A isolates, 42 (87.5%) did not show heterogeneous positions along the sequence and were unambiguously identified. Only two isolates (11/13 and 10/H12) were of sub-assemblage AI (of which 10/H12 showing a novel polymorphism compared to the AI reference sequence and was confirmed by sequencing in both directions). The vast majority of assemblage A isolates belonged to two previously described sub-assemblage AII subtypes. Subtype A2 was found in 17 isolates (of which 13/C146 showing a novel polymorphism compared to the AII reference sequence and confirmed by sequencing in both directions) and subtype A3 in 23 isolates. Five assemblage A isolates showed at least one heterogeneous position in their chromatogram. Of the three isolates that were confirmed by sequencing in both directions (10/20, 12/C89, 10/H8), heterogeneous positions were observed at one (two isolates) or two (one isolate) different positions along the *bg* gene sequence. The isolate showing two heterogeneous positions were hypothesized to represent a mixed A2/A3 subtype since the polymorphisms were observed at the two positions (383 and 391) corresponding to the nucleotide substitutions differentiating between the subtypes A2 and A3.

Out of 108 isolates identified as belonging to assemblage B, 89 (82.4%) did not show heterogeneous positions in their *bg* sequence and were unambiguously identified. Overall, 16 different subtypes were identified. The majority of the isolates (70 out of 89, 77.5%) were identical to five previously reported B1 subtypes. Among them, the most represented subtype was B1-3 (48 isolates), followed by B1-2 (11 isolates), B1-5 (six isolates), B1-1 (three isolates) and B1 (one isolate). The sequence of the isolate 13/CO (confirmed by sequencing in both directions) showed a novel polymorphism compared to the B1-2 subtype. Eight isolates were identified as belonging to sub-assemblage BIII: seven showed complete identity with the BIII reference sequence, whereas the sequence of the isolate 12/C75 showed a novel polymorphism compared to the same reference. The sequence of one isolate was identical to the sub-assemblage BIV reference sequence whereas the isolates 13/C151 and 12/C19 (confirmed by sequencing in both directions) showed one and two nucleotide substitutions compared to the same reference, respectively. The sequence of two isolates showed complete identity to the one of two previously described BIV-type isolates (Sweh198 and GU417). The sequences of six isolates were identical to those of three previously described subtypes (BG-Ber) (with four identical to BG-Ber2 and two identical to BG-Ber1 and BG-Ber6, respectively).

At least one heterogeneous position was found in the *bg* sequences of 19 assemblage B isolates (of which 18 were confirmed by sequencing in both directions). Overall SNPs occurred at 29 different positions along the *bg* gene sequence (**Table 4E**, **Appendix 4**), and the maximum number of SNPs observed per sequence was five (average 2.4). Two particular positions (497 and 623) showed a high frequency of SNPs occurrence, which were observed in six and five sequences respectively, however no sequences with the same pattern of SNPs were observed and some sequences also showed novel polymorphisms.

4.4.5.2 Glutamate dehydrogenase

Out of 153 samples amplifying at the *gdh* locus, 146 (95.4%) were successfully sequenced and typed at the assemblage level. The full list of isolates genotyped at the *gdh* locus is reported in **Table 4G,** and the detailed description of the SNPs found in the *gdh* sequences is reported in **Tables 4H-I**, **Appendix 4**. A total of 119 isolates (of which 115 were confirmed by sequencing in both directions and four in only one direction) were unambiguously genotyped at the sub-assemblage level because they did not show any heterogeneous position. The overall typing results obtained at the *gdh* locus are shown in **Figure 4.6**, and the diversity of the *gdh* sequences of the 115 isolates sequenced in both directions without heterogeneous positions as determined by phylogenetic analysis is reported in **Figure 4.7**.

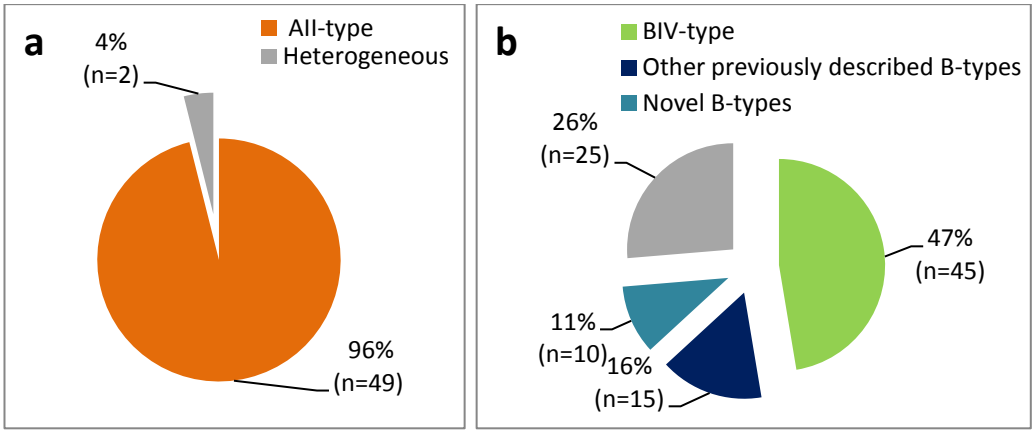


Figure 4.6: Overall sub-assemblage typing results at the *gdh* locus in 51 assemblage A (a) and 95 assemblage B isolates (b). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity or whether their sequence showed heterogeneous positions.

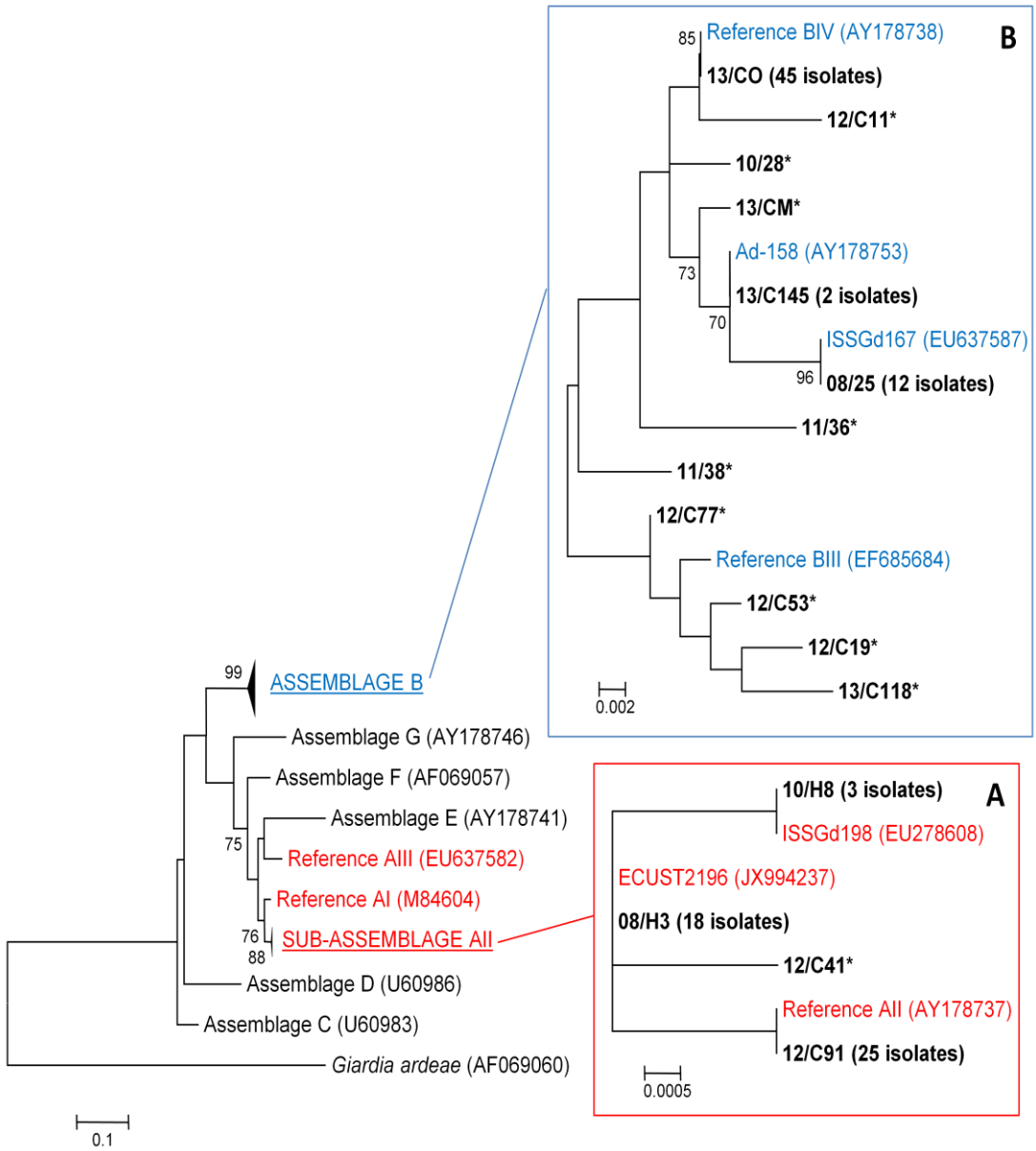


Figure 4.7: Maximum likelihood tree of the *gdh* gene sequence of 115 isolates without heterogeneous positions. One representative sequence from each identified genotype (in bold) is shown and it is indicated by its assigned study ID (with the total number of isolates sharing the same sequence reported in parentheses). The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates (in red for assemblage A, in blue for assemblage B and in black for the other assemblages) are indicated by their GenBank accession number. Optimal nucleotide substitution model: Tamura 3 parameter with Gamma distribution. Only bootstrap values ≥70% for bipartitions are reported.

All the 49 assemblage A isolates which *gdh* sequence did not show any heterogeneous position belonged to three previously described sub-assemblage AII subtypes, namely A2 (the reference AII subtype) (27 isolates), ECUST2196 (18 isolates) and ISSGd198 (three isolates). The sequence of one isolate (12/C41)

showed a novel polymorphism compared to the ECUST2196 subtype sequence. Two isolates confirmed by sequencing in both directions (8/11 and 12/C56) each showed a single heterogeneous position in their chromatograms (**Table 4H**, **Appendix 4**).

Out of 95 assemblage B isolates, 70 (73.7%) did not show any heterogeneous position in their *gdh* sequence and belonged to 13 different subtypes. A total of 45 isolates (64.3%) belonged to the reference BIV subtype, 13 (18.6%) to the previously described ISSGd167 subtype, and two isolates (2.8%) to the previously described Ad-158 subtype. With the exclusion of one isolate (13/C155) that was not confirmed by sequencing in both directions, nine isolates sequenced in both directions represented subtypes that differed from a minimum of two up to eight positions from either the reference sequences or any other sequence deposited in GenBank. The *gdh* sequence of these isolates showed from a minimum of one up to four novel polymorphisms (**Table 4I**, **Appendix 4**). Heterogeneous positions were found in the chromatograms of 25 assemblage B isolates confirmed by sequencing in both directions. Up to 41 different heterogeneous positions were observed along the *gdh* sequence, with a maximum of nine SNPs observed per sequence (average 4.6) (**Table 4L**, **Appendix 4**). Three positions (644, 746, 857 and 929) showed a high frequency of overlapping nucleotides occurrence, which were observed in eight, eight, eight and nine sequences respectively. The nucleotide substitutions differentiating between the sub-assemblages BIII and BIV along the *gdh* gene occurred at these positions (**Table 4L**, **Appendix 4**). However, only three sequences shared the same pattern of SNPs (at six different positions) and no isolate could be unambiguously identified as a genuine mixed sub-assemblage infection due to the occurrence of both heterogeneous positions and novel polymorphisms.

4.4.5.3 Triose phosphate isomerase

Out of 186 *tpi*-positive specimens that were re-amplified for the 530 bp fragment of the gene, 150 (80.6%) successfully amplified and 139 samples were successfully sequenced. In addition, 24 products amplified using the assemblage-specific PCR assays were successfully sequenced. In order to identify all the isolates sequenced at this locus in a consistent way, the 530 bp sequences were in the end trimmed down to the length of the shorter products amplified with the assemblage-specific PCR assay. In summary, 163 isolates were successfully sequenced and typed at the assemblage level. The full list of isolates genotyped at the *tpi* locus is reported in **Table 4M**, and the detailed description of the SNPs found in the *tpi* sequences is reported in **Tables 4N-P**, **Appendix 4**. In total 138 isolates (of which 124 were confirmed by sequencing in both directions and 14 in only one direction) did not show heterogeneous positions along the gene sequence and they were unambiguously genotyped at the sub-assemblage level. The overall typing results at the *tpi* locus are shown in **Figure 4.8**. The diversity of the *tpi* sequences of the 124 isolates successfully sequenced in both directions and without heterogeneous positions as determined by phylogenetic analysis is reported in **Figure 4.9**.

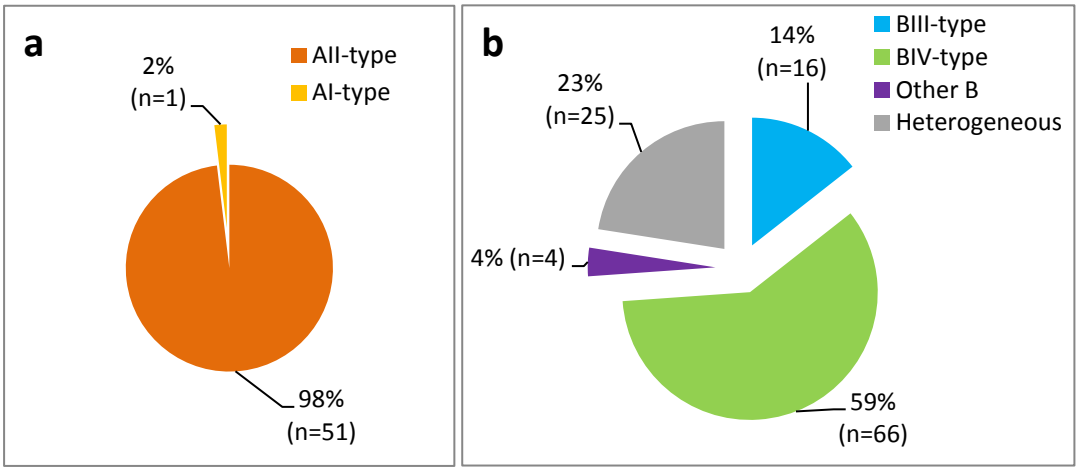


Figure 4.8: Overall sub-assemblage typing results at the *tpi* locus in 52 assemblage A (a) and 111 assemblage B isolates (b). Isolates (with number and percentage reported) are grouped by sub-assemblage similarity or whether their sequence showed heterogeneous positions.

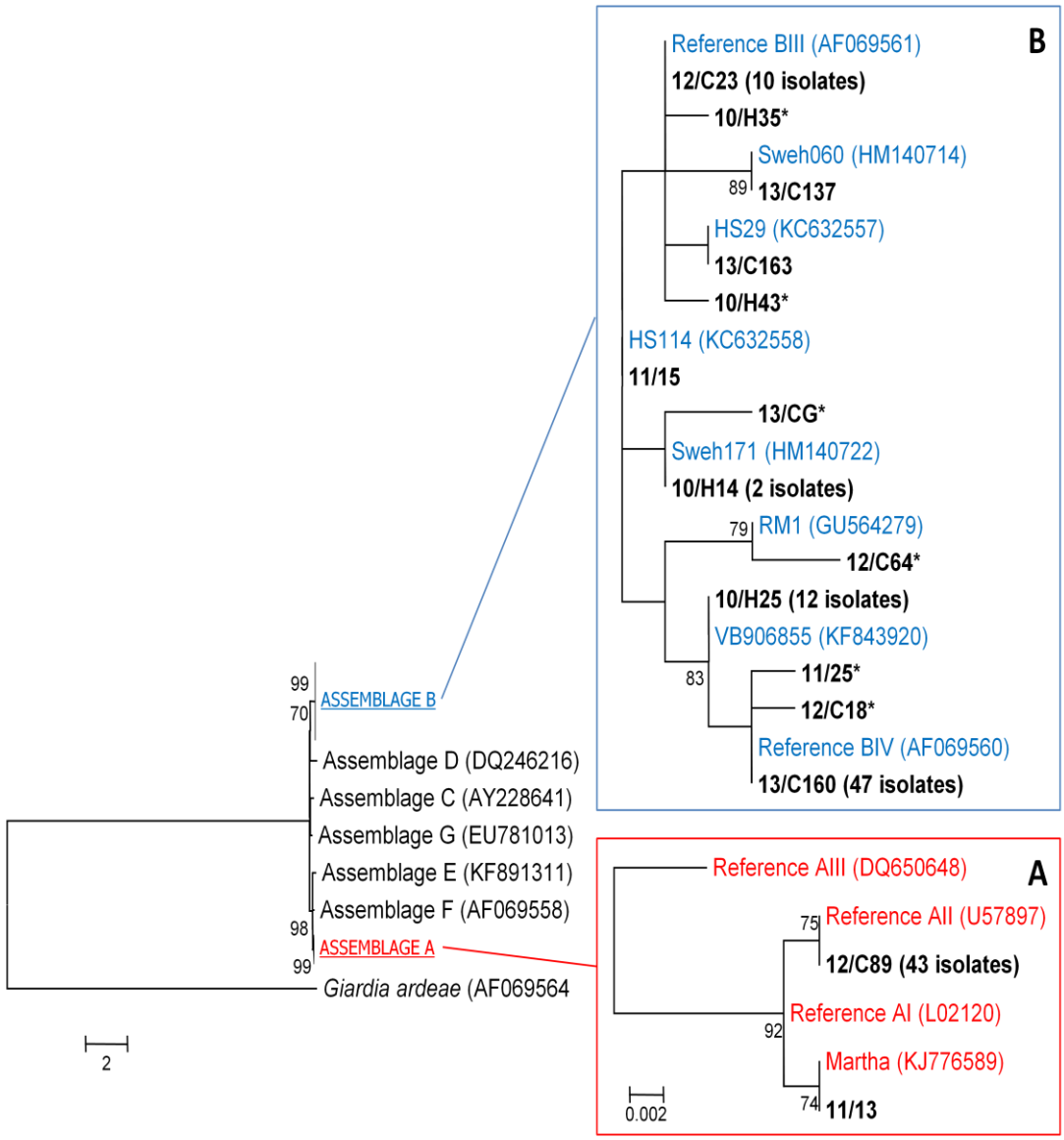


Figure 4.9: Maximum likelihood tree of the *tpi* gene sequence of 124 isolates without heterogeneous positions. One representative sequence from each identified genotype (in bold) is shown and it is indicated by its assigned study ID (with the total number of isolates sharing the same sequence reported in parentheses). The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates (in red for assemblage A, in blue for assemblage B and in black for the other assemblages) are indicated by their GenBank accession number. Optimal nucleotide substitution model: Kimura 2 parameter with gamma distribution and invariant sites. Only bootstrap values ≥70% for bipartitions are reported.

None of the 52 isolates identified as belonging to assemblage A showed any SNP in their *tpi* sequences. With the exception of one isolate (11/13) which sequence was

identical to the one of a previously described AI subtype, the *tpi* sequence of all the assemblage A isolates was identical to the sub-assemblage AII reference subtype.

Out of 111 isolates identified as belonging to assemblage B, 86 (77.5%) did not show any heterogeneous position in their *tpi* sequence. Overall, 13 different subtypes were found. The majority (50/86 isolates, 58.1%) were identical to the sub-assemblage BIV reference subtype, and 13 (16.2%) were identical to another previously described BIV-type subtype (VB906855). The sequences of two isolates (11/25 and 12/C18) each showed a novel polymorphism compared to the BIV subtype reference. A total of eleven isolates (12.8%) were identical to the sub-assemblage BIII reference subtype. Three isolates were identical to three different BIII subtypes previously described (HS29, Sweh060 and HS114). The sequences of two isolates (10/H35 and 10/H43) each showed a novel polymorphism compared to the BIII subtype reference. Four isolates all sequenced in both directions did not clearly group with either of the two major sub-assemblages reference sequences. Two isolates (10/H14 and 13/C149) showed complete sequence identity with a previously described isolate (Sweh171). One isolate (13/CG) showed two novel polymorphisms compared to the isolate Sweh171, whereas another one (12/C64) showed two novel polymorphisms compared with a previously described isolate (RM1). Another isolate (12/C50) produced a potentially novel sequence but the PCR product was successfully sequenced in only one direction so it could not be confirmed.

A total of 25 assemblage B isolates could not be unambiguously identified because their chromatograms showed heterogeneous positions along the *tpi* gene sequence. Overall, in 23 isolates confirmed by sequencing in both directions overlapping nucleotides occurred at 14 different positions along the *tpi* gene sequence (**Table 4P**, **Appendix 4**). The maximum number of heterogeneous positions observed per sequence was four (average 2.1). Four particular positions (628, 847, 889 and 892) showed a high frequency of SNPs occurrence, which were observed in five, eight, nine and nine sequences respectively. Four isolates (12/C21, 12/C30, 12/C77 and 13/C140) showed the same SNPs at these four positions and were hypothesized to represent a mixed BIII/BIV subtypes infection, since these particular positions corresponded to the nucleotide substitutions differentiating between the sub-

assemblages BIII and BIV (**Table 4P**, **Appendix 4**). No other sequences with the same pattern of SNPs were observed.

4.4.5.4 Small subunit ribosomal RNA

Out of 77 samples that amplified the *ssu-rRNA* gene, 34 were successfully sequenced (of which 32 in both directions). Overall, assemblage A was found in 12 samples (35.3%) and B in 22 (64.7%). All isolates showed complete sequence identity with the relative reference sequences, and no distinction between the major sub-assemblages was possible (**Figure 4.10**). The two isolates that were successfully sequenced only in one direction were identical to the others. None of the *ssu-rRNA* sequences showed heterogeneous positions.

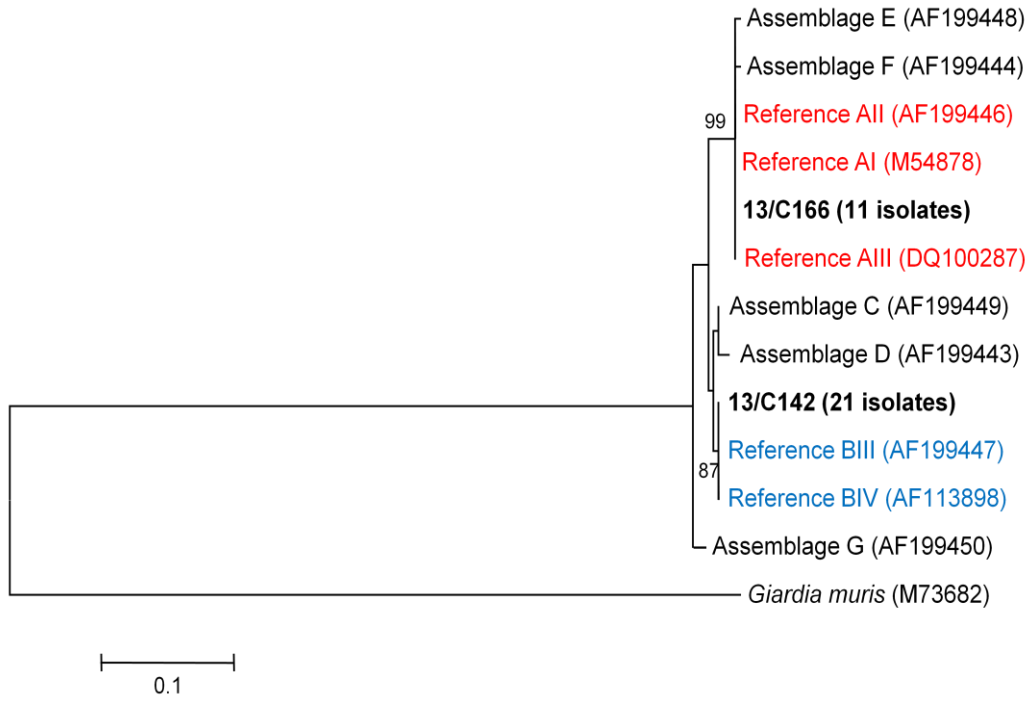


Figure 4.10: Maximum likelihood tree of the *ssu-rRNA* gene sequence of 32 isolates. One representative sequence from each identified genotype (in bold) is shown and it is indicated by its assigned study ID (with the total number of isolates sharing the same sequence reported in parentheses). The major assemblage and sub-assemblage reference sequences along with the sequences from previously described isolates (in red for assemblage A, in blue for assemblage B and in black for the other assemblages) are indicated by their GenBank accession number. Optimal nucleotide substitution model: Tamura 3-parameter. Only bootstrap values ≥70% for bipartitions are reported.

4.4.6 Multi-locus genotypes

A total of 34 assemblage A and 76 assemblage B isolates were successfully genotyped at all the three MLST loci.

4.4.6.1 Assemblage A multi-locus genotypes

A total of 31 assemblage A isolates showed sequences without heterogeneous positions and their multi-locus genotype (MLG) was unambiguously identified (**Table 4.7**). All MLGs belonged to the sub-assemblage AII. Two previously identified MLGs (AII-1 and AII-2) (Cacciò *et al.*, 2008) were found in nine (29% of the total) and four isolates (12.9% of the total), respectively. Four novel MLGs were also identified, with one of them accounting for nine isolates (29% of the total). Amongst the three isolates showing SNPs in the sequence of at least one of the three loci, one isolate showed a potentially mixed AII-1/AII-2 MLG whereas the other two were not clearly identifiable.

Table 4.7: Assemblage A multi-locus genotypes in 31 isolates unambiguously genotyped at the *bg*, *gdh* and *tpi* loci.

No. isolates	%*	*Bg*	*Gdh*	*Tpi*	MLG
9	29	A2	AII	AII	AII-1**
9	29	A3	ECUST2196	AII	AII-novel
6	19.3	A2	ECUST2196	AII	AII-novel
4	12.9	A3	AII	AII	AII-2**
2	6.4	A3	ISSGd198	AII	AII-novel
1	3.2	A2	ECUST2196a	AII	AII-novel

percentage of the 31 isolates;same MLGs reported by Cacciò et al. (2008)*

4.4.6.2 Assemblage B multi-locus genotypes

The *Giardia* multi-locus genotype was unambiguously identified in a total of 45 assemblage B isolates which sequences did not show any heterogeneous position (**Table 4.8**). Two previously identified MLGs (MLG 1 and 7) (Lebbad *et al.*, 2011) were found in 21 (46.7%) and one (2.2%) isolates, respectively. Up to 15 novel assemblage B MLGs were identified, each one accounting for one to three isolates only.

Table 4.8: Assemblage B multi-locus genotypes in 45 isolates unambiguously genotyped at the *bg*, *gdh* and *tpi* loci.

No. isolates	**% ***	***Bg***	***Gdh***	***Tpi***	**MLG**
21	46.7	B1-3	BIV	BIV	1**
3	6.7	B1-2	BIV	BIV	B novel
3	6.7	B1-5	BIV	VB906855	B novel
3	6.7	B1-5	ISSGd167	VB906856	B novel
2	4.4	B1-2	BIV	VB906855	B novel
2	4.4	B1-2	ISSGd167	BIV	B novel
1	2.2	B1-1	BIV	BIV	7**
1	2.2	B1-1	ISSGd167	BIV	B novel
1	2.2	B1-2a	BIV	BIV	B novel
1	2.2	B1-2	Ad-158	Sweh171	B novel
1	2.2	B1-2	Ad-158	VB906855	B novel
1	2.2	B1-3	Bb	BIV	B novel
1	2.2	B1-3	ISSGd167	BIV	B novel
1	2.2	B1	ISSGd167	BIV	B novel
1	2.2	BIII	Bc	VB906855	B novel
1	2.2	BIII	BIV	BIV	B novel
1	2.2	BIVb	Bg	BIII	B novel

percentage of the 45 isolates;same MLGs reported by Lebbad et al.* (2011)

4.5 DISCUSSION

We extracted the DNA from faecal specimens collected from patients diagnosed with giardiasis in North West England and we used PCR-based genotyping methods to characterise the molecular diversity of the *Giardia* parasites causing infection.

Overall, the combined use of the four different loci allowed the successful PCR amplification in more than 70% of the extracted specimens. In particular, the use of the *ssu-rRNA* locus ensured the amplification in specimens that were negative following PCRs at the *bg*, *gdh* and *tpi* loci. A greater success rate of amplification of the *ssu-rRNA* locus compared to the other three was also confirmed in the specimens with *Giardia* cysts confirmed by immunofluorescence microscopy. The small-subunit ribosomal RNA gene has a high copy number in the *Giardia* genome: this increased availability of starting DNA template determines an increased success of PCR amplification, particularly in specimens (like faeces or environmental samples) with a low quantity of target DNA or a higher quantity of PCR inhibitors (Wielinga & Thompson, 2007). The other loci also showed differences in their amplification success: the highest was observed for the *tpi* assemblage-specific product whereas it was lower and comparable between the *bg* and *gdh* loci. The differential amplification success of these loci has been commonly reported in studies using the same primers, although results vary. The use of the *tpi* assemblage A and B specific PCR assay allowed the successful amplification in 67 out of the 72 (93%) extracted human specimens (including also 13 that did not amplify with the *tpi* generic product PCR assay), and all the specimens successfully amplified at the *bg* locus (Geurden *et al.*, 2009). The amplification success rate reported by a study including 66 symptomatic patients from Germany was 92.4% for the *bg* locus, 44% for the generic *tpi* product and only 25.7% for the *gdh* locus (Broglia *et al.*, 2013). Failure of amplification of assemblage B at multiple loci including those used in this study was also extensively observed in dog samples (Beck *et al.*, 2012), with the phenomenon involving also the two canine assemblages C and D. Since the effect of a difference in the gene copy number between the *bg*, *gdh* and *tpi* markers could be reasonably excluded (all of them being single-copy genes), it has been proposed that the lack of amplification of particular loci could be due to the presence of nucleotide mismatches between the PCR primers and the genomic sequences (Broglia *et al.*,

2013). The presence of mismatches could then lead to the inability of certain primers to amplify particular sub-genotypes, as recently suggested the beta-giardin primers in respect to certain assemblage B sub-types (Robertson *et al.*, 2007a).

In order to check for the potential effect of PCR inhibition, samples that did not amplify at any of the four loci were spiked with *Giardia* DNA and re-amplified for both the *ssu-rRNA* and *bg* loci. All samples successfully amplified (data not shown). The inhibition of PCR was also tested and ruled out in a study from Peru, where spiking samples positive by microscopy but negative by PCR with low amounts of parasite DNA resulted in successful amplification (Lebbad *et al.*, 2008).

The parasite load and the extent of cyst and DNA degradation due to the storage of the specimens could have influenced the PCR amplification success too. The specimens originated from patients where *Giardia* infection was confirmed using a commercially available enzyme-linked immunoassay method (*GIARDIA/ CRYPTOSPORIDIUM CHEK*®, Techlab). It has been shown that this assay detects the presence of the cyst wall protein 1 (CWP1), a soluble antigen that is released in large quantities during the encystation of *Giardia* trophozoites *in vitro* (Boone *et al.*, 1999). It is then possible that in some of the specimens only the soluble antigen but either no cysts (or a very little amount of them) was present, resulting in the absence of *Giardia* DNA available for PCR amplification. However, although the analysis involved only a small number of samples, the presence of cysts was verified in all the specimens that were checked by immunofluorescence microscopy and the PCR success rate was not correlated with the parasite load. Furthermore, the two samples that tested negative at all the four loci contained parasite cysts.

The time passed from sample collection to DNA extraction seemed to significantly and negatively influence the amplification success in unpreserved specimens. This effect was likely due to the degradation of cysts and parasite DNA over time, particularly in samples collected in the years 2008-2010 that were stored unpreserved for several months or even years before being extracted. The effect of ethanol preservation on the PCR success could not be clearly explained. The preservative was added to the oldest specimens by the hospital laboratory only a few months before DNA was extracted, so it was not possible to assess its actual effect on the

preservation of DNA in samples that likely were already degraded when ethanol was added. Nevertheless, a certain number of specimens that were a few years old still amplified successfully at multiple loci.

The prevalence of the two assemblages found in the specimens reflected what has been reported in symptomatic patients from Europe (Geurden *et al.*, 2009; Lebbad *et al.*, 2011; Broglia *et al.*, 2013) and corresponded to what has been previously published from the UK (Amar *et al.*, 2002; Breathnach *et al.*, 2010; Elwin *et al.*, 2013). Assemblage B was responsible for the majority of infections, and the prevalence of mixed assemblage A and B infections was low (<5%), whereas none of the other assemblage (C-F) was found.

The molecular diversity of assemblage A isolates at the sub-assemblage level was successfully characterised in the majority of cases. Sequences with heterogeneous positions were observed only in a small proportion of isolates (and they were mostly observed at the *bg* locus), confirming the rare occurrence of sequences with overlapping nucleotides consistently reported in assemblage A parasites (Sprong *et al.*, 2009). Regardless of the gene analysed, the vast majority of the sequences matched those of previously described isolates and sequences showing novel polymorphisms were observed in very few cases. The vast majority of assemblage A isolates were assigned to the sub-assemblage AII consistently across the three loci, whereas sequences belonging to the sub-assemblage AI were found only in two (at the *bg* locus) and one (at the *tpi* locus) cases. The results confirmed the preponderance of sub-assemblage AII in humans compared to sub-assemblage AI, as previously shown in other studies using multi-locus sequence typing (Cacciò *et al.*, 2008; Geurden *et al.*, 2009; Lebbad *et al.*, 2011).

The degree of polymorphism (e.g the number of different sub-types identified) differed between the loci and showed remarkable similarities with other studies. At the *bg* locus the subtype A3 was found in the majority of isolates followed by A2, as reported from symptomatic patients in Belgium (Geurden *et al.*, 2009) and Germany (Broglia *et al.*, 2013). Two subtypes were found at the *gdh* locus, with AII being the most frequent as observed in both Sweden (Lebbad *et al.*, 2011) and Belgium (Geurden *et al.*, 2009), whereas only one AII sub-type was identified at the *tpi* locus. Several patients were infected with two previously described sub-assemblage AII

multi-locus genotypes (AII-1 and AII-2). These MLGs have been commonly reported in symptomatic patients from Italy (Cacciò *et al.*, 2008), Belgium (Geurden *et al.*, 2009) and Sweden (Lebbad *et al.*, 2011). As reported in the aforementioned studies, also novel MLGs showing different combinations of *bg*, *gdh* and *tpi* were observed suggesting that humans can be infected with an array of very diverse assemblage A genotypes.

The sub-assemblage diversity of assemblage B isolates was much more complex compared to assemblage A isolates. The analysis was complicated by the high frequency of occurrence of sequences with heterogeneous positions, with overlapping nucleotides observed on average in 20% of the sequences across the three loci. A few assemblage B isolates showing heterogeneous templates potentially represented mixed sub-assemblage infections but the identification could not be unequivocally resolved. Mixed templates are a common occurrence in assemblage B parasites (Sprong *et al.*, 2009), and their presence has been observed following DNA extraction and DNA sequencing even from single *Giardia* cysts and between different cysts isolated from the same patient (Ankarklev *et al.*, 2012).

The assignment to a particular B sub-assemblage of isolates that were unambiguously identified was not immediate due to discrepancies between the different markers. At the *bg* locus the majority of the isolates belonged to subtypes that were part of the previously identified B1 group (Geurden *et al.*, 2009), as it has also been observed in human isolates from Sweden (Lebbad *et al.*, 2011). The B1 group clustered nearby the sub-assemblage BIII by phylogenetic analysis, but whether it represents an actual sub-assemblage other than BIII and BIV has to be verified. Conversely, the analysis of the *gdh* and *tpi* markers assigned the majority of isolates to the sub-assemblage BIV. The genotypic diversity observed in assemblage B isolates was much higher compared with the one observed within assemblage A, consisting in 21 novel sub-types (the majority of which at the *gdh* and *tpi* loci). Similar levels of diversity at the three markers were reported in assemblage B parasites in human patients from Europe (Lebbad *et al.*, 2011; Broglia *et al.*, 2013). Following multi-locus analysis of isolates successfully characterised at all the three markers, the majority of patients were infected with the same multi-locus genotype (ML-1) previously reported in the majority of patients of a study in Sweden (Lebbad *et al.*, 2011). Similarly to what was observed for assemblage A, a high

number of assemblage B MLGs showing different combinations of *bg*, *gdh* and *tpi* never reported before were found.

Our results confirmed the need for a novel DNA sequence-based classification system applicable for the genotyping of *Giardia* parasites below the assemblage level. Although the use of the *bg*, *gdh* and *tpi* loci seems to produce relatively consistent typing results for assemblage A sub-assemblages (supported by both phylogenetic and multi-locus analysis), it is clearly limited in its applicability for the assemblage B sub-typing. The extremely high sequence variation often observed in assemblage B parasites (often involving a single nucleotidic difference between isolates) and the poor resolution of phylogenetic analyses complicate enormously the assignment of isolates to a specific sub-assemblage. Furthermore, the sub-groups BIII and BIV identified by allozyme electrophoretic studies are not supported by DNA sequence analysis (Feng & Xiao, 2011). The existence of several sub-assemblages within B (as suggested by the high levels of genetic variation) should be re-evaluated and confirmed following a more comprehensive biochemical and genetic characterization, in order to determine whether these subgroups are consistent and discrete. There is also the need of increasing the number of loci in order to improve the reliability and resolution of the multi-locus typing scheme. The use of next generation sequencing (NGS) technologies may help in determining the extent of genetic sub-structuring within both assemblage A and B at a whole-genome level: these data would be crucial in the identification of sub-assemblage specific genes or DNA sequence polymorphisms that could then be used as diagnostic markers. The comparison of a large set of clinical samples from different parts of the world would be pivotal in order to validate the reliability and stability of the new markers across different parasite populations.

To sum up, the molecular diversity of *Giardia* in symptomatic patients from North West England showed evident similarities with other areas of Europe. The majority of infections were caused by assemblage B alone, followed by infections with sub-assemblage AII. Molecular analysis at both the single-locus and multi-locus level revealed the occurrence of both novel and already described sub-types and multi-locus genotypes, with assemblage B parasites showing a significantly higher genotypic diversity and occurrence of heterogeneous templates compared with assemblage A parasites.

CHAPTER FIVE: DIFFERENCES IN THE CLINICAL OUTCOME AND TRANSMISSION ROUTES BETWEEN *G. DUODENALIS* ASSEMBLAGES

5.1 INTRODUCTION

An increasing amount of experimental evidences suggest that the two *Giardia* assemblages causing infection and disease in humans are truly different species (Thompson & Monis, 2012). However, the biological characteristics of *Giardia* assemblage A and B in terms of host-parasite interactions have been mostly investigated at the molecular or phenotypic level *in vitro*. Robust population data are lacking on whether the clinical course of infection in humans with one or the other assemblage significantly differ, or whether the two assemblages are transmitted preferentially via different routes.

The correlation between the infecting *Giardia* assemblage and the observed clinical outcome in the patients has been evaluated previously, but results are not consistent. A significant association between infection with assemblage A and the presence of diarrhoea in people of various ages has been reported in numerous studies (Read *et al.*, 2002; Aydin *et al.*, 2004; Haque *et al.*, 2005; Sahagún *et al.*, 2008), but the opposite finding (with infection with assemblage B being associated with diarrhoea and other intestinal symptoms) has also been shown in other occasions (Gelanew *et al.*, 2007; Al-Mohammed, 2011; Puebla *et al.*, 2014). In studies that included only symptomatic cases of giardiasis, the type and severity of symptoms reported by patients infected with the two assemblages have been described. Assemblage A appeared to be associated with intermittent diarrhoea and B with persistent diarrhoea in two studies (Homan & Mank, 2001; Helmy *et al.*, 2009). In the UK study by Breathnach *et al.* (2010) cases infected with the two assemblages showed a largely comparable clinical picture, although those infected with assemblage A reported suffering from fever more frequently. In a study from Sweden, infection with assemblage B was significantly associated with flatulence in children (Lebbad *et al.*, 2011). Data are lacking regarding the presence of epidemiological differences between the two *Giardia* assemblages. No differences in the socio-demographics (gender or age distribution) or in the travel history of patients infected with assemblage A or B was found in two studies involving the genotyping of a large number of samples from the UK (Breathnach *et al.*, 2010) and Sweden (Lebbad *et al.*, 2011). In these two studies information about other exposures (such as contact with recreational waters or with animals) was not collected from the patients, and the

comparison was made between cases and not against a control group (e.g. people without giardiasis). So far, the only study that attempted to determine the *Giardia* assemblage-specific risk factors using people that tested negative for the parasite as controls is the study of Anuar *et al.* (2014) performed in the Orang Asli community in Malaysia. The study revealed that infection with assemblage A was significantly associated with having close contact with pets in the household, whereas assemblage B infection was associated with the presence in the household of children below 15 years of age and other family members with *Giardia* infection, other than with consuming raw vegetables.

More data are needed in order to understand the biological and epidemiological differences between assemblage A and B in a developed country context. The presence of assemblage-related differences may have practical consequences for the clinical management and prevention of giardiasis. A precise diagnosis at the assemblage level would be required if a differential response of the parasites to drug treatment is suspected, or whether a different propensity of the two assemblages towards the development of either chronic infections or post-infectious gastrointestinal syndromes has been observed. The identification of assemblage-specific risk factors could lead to more specific insights into the epidemiology of *Giardia*, particularly in assessing the actual importance of zoonotic transmission, as it has been successfully shown for the two parasite species responsible for cryptosporidiosis (Hunter *et al.*, 2004).

5.2 AIMS OF THE STUDY

The aim of this analysis was to compare the reported clinical outcomes and exposure histories between patients found to be infected with *Giardia* assemblage A or B, in order to test whether the two parasite assemblages significantly differ in terms of their interaction with the human host or in their transmission routes. In particular, the epidemiological and molecular data generated in the previous two chapters were combined to investigate the assemblage-specific risk factors following an appropriate case-control study design.

5.3 MATERIALS AND METHODS

5.3.1 General statistical methods

Statistical analyses were performed using IBM® SPSS® Statistics 20 (IBM, USA). Cross-tabulations, odds ratio (OR) estimates and 95% confidence interval (CI) and the Pearson's Chi-Square (or Fisher's Exact when data were sparse) test were used to explore associations between categorical variables in univariable analysis. The Mann-Whitney's U test (and univariable logistic regression whenever appropriate) was used as tests of significance for continuous variables. Whenever variables were treated both as continuous and categorical the results from both were reported. All tests were two-sided and *p*-values of less than 5% were considered to be statistically significant. Multivariable logistic regression models were built by entering manually the variables returning a $p \leq 0.2$ in the univariable analysis. The importance of each factor and its effect on the overall model fit were assessed using likelihood ratio tests, and factors without a significant effect on the model fit were dropped. Gender and age were always retained in the models regardless of their significance and interaction terms including these two variables were included to control for confounding.

5.3.2 Analysis outline

The socio-demographic, clinical and epidemiological characteristics of the patients infected with either assemblage A or B were described and compared at different levels using distinct datasets as specified below. Only cases with a single-assemblage infection (e.g. either assemblage A or B) were included in the analyses.

5.3.2.1 Gender, age and spatio-temporal distribution of assemblage A and B

Differences in the distribution of the two *Giardia* assemblages in relation to the gender and age of the cases were explored first. This analysis included all the 240 cases that were genotyped and described in the previous chapter and that were part of the surveillance (Chapter 2) and case-control study (Chapter 3) datasets. Additionally, 56 cases that were genotyped at the assemblage level using the *ssu*-

rRNA gene in a previous study (Weerapol Taweenan, PhD thesis) were also included in this analysis because they were also part of the surveillance study. In total, 296 genotyped cases were included in the analysis on the gender and age distribution of the two assemblages. Differences in the distribution of the two *Giardia* assemblages in relation to the cases' month of report and local authority were specifically explored in the area of Central Lancashire in the timeframe of the five years with complete data from each month. This analysis was restricted to 179 genotyped cases resident in Central Lancashire and reported from 2008 to 2012.

5.3.2.2 Clinical outcomes of patients infected with assemblage A or B

The self-reported clinical outcomes reported by the patients infected with *Giardia* assemblage A and B were described and compared. Cases that were asymptomatic were excluded. The data collected in the enhanced surveillance study and in the case-control study were analysed separately since the questionnaires used differed in terms of the questions asked. In total, this analysis included 197 genotyped cases that returned the respective study questionnaire and with clinical information reported, of which 103 from the enhanced surveillance study (36 assemblage A and 76 assemblage B) and 85 from the case-control study (28 assemblage A and 57 assemblage B). The clinical outcomes of cases infected with assemblage A were compared to those reported by cases infected with B (chosen as the baseline category) by calculating the odds ratio (OR) estimate and 95% confidence interval (CI), and multivariable analysis was then used to assess which outcomes were independently associated with infection with either assemblage A or B.

5.3.2.3 Exposure profile of patients with assemblage A or B infection and assemblage-specific risk factors

Differences in the reported exposures between cases infected with assemblage A and B were analysed at two levels and separately in the enhanced surveillance and case-control study datasets. Due to the absence of exposure information collected from control subjects in the surveillance study, only case-case analysis was performed in this dataset to compare the exposures of the 36 assemblage A cases with those

reported by the 76 assemblage B cases. The group of cases infected with assemblage B was chosen as the baseline category for comparisons.

The risk factors associated to the infection with a specific *Giardia* assemblage were explored using the genotyped cases from the case-control study. The exposures reported by the cases (28 infected with assemblage A and 57 infected with B) were analysed and compared with those of the 226 control subjects. Comparisons were made separately for the two assemblages in univariable analysis, and the assemblage-specific multivariable models were built. The multivariable models were fitted before and after multiple imputation of missing values.

5.4 RESULTS

5.4.1 Gender and age distribution of *Giardia* assemblages

In the 296 cases typed at the assemblage level and included in the gender and age analysis, there was no difference in the prevalence of the two assemblages between males and females (Pearson's χ^2, p=0.722). Assemblage B was more prevalent than A in both genders (in 180 males: 119 B, 66.1% and 61 A, 33.9%; in 116 females: 79 B, 68.1% and 37 A, 31.9%).

The median age of the cases infected with A was significantly higher (49 years, range 11 months to 94 years) than the one of the cases infected with B (38.5 years, range one to 83 years) (Mann-Whitney U, p=0.002). The different age pattern of the two assemblages was confirmed when cases were grouped in 10 years age bands (Fisher's Exact, p<0.001): the prevalence of assemblage A was at its lowest in people from their 20s to their 40s and peaked in people in their 70s and over, whereas the prevalence of B was higher in people from their 20s to their 40s and decreased in the elderly (**Figure 5.1**).

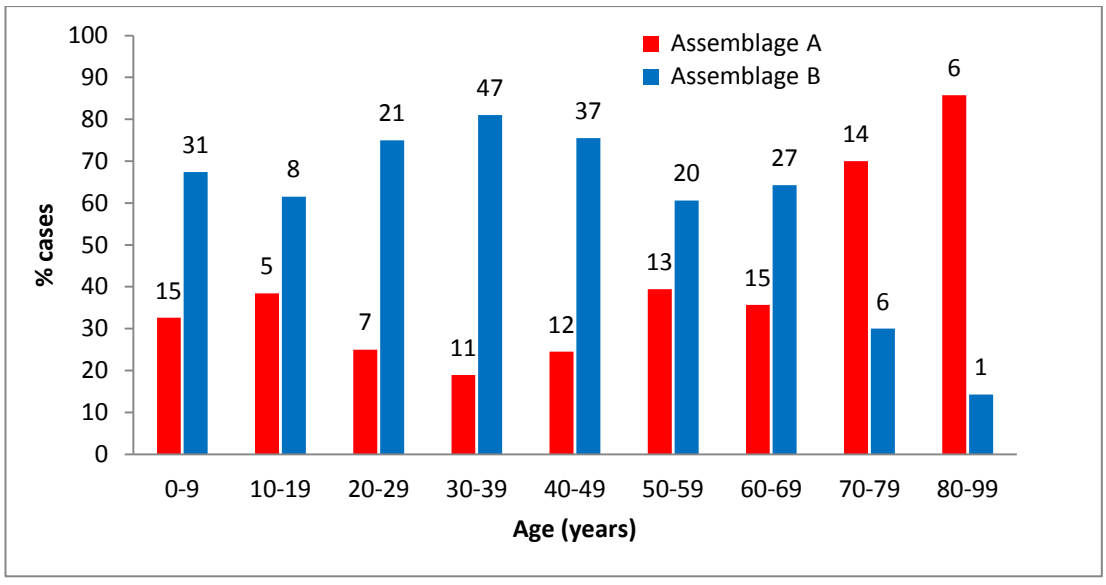


Figure 5.1: Age prevalence of assemblage A and B cases (n=296). The number of cases is reported above the bars. The last two age categories (80-89y and 90-99y) were merged because in people over 90 years of age only one assemblage A and none assemblage B cases were found.

5.4.2 *Giardia* assemblages distribution in Central Lancashire by month of report and local authority

By considering the 179 genotyped cases reported from 2008 to 2012 and resident in Central Lancashire, there was no significant difference in the distribution of the two assemblages by month of report (Fisher's Exact, p=0.200), and considerable variation was observed in the proportion of cases reported by month in both assemblage A and B (**Figure 5.2a**). The distribution of cases in the four seasons did not significantly differ between the two assemblages (Pearson's χ^2, p=0.276), although assemblage B cases seemed to occur more frequently in Winter and assemblage A in the other seasons (**Figure 5.2b**).

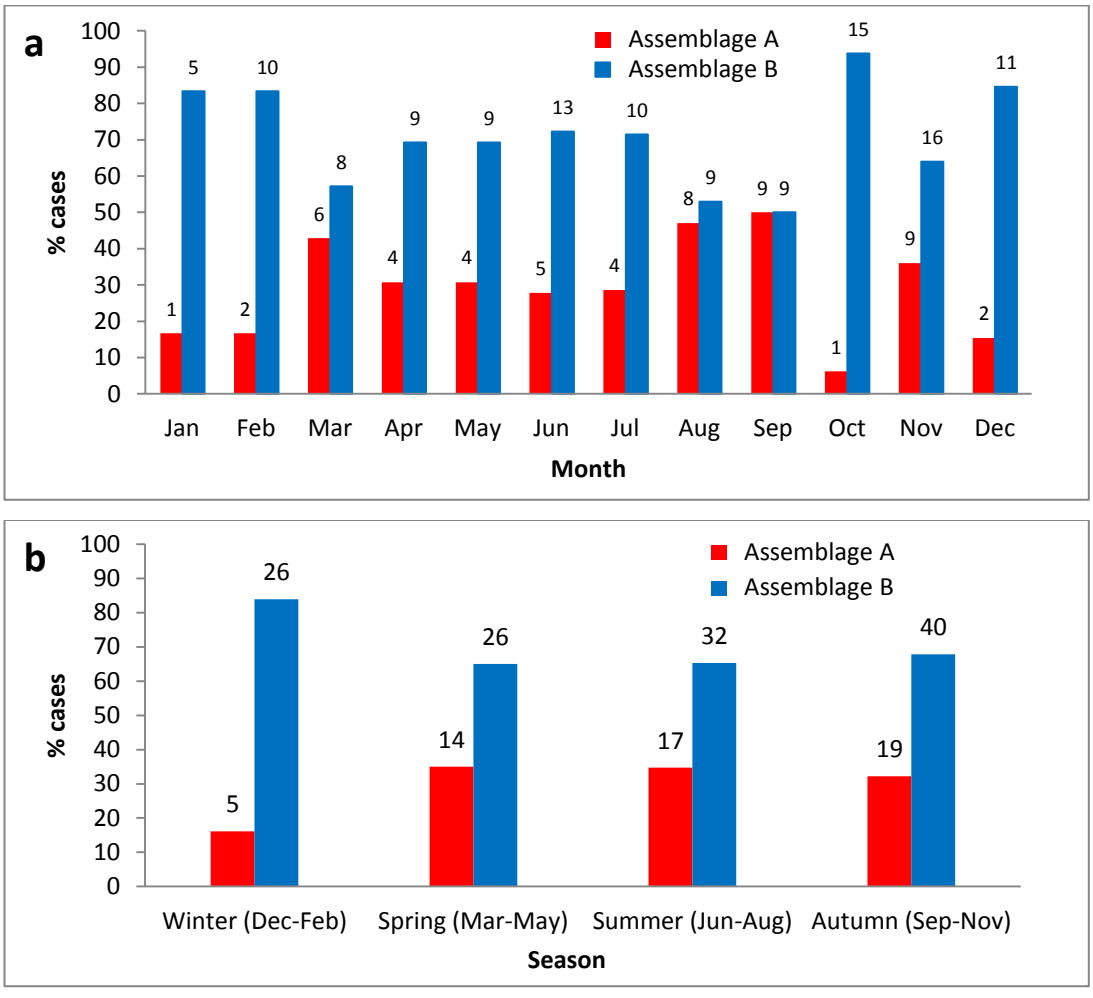


Figure 5.2: Distribution by month (A) and season (B) of report of assemblage A and B in 179 genotyped cases resident in Central Lancashire, 2008-2012. The number of cases is reported above the bars.

Within Central Lancashire, a significant difference in the prevalence of the two assemblages was found between the three local authorities (Pearson's χ^2, p=0.014): although overall assemblage B was more prevalent than A, in South Ribble the prevalence of assemblage A was higher and the prevalence of B lower if compared with both Preston and Chorley (**Figure 5.3**).

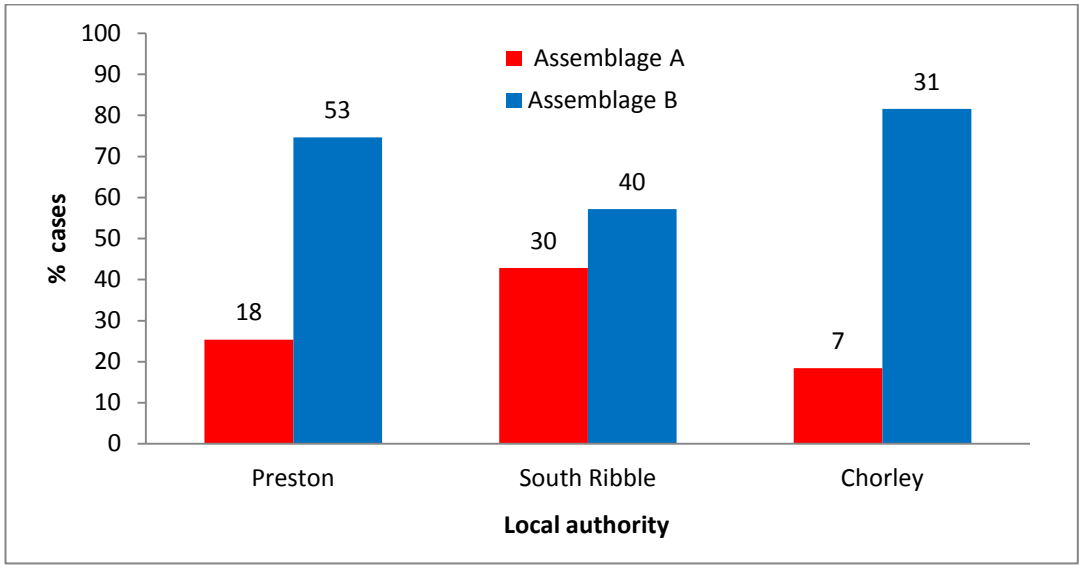


Figure 5.3: Distribution by local authority of assemblage A and B in 179 genotyped cases resident in Central Lancashire, 2008-2012. The raw number of cases is reported above the bars.

5.4.3 Clinical outcomes of patients infected with assemblage A or B

In the 103 symptomatic cases genotyped from the enhanced surveillance study, assemblage B cases reported more frequently vomiting and a higher number of symptoms, a longer illness and more days of normal activity prevented compared with assemblage A cases but the differences were not significant (**Table 5.1**). Assemblage A cases reported being admitted to hospital due to illness significantly more frequently than those infected with assemblage B (OR 4.38, 95% CI 1.02-18.91, p=0.034) (**Table 5.1**). The association between assemblage A and hospitalisation was confounded by the age of the cases (multivariable logistic regression model with age entered as predictor: admitted to hospital OR 3.59, 95% CI 0.79-16.27, p=0.097): amongst hospitalised cases, those infected with assemblage A were older (median 71.5 years, range 33 to 79) than those infected with assemblage B (median 13 years, range six to 75).

In the 85 symptomatic cases genotyped from the case-control study, assemblage B cases reported significantly more symptoms (including vomiting, abdominal pain, swollen stomach and loss of appetite) compared with assemblage A cases (**Table 5.2**). Assemblage B cases seemed to experience weight loss more frequently than assemblage A cases also, but the difference was only marginally significant (p=0.088). As observed in enhanced surveillance cases, hospitalisation due to illness was observed more frequently in assemblage A cases (18% compared to 7% in assemblage B cases) but the difference was not statistically significant (p=0.157). Following multivariable analysis, only swollen stomach was significantly and independently associated with infection with assemblage B (multivariable logistic regression model adjusted for age and gender: swollen stomach OR 0.26, 95% CI 0.09-0.77, p=0.015).

Table 5.1: Comparison of the self-reported clinical outcomes between symptomatic patients infected with assemblage A (n=36) and with assemblage B (n=67) from the enhanced surveillance study. Odds ratios were calculated using assemblage B as the baseline category.

			Assemblage n (%)**			
Variable	**Valid (% missing)***	**Category**	**A**	**B**	**OR (95% CI)*****	***p*-value**
Diarrhoea	103	No	3 (8.3)	4 (6)	Ref.	0.693
		Yes	33 (91.7)	63 (94)	0.70 (1.15-3.31)	
Blood in stools	103	No	34 (94.4)	63 (94)	Ref.	1.000
		Yes	2 (5.6)	4 (6)	0.93 (0.16-5.32)	
Vomiting	103	No	23 (63.9)	33 (49.3)	Ref.	0.155
		Yes	13 (36.1)	34 (50.7)	0.55 (0.24-1.26)	
Abdominal pain (cramps)	103	No	8 (22.2)	19 (28.4)	Ref.	0.500
		Yes	28 (77.8)	48 (71.6)	1.38 (0.54-3.58)	
Fever	103	No	20 (55.6)	39 (58.2)	Ref.	0.795
		Yes	16 (44.4)	28 (41.8)	1.11 (0.49-2.52)	
Median number of symptoms	103	Median (range)	2 (1-5)	3 (1-4)	0.91 (0.60-1.39)	$0.667^{C}/0.593^{M}$
Feeling still ill when filling the questionnaire	102 (1)	No	24 (66.7)	46 (69.7)	Ref.	0.753
		Yes	12 (33.3)	20 (30.3)	1.15 (0.48-2.74)	
Median length of illness (no. days) in patients no longer ill	68	Median (range)	14 (1-180)	15 (4-80)	1.00 (0.98-1.02)	$0.810^{C}/0.183^{M}$
Median normal activity prevented (no. days)	48 (53.4)	Median (range)	2 (0-90)	5 (0-40)	1.02 (0.98-1.06)	$0.444^{C}/0.261^{M}$
Admitted to hospital due to illness	92 (10.7)	No	26 (81.3)	57 (95)	Ref.	0.034
		Yes	6 (18.8)	3 (5)	4.38 (1.02-18.91)	

*percentages refer to the proportion of participants with missing information for the variable; **number of cases genotyped as assemblage A or B with percentages indicating the proportion among cases that answered the question; ***Odds ratio with 95% confidence interval; ^{C} logistic regression on the continuous variable; ^{M}Mann-Whitney U-test

Table 5.2: Comparison of the self-reported clinical outcomes between symptomatic patients infected with assemblage A (n=28) and with assemblage B (n=57) in the case-control study. Odds ratios were calculated using assemblage B as the baseline category.

			Assemblage n (%)*			
Variable	**No. valid (% missing)**	**Category**	**A**	**B**	**OR (95% CI)*** **	***p*-value**
Diarrhoea	83 (2.4)	No	0 (0)	1 (1.8)	Ref.	1.000
		Yes	28 (100)	54 (98.2)	n/a****	
Type of diarrhoea	75 (11.8)	Persistent	13 (46.4)	26 (55.3)	Ref.	0.456
		Intermittent	15 (53.6)	21 (44.7)	1.43 (0.56-3.66)	
Blood in stools	71 (16.5)	No	19 (86.4)	49 (100)	Ref.	0.027
		Yes	3 (13.6)	0 (0)	n/a****	
Vomiting	72 (15.3)	No	14 (63.6)	18 (36)	Ref.	0.030
		Yes	8 (36.4)	32 (64)	0.32 (0.11-0.91)	
Abdominal pain	77 (9.4)	No	9 (36)	8 (15.4)	Ref.	0.041
		Yes	16 (64)	44 (84.6)	0.32 (0.11-0.98)	
Fever	68 (20)	No	12 (63.2)	23 (46.9)	Ref.	0.230
		Yes	7 (36.8)	26 (53.1)	0.52 (0.17-1.53)	
Flatulence	72 (15.3)	No	5 (22.7)	6 (12)	Ref.	0.293
		Yes	17 (77.3)	44 (88)	0.46 (0.12-1.72)	
Swollen stomach	72 (15.3)	No	16 (66.7)	17 (35.4)	Ref.	0.012
		Yes	8 (33.3)	31 (64.9)	0.27 (0.10-0.77)	
Tiredness	77 (9.4)	No	7 (26.9)	13 (25.5)	Ref.	0.892
		Yes	19 (73.1)	38 (74.5)	0.93 (0.32-2.71)	
Loss of appetite	77 (9.4)	No	10 (41.7)	9 (17)	Ref.	0.020
		Yes	14 (58.3)	44 (83)	0.29 (0.10-0.85)	
Loss of weight	74 (12.9)	No	12 (52.2)	16 (31.4)	Ref.	0.088
		Yes	11 (47.8)	35 (68.6)	0.42 (0.15-1.15)	
Amount of weight loss (kilograms) in patients reporting it	38	Median (range)	5 (1- 9.5)	3.6 (1.8-10)	1.00 (0.75-1.32)	$0.987^{C}/0.899^{M}$
Median number of symptoms	83 (2.4)	Median (range)	5 (1-9)	7 (1-9)	0.70 (0.56-0.88)	$0.003^{C}/0.001^{M}$
Feeling still ill when filling the questionnaire	81 (4.7)	No	17 (63)	35 (64.8)	Ref.	0.870
		Yes	10 (37)	19 (35.2)	1.08 (0.41-2.83)	

Median length of illness (no. days) in patients no longer ill	41	Median (range)	11 (4-50)	20.5 (6-73)	0.95 (0.88-1.02)	0.160[C]/0.043[M]
Admitted to hospital due to illness	83 (2.4)	No	23 (82.1)	51 (92.7)	Ref.	0.157
		Yes	5 (17.9)	4 (7.3)	2.77 (0.68-11.28)	

*percentages refer to the proportion of participants with missing information for the variable; **number of cases genotyped as assemblage A or B with percentages indicating the proportion among cases that answered the question; ***Odds ratio with 95% confidence interval; ****Odds ratio not calculated because no variation present in the data; [C]Logistic regression on the continuous variable; [M]Mann-Whitney U-test.

5.4.4 Comparison of reported exposures of *Giardia* assemblage A and B cases from the surveillance dataset

The exposures reported in the month prior to illness by cases from the enhanced surveillance study were compared by case-case analysis between the 39 cases infected with assemblage A and 69 cases infected with assemblage B that returned the study questionnaire. Exposures that were found to be different between the two groups of cases with a significance level of $p \leq 0.2$ are shown in **Table 5.3**, whereas all the remaining comparisons ($p > 0.2$) are reported in **Table 5A**, **Appendix 5**. The only strictly significant differences ($p < 0.05$) were found in two variables: assemblage B cases reported more frequently both touching animals other than livestock, horses, birds or a deer while visiting a premise with animals (exposure that was not reported by any of the assemblage A cases), and drinking fruit juice. The majority (85%, 17 out of 20) assemblage B cases that reported travel outside the UK travelled to a destination potentially at-risk for *Giardia*, whereas the assemblage A cases that travelled abroad were more evenly distributed between destinations at-risk (53.8%, seven out of 13) and destination not at-risk (46.1%, six out of 13). Amongst the cases answering the question, 44% of assemblage B cases (28 out of 63) had a child in the household against the 27% (10 out of 37) of assemblage A cases.

Table 5.3: Exposure variables different at $p \leq 0.2$ in single variable analysis between cases infected with assemblage A (n=39) and B (n=69) that returned the surveillance study questionnaire. Odds ratios were calculated using assemblage B as the baseline category.

			Assemblage n (%)**			
Variable	**No. valid (% missing)***	**Category**	**A**	**B**	**OR (95% CI)*****	***p*-value**
Travelling abroad to an at-risk destination	105 (2.8)	Not travelling abroad	25 (65.8)	47 (70.1)	Ref.	0.133
		Travelling abroad to a not at-risk destination	6 (15.8)	3 (4.5)	3.76 (0.87-16.33)	
		Travelling abroad to any at-risk destination	7 (18.4)	17 (25.4)	0.77 (0.28-2.11)	
Practised fieldsports	108	No	35 (89.7)	67 (97.1)	Ref.	0.186
		Yes	4 (10.3)	2 (2.9)	3.83 (0.67-21.94)	
Went caravanning	108	No	35 (89.7)	67 (97.1)	Ref.	0.186
		Yes	4 (10.3)	2 (2.9)	3.83 (0.67-21.94)	
Keeping any other pet	102 (5.5)	No	36 (100)	61 (92.4)	Ref.	0.158
		Yes	0 (0)	5 (7.6)	n/a****	
Touching any other animal while visiting a premise with animals	91 (15.7)	No	32 (100)	51 (86.4)	Ref.	0.047
		Yes	0 (0)	8 (13.6)	n/a****	
Eating any uncooked vegetables	94 (13)	No	36 (60)	25 (73.5)	Ref.	0.187
		Yes	24 (40)	9 (26.5)	0.54 (0.21-1.36)	
Drinking any fruit or vegetable juice	97 (10.2)	No	13 (37.1)	10 (16.1)	Ref.	0.019
		Yes	22 (62.9)	52 (83.9)	0.32 (0.12-0.85)	
No. of children (<16y) in the house	104 (3.7)	0	27 (73)	35 (52.2)	Ref.	$0.146^{CT}/0.012^{C}$
		1	5 (13.5)	9 (13.4)	0.72 (0.22-2.40)	
		2	4 (10.8)	16 (23.9)	0.32 (0.10-1.08)	
		3	1 (2.7)	5 (7.5)	0.26 (0.03-2.35)	
		4+	0 (0)	2 (3)	n/a****	
Any child in nappies	100 (7.4)	No children in the house	27 (73)	35 (55.6)	Ref.	0.164
		At least one child but not in nappies	4 (10.8)	15 (23.8)	0.35 (0.10-1.16)	
		At least one child in nappies	6 (16.2)	13 (20.6)	0.60 (0.20-1.78)	

*percentage refers to the proportion of participants with missing information for the variable; **percentages refer to the proportion among the cases that answered the question; ***Odds ratio with 95% confidence interval; **** Odds ratio not calculated because no variation present in the data; ^{C}Logistic regression on the continuous variable; ^{CT} Logistic regression on the categorized variable

5.4.5 Risk factors for *Giardia* assemblage A and B infection

The risk factors for infection with the two assemblages were explored using the case-control study dataset. Socio-demographic and exposure variables that showed an association with infection with either assemblage A or B with a *p*-value ≤0.2 in univariable analysis are reported in **Table 5.4**. The results of the remaining variables (*p*>0.2) are shown in **Table 5B**, **Appendix 5**.

Taking medicines for indigestion was positively associated with assemblage B infection (OR 2.44, 95% CI 1.31-4.53, *p*=0.004).

A negative association between keeping a pet and assemblage B infection was found (OR 0.49, 95% CI 0.26-0.92, *p*=0.025). By considering the type of pet owned, dogs were the animals most frequently reported by both assemblage A (22.2%) and B (10.7%). However, owning a dog was positively associated with assemblage A infection only (OR 2.47, 95% CI 1.08-5.66, *p*=0.029). Similarly to pet ownership, touching pets was negatively associated with assemblage B infection (OR 0.47, 95% CI 0.25-0.87, *p*=0.015). Visiting a wildlife park or zoo was reported more frequently in both assemblage A and B cases (10.7%) compared to controls (3.6%), however the association was significant only for assemblage B (OR 3.25, 95% CI 1.08-9.80, *p*=0.039). Travelling abroad was reported more frequently in both assemblage A and B cases compared to controls, although the association was significant for assemblage B only (OR 2.97, 95% CI 1.34-6.56, *p*=0.005). All the assemblage B cases travelled to at-risk countries (e.g. Middle East and Asia including Turkey, Africa including Canary Islands, Central and South America).

Swimming in a swimming pool was significantly and positively associated with infection with both assemblage A (OR 2.89, 95% CI 1.27-6.58, *p*=0.009) and B (OR 2.23, 95% CI 1.19-4.19, *p*=0.011). Using a Jacuzzi or a hot tub was significantly and positively associated with assemblage B infection (OR 4.70, 95% CI 2.02-10.97, *p*<0.001). A negative association between practising fieldsports and infection with assemblage B was found (OR 0.14, 95% CI 0.02-1.08, *p*=0.030). A negative association between walking in the countryside and infection with assemblage A was found (OR 0.34, 95% CI 0.12-0.95, *p*=0.033). Doing gardening was reported more frequently in controls (48.4%) than in both assemblage A (32%) or B (25.5%) cases but the association was significant for assemblage B only (OR 0.36, 95% CI 0.18-0.72, *p*=0.003).

Assemblage B cases reported significantly more frequently drinking bottled water than controls (OR 1.84, 95% CI 1.00-3.39, p=0.047). A positive association was found between assemblage B infection and reporting household tap water both having unusual taste (OR 5.70, 95% CI 1.24-26.26, p=0.031) or being discoloured (OR 5.42, 95% CI 1.41-20.93, p=0.018).

Eating salad or raw vegetables was more frequently reported by controls compared with both assemblage A and B cases, but the negative association was significant for assemblage B only (OR 0.43, 95% CI 0.19-1.00, p=0.045). .

A significant association was found between assemblage B infection and having a child attending nursery or a playgroup in the household (OR 2.70, 95% CI 1.39-5.25, p=0.004). A significant association was found only between assemblage B infection and nappy changing (OR 2.93, 95% CI 1.32-6.51, p=0.034) and between assemblage B infection and reporting having another person ill with similar symptoms in the household (OR 2.19, 95% CI 0.99-4.85, p=0.048).

Table 5.4: Variables showing an association with a $p \leq 0.2$ in single variable analysis with infection with either assemblage A (n=28 cases) or assemblage B (n=57 cases) compared with healthy controls (n=226) that returned the case-control study questionnaire.

Variable	Assemblage subset	No. valid (% missing)*	Category	Cases n (%)**	Controls n (%)**	OR (95% CI)***	*p*-value
SOCIODEMOGRAPHICS AND AREA AND SEASON VARIABLES							
Age (years)	A	254	0-4y	5 (17.9)	26 (11.5)	Ref.	$0.167^{CT}/0.733^{C}$
			5-14y	2 (7.1)	12 (5.3)	0.87 (0.15-5.12)	
			15-44y	5 (17.9)	62 (27.4)	0.42 (0.11-1.57)	
			45-64y	5 (17.9)	75 (33.2)	0.35 (0.09-1.29)	
			65+y	11 (39.3)	51 (22.6)	1.12 (0.35-3.57)	
	B	283	0-4y	8 (14)	26 (11.5)	Ref.	$0.017^{CT}/0.047^{C}$
			5-14y	2 (3.5)	12 (5.3)	0.54 (0.10-2.94)	
			15-44y	28 (49.1)	62 (27.4)	1.47 (0.59-3.64)	
			45-64y	13 (22.8)	75 (33.2)	0.56 (0.21-1.51)	
			65+y	6 (10.5)	51 (22.6)	0.38 (0.12-1.22)	
HEALTH DETAILS							
Taking any medicine for indigestion	A	252 (0.8)	No	21 (77.8)	175 (77.8)	Ref.	1.000
			Yes	6 (22.2)	50 (22.2)	1.00 (0.38-2.61)	
	B	281 (0.7)	No	33 (58.9)	175 (77.8)	Ref.	0.004
			Yes	23 (41.1)	50 (22.2)	2.44 (1.31-4.53)	
ANIMAL CONTACT							
Keeping a pet (any type)	A	253 (0.4)	No	12 (44.4)	120 (53.1)	Ref.	0.395
			Yes	15 (55.6)	106 (46.9)	1.41 (0.63-3.16)	
	B	282 (0.4)	No	39 (69.6)	120 (53.1)	Ref.	0.025
			Yes	17 (30.4)	106 (46.9)	0.49 (0.26-0.92)	
Keeping a pet by species	A	252 (0.8)	Not keeping pets	12 (44.4)	120 (53.3)	Ref.	0.347
			Keeping dogs	6 (22.2)	34 (15.1)	1.76 (0.62-5.05)	
			Keeping cats	3 (11.1)	25 (11.1)	1.20 (0.31-4.57)	
			Keeping birds	0 (0)	7 (3.1)	n/a****	
			Keeping horses	0 (0)	2 (0.9)	n/a****	
			Keeping rodents	0 (0)	7 (3.1)	n/a****	

			Keeping a pet other than those above (fish, amphibians, stick insects, spiders)	0 (0)	5 (2.2)	n/a****	
			Keeping more than one type of pet	6 (22.2)	25 (11.1)	2.40 (0.82-7.00)	
	B	281 (0.7)	Not keeping pets	39 (69.6)	120 (53.3)	Ref.	0.224
			Keeping dogs	6 (10.7)	34 (15.1)	0.54 (0.21-1.39)	
			Keeping cats	2 (3.6)	25 (11.1)	0.25 (0.06-1.09)	
			Keeping birds	2 (3.6)	7 (3.1)	0.88 (0.17-4.41)	
			Keeping horses	0 (0)	2 (0.9)	n/a****	
			Keeping rodents	1 (1.8)	7 (3.1)	0.44 (0.05-3.68)	
			Keeping a pet other than those above (fish, amphibians, stick insects, spiders)	0 (0)	5 (2.2)	n/a****	
			Keeping more than one type of pet	6 (10.7)	25 (11.1)	0.74 (0.28-1.93)	
Keeping a dog	A	252 (0.8)	No	16 (59.3)	176 (78.2)	Ref.	0.029
			Yes	11 (40.7)	49 (21.8)	2.47 (1.08-5.66)	
	B	281 (0.7)	No	45 (80.4)	176 (78.2)	Ref.	0.727
			Yes	11 (19.6)	49 (21.8)	0.88 (0.42-1.82)	
No. of dogs kept	A	252 (0.8)	0	16 (59.3)	176 (78.2)	Ref.	$0.054^{CT}/0.067^{C}$
			1	9 (33.3)	44 (19.6)	2.25 (0.93-5.43)	
			2	2 (7.4)	2 (0.9)	11.00 (1.45-83.39)	
			3+	0 (0)	3 (1.3)	n/a****	
	B	280 (1.1)	0	45 (81.8)	176 (78.2)	Ref.	$0.021^{CT}/0.970^{C}$
			1	6 (10.9)	44 (19.6)	0.53 (0.21-1.33)	
			2	4 (7.3)	2 (0.9)	7.82 (1.39-44.06)	
			3+	0 (0)	3 (1.3)	n/a****	
No. of cats kept	A	252 (0.8)	0	22 (78.6)	182 (81.2)	Ref.	$0.840^{CT}/0.758^{C}$
			1	4 (14.3)	30 (13.4)	1.10 (0.35-3.43)	
			2	2 (7.1)	10 (4.5)	1.65 (0.34-8.04)	
			3+	0 (0)	2 (0.9)	n/a****	
	B	279 (1.4)	0	51 (92.7)	182 (81.2)	Ref.	$0.058^{CT}/0.011^{C}$
			1	4 (7.3)	30 (13.4)	0.48 (0.16-1.41)	
			2	0 (0)	10 (4.5)	n/a****	
			3+	0 (0)	2 (0.9)	n/a****	

No. of different pet species owned	A	252 (0.8)	0	12 (44.4)	120 (53.3)	Ref.	0.067[CT]/0.027[C]
			1	9 (33.3)	79 (35.1)	1.14 (0.46-2.83)	
			2	2 (7.4)	21 (9.3)	0.95 (0.20-4.56)	
			3+	4 (14.8)	5 (2.2)	8.00 (1.89-33.85)	
	B	281 (0.7)	0	39 (69.6)	120 (53.3)	Ref.	0.121[CT]/0.146[C]
			1	11 (19.6)	79 (35.1)	0.43 (0.21-0.89)	
			2	5 (8.9)	21 (9.3)	0.73 (0.26-2.07)	
			3+	1 (1.8)	5 (2.2)	0.61 (0.07-5.43)	
Cleaning up (touching) pets' faeces	A	252 (0.8)	Not keeping pets	12 (44.4)	120 (53.3)	Ref.	0.532
			Keeping pets but not cleaning up their faeces	10 (37)	60 (26.7)	1.67 (0.68-4.08)	
			Keeping pets and cleaning up their faeces	5 (18.5)	45 (20)	1.11 (0.37-3.33)	
	B	281 (0.7)	Not keeping pets	39 (69.6)	120 (53.3)	Ref.	0.039
			Keeping pets but not cleaning up their faeces	7 (12.5)	60 (26.7)	0.36 (0.15-0.85)	
			Keeping pets and cleaning up their faeces	10 (17.9)	45 (20)	0.68 (0.31-1.48)	
Any pet with diarrhoea in the house	A	252 (0.8)	Not keeping pets	12 (44.4)	120 (53.3)	Ref.	0.345
			Keeping pets but none with diarrhoea	15 (55.6)	100 (44.4)	1.50 (0.67-3.35)	
			Keeping pets and at least one with diarrhoea	0 (0)	5 (2.2)	n/a****	
	B	281 (0.7)	Not keeping pets	39 (69.6)	120 (53.3)	Ref.	0.081
			Keeping pets but none with diarrhoea	16 (28.6)	100 (44.4)	0.49 (0.26-0.93)	
			Keeping pets and at least one with diarrhoea	1 (1.8)	5 (2.2)	0.61 (0.07-5.43)	
Touching any pet (either own or other people's)	A	234 (7.9)	No	12 (44.4)	77 (37.2)	Ref.	0.466
			Yes	15 (55.6)	130 (62.8)	0.74 (0.33-1.66)	
	B	259 (8.5)	No	29 (55.8)	77 (37.2)	Ref.	0.015
			Yes	23 (44.2)	130 (62.8)	0.47 (0.25-0.87)	
Visiting or working at a wildlife park or zoo	A	253 (0.4)	No	25 (89.3)	217 (96.4)	Ref.	0.080
			Yes	3 (10.7)	8 (3.6)	3.25 (0.81-13.07)	
	B	281 (0.7)	No	50 (89.3)	217 (96.4)	Ref.	0.039
			Yes	6 (10.7)	8 (3.6)	3.25 (1.08-9.80)	
Visiting any other	A	252 (0.8)	No	25 (92.6)	211 (93.8)	Ref.	0.684

premise(s) with animals			Yes	2 (7.4)	14 (6.2)	1.21 (0.26-5.62)	
	B	280 (1.1)	No	55 (100)	211 (93.8)	Ref.	0.080
			Yes	0 (0)	14 (6.2)	n/a****	
Touching animals in the wild	A	213 (16.1)	No	26 (100)	182 (97.3)	Ref.	1.000
			Yes	0 (0)	5 (2.7)	n/a****	
	B	239 (15.5)	No	48 (92.3)	182 (97.3)	Ref.	0.107
			Yes	4 (7.7)	5 (2.7)	3.03 (0.78-11.73)	
TRAVEL DETAILS							
Travelling abroad (outside the UK)	A	252 (0.8)	No	21 (80.8)	207 (91.6)	Ref.	0.084
			Yes	5 (19.2)	19 (8.4)	2.59 (0.88-7.66)	
	B	282 (0.4)	No	44 (78.6)	207 (91.6)	Ref.	0.005
			Yes	12 (21.4)	19 (8.4)	2.97 (1.34-6.56)	
Travelling abroad to an at-risk destination	A	252 (0.8)	Not travelling abroad	21 (80.8)	207 (91.6)	Ref.	0.167
			Travelling abroad to a not at-risk destination	5 (19.2)	17 (7.5)	2.90 (0.97-8.65)	
			Travelling abroad to any at-risk *destination (e.g. Middle East and Asia including Turkey, Africa including Canary Islands, Central and South America)*	0 (0)	2 (0.9)	n/a****	
	B	282 (0.4)	Not travelling abroad	44 (78.6)	207 (91.6)	Ref.	<0.001
			Travelling abroad to a not at-risk destination	0 (0)	17 (7.5)	n/a****	
			Travelling abroad to any at-risk *destination (e.g. Middle East and Asia including Turkey, Africa including Canary Islands, Central and South America)*	12 (21.4)	2 (0.9)	28.23 (6.1-130.61)	
Travelling in the UK (England, Wales, Scotland)	A	251 (1.2)	No	19 (70.4)	155 (69.2)	Ref.	0.901
			Yes	8 (29.6)	69 (30.8)	0.95 (0.39-2.26)	
	B	278 (1.8)	No	44 (81.5)	155 (69.2)	Ref.	0.072
			Yes	10 (18.5)	69 (30.8)	0.51 (0.24-1.07)	
RECREATIONAL ACTIVITIES							
Swimming or paddling in a swimming pool	A	253 (0.4)	No	15 (55.6)	177 (78.3)	Ref.	0.009
			Yes	12 (44.4)	49 (21.7)	2.89 (1.27-6.58)	

	B	281 (0.7)	No	34 (61.8)	177 (78.3)	Ref.	0.011
			Yes	21 (38.2)	49 (21.7)	2.23 (1.19-4.19)	
Immersing the head underwater while swimming or paddling in a swimming pool	A	240 (5.5)	Not going to swimming pool	15 (62.5)	177 (81.9)	Ref.	0.109[CT]
			Swimming or paddling without immersing the head	1 (4.2)	4 (1.9)	2.95 (0.31-28.09)	
			Swimming or paddling immersing the head	8 (33.3)	35 (16.2)	2.70 (1.06-6.85)	
	B	265 (6.4)	Not going to swimming pool	34 (69.4)	177 (81.9)	Ref.	0.025
			Swimming or paddling without immersing the head	5 (10.2)	4 (1.9)	6.51 (1.66-25.48)	
			Swimming or paddling immersing the head	10 (20.4)	35 (16.2)	1.49 (0.67-3.29)	
Using a Jacuzzi or a hot tub	A	241 (5.1)	No	24 (100)	203 (93.5)	Ref.	1.000
			Yes	0 (0)	14 (6.5)	n/a****	
	B	266 (6)	No	37 (75.5)	203 (93.5)	Ref.	<0.001
			Yes	12 (24.5)	14 (6.5)	4.70 (2.02-10.97)	
Practising fieldsports (e.g. football, golf etc)	A	238 (6.3)	No	23 (88.5)	186 (87.7)	Ref.	1.000
			Yes	3 (11.5)	26 (12.3)	0.93 (0.26-3.33)	
	B	263 (7.1)	No	50 (98)	186 (87.7)	Ref.	0.030
			Yes	1 (2)	26 (12.3)	0.14 (0.02-1.08)	
Walking in the countryside	A	246 (3.1)	No	20 (80)	128 (57.9)	Ref.	0.033
			Yes	5 (20)	93 (42.1)	0.34 (0.12-0.95)	
	B	272 (3.9)	No	28 (54.9)	128 (57.9)	Ref.	0.695
			Yes	23 (45.1)	93 (42.1)	1.13 (0.61-2.09)	
Frequency of walking in the countryside (no. times)	B	243 (14.1)	0	28 (71.8)	128 (62.7)	Ref.	0.087[CT]/0.234[C]
			1-2	2 (5.1)	38 (18.6)	0.24 (0.05-1.06)	
			3-4	7 (17.9)	19 (9.3)	1.68 (0.65-4.39)	
			5-6	1 (2.6)	6 (2.9)	0.76 (0.09-6.58)	
			7+	1 (2.6)	13 (6.4)	0.35 (0.04-2.80)	
Doing gardening	A	242 (4.7)	No	17 (68)	112 (51.6)	Ref.	0.120
			Yes	8 (32)	105 (48.4)	0.50 (0.21-1.21)	
	B	268 (5.3)	No	38 (74.5)	112 (51.6)	Ref.	0.003
			Yes	13 (25.5)	105 (48.4)	0.36 (0.18-0.72)	

Frequency of doing gardening (no.times)	B	235 (17)	0	38 (90.5)	112 (58)	Ref.	<0.001[CT,C]
			1-2	3 (7.1)	31 (16.1)	0.28 (0.08-0.99)	
			3-4	0 (0)	27 (14)	n/a****	
			5-6	0 (0)	11 (5.7)	n/a****	
			7+	1 (2.4)	12 (6.2)	0.25 (0.03-1.95)	
WATER CONSUMPTION							
Drinking bottled water	A	239 (5.9)	No	15 (57.7)	115 (54)	Ref.	0.721
			Yes	11 (42.3)	98 (46)	0.86 (0.38-1.96)	
	B	267 (5.7)	No	21 (38.9)	115 (54)	Ref.	0.047
			Yes	33 (61.1)	98 (46)	1.84 (1.00-3.39)	
Water from the tap reported having an unusual taste	A	247 (2.8)	No	26 (100)	218 (98.6)	Ref.	1.000
			Yes	0 (0)	3 (1.4)	n/a****	
	B	276 (2.5)	No	51 (92.7)	218 (98.6)	Ref.	0.031
			Yes	4 (7.3)	3 (1.4)	5.70 (1.24-26.26)	
Water from the tap reported being discoloured	A	247 (2.8)	No	26 (100)	217 (98.2)	Ref.	1.000
			Yes	0 (0)	4 (1.8)	n/a****	
	B	276 (2.5)	No	50 (90.9)	217 (98.2)	Ref.	0.018
			Yes	5 (9.1)	4 (1.8)	5.42 (1.41-20.93)	
FOOD CONSUMPTION							
FOOD CONSUMPTION HABITS							
No. times per week eating raw fruit	A	244 (3.9)	0	4 (14.8)	15 (6.9)	Ref.	0.063[CT]/0.095[C]
			1-2	5 (18.5)	27 (12.4)	0.69 (0.16-2.99)	
			3-4	7 (25.9)	29 (13.4)	0.90 (0.23-3.59)	
			5+	11 (40.7)	146 (67.3)	0.28 (0.08-0.99)	
	B	273 (3.5)	0	6 (10.7)	15 (6.9)	Ref.	0.318[CT]/0.101[C]
			1-2	5 (8.9)	27 (12.4)	0.46 (0.12-1.78)	
			3-4	12 (21.4)	29 (13.4)	1.03 (0.32-3.30)	
			5+	33 (58.9)	146 (67.3)	0.56 (0.20-1.57)	
FOOD CONSUMPTION DURING THE EXPOSURE WINDOW							
Eating salads or raw vegetables	A	247 (2.8)	No	5 (18.5)	19 (8.6)	Ref.	0.157
			Yes	22 (81.5)	201 (91.4)	0.42 (0.14-1.22)	
	B	276 (2.5)	No	10 (17.9)	19 (8.6)	Ref.	0.045

			Yes	46 (82.1)	201 (91.4)	0.43 (0.19-1.00)	
Provenience of salads or raw vegetables	A	245 (3.5)	Did not eat salad or raw vegetables	5 (18.5)	19 (8.7)	Ref.	0.532
			Market	1 (3.7)	8 (3.7)	0.47 (0.05-4.74)	
			Greengrocers	0 (0)	7 (3.2)	n/a****	
			Supermarket	18 (66.7)	141 (64.7)	0.48 (0.16-1.46)	
			Homegrown	0 (0)	2 (0.9)	n/a****	
			Other place (e.g. restaurant etc.)	0 (0)	1 (0.5)	n/a****	
			Multiple proveniences (at least two different)	3 (11.1)	40 (18.3)	0.28 (0.06-1.32)	
	B	270 (4.6)	Did not eat salad or raw vegetables	10 (19.2)	19 (8.7)	Ref.	0.001
			Market	5 (9.6)	8 (3.7)	1.19 (0.31-4.60)	
			Greengrocers	2 (3.8)	7 (3.2)	0.54 (0.09-3.12)	
			Supermarket	28 (53.8)	141 (64.7)	0.38 (0.16-0.90)	
			Homegrown	0 (0)	2 (0.9)	n/a****	
			Other place (e.g. restaurant etc.)	4 (7.7)	1 (0.5)	7.60 (0.75-77.43)	
			Multiple proveniences (at least two different)	3 (5.8)	40 (18.3)	014 (0.03-0.58)	
Provenience of cooked vegetables	A	242 (4.7)	Did not eat cooked vegetables	1 (3.7)	14 (6.5)	Ref.	0.877
			Market	2 (7.4)	11 (5.1)	2.54 (0.20-31.86)	
			Greengrocers	1 (3.7)	8 (3.7)	1.75 (0.10-31.96)	
			Supermarket	20 (74.1)	142 (66)	1.97 (0.25-15.82)	
			Homegrown	0 (0)	3 (1.4)	n/a****	
			Other place (e.g. restaurant etc.)	0 (0)	2 (0.9)	n/a****	
			Multiple proveniences (at least two different)	3 (11.1)	35 (16.3)	1.20 (0.11-12.54)	
	B	264 (6.7)	Did not eat cooked vegetables	4 (8.2)	14 (6.5)	Ref.	0.117
			Market	3 (6.1)	11 (5.1)	0.95 (0.18-5.19)	
			Greengrocers	2 (4.1)	8 (3.7)	0.87 (0.13-5.90)	
			Supermarket	32 (65.3)	142 (66)	0.79 (0.24-2.56)	
			Homegrown	0 (0)	3 (1.4)	n/a****	
			Other place (e.g. restaurant etc.)	4 (8.2)	2 (0.9)	7.00 (0.92-53.23)	
			Multiple proveniences (at least two different)	4 (8.2)	35 (16.3)	0.40 (0.09-1.83)	

Provenience of fruit juice	A	241 (5.1)	Did not drink fruit juice	8 (29.6)	58 (27.1)	Ref.	0.217
			Market	0 (0)	1 (0.5)	n/a****	
			Greengrocers	1 (3.7)	0 (0)	n/a****	
			Supermarket	18 (66.7)	145 (67.8)	0.90 (0.37-2.18)	
			Other place (e.g. restaurant etc.)	0 (0)	2 (0.9)	n/a****	
			Multiple proveniences (at least two different)	0 (0)	8 (3.7)	n/a****	
	B	266 (6)	Did not drink fruit juice	14 (26.9)	58 (27.1)	Ref.	0.079
			Market	1 (1.9)	1 (0.5)	4.14 (0.24-70.38)	
			Supermarket	34 (65.4)	145 (67.8)	0.97 (0.49-1.94)	
			Other place (e.g. restaurant etc.)	3 (5.8)	2 (0.9)	6.21 (0.95-40.81)	
			Multiple proveniences (at least two different)	0 (0)	8 (3.7)	n/a****	
Eating beef	A	245 (3.5)	No	3 (12)	56 (25.5)	Ref.	0.136
			Yes	22 (88)	164 (74.5)	2.50 (0.72-8.69)	
	B	273 (3.5)	No	9 (17)	56 (25.5)	Ref.	0.194
			Yes	44 (83)	164 (74.5)	1.67 (0.77-3.64)	
Provenience of chicken	A	247 (2.8)	Did not eat chicken	2 (8)	18 (8.1)	Ref.	0.736
			Market stall	0 (0)	4 (1.8)	n/a****	
			Butcher	2 (8)	29 (13.1)	0.62 (0.08-4.80)	
			Supermarket	17 (68)	149 (67.1)	1.03 (0.22-4.81)	
			Other place (e.g. restaurant etc.)	0 (0)	2 (0.9)	n/a****	
			Multiple proveniences (at least two different)	4 (16)	20 (9)	1.80 (0.29-11.03)	
	B	271 (4.2)	Did not eat chicken	2 (4.1)	18 (8.1)	Ref.	0.054
			Market stall	1 (2)	4 (1.8)	2.25 (0.16-31.33)	
			Butcher	6 (12.2)	29 (13.1)	1.86 (0.34-10.24)	
			Supermarket	30 (61.2)	149 (67.1)	1.81 (0.40-8.22)	
			Other place (e.g. restaurant etc.)	5 (10.2)	2 (0.9)	22.50 (2.50-202.29)	
			Multiple proveniences (at least two different)	5 (10.2)	20 (9)	2.25 (0.39-13.07)	
Eating shellfish	A	212 (16.5)	No	17 (81)	139 (72.8)	Ref.	0.420

			Yes	4 (19)	52 (27.2)	0.63 (0.20-1.96)	
	B	237 (16.3)	No	39 (84.8)	139 (72.8)	Ref.	0.091
			Yes	7 (15.2)	52 (27.2)	0.48 (0.20-1.14)	
Provenience of shellfish	A	212 (16.5)	Did not eat shellfish	17 (81)	139 (72.8)	Ref.	0.194
			Fishmonger	1 (4.8)	3 (1.6)	2.72 (0.27-27.69)	
			Supermarket	2 (9.5)	41 (21.5)	0.40 (0.09-1.80)	
			Other place (e.g. restaurant etc.)	1 (4.8)	1 (0.5)	8.18 (0.49-136.79)	
			Multiple proveniences (at least two different)	0 (0)	7 (2.5)	n/a****	
	B	237 (16.3)	Did not eat shellfish	39 (84.8)	139 (72.8)	Ref.	0.031
			Fishmonger	0 (0)	3 (1.6)	n/a****	
			Supermarket	5 (10.9)	41 (21.5)	0.43 (0.16-1.17)	
			Other place (e.g. restaurant etc.)	2 (4.3)	1 (0.5)	7.13 (0.63-80.69)	
			Multiple proveniences (at least two different)	0 (0)	7 (2.5)	n/a****	
GENERAL HOUSEHOLD DETAILS							
No. of children (<16y) in the house	A	254	0	16 (57.1)	134 (59.3)	Ref.	$0.317^{CT}/0.223^{C}$
			1	5 (17.9)	43 (19)	0.97 (0.34-2.81)	
			2	2 (7.1)	34 (15)	0.49 (0.11-2.25)	
			3	3 (10.7)	11 (4.9)	2.28 (0.58-9.06)	
			4+	2 (7.1)	4 (1.8)	4.19 (0.71-24.70)	
	B	282 (0.4)	0	26 (46.4)	134 (59.3)	Ref.	$0.177^{CT}/0.021^{C}$
			1	9 (16.1)	43 (19)	1.08 (0.47-2.48)	
			2	13 (23.2)	34 (15)	1.97 (0.92-4.23)	
			3	5 (8.9)	11 (4.9)	2.34 (0.75-7.31)	
			4+	3 (5.4)	4 (1.8)	3.86 (0.82-18.30)	
Any child attending a childcare nursery or a playgroup	A	254	No children in the house	16 (57.1)	134 (59.3)	Ref.	0.938
			Children not attending nursery or playgroup	6 (21.4)	50 (22.1)	1.00 (0.37-2.71)	
			At least one child attending nursery or playgroup	6 (21.4)	42 (18.6)	1.19 (0.44-3.25)	
	B	280 (1.1)	No children in the house	26 (48.1)	134 (59.3)	Ref.	0.002

			Children not attending nursery or playgroup	6 (11.1)	50 (22.1)	0.62 (0.24-1.59)	
			At least one child attending nursery or playgroup	22 (40.7)	42 (18.6)	2.70 (1.39-5.25)	
Any child in nappies	A	252 (0.8)	No children in the house	16 (57.1)	134 (59.8)	Ref.	0.092
			At least one child but not in nappies	3 (10.7)	52 (23.2)	0.48 (0.13-1.73)	
			At least one child in nappies	9 (32.1)	38 (17)	1.98 (0.81-4.84)	
	B	278 (1.8)	No children in the house	26 (48.1)	134 (59.8)	Ref.	0.116
			At least one child but not in nappies	12 (22.2)	52 (23.2)	1.19 (056-2.53)	
			At least one child in nappies	16 (29.6)	38 (17)	2.17 (1.06-4.46)	
Changing nappies	A	250 (1.6)	No children in the house or children not in nappies	19 (70.4)	186 (83.4)	Ref.	0.285
			Child in nappies but not changing nappies	4 (14.8)	17 (7.6)	2.30 (0.70-7.55)	
			Child in nappies and changing nappies	4 (14.8)	20 (9)	1.96 (0.61-6.33)	
	B	276 (2.5)	No children in the house or children not in nappies	38 (71.7)	186 (83.4)	Ref.	0.034
			Child in nappies but not changing nappies	3 (5.7)	17 (7.6)	0.86 (0.24-3.09)	
			Child in nappies and changing nappies	12 (22.6)	20 (9)	2.93 (1.32-6.51)	
Any other person with diarrhoea in the house	A	236 (7.1)	No	21 (84)	188 (89.1)	Ref.	0.502
			Yes	4 (16)	23 (10.9)	1.56 (0.49-4.93)	
	B	263 (7.1)	No	41 (78.8)	188 (89.1)	Ref.	0.048
			Yes	11 (21.2)	23 (10.9)	2.19 (0.99-4.85)	

*percentages refer to the proportion of participants with missing information for the variable in the specific dataset; **percentages refer to the proportion among cases and controls that answered the question; ***Odds ratio with 95% confidence interval; *** *Odds ratio not calculated because no variation present in the data; ^{CT} logistic regression on the categorized variable; ^{C} logistic regression on the continuous variable

Following multivariable logistic regression analysis, both keeping a dog and swimming at a swimming pool were found to be independently associated with assemblage A infection (**Table 5.5**). After the exclusion of two suspected secondary household cases, the odds ratio of swimming in a swimming pool showed a reduction from 4.36 (95% CI 1.34-9.52, *p*=0.011) to 3.20 (95% CI 0.87-11.78, *p*=0.080).

Table 5.5: Constructed multivariable model for the risk factors for *Giardia* assemblage A infection based on 243 valid (e.g. without missing data) observations (23 assemblage A cases and 220 controls). Estimates following missing data multiple imputation (100 imputed datasets) are reported in parentheses.

Variables included		**B**	**Adjusted OR**	**95% CI**	***p*-value**
Keeping a dog		1.27 (1.15)	3.57 (3.15)	1.34-9.52 (1.25-7.97)	0.011 (0.015)
Swimming or paddling in a swimming pool		1.47 (1.88)	4.36 (6.57)	1.24-15.36 (2.07-20.84)	0.022 (0.001)
Walking in the countryside		-1.52 (-1.34)	0.22 (0.26)	0.07-0.69 (0.09-0.79)	0.010 (0.017)
Travelling abroad (outside the UK)		0.75 (0.49)	2.12 (1.63)	0.59-7.63 (0.46-5.77)	0.250 (0.444)
Male gender		0.26	1.30	0.49-3.49	0.598
Age	0-4y		Reference		0.662
	5-14y	0.28	1.32	0.18-9.49	0.784
	15-44y	-0.22	0.80	0.16-4.09	0.793
	45-64y	-0.01	0.99	0.21-4.67	0.986
	65+y	0.79	2.20	0.41-11.66	0.356
Constant		*-3.14*			

Model χ^2 (12) = 23.82, *p*=0.005; Nagelkerke R^2 = 0.20; Hosmer and Lemeshow χ^2 (8) = 7.03, *p*=0.533

Following multivariable logistic regression analysis both taking medicines for indigestion prior to illness and using a Jacuzzi or a hot tub were found to be significantly and independently associated with assemblage B infection (**Table 5.6**), and a positive association was also found for having children attending a nursery or playgroup in the household and reporting the water from the tap as being discoloured. After the exclusion of one suspected secondary household case, no relevant changes were observed in the odds ratios of any of the variables.

Table 5.6: Constructed multivariable model for the risk factors for *Giardia* assemblage B based on 257 valid (e.g. without missing data) observations (44 assemblage B cases and 213 controls). Estimates following missing data multiple imputation (100 imputed datasets) are reported in parentheses.

Variables included		B	Adjusted OR	95% CI	*p*-value
Taking medicines for indigestion		1.22 (0.45)	3.40 (1.57)	1.46-7.90 (0.67-3.70)	0.005 (0.104)
Using a Jacuzzi or hot tub		1.42 (0.76)	4.12 (2.13)	1.46-11.67 (0.90-5.06)	0.008 (0.086)
Water coming from the tap reported discoloured		1.63 (0.49)	5.13 (1.63)	0.84-31.35 (0.47-5.68)	0.077 (0.439)
Any child at nursery of playgroup in the household	No children in the house		Reference		0.055
	Children not attending nursery or playgroup	-0.72 (-0.21)	0.49 (0.81)	0.15-1.61 (0.39-1.68)	0.238 (0.577)
	At least one child attending nursery or playgroup	0.77 (0.58)	2.17 (1.79)	0.80-5.92 (0.87-3.72)	0.129 (0.116)
Travelling abroad (outside the UK)		0.88 (1.28)	2.42 (3.58)	0.81-7.25 (1.91-6.73)	0.114 (<0.001)
Male gender		0.19	1.21	0.57-2.56	0.624
Age	0-4y		Reference		0.045
	5-14y	-18.76	n/a***	n/a***	1.000
	15-44y	0.77	2.15	0.59-7.81	0.244
	45-64y	-0.40	0.67	0.15-2.89	0.589
	65+y	-0.78	0.46	0.08-2.70	0.389
Constant		*-2.59*			

***Odds ratio not calculated because no variation present in the data; Model χ^2 (11) = 48.79, p<0.001; Nagelkerke R^2 = 0.29; Hosmer and Lemeshow χ^2 (8) = 12.88, p=0.116

5.4.6 *Giardia* assemblages and multi-locus genotypes in household clusters

Of the 13 patients part of the five distinct household clusters identified in the case-control study, 12 had the infecting *Giardia* assemblage successfully determined (**Table 5.7**). Assemblage A was found in two clusters, and in one of the two (cluster 1) where at least two family members were successfully genotyped the *Giardia* MLG causing infection was the same. Assemblage B was found in three clusters. In the cluster involving four people (cluster 3), the same *Giardia* assemblage B MLG was found in all the cases. In another cluster (cluster 4), the mother and the son were presumably infected with the same MLG (MLG-1) although in the son the *tpi* locus was not successfully amplified and the comparison was possible only at the *bg* and *gdh* locus.

Table 5.7: *Giardia* assemblages and multi-locus genotypes found in 13 cases part of five household clusters.

Cluster	Number of cases involved	Age, gender (family role)	*Giardia* MLG
1	3	1y, male (son)	Assemblage A (*tpi*) – sub-type and MLG not available (*bg* and *gdh* failed)
		6y, male (son)	AII-1 (*bg* A2, *gdh* AII, *tpi* AII)
		33y, male (father)	AII-1 (*bg* A2, *gdh* AII, *tpi* AII)
2	2	3y, male (son)	AII-novel (*bg* A3, *gdh* ECUST2196, *tpi* AII)
		31y, female (mother)	Faecal specimen not available
3	4	1y, male (son)	B-novel (*bg* B1-5, *gdh* BIV, *tpi* VB906855)
		3y, male (son)	B-novel (*bg* B1-5, *gdh* BIV, *tpi* VB906855)
		33y, male (father)	B-novel (*bg* B1-5, *gdh* BIV, *tpi* VB906855)
		34y, female (mother)	B-novel (*bg* B1-5, *gdh* BIV, *tpi* VB906855)
4	2	30y, female (mother)	B MLG 1 (*bg* B1-3, *gdh* BIV, *tpi* BIV)
		1y, male (son)	Full MLG not available (*bg* B1-3, *gdh* BIV, *tpi* failed)
5	2	42y, female (wife)	Assemblage B, MLG not available only *bg* amplified and with heterogeneous positions
		64y, male (husband)	Assemblage B, MLG not available all three loci with heterogeneous positions

5.5 DISCUSSION

We examined potential differences between the *Giardia* assemblages A and B in the clinical outcome of infection or transmission routes. The epidemiological information collected from the cases of giardiasis and the information on the *Giardia* assemblage responsible for infection obtained following molecular typing of the parasite from stools were integrated.

No association between the infecting assemblages and gender was found. Although the gender distribution of assemblage A and B has been rarely compared, our results correspond to those reported in other large surveys of symptomatic patients from developed countries involving genotyping of the parasites (Breathnach *et al.*, 2010; Lebbad *et al.*, 2011). The two assemblages showed a different distribution in relation with the patients' age: the age of the cases infected with B was lower than in assemblage A cases. In particular, following the analysis of specific age groups assemblage B seemed to be more common in young adults, whereas the occurrence of assemblage A was comparatively lower in this group and increased with the patient age (particularly in people 70 years old and over). Although the analysis was done on a relatively small amount of data and the number of genotyped cases was uneven between age groups, this was nevertheless an interesting finding. The age prevalence of *Giardia* assemblages in humans has not been investigated thoroughly in large samples, so no direct comparison of our results with other studies was possible. In the study by Breathnach *et al.* (2010) although both assemblages showed a bimodal age distribution (with a major peak in adults and in children below ten years of age), assemblage B was more prevalent than A in children and adults in their 30s whereas the opposite was seen in people in their forties and particularly in the elderly (60+ years of age). Whether this pattern is related to differences between the age groups in either the frequency of exposure to one or the other assemblage or the development of immunity to one or the other requires further investigation.

No relationship between infection with one of the two assemblages and the month in which the cases were reported was apparent. The variation observed could due to the relatively small numbers of genotyped cases in certain months, and more data are

needed for comparison since the seasonality of occurrence of the two *Giardia* assemblages has never been explored before. An apparent seasonal pattern was observed though: assemblage A infections were lower in the Winter months and increased in the other seasons, whereas assemblage B infections peaked during winter. The significance of this pattern should be confirmed on a larger sample size and in different populations, and whether this pattern reflects true differences in the transmission of the two parasite assemblages remain to be seen.

The self-reported clinical outcomes of infection differed between cases infected with the two assemblages. At least five symptoms were reported more frequently by the cases infected with assemblage B, which also reported a higher number of experienced symptoms at the same time and a longer duration of illness compared to cases infected with assemblage A. Conversely, hospitalisation and the presence of blood in the stools were reported more frequently by assemblage A cases. This finding was probably due to an effect of age: overall assemblage A cases were older than B cases, and cases reporting blood in stools and hospitalisation were older than those not reporting these two outcomes as previously shown in Chapters 2 and 3. Differently from what has been previously reported by Helmy *et al.* (2009) and Homan and Mank (2001), no association between the infecting assemblage and the type of diarrhoea (intermittent or persistent) was observed in our sample. Overall, our results suggest that infection with assemblage B may be associated with a more severe illness. In children from Cuba assemblage B was associated with a higher frequency of diarrhoea, flatulence and abdominal pain (Puebla *et al.*, 2014), and an association between assemblage B infection and flatulence was also observed in children from Sweden (Lebbad *et al.*, 2011). Only symptomatic cases of infection were included in our study. If assemblage B causes a more severe illness in humans, then people infected with this assemblage will refer to their general practitioner more frequently and they will then represent the majority of notified cases. This may explain the higher prevalence of assemblage B infections compared to those caused by assemblage A observed in symptomatic patients from the UK and other developed countries. More data are needed to determine whether assemblage A commonly occurs in people that are asymptomatic or show a relatively mild symptomatology in developed countries, as it has been reported in a few occasions from Ethiopia

(Gelanew *et al.*, 2007), Saudi Arabia (Al-Mohammed, 2011) and Cuba (Puebla *et al.*, 2014).

The successful multi-locus molecular typing of the parasites from the cases part of two household clusters confirmed that the same multi-locus genotype of *Giardia* was responsible for infection amongst the family members. This result confirms that these cases were infected either from the same original source or by person-to-person transmission within the household. Our results confirmed that both assemblages can be implicated in household outbreaks, as shown in previous studies (Breathnach *et al.*, 2010; Lebbad *et al.*, 2011).

The analysis of the exposures reported by the cases infected with the two assemblages revealed both similarities and differences between the two groups. With the exception of the study by Anuar *et al.* (2014) from Malaysia, the epidemiology of the two assemblages has never been thoroughly investigated before. We first used case-case analysis to compare directly the exposure of the cases infected with the two assemblages, but no major differences were found.

Although case-case analysis has been shown to be able to highlight the major risk factors for infection without the need for enrolling proper controls in some instances (Gillespie *et al.*, 2012), the applicability of this approach to detect differences when comparing directly two groups of patients infected with two or more variants of the same pathogens must be considered carefully. It is important to note that using a group of patients infected with a similar pathogen as a "control" group can obscure certain associations with infection: this is due to the fact that the "controls" may share several exposures with the "cases" (Gillespie *et al.*, 2002). If the two *Giardia* assemblages have largely overlapping epidemiological characteristics, then a simple case-case analysis between assemblages can fail to detect differences.

Our findings following the case-control analysis further confirmed the importance of using a proper control group for the evaluation of risk factors associated to a particular variant of the pathogen. It is important to note that the statistical power of the case-control analysis was reduced by a drop in the number of cases successfully genotyped compared to the overall risk factor analysis, by the lower number of assemblage A cases compared to B cases and by the necessity of performing the

analysis using two subsets of cases. Nevertheless, despite these limitations important associations between reported exposures and the infecting assemblage were found.

Travelling abroad was associated with the acquisition of giardiasis regardless the parasite assemblage considered. The significance of travel was stronger for assemblage B infections though. Foreign travel was a highly significant risk factor for giardiasis in general (see Chapter 3). No differences in the travel history of assemblage A and B cases were found by Breathnach *et al.* (2010) either.

An association was observed between swimming in a swimming pool and infection in both assemblages, but following multivariable analysis this exposure was independently associated only with assemblage A infection. Similarly, assemblage B infection was associated with using a Jacuzzi or a hot tub but the exposure lost significance following multivariable analysis.

In case-control analysis, exposures related to the contact with children or other people were consistently and more frequently reported by assemblage B cases. Having children going to a nursery, changing nappies and the presence of another person ill with diarrhoea in the household were all associated with assemblage B infection. No association of these variables with assemblage A infection was apparent. Despite the fact that only having children going to a nursery was retained in the multivariable analysis and was only marginally significant, overall these results suggest that the transmission of assemblage B is predominantly human-to-human and that children play a major role in this. Our findings are in line with those of Anuar *et al.* (2014), which reported the presence of children and other family members infected with giardiasis as significant risk factors for assemblage B infection. In a nursery outbreak in North Wales, all the 21 children, child care workers and parents involved and successfully genotyped were infected with assemblage B (Amar *et al.*, 2002). The age prevalence of assemblage B in our sample was clearly bimodal, with adults and children representing most of B cases. Our data collectively support the hypothesis of assemblage B transmission from the children to their parents or relatives, which is likely to occur via contact with contaminated nappies.

Another important finding was that in our sample dog ownership was significantly and independently associated with assemblage A infection. Close contact with pets as dogs and cats was the only significant predictor for assemblage A infection in the study by Anuar *et al.* (2014). These finding seems to confirm the zoonotic potential of this assemblage compared to B, a hypothesis supported ted by the fact that assemblage A is the most frequent non-host-specific assemblage (Feng & Xiao, 2011). With the exception of one case, all people in our sample were infected with the sub-assemblage AII. Although dogs are more commonly infected by sub-assemblage AI (Sprong *et al.*, 2009; Ryan & Cacciò, 2013), infections with sub-assemblage AII have also been reported in these pets (Ponce-Macotela *et al.*, 2002; Traub *et al.*, 2004; Claerebout *et al.*, 2009). The analysis of the *Giardia* assemblage A multi-locus genotypes in the cases reporting owning a dog failed to find a common MLG between them. Furthermore, faecal samples were not collected from the pets of the *Giardia* cases and so no information on either presence of infection or the parasite genotype was available. No association was found between assemblage A infection and exposure to other animal species. However, one case (a male 53 years old) was found to be infected with sub-assemblage AI that was confirmed at two loci (*bg* and *tpi*) (isolate code 11/13). The *bg* AI sequence of this isolate was identified previously in a series of hosts including horses. Interestingly, the case reported visiting ten times a horse riding school and touching horses in the exposure window, whereas it lacked all the major significant exposures (e.g. travelling abroad, swimming, children in nappies). These results may indicate a zoonotic transmission event from horses. Sub-assemblage AI is not common in humans, and in our study was found only in this particular case. However, in a study from New Zealand the same sub-assemblage AI was found in both calves and farmers from the same area (Winkworth *et al.*, 2008).

To sum up, the results that emerged from integration of molecular and clinico-epidemiological information showed that the two *Giardia* assemblages partially differ in terms of their epidemiology and the clinical outcome of illness they cause in humans. Confirming the findings reported by some previous studies done in developed countries, assemblage B appeared to be slightly more virulent than A. The different age prevalence observed in the two assemblages is an interesting and new finding and it requires further investigation. More importantly, although the

epidemiological profile mostly overlapped between infection with the two assemblages, our analysis seemed to confirm the zoonotic transmission potential of assemblage A and the existence of a mostly human reservoir for assemblage B.

CHAPTER SIX: FINAL CONCLUSIONS

This study was the first of this kind describing the epidemiology and molecular epidemiology of symptomatic giardiasis in the UK, integrating clinico-epidemiological data from cases of disease with the information about the molecular diversity of the parasite. Such comprehensive approach is important to understand the true burden of disease and to highlight important epidemiological differences associated with the biological and genetic variability of the pathogen. The latter aspect, in particular, has not been systematically investigated in the case of giardiasis in developed countries.

We confirmed that the burden of giardiasis can be greatly underestimated if community faecal specimens are not screened systematically using a sensitive method. We detected high rates of disease in adults and the majority of the cases did not report foreign travel: these findings challenged the common view that *Giardia* affects children and is mostly associated with travel abroad. Using a case-control study design we revealed that transmission of *Giardia* within the UK is likely to occur mostly through the human-to-human route in two contexts: swimming pools and households. We confirmed the importance of young children and changing nappies in the transmission of the parasite, particularly within households. Our data have important public health implications, since they revealed that the burden of disease can be higher than expected and that control of this parasitic infection should be focused on swimming pools and in avoiding transmission within the household. Using a multi-locus genotyping approach, we also described for the first time the genetic diversity of *Giardia* in a large number of UK patients, showing that the prevalence and levels of molecular diversity of parasite assemblages in patients from North West England correspond to what has been observed both worldwide. Our most important finding was the association between infection with the two assemblages and two distinct exposures, namely the presence of children and changing nappies for assemblage B and dog ownership for assemblage A. These results seemed to confirm that the two assemblages may have a preferential reservoir: animal for assemblage A and human for assemblage B. The presence of assemblage-specific risk factors has important implications for the biology and control of *Giardia*. First, the infecting *Giardia* assemblage should be determined and considered in all future epidemiological and risk factors studies on giardiasis to account for variation in exposure and to highlight other potential assemblage-specific

exposures Second, our findings give a stronger support to the notion of assemblages as truly separate species and so they are important for the whole *Giardia* and scientific community.

Our work is not without weaknesses, mostly due to the logistical limitations that are inherent to case-control studies. Although our patients were checked for co-infection with other major pathogens, not all the potential aetiological causes of diarrhoea (including chronic conditions such as IBS) were taken into account. Although the presence of *Giardia* was confirmed in all of them, it is possible that in some instances the parasite was not the direct cause of gastroenteritis. This aspect could have caused a certain degree of case misclassification resulting in both the reduction of the strength of certain associations or in the appearance of spurious associations, such as the one we observed between giardiasis and IBS. The case-control study recruitment also suffered from a low response rate. This may have had an effect on our ability to detect additional risk factors for disease due to a reduced study power. If this study had to be repeated, this limitation could be resolved by including additional areas in the study, extending the period of recruitment or contacting the participants through phone in an attempt to increase the response rates. Regarding the genotyping of the parasites, the PCR success rate was not optimal resulting in the loss of several samples and in the reduction of the number of the genotyped cases available for statistical comparisons. The optimization of DNA extraction and a full re-evaluation of the currently available primers (in terms of both their sensitivity and specificity) could allow a better genotyping success at the multi-locus level. As a consequence of the loss of some samples, the analysis of the *Giardia* assemblage-specific risk factors (although it was able to highlight the zoonotic transmission potential of assemblage A) was based on a relative small number of genotyped cases. This may have had an impact on our ability to detect other significant exposures associated with the two assemblages. If the samples collected and genotyped in our study will be re-evaluated in the near future, the use of a more sensitive real-time PCR assay and the inclusion of more loci in the multi-locus typing scheme could be beneficial in both increasing the sample size of successfully genotyped cases in and assigning parasite isolates to specific sub-assemblages and subtypes with increased confidence. Due to logistical limitations, we were not able to determine the presence and the genotypes of *Giardia* in the pets of the case patients. In future studies,

sampling both people and their pets in the household to check for shared parasite genotypes should be attempted to obtain a better picture of the zoonotic transmission dynamics of *Giardia*.

Based on the limitations described above, future work on giardiasis should be focused on some key aspects that are pivotal for our better understanding of both the epidemiology and host-parasite interactions of this pathogen assuming the availability of funds. In public health terms, more extensive prevalence studies (including also the adult population) are needed to determine the actual burden of asymptomatic infections: these may represent a significant public health risk leading to sustained transmission in the community and re-infection within households. As part of this approach, the characteristics of asymptomatic and symptomatic cases should be compared to highlight host factors that may be related to a different clinical course of infection. These data could potentially help explaining the excess of giardiasis that we observed in the male population, and whether immunological and/or epidemiological factors are involved in this (Klein, 2004). More data are also needed to determine the levels of exposure to the parasite at the population level, whether short-term or long-term immunity to re-infection can develop and how these two aspects affect in turn the clinical course of infection (particularly in relation to the type of parasite assemblage causing infection). Large population–based serological studies can help in the elucidation of these aspects, but more data and more reliable antibody-detection methods are needed (Casemore, 2006). The presumptive association between IBS and the presence of *Giardia* also warrants further investigation: more data are needed in assessing the prevalence of *Giardia* infection in people diagnosed with IBS in the UK and detailed studies on the pathophysiology of giardiasis may help in determining whether this parasite can act as a trigger for the development of IBS.

We desperately need a better understanding of the genetics of *Giardia* in order to exploit this information for the development of more powerful molecular markers for studying its molecular epidemiology. The application of new methods is particularly crucial for the accurate genotyping of assemblage B parasites and to overcome the problem of the common occurrence of mixed templates in the analysis of DNA sequences. This goal could be achieved by using genes that are less affected by allele sequence heterozygosity or genes that are unique to particular sub-assemblages: these markers could then be incorporated in multiplex real-time PCR

assays allowing the sensitive and univocal identification of assemblage B sub-assemblages in faecal DNA samples. The comparative analysis of *Giardia* genomes is crucial for the identification of such markers (Ryan & Cacciò, 2013). More assemblage B isolates need to be sequenced to allow for a comprehensive comparative genomics analyses. In order to achieve this, we should focus on the optimization and application of whole-genome sequencing (WGS) tecniques of parasites isolated directly from faeces. This approach would allow the study of the actual genetic variability of *Giardia* in large numbers of specimens (representative of different hosts and geographic areas), overcoming at the same time the need for culturing the parasites and so the selection of particular genotypes. Although this approach is still in its infancy in parasitology, it has already been used extensively on bacterial and viral pathogens showing great potential in the detection and tracking of outbreaks in a public health microbiology context (Köser *et al.*, 2012).

APPENDIX 1: Surveillance study case-case comparisons

Table 2.4.1: Comparisons ($p \geq 0.05$) of the clinical and exposure characteristics between the *Giardia* cases that did not travel abroad (n=162) and those that travelled abroad (n=76) in the month prior to illness and that returned the surveillance questionnaire.

Variable	No. valid (% missing)	Category	No travel abroad n (%)*	Travel abroad n (%)*	*p*-value
Median length of illness (days)	165 (30.7)	-	16.5 (1-180)	21 (4-362)	0.330[M]
Median days of normal activity prevented	118 (50.4)	-	7 (0-42)	5 (0-90)	0.421[M]
Feeling still ill when filling the questionnaire	232 (2.5)	No	114 (72.6)	53 (70.7)	0.758
		Yes	43 (27.4)	22 (29.3)	
Median number of symptoms reported	236 (0.8)	-	3 (0-5	3 (0-5)	0.680[M]
Reported diarrhoea	236 (0.8)	No	10 (6.3)	4 (5.3)	1.000
		Yes	150 (93.8)	72 (94.7)	
Reported blood in stools	236 (0.8)	No	142 (88.8)	69 (90.8)	0.634
		Yes	18 (11.3)	7 (9.2)	
Reported vomiting	236 (0.8)	No	82 (51.3)	48 (63.2)	0.086
		Yes	78 (48.8)	28 (36.8)	
Reported abdominal pain	236 (0.8)	No	41 (25.6)	14 (18.4)	0.221
		Yes	119 (74.4)	62 (81.6)	
Reported fever	236 (0.8)	No	90 (56.3)	44 (57.9)	0.812
		Yes	70 (43.8)	32 (42.1)	
Median number of glasses of un-boiled tap water drank per day	206 (13.5)	-	3 (0-20)	3 (0-10)	0.173[M]
Drinking from a private water source	231 (2.9)	No	157 (99.4)	72 (98.6)	0.533
		Yes	1 (0.6)	1 (1.4)	
Drinking from a river, spring or pond	231 (2.9)	No	155 (98.1)	73 (100)	0.553
		Yes	3 (1.9)	0 (0)	
Drinking from a water filter	231 (2.9)	No	133 (84.2)	57 (78.1)	0.260
		Yes	25 (15.8)	16 (21.9)	
Walking in the countryside	238	No	108 (66.7)	49 (64.5)	0.739
		Yes	54 (33.3)	27 (35.5)	

Going picnicking	238	No	146 (90.1)	67 (88.2)	0.645
		Yes	16 (9.9)	9 (11.8)	
Practising fieldsports	238	No	150 (92.6)	74 (97.4)	0.236
		Yes	12 (7.4)	2 (2.6)	
Going camping	238	No	154 (95.1)	72 (94.7)	1.000
		Yes	8 (4.9)	4 (5.3)	
Going caravanning	238	No	155 (95.7)	69 (90.8)	0.148
		Yes	7 (4.3)	7 (9.2)	
Keeping a pet	224 (5.9)	No	72 (47.4)	44 (61.1)	0.055
		Yes	80 (52.6)	28 (38.9)	
Keeping dogs	224 (5.9)	No	104 (68.4)	56 (77.8)	0.148
		Yes	48 (31.6)	16 (22.2)	
Keeping cats	224 (5.9)	No	117 (77)	63 (87.5)	0.064
		Yes	35 (23)	9 (12.5)	
Keeping rabbits	224 (5.9)	No	146 (96.1)	66 (91.7)	0.207
		Yes	6 (3.9)	6 (8.3)	
Keeping birds	224 (5.9)	No	141 (92.8)	67 (93.1)	0.937
		Yes	11 (7.2)	5 (6.9)	
Keeping horses	224 (5.9)	No	151 (99.3)	71 (98.6)	0.541
		Yes	1 (0.7)	1 (1.4)	
Visiting a premise with animals	211 (11.3)	No	103 (73.6)	49 (69)	0.518
		Yes	37 (26.4)	22 (31)	
Touching cattle	205 (13.9)	No	131 (96.3)	65 (94.2)	0.489
		Yes	5 (3.7)	4 (5.8)	
Touching pigs	205 (13.9)	No	134 (98.5)	68 (98.6)	1.000
		Yes	2 (1.5)	1 (1.4)	
Touching sheep	205 (13.9)	No	129 (94.9)	66 (95.7)	1.000
		Yes	7 (5.1)	3 (4.3)	
Touching horses	205 (13.9)	No	126 (92.6)	68 (98.6)	0.103
		Yes	10 (7.4)	1 (1.4)	
Touching goats	205 (13.9)	No	132 (97.1)	65 (94.2)	0.446
		Yes	4 (2.9)	4 (5.8)	

Touching chickens	205 (13.9)	No	131 (96.3)	65 (94.2)	0.489
		Yes	5 (3.7)	4 (5.8)	
Touching any other animal	202 (15.1)	No	129 (96.3)	61 (89.7)	0.110
		Yes	5 (3.7)	7 (10.3)	
Eating any fresh fruit	221 (7.1)	No	14 (9.2)	3 (4.3)	0.209
		Yes	138 (90.8)	66 (95.7)	
Eating any salad item	223 (6.3)	No	25 (16.2)	6 (8.7)	0.133
		Yes	129 (83.8)	63 (91.3)	
Eating any uncooked vegetable	213 (10.5)	No	81 (55.5)	34 (50.7)	0.520
		Yes	65 (44.5)	33 (49.3)	
Drinking any fruit or vegetable juice	216 (9.2)	No	40 (27)	14 (20.6)	0.310
		Yes	108 (73)	54 (79.4)	
Frequency of eating fresh fruit	221 (7.1)	Not at all	14 (9.2)	3 (4.3)	0.559
		Less than once a week	13 (8.6)	5 (7.2)	
		Once or twice a week	32 (21.1)	12 (17.4)	
		Between three and six times a week	29 (19.1)	15 (21.7)	
		Once a day	30 (19.7)	20 (29)	
		More than once a day	34 (22.4)	14 (20.3)	
Frequency of eating salad items	223 (6.3)	Not at all	25 (16.2)	6 (8.7)	0.167
		Less than once a week	17 (11)	8 (11.6)	
		Once or twice a week	51 (33.1)	18 (26.1)	
		Between three and six times a week	39 (25.3)	23 (33.3)	
		Once a day	17 (11)	7 (10.1)	
		More than once a day	5 (3.2)	7 (10.1)	
Frequency of eating uncooked vegetables	213 (10.5)	Not at all	81 (55.5)	34 (50.7)	0.332
		Less than once a week	24 (16.4)	11 (16.4)	
		Once or twice a week	22 (15.1)	18 (26.9)	
		Between three and six times a week	11 (7.5)	2 (3)	
		Once a day	7 (4.8)	2 (3)	
		More than once a day	1 (0.7)	0 (0)	

Frequency of drinking fruit or vegetable juice	216 (9.2)	Not at all	40 (27)	14 (20.6)	0.509
		Less than once a week	14 (9.5)	6 (8.8)	
		Once or twice a week	37 (25)	19 (27.9)	
		Between three and six times a week	24 (16.2)	11 (16.2)	
		Once a day	22 (14.9)	16 (23.5)	
		More than once a day	11 (7.4)	2 (2.9)	
Eating fresh fruit from a greengrocer	221 (7.1)	No	142 (93.4)	64 (92.8)	1.000
		Yes	10 (6.6)	5 (7.2)	
Eating fresh fruit from a supermarket	221 (7.1)	No	27 (17.8)	18 (26.1)	0.154
		Yes	125 (82.2)	51 (73.9)	
Eating homegrown fresh fruit	221 (7.1)	No	150 (98.7)	67 (97.1)	0.591
		Yes	2 (1.3)	2 (2.9)	
Eating salad items from a greengrocer	223 (6.3)	No	144 (93.5)	65 (94.2)	1.000
		Yes	10 (6.5)	4 (5.8)	
Eating salad items from a supermarket	223 (6.3)	No	39 (25.3)	19 (27.5)	0.728
		Yes	115 (74.7)	50 (72.5)	
Eating homegrown salad items	223 (6.3)	No	148 (96.1)	67 (97.1)	1.000
		Yes	6 (3.9)	2 (2.9)	
Eating uncooked vegetables from a market	214 (10.1)	No	146 (99.3)	65 (97)	0.232
		Yes	1 (0.7)	2 (3)	
Eating uncooked vegetables from a greengrocer	214 (10.1)	No	140 (95.2)	63 (94)	0.743
		Yes	7 (4.8)	4 (6)	
Eating uncooked vegetables from supermarket	214 (10.1)	No	87 (59.2)	40 (59.7)	0.943
		Yes	60 (40.8)	27 (40.3)	
Eating homegrown uncooked vegetables	214 (10.1)	No	145 (98.6)	67 (100)	1.000
		Yes	2 (1.4)	0 (0)	
Drinking juice from a greengrocer	216 (9.2)	No	145 (98)	66 (97.1)	0.651
		Yes	3 (2)	2 (2.9)	
Drinking juice from a supermarket	216 (9.2)	No	49 (33.1)	25 (36.8)	0.599
		Yes	99 (66.9)	43 (63.2)	
Drinking homemade juice	216 (9.2)	No	148 (100)	66 (97.1)	0.098

		Yes	0 (0)	2 (2.9)	
Any child in the household	227 (4.6)	No	86 (55.5)	34 (47.2)	0.246
		Yes	69 (44.5)	38 (52.8)	
Median number of children	227 (4.6)	-	0 (0-5)	1 (0-4)	0.413[M]
Any child in nappies	225 (5.5)	No	113 (74.3)	56 (76.7)	0.700
		Yes	39 (25.7)	17 (23.3)	
Changing nappies	218 (8.4)	No	121 (82.3)	58 (81.7)	0.910
		Yes	26 (17.7)	13 (18.3)	
Any other person ill in the household with similar symptoms	222 (6.7)	No	136 (88.9)	56 (81.2)	0.119
		Yes	17 (11.1)	13 (18.8)	
Any contact outside the household with someone ill with similar symptoms	208 (12.6)	No	126 (88.1)	58 (89.2)	0.815
		Yes	17 (11.9)	7 (10.8)	

*percentages refer to the proportion among cases that answered the question; [M] Mann-Whitney U test

Table 2.5.1: Comparisons ($p \geq 0.05$) of the exposure history reported in the month prior to illness between male (n=138) and female (n=108) *Giardia* cases than returned the surveillance questionnaire.

Variable	**No. valid (% missing)**	**Category**	**Males n (%)***	**Females n (%)***	***p*-value**
Median number of glasses of un-boiled tap water drank per day	212 (13.8)	-	3 (0-12)	3 (0-20)	0.589 [M]
Drinking un-boiled tap water from a mains water supply	237 (3.6)	No	19 (14.4)	16 (15.2)	0.856
		Yes	113 (85.6)	89 (84.8)	
Drinking from a private water source	237 (3.6)	No	132 (100)	102 (97.1)	0.086
		Yes	0 (0)	3 (2.9)	
Drinking bottled water	237 (3.6)	No	56 (42.4)	37 (35.2)	0.260
		Yes	76 (57.6)	68 (64.8)	
Drinking from a river/spring/pond	237 (3.6)	No	131 (99.2)	103 (98.1)	0.586
		Yes	1 (0.8)	2 (1.9)	
Drinking from a water filter	237 (3.6)	No	108 (81.8)	87 (82.9)	0.835
		Yes	24 (18.2)	18 (17.1)	

Eating any fresh fruit	225 (8.5)	No	11 (8.8)	7 (7)	0.621
		Yes	114 (91.2)	93 (93)	
Eating any salad item	227 (7.7)	No	21 (16.8)	10 (9.8)	0.127
		Yes	104 (83.2)	92 (90.2)	
Eating any uncooked vegetable	218 (11.4)	No	66 (54.5)	51 (52.6)	0.772
		Yes	55 (45.5)	46 (47.4)	
Drinking any fruit or vegetable juice	221 (10.2)	No	28 (22.8)	27 (27.6)	0.414
		Yes	95 (77.2)	71 (72.4)	
Eating fresh fruit from a market	225 (8.5)	No	112 (89.6)	95 (95)	0.138
		Yes	13 (10.4)	5 (5)	
Eating fresh fruit from a greengrocer	225 (8.5)	No	116 (92.8)	94 (94)	0.720
		Yes	9 (7.2)	6 (6)	
Eating fresh fruit from supermarket	225 (8.5)	No	27 (21.6)	20 (20)	0.769
		Yes	98 (78.4)	80 (80)	
Eating homegrown fresh fruit	225 (8.5)	No	122 (97.6)	99 (99)	0.631
		Yes	3 (2.4)	1 (1)	
Eating fresh fruit from a place other than market supermarket, greengrocers or homegrown	225 (8.5)	No	115 (92)	93 (93)	0.778
		Yes	10 (8)	7 (7)	
Eating salad items from a greengrocer	228 (7.3)	No	118 (93.7)	95 (93.1)	1.000
		Yes	8 (6.3)	7 (6.9)	
Eating salad items from supermarket	228 (7.3)	No	40 (31.7)	21 (20.6)	0.058
		Yes	86 (68.3)	81 (79.4)	
Eating homegrown salad items	228 (7.3)	No	120 (95.2)	100 (98)	0.302
		Yes	6 (4.8)	2 (2)	
Eating salad items from a place other than market supermarket, greengrocers or homegrown	228 (7.3)	No	117 (92.9)	93 (91.2)	0.640
		Yes	9 (7.1)	9 (8.8)	
Eating uncooked vegetables from a market	219 (11)	No	119 (98.3)	97 (99)	1.000
		Yes	2 (1.7)	1 (1)	
Eating uncooked vegetables from a greengrocer	219 (11)	No	115 (95)	93 (94.9)	1.000
		Yes	6 (5)	5 (5.1)	
Eating uncooked vegetables from supermarket	219 (11)	No	75 (62)	56 (57.1)	0.468
		Yes	46 (38)	42 (42.9)	

Eating homegrown uncooked vegetables	219 (11)	No	119 (98.3)	98 (100)	0.503
		Yes	2 (1.7)	0 (0)	
Eating uncooked vegetables from a place other than market supermarket, greengrocers or homegrown	219 (11)	No	116 (95.9)	95 (96.9)	0.734
		Yes	5 (4.1)	3 (3.1)	
Drinking juice from a market	221 (10.2)	No	118 (95.9)	97 (99)	0.230
		Yes	5 (4.1)	1 (1)	
Drinking juice from a greengrocer	221 (10.2)	No	118 (95.9)	97 (99)	0.230
		Yes	5 (4.1)	1 (1)	
Drinking juice from supermarket	221 (10.2)	No	38 (30.9)	37 (37.8)	0.285
		Yes	85 (69.1)	61 (62.2)	
Drinking homemade juice	221 (10.2)	No	121 (98.4)	98 (100)	0.504
		Yes	2 (1.6)	0 (0)	
Drinking juice from a place other than market supermarket, greengrocers or homegrown	221 (10.2)	No	115 (93.5)	92 (93.9)	0.908
		Yes	8 (6.5)	6 (6.1)	
Frequency of eating fresh fruit	225 (8.5)	Not at all	11 (8.8)	7 (7)	0.573
		Less than once a week	11 (8.8)	7 (7)	
		Once or twice a week	21 (16.8)	24 (24)	
		Between three and six times a week	25 (20)	19 (19)	
		Once a day	33 (26.4)	19 (19)	
		More than once a day	24 (19.2)	24 (24)	
Frequency of eating salad items	227 (7.7)	Not at all	21 (16.8)	10 (9.8)	0.328
		Less than once a week	12 (9.6)	14 (13.7)	
		Once or twice a week	34 (27.2)	37 (36.3)	
		Between three and six times a week	34 (27.2)	28 (27.5)	
		Once a day	16 (12.8)	9 (8.8)	
		More than once a day	8 (6.4)	4 (3.9)	
Frequency of eating uncooked vegetables	218 (11.4)	Not at all	66 (54.5)	51 (52.6)	0.463
		Less than once a week	16 (13.2)	21 (21.6)	
		Once or twice a week	25 (20.7)	16 (16.5)	
		Between three and six times a week	8 (6.6)	5 (5.2)	
		Once a day	6 (5)	3 (3.1)	
		More than once a day	0 (0)	1 (1)	

Frequency of drinking fruit or vegetable juice	221 (10.2)	Not at all	28 (22.8)	27 (27.6)	0.417
		Less than once a week	8 (6.5)	12 (12.2)	
		Once or twice a week	35 (28.5)	23 (23.5)	
		Between three and six times a week	18 (14.6)	17 (17.3)	
		Once a day	25 (20.3)	15 (15.3)	
		More than once a day	9 (7.3)	4 (4.1)	
Going swimming at a swimming pool	246	No	94 (68.1)	65 (60.2)	0.197
		Yes	44 (31.9)	43 (39.8)	
Swimming in a lake, pond or river	246	No	129 (93.5)	99 (91.7)	0.588
		Yes	9 (6.5)	9 (8.3)	
Practising watersports in freshwater	246	No	136 (98.6)	104 (96.3)	0.409
		Yes	2 (1.4)	4 (3.7)	
Practising watersports at sea	246	No	127 (92)	98 (90.7)	0.720
		Yes	11 (8)	10 (9.3)	
Went fishing	246	No	132 (95.7)	106 (98.1)	0.472
		Yes	6 (4.3)	2 (1.9)	
Walking in the countryside	246	No	95 (68.8)	70 (64.8)	0.505
		Yes	43 (31.2)	38 (35.2)	
Going picnicking	246	No	123 (89.1)	98 (90.7)	0.678
		Yes	15 (10.9)	10 (9.3)	
Practising fieldsports	246	No	127 (92)	105 (97.2)	0.081
		Yes	11 (8)	3 (2.8)	
Going camping	246	No	132 (95.7)	102 (94.4)	0.663
		Yes	6 (4.3)	6 (5.6)	
Going caravanning	246	No	129 (93.5)	103 (95.4)	0.525
		Yes	9 (6.5)	5 (4.6)	
Keeping a pet	229 (6.9)	No	65 (51.6)	54 (52.4)	0.899
		Yes	61 (48.4)	49 (47.6)	
Keeping dogs	229 (6.9)	No	85 (67.5)	78 (75.7)	0.169
		Yes	41 (32.5)	25 (24.3)	
Keeping rabbits	229 (6.9)	No	118 (93.7)	99 (96.1)	0.405
		Yes	8 (6.3)	4 (3.9)	

Keeping birds	229 (6.9)	No	115 (91.3)	97 (94.2)	0.404
		Yes	11 (8.7)	6 (5.8)	
Keeping horses	229 (6.9)	No	124 (98.4)	103 (100)	0.503
		Yes	2 (1.6)	0 (0)	
Visiting a premise with animals	215 (12.6)	No	87 (73.1)	67 (69.8)	0.592
		Yes	32 (26.9)	29 (30.2)	
Touching cattle	209 (15)	No	109 (93.2)	90 (97.8)	0.191
		Yes	8 (6.8)	2 (2.2)	
Touching pigs	209 (15)	No	114 (97.4)	91 (98.9)	0.632
		Yes	3 (2.6)	1 (1.1)	
Touching sheep	209 (15)	No	110 (94)	88 (95.7)	0.759
		Yes	7 (6)	4 (4.3)	
Touching horses	209 (15)	No	109 (93.2)	88 (95.7)	0.442
		Yes	8 (6.8)	4 (4.3)	
Touching goats	209 (15)	No	111 (94.9)	89 (96.7)	0.734
		Yes	6 (5.1)	3 (3.3)	
Touching chickens	209 (15)	No	111 (94.9)	89 (96.7)	0.734
		Yes	6 (5.1)	3 (3.3)	
Touching any other animal	206 (16.3)	No	110 (94.8)	82 (91.1)	0.293
		Yes	6 (5.2)	8 (8.9)	
Any child in the household	235 (4.5)	No	71 (53)	52 (51.5)	0.820
		Yes	63 (47)	49 (48.5)	
Median number of children	235 (4.5)	-	0 (0-5)	0 (0-4)	0.902[M]
Any child in nappies	233 (5.3)	No	103 (78)	71 (70.3)	0.179
		Yes	29 (22)	30 (29.7)	
Any other person ill in the household with similar symptoms	230 (6.5)	No	111 (86)	89 (88.1)	0.643
		Yes	18 (14)	12 (11.9)	
Any contact outside the household with someone ill with similar symptoms	216 (12.2)	No	109 (87.9)	83 (90.2)	0.593
		Yes	15 (12.1)	9 (9.8)	

*percentages refer to the proportion among cases that answered the question; [M] Mann-Whitney U test

APPENDIX 2: Case-control study forms

Figure A: Case-control study invitation letter.

INVITATION LETTER:
Adult Case

Health Protection Unit Headed Paper

«title» «First_name» «Surname»
«ADDRESS1»
«ADDRESS2»
«TOWN»
«POSTCODE» Day month year

NORTH WEST STUDY OF *GIARDIA* INFECTION

Dear «title» «First name» «Surname»

We would like to invite you to take part in our research study on the spread of a common tummy bug called *Giardia* in North West England. The study involves completion of the questionnaire enclosed with this letter. We have enclosed an information sheet explaining about *Giardia* and why our research is important to your community. Please read the information sheet carefully before deciding whether to take part in the study.

You have been asked to take part in this study because you have been found to have *Giardia* infection following your recent visit to your doctor. The information we get from this study will help prevent *Giardia* infection in the community, so we would be very grateful if you could help us by agreeing to take part in the study.

The study has been fully approved according to the ethical requirements of the NHS.

Participation in the study is entirely voluntary and is one-off. If you agree to take part please complete the ***Consent Form*** and the ***Questionnaire*** enclosed and return them in the pre-paid envelope. It should take just 10 minutes to complete the questionnaire. Most questions can be answered by putting ticks or numbers in boxes, or by writing in details where appropriate. Please answer all the questions. If you would like to know the results of the study when they become available please tick the box on the Consent Form.

If you have any concerns or additional questions about this study please phone the principal investigator Dr Kenneth Lamden (tel. XXXXX XXXXXX). He will do his best to answer your questions. If you are unhappy for any reason and wish to complain formally you can do this via the NHS complaints procedure.

Please try to return the forms and the questionnaire **within 2 weeks of the date on this letter.**

Yours sincerely

Dr Kenneth Lamden FFPH
Consultant in Communicable Disease Control

Figure B: Case-control study participant information leaflet.

Health Protection Agency

NORTH WEST STUDY OF *GIARDIA* INFECTION

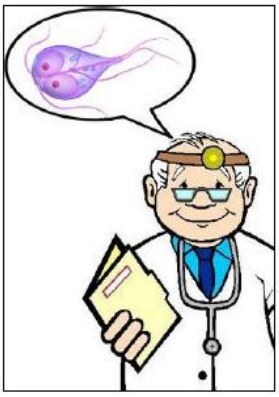
CASE

INFORMATION SHEET

We would like to invite you to take part in a research study aiming to find out **how the tummy bug *Giardia* is spread in North West England**. Before you decide if you want to join in, it's important to understand why the study is being done and what it will involve for you. So please read this leaflet carefully. We'd suggest this should take about 5 minutes. Talk to your family or friends if you want to.

What is Giardia and how can you get it?

- *Giardia* is a very common bug living in the intestine of both people and animals.
- If you have *Giardia* you usually suffer diarrhoea, wind and tummy pain.
- You mainly get *Giardia* from drinking unclean water or eating raw food. You can also get it from touching animals without washing your hands.
- A lot of people are affected by it every year worldwide and it can be quite common in children and young people.
- You can easily treat this illness with an antibiotic.



Why are we doing this study?

- About 3000 people per year suffer from *Giardia* infection in England and Wales. However, we don't know a lot about *Giardia* in this country.
- We want to understand more about how this bug is spread in North West England, so we can protect people better in the future.
- This study has been organised and funded by the UK Health Protection Agency, along with the University of Liverpool and Royal Preston and Central Manchester University Hospitals.

- This study has been reviewed by a Research Ethics Committee (REC) and will be run in complete accordance with NHS ethical procedure to protect the interests of all the participants.

Why have I been asked to take part in the study?

You have been asked to take part in this study because you have been found to have *Giardia* infection following your recent visit to your doctor. To understand the causes of *Giardia* we need to collect information from people like yourself who have had the bug and compare it to people who haven't had *Giardia*. We are also conducting laboratory tests on the poo samples given by people with *Giardia* to find the particular type of *Giardia* that caused the infection.

What will I have to do if I decide to take part?

If you decide to take part (your son/daughter to take part) you will just need to fill in the enclosed questionnaire and the consent form, and return them to us in the pre-paid envelope provided. You can ask someone else to help you if you need to. The questionnaire asks about activities and behaviours that are related to *Giardia*. This information is essential to find out the possible causes of *Giardia*.

Do I have to take part?

No, it is completely up to you to decide if you want to take part.

Is there any risk to be worried about if I take part?

No, there is no disadvantage or risk associated with taking part in this study. The information from the study will only be used for research purposes and won't have any consequence for your health.

What are the potential benefits of taking part?

You won't receive any direct benefit from taking part in this study, but the information we get might help to reduce the number of people who get the disease in the future.

Will anyone else know I'm doing this?

We will follow ethical and legal practice to keep your information private, and only people doing the research will see it.

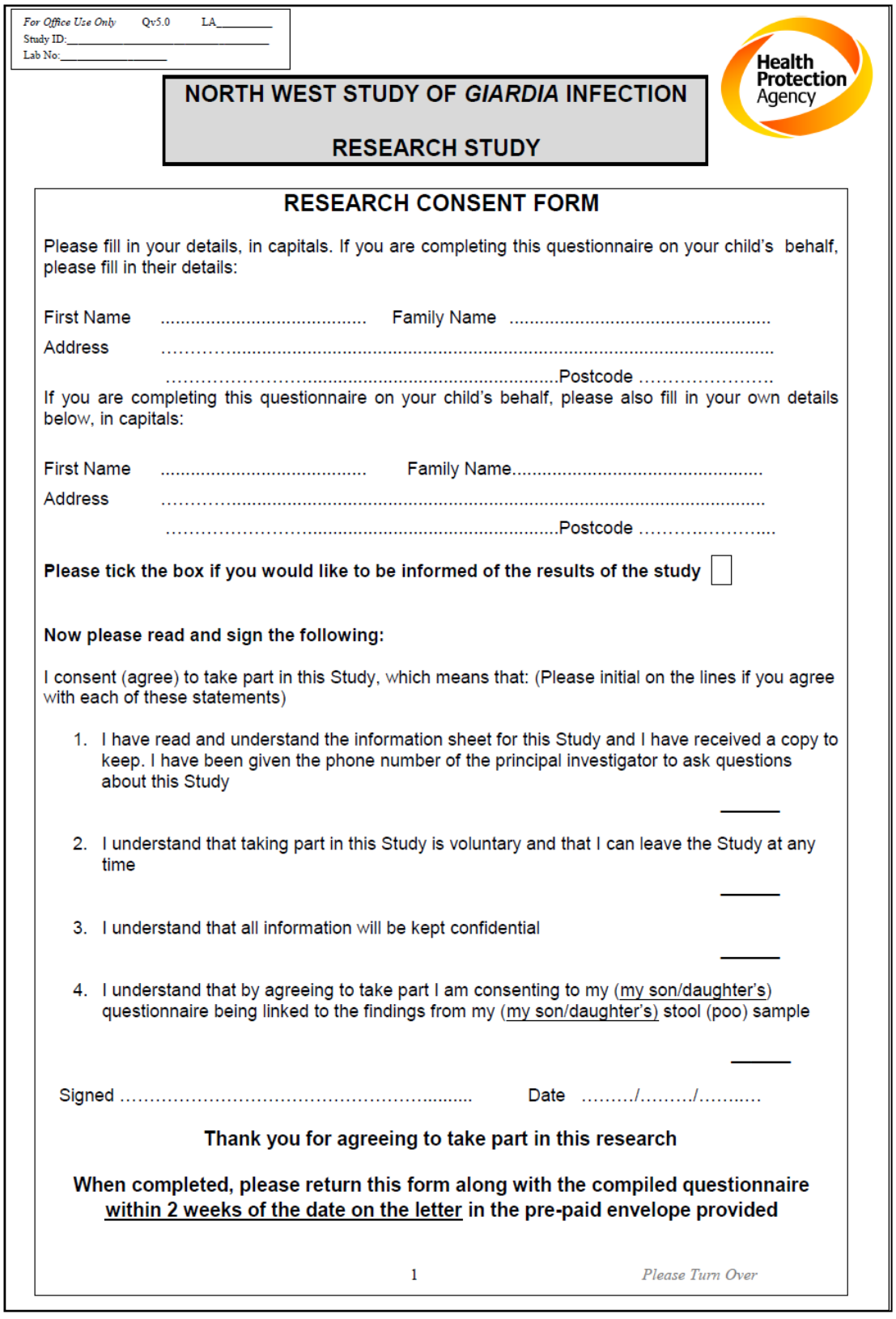
What will happen to the results of the study?

Results will be published in scientific journals and as part of a PhD thesis. You won't be identified in any publication. At the end of the study you can receive a summary of the results by declaring your interest at the end of the consent form

What if I have additional questions about the study or there is a problem?

If you have any concern or additional question about this study, please phone the principal investigator Dr Kenneth Lamden (tel. 01257 246450). He will do his best to answer your questions. If you are unhappy for any reason and wish to complain formally, you can do this via the NHS complaints procedure.

Figure C: Case-control study consent form and case questionnaire.



For Office Use Only Qv5.0 LA_____
Study ID:_____
Lab No:_____

Health Protection Agency

NORTH WEST STUDY OF *GIARDIA* INFECTION

RESEARCH STUDY

RESEARCH CONSENT FORM

Please fill in your details, in capitals. If you are completing this questionnaire on your child's behalf, please fill in their details:

First Name ..... Family Name .....

Address .....

.....Postcode .....

If you are completing this questionnaire on your child's behalf, please also fill in your own details below, in capitals:

First Name ..... Family Name.....

Address .....

.....Postcode .....

Please tick the box if you would like to be informed of the results of the study ☐

Now please read and sign the following:

I consent (agree) to take part in this Study, which means that: (Please initial on the lines if you agree with each of these statements)

1. I have read and understand the information sheet for this Study and I have received a copy to keep. I have been given the phone number of the principal investigator to ask questions about this Study

   _____

2. I understand that taking part in this Study is voluntary and that I can leave the Study at any time

   _____

3. I understand that all information will be kept confidential

   _____

4. I understand that by agreeing to take part I am consenting to my (my son/daughter's) questionnaire being linked to the findings from my (my son/daughter's) stool (poo) sample

   _____

Signed ..... Date ...../...../.....

Thank you for agreeing to take part in this research

When completed, please return this form along with the compiled questionnaire within 2 weeks of the date on the letter in the pre-paid envelope provided

1 *Please Turn Over*

For Office Use Only Qv5.0 LA_____
Study ID:_____

Lab No:_____

STRICTLY CONFIDENTIAL
NORTH WEST STUDY OF *GIARDIA* INFECTION

Case Questionnaire

WHAT DO I HAVE TO DO?

If you have decided to take part in this study please complete this questionnaire on behalf of yourself, or your child, as requested in the invitation letter. It enquires about activities and behaviours that may be linked to *Giardia* infection. This kind of information is essential for us to understand how *Giardia* is transmitted.

IT IS VERY IMPORTANT TO ANSWER ALL THE QUESTIONS EVEN IF THE ANSWER IS NO

Please go through all the sections and tick the boxes and/or fill in the details as instructed. Please fill in the details in block capitals.

Today's date: ____/_____/____

SECTION 1: PERSONAL DETAILS *(of person who was ill)*

1.1 Age: _____ 1.2 Sex: Male ☐ Female ☐

1.3 Postcode: _____

1.4 **What is your ethnicity?**

White ☐ Asian ☐ Chinese ☐

Black African ☐ Black Caribbean ☐

Other ☐ **If other**, please specify _____

1.5 **What is your occupational status?** Currently in work ☐ Student ☐
Unemployed ☐ Retired ☐ Child below working age ☐

1.6 **What is your occupation** (if currently unemployed, what was your most recent occupation; if retired, what was your main occupation)?

Job title _____

What do/did you actually do? _____

_____

1.7 How old were you when you first left full-time education (school, college, university)? _____years

Still in full time education ☐ Pre-school child ☐

SECTION 2: HEALTH DETAILS

2.1 On what date did you start feeling ill? (if unsure, please use the best estimate) ____/____/_____ (dd/mm/yyyy)

2.2 While you were ill with *Giardia*, did you experience any of the following symptoms? Please tick **Yes** or **No** for each symptom listed. For each symptom that you experienced, please indicate how many days you suffered from it and whether the symptom is still present

	Yes	No	**If yes**, no. of days	Still present
Diarrhoea (3 or more loose or watery stools in 24 hour period)	☐	☐	_____	☐
If you had diarrhoea, how many times a day did you visit the toilet on average?	_____ (number)			
Was the diarrhoea persistent ☐ *or intermittent*		☐		

Did you have other symptoms?

	Yes	No	**If yes**, no. of days	Still present
Flatulence (wind)	☐	☐	_____	☐
Blood in stools	☐	☐	_____	☐
Swollen stomach	☐	☐	_____	☐
Excessive tiredness	☐	☐	_____	☐
Loss of appetite	☐	☐	_____	☐
Vomiting	☐	☐	_____	☐
Abdominal pain (cramps)	☐	☐	_____	☐
Fever (feeling hot or having a temperature)	☐	☐	_____	☐
Weight loss	☐	☐	_____	☐

If you lost weight, about how much did you lose? _____kg or _____ pounds

2.3 Are you still ill? Yes☐ No ☐

If you are no longer ill, how many days were you ill for? _____ days

2.4 Did you visit your doctor as a result of your illness? Yes☐ No ☐

2.5 Were you admitted to hospital as a result of your illness? Yes☐ No ☐

If yes, how many days did you spend in hospital? _____ days

2.6 Did you take any medicine for indigestion in the 3 weeks before you first started feeling ill? Yes ☐ No ☐

If yes, please specify which medicine(s)	Yes	No
-Antacids (e.g. Rennie, Milk of Magnesia)	☐	☐
-Ranitidine (Zantac)	☐	☐
-Omeprazole (losec) or lansoprazole (zoton)	☐	☐

-Other, Yes ☐ No ☐

If yes, name of medicine_____

2.7 Do you suffer from any of the following conditions?

	Yes	No
-Inflammatory Bowel Disease (IBD) e.g. Crohn's disease, ulcerative colitis	☐	☐
-Irritable Bowel Syndrome (IBS)	☐	☐

SECTION 3: OCCUPATIONAL DETAILS

3.1 **If you are currently in work (full or part-time)**, does your job involve any of the following:

	Yes	No
- working with animals	☐	☐
- working with human faeces (e.g. nappy changing)	☐	☐
- working with animal manure/slurry	☐	☐
- direct contact with sewage	☐	☐
- direct contact with freshwaters (river, lake, pond)	☐	☐

SECTION 4: ANIMAL CONTACT

A. Pets

4.1 Do you keep any animals? Yes ☐ No ☐

If no, please go to **Question 4.4**
If yes, do you keep any of the following animals? Please also indicate how many.

	Yes	No	Number		Yes	No	Number
Dog	☐	☐	_____	Cat	☐	☐	_____
Bird	☐	☐	_____	Horse	☐	☐	_____

If you keep any pets other than those already mentioned, please specify below which type(s) and how many: *(e.g. horse, hamster, reptiles etc)*

_____

4.2 **If you keep any pets** did you / your child clean up or touch any of these animals' faeces (poo) in the **3 weeks** before you first started feeling ill? Yes ☐ No ☐

4.3 **If you keep any pets** were any of these animals ill with diarrhoea in the 3 weeks before you first started feeling ill? Yes ☐ No ☐

Please Turn Over

4.4 Did you touch any pet animals (yours or other people's) in the **3 weeks?** before you first started feeling ill? Yes ☐ No ☐

If yes, please specify below which animals you touched (*e.g. dog, cat, rabbit, horse etc):*

_____

B. Farms

4.5 Did you visit or work on any farms in the 3 weeks before you first started feeling ill? Yes ☐ No ☐

If yes, did you touch any of these animals?

	Yes	No		Yes	No		Yes	No
Cattle	☐	☐	Pigs	☐	☐	Sheep	☐	☐
Goats	☐	☐	Horses/ponies	☐	☐	Chickens/poultry	☐	☐

If you touched animals other than those above, please specify below which types of animals:

_____

C. Wildlife Parks / Zoos

4.6 Did you work on or visit any zoos or wildlife parks in the 3 weeks before you first started feeling ill? Yes ☐ No ☐

If yes, did you touch any animals? Yes ☐ No ☐

***If yes**, please specify below which types of animals you touched*: Yes ☐ No ☐

_____

D. Other Animal Contact

4.7 Did you visit or work on any other animal premises (apart from those listed above) in the 3 weeks before you first started feeling ill? Yes ☐ No ☐

If yes, *please specify below which type of premise(s) you visited*: Yes ☐ No ☐

_____

If yes, did you touch any animals during a visit? Yes ☐ No ☐

If yes, *please specify below which types of animals you touched*: Yes ☐ No ☐

_____

4.8 Did you touch any animals in the wild in the 3 weeks before you first started? feeling ill? Yes ☐ No ☐

If yes, please specify below which animals you touched (*e.g. deer, foxes, rabbits, squirrels, etc*):

_____

SECTION 5: TRAVEL DETAILS

5.1 Did you travel **abroad (outside the UK)** in the 3 weeks before you started feeling ill? Yes ☐ No ☐

If yes, where *(please specify below the countries in the order you visited them)*:

_____

5.2 Did you travel **in the UK** in the 3 weeks before you started feeling ill? Yes ☐ No ☐

If yes, where (please specify below the places in the order you visited them):

_____

SECTION 6: RECREATIONAL ACTIVITIES

6.1 Did you take part in any of the following activities in the **3 weeks** before you first started feeling ill? For each type, please also indicate how many times.

Water activities	Yes	No	**If yes**, *no. of times*	**If yes**, *is it likely you put your head under the water (please write yes/no)?*
Swimming/paddling in a swimming pool	☐	☐	_____	_____
If yes, *name and location of swimming pool*	_____			
Jacuzzi or hot tub	☐	☐	_____	_____
If yes, *name and location of Jacuzzi or hot tub*	_____			
Swimming/paddling in a lake/pond or river	☐	☐	_____	_____
Swimming/paddling in the sea	☐	☐	_____	_____
Water sports in fresh water *(e.g. sailing, windsurfing or canoeing)*	☐	☐	_____	_____
Water sports in the sea *(e.g. sailing, windsurfing or canoeing)*	☐	☐	_____	_____

Other outdoor activities	Yes	No	**If yes**, no. of times
Fishing	☐	☐	_____
Hunting	☐	☐	_____
Picnicking	☐	☐	_____
Camping	☐	☐	_____
Caravanning	☐	☐	_____
Field sports (e.g. football etc)	☐	☐	_____
Walking in the country	☐	☐	_____
Gardening	☐	☐	_____

6.2 If you took part in any other outdoor activity not mentioned above, please specify activity and number of times below:

_____

SECTION 7: WATER CONSUMPTION

7.1 **From which of the following sources do you receive your household water supply?**

-Water company (mains water) Yes ☐ No ☐
-Private water supply (e.g. spring, well) Yes ☐ No ☐
-Both Yes ☐ No ☐

7.2 **Did you drink any of the following types of water in the 3 weeks before you first became ill?**

-Tap water (*includes any water taken straight from the tap or mixed with cordial, squash, juice)* Yes ☐ No ☐

If yes, *on average,* how many glasses of water straight from the tap do you / your child normally drink each day? _____glasses

-Bottled water Yes ☐ No ☐
-Unboiled water straight from lake, river or stream Yes ☐ No ☐

7.3 **Did you notice any of the following changes in your water supply in the 3 weeks before you first became ill?**

-Water had unusual taste Yes ☐ No ☐
-Water had unusual smell Yes ☐ No ☐
-Water was discoloured Yes ☐ No ☐

SECTION 8: FOOD CONSUMPTION

8.1 Approximately how many times a **week** do you usually eat the following types of foods?
If none, please write **0**.

	No. of times a week
-Salads or raw vegetables *(including coleslaw etc.)*	_____
-Cooked vegetables	_____
-Raw fresh fruit	_____

8.2 Do you peel raw fruit before eating it? *(e.g apples, pears)* Yes ☐ No ☐

8.3 Do you wash raw fruit before eating it? Yes ☐ No ☐

8.4 In the 3 weeks before you became ill, is it likely that you ate any of the following types of food?

If yes, *please specify where you got the food from.*

	Yes	No	Market	Green-grocer	Super-market	Home grown	Other place *(specify)*
Salads and/or raw vegetables	☐	☐	☐	☐	☐	☐	_____
Cooked vegetables	☐	☐	☐	☐	☐	☐	_____
Fruit juice	☐	☐	☐	☐	☐		_____
Pre-packed sandwiches	☐	☐	☐	☐	☐		_____

	Yes	No	Market stall	Butcher/ fish monger	Super-market	
Beef	☐	☐	☐	☐	☐	_____
Lamb	☐	☐	☐	☐	☐	_____
Pork	☐	☐	☐	☐	☐	_____
Chicken	☐	☐	☐	☐	☐	_____
Fish (**not** shellfish)	☐	☐	☐	☐	☐	_____
Shellfish	☐	☐	☐	☐	☐	_____

8.5 In the 3 weeks before you became ill, did you eat in any of the following places?

	Yes	No	No of times
-Pub or restaurant	☐	☐	_____
-Canteen	☐	☐	_____
-Takeaway or fast food	☐	☐	_____
-Barbecue	☐	☐	_____

If you ate in a place other than those listed above, please specify place and number of times:

_____

Please Turn Over

SECTION 9: GENERAL HOUSEHOLD DETAILS

9.1 How many people, **including yourself**, live in your household?

Adults *(16 years and over)* _____ Children *(under 16 years)* _____

If there are no children in your household please go to Q 9.4

9.2 Do any children in your household attend a childcare nursery or a playgroup? Yes ☐ No ☐

9.3 Are any of the children in nappies in your household? Yes ☐ No ☐

If yes, were you involved in changing nappies in the 3 weeks before you? became ill? Yes ☐ No ☐

9.4 Was anyone in your household ill with diarrhoea in the 3 weeks before? you became ill? Yes ☐ No ☐

9.5 Did you have any contact with anyone else, other than those in your household, with diarrhoea in the 3 weeks before you became ill? Yes ☐ No ☐

If yes, please specify below where you had contact with them
(*e.g. school, day nursery, day care centre, occupational or social contact etc*)

_____

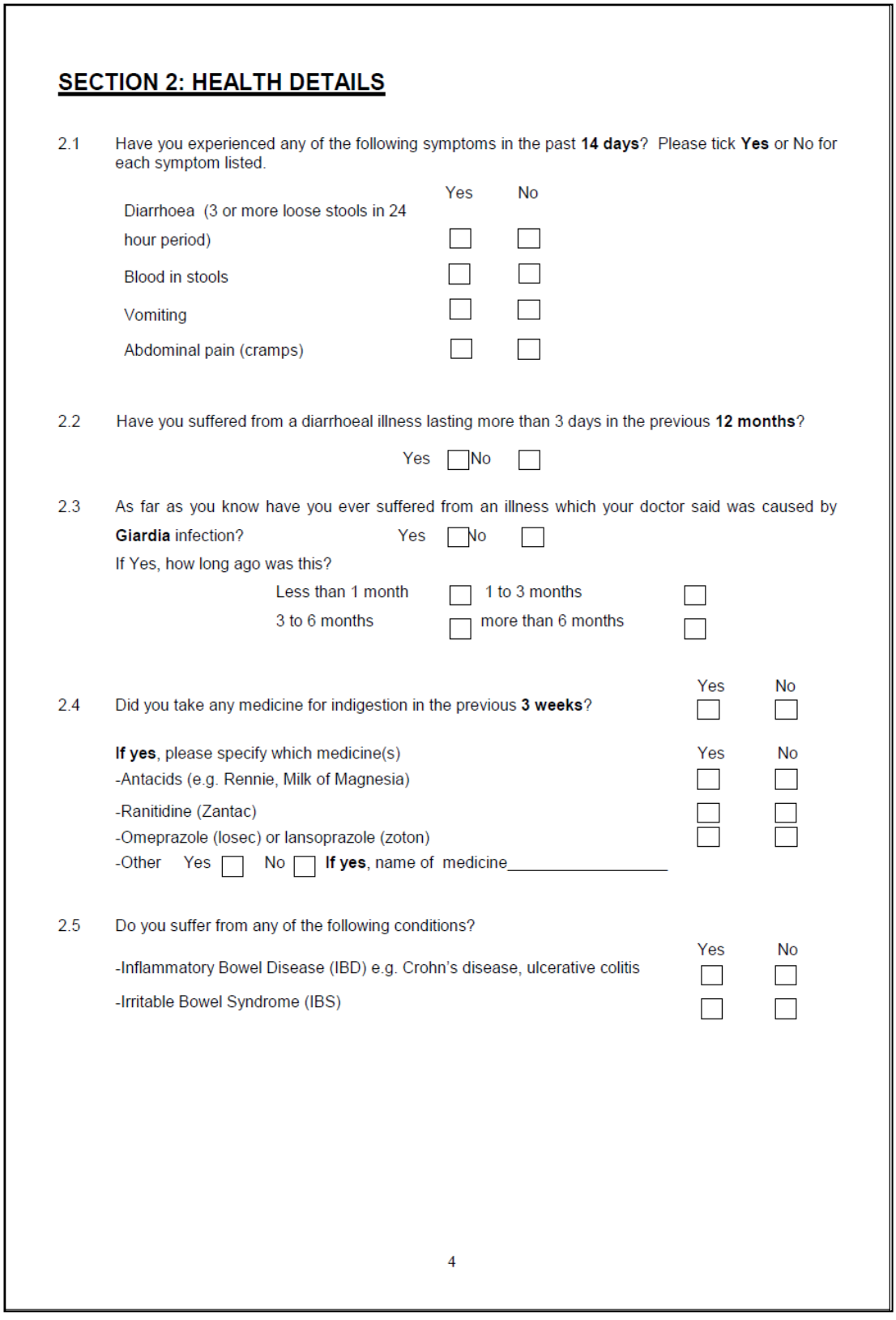
9.6 How would you describe the area in which you live?

City ☐ Town ☐ Village ☐

9.7 Finally, what do you think caused your illness?

THANK YOU FOR YOUR TIME AND COOPERATION IN COMPLETING THIS QUESTIONNAIRE. YOUR ANSWERS WILL HELP US GREATLY IN OUR EFFORTS TO PREVENT THE SPREAD OF *GIARDIA* IN THE COMMUNITY.
PLEASE RETURN THE QUESTIONNAIRE WITHIN 2 WEEKS OF THE DATE ON THE LETTER IN THE PRE-PAID ENVELOPE PROVIDED

Figure D: Case-control study control questionnaire health details section.



SECTION 2: HEALTH DETAILS

2.1 Have you experienced any of the following symptoms in the past **14 days**? Please tick **Yes** or No for each symptom listed.

	Yes	No
Diarrhoea (3 or more loose stools in 24 hour period)	☐	☐
Blood in stools	☐	☐
Vomiting	☐	☐
Abdominal pain (cramps)	☐	☐

2.2 Have you suffered from a diarrhoeal illness lasting more than 3 days in the previous **12 months**?

Yes ☐ No ☐

2.3 As far as you know have you ever suffered from an illness which your doctor said was caused by **Giardia** infection? Yes ☐ No ☐

If Yes, how long ago was this?

Less than 1 month ☐ 1 to 3 months ☐

3 to 6 months ☐ more than 6 months ☐

2.4 Did you take any medicine for indigestion in the previous **3 weeks**? Yes ☐ No ☐

If yes, please specify which medicine(s)	Yes	No
-Antacids (e.g. Rennie, Milk of Magnesia)	☐	☐
-Ranitidine (Zantac)	☐	☐
-Omeprazole (losec) or lansoprazole (zoton)	☐	☐

-Other Yes ☐ No ☐ **If yes**, name of medicine_____

2.5 Do you suffer from any of the following conditions?

	Yes	No
-Inflammatory Bowel Disease (IBD) e.g. Crohn's disease, ulcerative colitis	☐	☐
-Irritable Bowel Syndrome (IBS)	☐	☐

4

Figure E: Case-control study control reminder letter.

REMINDER LETTER

Health Protection Unit Headed Paper

«title» «First_name» «Surname»
«ADDRESS1»
«ADDRESS2»
«TOWN»
«POSTCODE» Day month year

NORTH WEST STUDY OF *GIARDIA* INFECTION

Dear «title» «First name» «Surname»

You may remember that approximately 2 weeks ago we sent you a letter inviting you take part in a study of *Giardia* transmission in the North West. We sent you a questionnaire requesting some details about your recent activities, food and water consumption and animal contact. As explained in the information sheet you received, it is very important that we collect this information so that we can find out the causes of this major illness and reduce its burden in the community.

We would, therefore, be very grateful if you would take the time to complete the attached questionnaire and the consent/decline form and return them to us by post in the pre-paid envelope provided. We would like to remind you that any information that you give us will remain strictly confidential. If you have returned the questionnaire in the last few days, please ignore this reminder.

Should you require any further information about the study, please do not hesitate to contact the study principal investigator, Dr Kenneth Lamden (tel. XXXXX XXXXXX). We would like to thank you in advance for your help with this study.

Yours sincerely

Dr Kenneth Lamden FFPH
Consultant in Communicable Disease Control

APPENDIX 3: Case-control study univariable analysis (variables with *p*>0.2)

Table 3.3.1: List of variables that returned a *p*-value>0.2 in univariable analysis. Results of both the general risk factor analysis (e.g. including all the eligible118 cases and 226 controls) and the indigenous risk factor analysis (e.g. including only the 86 cases and 207 controls that did not report travelling abroad in the exposure window) are shown for comparison.

Variable	Data subset	No. valid (% missing)*	Category	Cases n (%)**	Controls n (%)**	OR (95% CI)***	*p*-value
SOCIODEMOGRAPHICS, AREA AND SEASON VARIABLES							
Gender	All	344	Female	53 (44.9)	110 (48.7)	Ref.	0.508
			Male	65 (55.1)	116 (51.3)	1.16 (0.61-2.32)	
	Indigenous	293	Female	36 (41.9)	104 (50.2)	Ref.	0.191
			Male	50 (58.1)	103 (49.8)	1.40 (0.84-2.33)	
Age (years)	All	344	0-4y	16 (13.5)	26 (11.5)	Ref.	$0.230^{CT}/0.272^{C}$
			5-14y	6 (5.1)	12 (5.3)	0.81 (0.25-2.59)	
			15-44y	44 (37.3)	62 (27.4)	1.15 (0.55-2.40)	
			45-64y	27 (22.9)	75 (33.2)	0.58 (0.27-1.25)	
			65+y	25 (21.2)	51 (22.6)	0.80 (0.36-1.75)	
	Indigenous	293	0-4y	10 (11.6)	23 (11.1)	Ref.	$0.330^{CT}/0.422^{C}$
			5-14y	5 (5.8)	12 (5.8)	0.96 (0.27-3.45)	
			15-44y	34 (39.5)	59 (28.5)	1.32 (0.56-3.11)	
			45-64y	19 (22.1)	67 (32.4)	0.65 (0.26-1.60)	
			65+y	18 (20.9)	46 (22.2)	0.90 (0.36-2.26)	
Ethnicity	All	340 (1.2)	Non-white *(e.g. Asian, Chinese, Black African or Caribbean or other)*	16 (13.7)	21 (9.4)	Ref.	0.231
			White	101 (86.3)	202 (90.6)	0.66 (0.33-1.31)	
	Indigenous	291 (0.7)	Non-white *(e.g. Asian, Chinese, Black African or Caribbean or other)*	8 (9.3)	19 (9.3)	Ref.	0.993
			White	78 (90.7)	186 (90.7)	1.00 (0.42-2.37)	
Occupational status	All	342 (0.6)	Currently in work	50 (42.7)	96 (42.7)	Ref.	0.908
			Student	5 (4.3)	11 (4.9)	0.87 (0.29-2.65)	

			Unemployed	9 (7.7)	12 (5.3)	1.44 (0.57-3.65)	
			Retired	33 (28.2)	70 (31.1)	0.90 (0.53-1.55)	
			Child below working age	20 (17.1)	36 (16)	1.07 (0.56-2.03)	
	Indigenous	292 (0.3)	Currently in work	36 (41.9)	87 (42.2)	Ref.	0.965
			Student	5 (5.8)	11 (5.3)	1.10 (0.36-3.39)	
			Unemployed	7 (8.1)	12 (5.8)	1.41 (0.51-3.87)	
			Retired	25 (29.1)	63 (30.6)	0.96 (0.52-1.76)	
			Child below working age	13 (15.1)	33 (16)	0.95 (0.45-2.02)	
Educational status	All	340 (1.2)	Left full-time education	90 (78.3)	174 (77.3)	Ref.	0.358
			Still in full-time education	8 (7)	25 (11.1)	0.62 (0.27-1.43)	
			Pre-school child	17 (14.8)	26 (11.6)	1.26 (0.65-2.45)	
	Indigenous	290 (1)	Left full-time education	66 (78.6)	158 (76.7)	1.00	0.595
			Still in full-time education	7 (8.3)	25 (12.1)	0.67 (0.28-1.63)	
			Pre-school child	11 (13.1)	23 (11.2)	1.14 (0.53-2.48)	
Level of neighbourhood deprivation (IMD rank)	All	343 (0.3)	Very low (24,363-32,482)	26 (22)	58 (25.8)	Ref.	0.876^{CT} / 0.712^{C}
			Low (16,242-24,362)	32 (27.1)	58 (25.8)	1.23 (0.65-2.32)	
			Moderate (8,121-16,241)	24 (20.3)	41 (18.2)	1.31 (0.66-2.59)	
			Very high (1-8,120)	36 (30.5)	68 (30.2)	1.18 (0.64-2.18)	
	Indigenous	292 (0.3)	Very low (24,363-32,482)	18 (20.9)	50 (24.3)	Ref.	0.817^{CT}/0.932^{C}
			Low (16,242-24,362)	26 (30.2)	52 (25.2)	1.39 (0.68-2.84)	
			Moderate (8,121-16,241)	17 (19.8)	40 (19.4)	1.18 (0.54-2.58)	
			Very high (1-8,120)	25 (29.1)	64 (31.1)	1.08 (0.53-2.21)	
Reported area of living	All	337 (2)	City	20 (17.5)	42 (18.8)	Ref.	0.373
			Town	62 (54.4)	104 (46.6)	1.25 (0.67-2.32)	
			Village	32 (28.1)	77 (34.5)	0.87 (0.44-1.71)	
	Indigenous	287 (2)	City	15 (18.3)	38 (18.5)	Ref.	0.737
			Town	43 (52.4)	98 (47.8)	1.11 (0.55-2.23)	
			Village	24 (29.3)	69 (33.7)	0.88 (0.41-1.88)	
Study catchment area	All	344	Central Lancashire	67 (56.8)	136 (60.2)	Ref.	0.738
			East Lancashire	39 (33.1)	72 (31.9)	1.10 (0.67-1.79)	
			Greater Manchester	12 (10.2)	18 (8)	1.35 (0.62-2.97)	
	Indigenous	293	Central Lancashire	55 (64)	126 (60.9)	Ref.	0.162

			East Lancashire	21 (24.4)	68 (32.9)	0.71 (0.39-1.27)	
			Greater Manchester	10 (11.6)	13 (6.3)	1.76 (0.73-4.26)	
Season	All	344	Winter (December-February)	24 (20.3)	48 (21.2)	Ref.	0.753
			Spring (March-May)	35 (29.7)	70 (31)	1.00 (0.53-1.89)	
			Summer (June-August)	30 (25.4)	46 (20.4)	1.30 (0.67-2.55)	
			Autumn (September-November)	29 (24.6)	62 (27.4)	0.93 (0.48-1.81)	
	Indigenous	293	Winter (December-February)	15 (17.4)	44 (21.3)	Ref.	0.786
			Spring (March-May)	27 (31.4)	67 (32.4)	1.18 (0.57-2.47)	
			Summer (June-August)	21 (24.4)	41 (19.8)	1.50 (0.68-3.30)	
			Autumn (September-November)	23 (26.7)	55 (26.6)	1.23 (0.57-2.63)	
HEALTH DETAILS							
Suffering from inflammatory bowel disease (IBD)	All	326 (5.2)	No	106 (99.1)	217 (99.1)	Ref.	1.000
			Yes	1 (0.9)	2 (0.9)	1.02 (0.09-11.42)	
	Indigenous	279 (4.8)	No	78 (98.7)	198 (99)	Ref.	1.000
			Yes	1 (1.3)	2 (1)	1.27 (0.11-14.20)	
OCCUPATIONAL DETAILS							
Having a potentially at-risk occupation	All	341 (0.9)	Not currently working	67 (57.8)	129 (57.3)	Ref.	0.864
			Working but not with any at-risk occupation	43 (37.1)	83 (36.9)	1.00 (0.62-1.60)	
			Working with animals	0 (0)	1 (0.4)	n/a***	
			Working with manure/slurry	0 (0)	1 (0.4)	n/a***	
			Working with human faeces	3 (2.6)	7 (3.1)	0.82 (0.21-3.29)	
			Contact with sewage	0 (0)	1 (0.4)	n/a***	
			Contact with freshwaters	1 (0.9)	1 (0.4)	1.92 (0.12-31.27)	
			More than one at-risk occupation	2 (1.7)	2 (0.9)	1.92 (0.26-13.97)	
	Indigenous	292 (0.3)	Not currently working	50 (58.1)	119 (57.8)	Ref.	0.748
			Working but not with any at-risk occupation	30 (34.9)	75 (36.4)	0.95 (0.56-1.63)	
			Working with animals	0 (0)	1 (0.5)	n/a***	
			Working with manure/slurry	0 (0)	1 (0.5)	n/a***	
			Working with human faeces	3 (3.5)	7 (3.4)	1.02 (0.25-4.10)	
			Contact with sewage	0 (0)	1 (0.5)	n/a***	

			Contact with freshwaters	1 (1.2)	1 (0.5)	2.38 (0.15-38.80)	
			More than one at-risk occupation	2 (2.3)	1 (0.5)	4.76 (0.42-53.69)	
ANIMAL CONTACT							
Keeping a pet by type	All	341 (0.9)	Not keeping pets	72 (62.1)	120 (53.3)	Ref.	0.229
			Keeping dogs	17 (14.7)	34 (15.1)	0.83 (0.43-1.60)	
			Keeping cats	6 (5.2)	25 (11.1)	0.40 (0.16-1.02)	
			Keeping birds	3 (2.6)	7 (3.1)	0.71 (0.18-2.85)	
			Keeping horses	0 (0)	2 (0.9)	n/a***	
			Keeping rodents	1 (0.9)	7 (3.1)	0.24 (0.03-1.97)	
			Keeping a pet other than those above (fish, amphibians, stick insects, spiders)	1 (0.9)	5 (2.2)	0.33 (0.04-2.91)	
			Keeping more than one type of pet	16 (13.8)	25 (11.1)	1.07 (0.53-2.13)	
	Indigenous	290 (1)	Not keeping pets	50 (59.5)	108 (52.4)	Ref.	0.253
			Keeping dogs	10 (11.9)	33 (16)	0.65 (0.30-1.43)	
			Keeping cats	4 (4.8)	22 (10.7)	0.39 (0.13-1.20)	
			Keeping birds	3 (3.6)	6 (2.9)	1.08 (0.26-4.49)	
			Keeping horses	0 (0)	2 (1)	n/a***	
			Keeping rodents	1 (1.2)	7 (3.4)	0.31 (0.04-2.58)	
			Keeping a pet other than those above (fish, amphibians, stick insects, spiders)	1 (1.2)	5 (2.4)	0.43 (0.05-3.79)	
			Keeping more than one type of pet	15 (17.9)	23 (11.2)	1.41 (0.68-2.93)	
No. of dogs kept	All	340 (1.2)	0	86 (74.8)	176 (78.2)	Ref.	$0.478^{T}/0.471^{C}$
			1	23 (20)	44 (19.6)	1.07 (0.61-1.88)	
			2	6 (5.2)	2 (0.9)	6.14 (1.21-31.05)	
			3+	0 (0)	3 (1.3)	n/a***	
	Indigenous	289 (1.4)	0	62 (74.7)	160 (77.7)	1.00	$0.557^{T}/0.539^{C}$
			1	16 (19.3)	41 (19.9)	1.00 (0.53-1.92)	
			2	5 (6)	2 (1)	6.45 (1.22-34.13)	
			3+	0 (0)	3 (1.5)	n/a***	
No. of different pet species owned	All	341 (0.9)	0	72 (62.1)	120 (53.3)	Ref.	$0.713^{T}/0.955^{C}$
			1	28 (24.1)	79 (35.1)	0.59 (0.35-0.99)	
			2	9 (7.8)	21 (9.3)	0.71 (0.31-1.64)	

			3+	7 (6)	5 (2.2)	2.33 (0.71-7.63)	
	Indigenous	290 (1)	0	50 (59.5)	108 (52.4)	1.00	0.693[T]/0.472[C]
			1	19 (22.6)	74 (35.9)	0.55 (0.30-1.02)	
			2	8 (9.5)	19 (9.2)	0.91 (0.37-2.22)	
			3+	7 (8.3)	5 (2.4)	3.02 (0.91-10.00)	
Cleaning up (touching) pets' faeces	All	341 (0.9)	Not keeping pets	72 (62.1)	120 (53.3)	Ref.	0.290
			Keeping pets but not cleaning up their faeces	24 (20.7)	60 (26.7)	0.67 (0.38-1.16)	
			Keeping pets and cleaning up their faeces	20 (17.2)	45 (20)	0.74 (0.41-1.35)	
	Indigenous	290 (1)	Not keeping pets	50 (59.5)	108 (52.4)	Ref.	0.540
			Keeping pets but not cleaning up their faeces	19 (22.6)	56 (27.2)	0.73 (0.39-1.36)	
			Keeping pets and cleaning up their faeces	15 (17.9)	42 (20.4)	0.77 (0.39-1.52)	
Any pet with diarrhoea in the house	All	340 (1.2)	Not keeping pets	72 (62.6)	120 (53.3)	Ref.	0.202
			Keeping pets but none with diarrhoea	42 (36.5)	100 (44.4)	0.70 (0.44-1.11)	
			Keeping pets and at least one with diarrhoea	1 (0.9)	5 (2.2)	0.33 (0.04-2.91)	
	Indigenous	289 (1.4)	Not keeping pets	50 (60.2)	108 (52.4)	Ref.	0.421
			Keeping pets but none with diarrhoea	32 (38.6)	93 (45.1)	0.74 (0.44-1.25)	
			Keeping pets and at least one with diarrhoea	1 (1.2)	5 (2.4)	0.43 (0.05-3.79)	
Visiting or working at a farm	All	340 (1.2)	No	101 (88.6)	192 (85)	Ref.	0.358
			Yes	13 (11.4)	34 (15)	0.73 (0.37-1.44)	
	Indigenous	292 (0.3)	No	74 (87.1)	176 (85)	Ref.	0.653
			Yes	11 (12.9)	31 (15)	0.84 (0.40-1.77)	
Touching animals while visiting or working at a farm	All	339 (1.5)	Neither visiting nor working at a farm	101 (88.6)	192 (85.3)	Ref.	0.659
			Not touching animals while at a farm	4 (3.5)	12 (5.3)	0.63 (0.20-2.01)	
			Touching animals while at a farm	9 (7.9)	21 (9.3)	0.81 (0.36-1.84)	
	Indigenous	291 (0.7)	Neither visiting nor working at a farm	74 (87.1)	176 (85.4)	1.00	0.936
			Not touching animals while at a farm	4 (4.7)	12 (5.3)	0.86 (0.27-2.80)	
			Touching animals while at a farm	7 (8.2)	19 (9.2)	0.88 (0.35-2.17)	
Touching animals while at a farm by type	All	339 (1.5)	Neither visiting nor working at a farm	101 (88.6)	192 (85.3)	Ref.	0.261
			Not touching animals while at a farm	4 (3.5)	12 (5.8)	0.58 (0.19-1.84)	

of animal			Touching pigs	1 (0.9)	0 (0)	n/a***	
			Touching sheep or goats	0 (0)	3 (1.3)	n/a***	
			Touching horses or ponies	1 (0.9)	4 (1.8)	0.47 (0.05-4.31)	
			Touching chickens or poultry	0 (0)	1 (0.4)	n/a***	
			Touching any other animal	0 (0)	2 (0.9)	n/a***	
			Touching more than one type of animal (cattle included)	7 (6.1)	10 (4.4)	1.33 (0.49-3.60)	
	Indigenous	291 (0.7)	Neither visiting nor working at a farm	74 (87.1)	176 (85.4)	Ref.	0.592
			Not touching animals while at a farm	4 (4.7)	12 (5.8)	0.79 (0.25-2.54)	
			Touching sheep or goats	0 (0)	2 (1)	n/a***	
			Touching horses or ponies	1 (1.2)	4 (1.9)	0.58 (0.06-5.41)	
			Touching chickens or poultry	0 (0)	1 (0.5)	n/a***	
			Touching any other animal	0 (0)	2 (1)	n/a***	
			Touching more than one type of animal (cattle included)	6 (7.1)	9 (4.4)	1.59 (0.54-4.61)	
Touching animals at a wildlife park or zoo	All	340 (1.2)	Not visiting a wildlife park or zoo	106 (92.2)	217 (96.4)	Ref.	0.235
			Visiting but not touching any animal	6 (5.2)	6 (2.7)	2.05 (0.64-6.50)	
			Visiting and touching an animal	3 (2.6)	2 (0.9)	3.07 (0.50-18.65)	
	Indigenous	290 (1)	Not visiting a wildlife park or zoo	79 (94)	198 (96.1)	Ref.	0.339
			Visiting but not touching any animal	2 (2.4)	6 (2.9)	0.83 (0.16-4.23)	
			Visiting and touching an animal	3 (3.6)	2 (1)	3.76 (0.62-22.93)	
Touching animals in the wild	All	293 (14.8)	No	100 (94.3)	182 (97.3)	Ref.	0.214
			Yes	6 (5.7)	5 (2.7)	2.18 (0.65-7.34)	
	Indigenous	253 (13.7)	No	78 (97.5)	170 (98.3)	Ref.	0.653
			Yes	2 (2.5)	3 (1.7)	1.45 (0.24-8.87)	
RECREATIONAL ACTIVITIES							
Frequency of swimming or paddling in a swimming pool (no. times)	All	324 (5.8)	0	75 (70.8)	177 (81.2)	Ref.	$0.106^{T}/0.413^{C}$
			1-2	15 (14.2)	17 (7.8)	2.08 (0.99-4.39)	
			3-4	7 (6.6)	12 (5.5)	1.38 (0.52-3.63)	
			5-6	4 (3.8)	6 (2.8)	1.57 (0.43-5.74)	
			7+	5 (4.7)	6 (2.8)	1.97 (0.58-6.64)	
	Indigenous	279 (4.8)	0	58 (72.5)	168 (84.4)	Ref.	$0.106^{T}/0.438^{C}$

			1-2	12 (15)	14 (7)	2.48 (1.09-5.68)	
			3-4	7 (8.8)	12 (6)	1.69 (0.63-4.50)	
			5-6	2 (2.5)	3 (1.5)	1.93 (0.31-11.85)	
			7+	1 (1.3)	2 (1)	1.45 (0.13-16.27)	
Frequency of using a Jacuzzi or a hot tub (no. times)	All	307 (10.8)	0	90 (94.7)	203 (95.8)	Ref.	0.976 [T]/0.672 [C]
			1-2	4 (4.2)	6 (2.8)	1.50 (0.41-5.46)	
			3+	1 (1.1)	3 (1.4)	0.75 (0.07-7.33)	
	Indigenous	261 (10.9)	0	64 (94.1)	186 (96.4)	Ref.	0.775 [T]/0.554 [C]
			1-2	3 (4.4)	4 (2.1)	2.18 (0.47-10.00)	
			3+	1 (1.5)	3 (1.6)	0.97 (0.10-9.48)	
Immersing the head underwater while using a Jacuzzi or hot tub	All	304 (11.6)	Not using a Jacuzzi or a hot tub	90 (96.8)	203 (96.2)	Ref.	0.935
			Using a Jacuzzi or a hot tub without immersing the head	2 (2.2)	6 (2.8)	0.75 (0.15-3.80)	
			Using Jacuzzi or a hot tub immersing the head	1 (1.1)	2 (0.9)	1.13 (0.10-12.60)	
	Indigenous	258 (11.9)	Not using a Jacuzzi or a hot tub	64 (97)	186 (96.9)	1.00	0.506
			Using a Jacuzzi or a hot tub without immersing the head	2 (3)	4 (2.1)	1.45 (0.26-8.12)	
			Using Jacuzzi or a hot tub immersing the head	0 (0)	2 (1)	n/a***	
Swimming in a lake, pond or river	All	316 (8.1)	No	99 (97.1)	209 (97.7)	Ref.	0.716
			Yes	3 (2.9)	5 (2.3)	1.27 (0.30-5.41)	
	Indigenous	272 (7.2)	No	75 (98.7)	192 (98)	Ref.	1.000
			Yes	1 (1.3)	4 (2)	0.64 (0.07-5.82)	
Swimming in the sea	All	321 (6.7)	No	93 (91.2)	206 (94.1)	Ref.	0.340
			Yes	9 (8.8)	13 (5.9)	1.53 (0.63-3.71)	
	Indigenous	276 (5.8)	No	74 (98.7)	197 (98)	Ref.	1.000
			Yes	1 (1.3)	4 (2)	0.67 (0.07-6.05)	
Frequency of swimming in the sea (no. times)	All	318 (7.6)	0	95 (94)	206 (94.1)	Ref.	0.622 [T]/0.243 [C]
			1-2	2 (2)	10 (4.6)	0.44 (0.09-2.06)	
			3+	4 (4)	3 (1.4)	2.95 (0.65-13.46)	
	Indigenous	276 (5.8)	0	74 (98.7)	197 (98)	Ref.	0.917 [T]/0.464 [C]
			1-2	0 (0)	4 (2)	n/a***	

			3+	1 (1.3)	0 (0)	n/a***	
Practising watersports in freshwater	All	324 (5.8)	No	104 (100)	219 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
	Indigenous	277 (5.5)	No	76 (100)	200 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
Practising watersports in the sea	All	324 (5.8)	No	103 (99)	219 (99.5)	Ref.	0.540
			Yes	1 (1)	1 (0.5)	2.13 (0.13-34.33)	
	Indigenous	277 (5.5)	No	76 (100)	200 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
Going fishing	All	319 (7.3)	No	110 (100)	206 (98.6)	Ref.	0.554
			Yes	0 (0)	3 (1.4)	n/a***	
	Indigenous	273 (6.8)	No	82 (100)	188 (98.4)	Ref.	0.556
			Yes	0 (0)	3 (1.6)	n/a	
Going hunting		Exposure not reported in either cases or controls					
Going camping	All	318 (7.6)	No	107 (98.2)	206 (98.6)	Ref.	1.000
			Yes	2 (1.8)	3 (1.4)	1.28 (0.21-7.80)	
	Indigenous	272 (7.2)	No	81 (100)	188 (98.4)	Ref.	0.557
			Yes	0 (0)	3 (1.6)	n/a***	
Going caravanning	All	320 (7)	No	106 (97.2)	199 (94.3)	Ref.	0.239
			Yes	3 (2.8)	12 (5.7)	0.47 (0.13-1.70)	
	Indigenous	274 (6.5)	No	79 (97.5)	181 (93.8)	Ref.	0.244
			Yes	2 (2.5)	12 (6.2)	0.38 (0.08-1.75)	
Walking in the countryside	All	329 (4.4)	No	69 (63.9)	128 (57.9)	Ref.	0.300
			Yes	39 (36.1)	93 (42.1)	0.78 (0.48-1.25)	
	Indigenous	282 (3.8)	No	50 (62.5)	120 (59.4)	Ref.	0.632
			Yes	30 (37.5)	82 (40.6)	0.88 (0.51-1.50)	
WATER CONSUMPTION							
No. glasses of un-boiled tap water per day	All	321 (6.7)	0	10 (9.6)	22 (10.1)	Ref.	$0.958^{T}/0.496^{C}$
			1-2	46 (44.2)	98 (45.2)	1.03 (0.45-2.36)	
			3-4	31 (29.8)	60 (27.6)	1.14 (0.48-2.70)	
			5+	17 (16.3)	37 (17.1)	1.01 (0.39-2.59)	
	Indigenous	277 (5.5)	0	4 (5.2)	21 (10.5)	Ref.	$0.548^{T}/0.341^{C}$

			1-2	36 (46.8)	91 (45.5)	2.08 (0.67-6.47)	
			3-4	25 (32.5)	54 (27)	2.43 (0.76-7.83)	
			5+	12 (15.6)	34 (17)	1.85 (0.53-6.50)	
Drinking un-boiled water from a lake, river or stream	All	307 (10.8)	No	104 (98.1)	199 (99)	Ref.	0.611
			Yes	2 (1.9)	2 (1)	1.91 (0.27-13.78)	
	Indigenous	263 (10.2)	No	77 (98.7)	183 (98.9)	Ref.	1.000
			Yes	1 (1.3)	2 (1.1)	1.19 (0.11-13.30)	
Water from the tap reported having an unusual smell	All	334 (2.9)	No	112 (98.2)	215 (97.7)	Ref.	1.000
			Yes	2 (1.8)	5 (2.3)	0.77 (0.15-4.02)	
	Indigenous	284 (3.1)	No	83 (100)	196 (97.5)	Ref.	0.326
			Yes	0 (0)	5 (2.5)	n/a***	
FOOD CONSUMPTION							
FOOD CONSUMPTION HABITS							
No. times per week eating cooked vegetables	All	329 (4.4)	0	3 (2.7)	6 (2.7)	Ref.	$0.926^{T}/0.156^{C}$
			1-2	12 (10.9)	26 (11.9)	0.92 (0.20-4.33)	
			3-4	34 (30.9)	63 (28.8)	1.08 (0.25-4.59)	
			5+	61 (55.5)	124 (56.6)	0.98 (0.24-4.07)	
	Indigenous	281 (4.1)	0	1 (1.2)	6 (3)	Ref.	$0.442^{T}/0.551^{C}$
			1-2	9 (11.1)	26 (13)	2.08 (0.22-19.68)	
			3-4	24 (29.6)	59 (29.5)	2.44 (0.28-21.37)	
			5+	47 (58)	109 (54.5)	2.59 (0.30-22.09)	
Peeling raw fruit before eating	All	333 (3.2)	No	81 (73)	173 (77.9)	Ref.	0.316
			Yes	30 (27)	49 (22.1)	1.31 (0.77-2.21)	
	Indigenous	284 (3.1)	No	59 (72.8)	157 (77.3)	Ref.	0.422
			Yes	22 (27.2)	46 (22.7)	1.27 (0.71-2.29)	
Washing raw fruit before eating	All	330 (4.1)	No	36 (33.3)	70 (31.5)	Ref.	0.742
			Yes	72 (66.7)	152 (68.5)	0.92 (0.56-1.50)	
	Indigenous	281 (4.1)	No	26 (33.3)	65 (32)	Ref.	0.833
			Yes	52 (66.7)	138 (68)	0.94 (0.54-1.64)	
FOOD CONSUMPTION DURING THE EXPOSURE WINDOW							
Eating cooked vegetables	All	333 (3.2)	No	8 (7.1)	14 (6.3)	Ref.	0.779
			Yes	104 (92.9)	207 (93.7)	0.88 (0.36-2.16)	

	Indigenous	284 (3.1)	No	5 (6.1)	13 (6.4)	Ref.	0.916
			Yes	77 (93.9)	189 (93.6)	1.06 (0.36-3.07)	
Drinking fruit juice	All	331 (3.8)	No	33 (29.5)	58 (26.5)	Ref.	0.566
			Yes	79 (70.5)	161 (73.5)	0.86 (0.52-1.43)	
	Indigenous	283 (3.4)	No	26 (31.3)	54 (27)	Ref.	0.462
			Yes	57 (68.7)	146 (73)	0.81 (0.46-1.42)	
Provenience of fruit juice	All	323 (6.1)	Did not drink fruit juice	33 (30.3)	58 (27.1)	Ref.	0.088
			Market	1 (0.9)	1 (0.5)	1.76 (0.11-29.03)	
			Greengrocers	1 (0.9)	0 (0)	n/a***	
			Supermarket	68 (62.4)	145 (67.8)	0.82 (0.49-1.38)	
			Other place (e.g. restaurant etc.)	5 (4.6)	2 (0.9)	4.39 (0.81-23.9)	
			Multiple proveniences (at least two different)	1 (0.9)	8 (3.7)	0.22 (0.03-1.83)	
	Indigenous	276 (5.8)	Did not drink fruit juice	26 (32.1)	54 (27.7)	Ref.	0.552
			Market	1 (1.2)	1 (0.5)	2.08 (0.12-34.53)	
			Greengrocers	1 (1.2)	0 (0)	n/a***	
			Supermarket	51 (63)	133 (68.2)	0.80 (0.45-1.41)	
			Other place (e.g. restaurant etc.)	1 (1.2)	2 (1)	1.04 (0.09-11.98)	
			Multiple proveniences (at least two different)	1 (1.2)	5 (2.6)	0.41 (0.05-3.74)	
Eating pre-packed sandwiches	All	321 (6.7)	No	74 (66.7)	144 (68.6)	Ref.	0.728
			Yes	37 (33.3)	66 (31.4)	1.09 (0.67-1.78)	
	Indigenous	275 (6.1)	No	56 (68.3)	134 (69.4)	Ref.	0.852
			Yes	26 (31.7)	59 (30.6)	1.05 (0.60-1.84)	
Provenience of pre-packed sandwiches	All	317 (7.8)	Did not eat pre-packed sandwiches	74 (67.9)	144 (69.2)	Ref.	0.318
			Greengrocers	0 (0)	1 (0.5)	n/a***	
			Supermarket	30 (27.5)	60 (28.8)	0.97 (0.58-1.64)	
			Other place (e.g. restaurant etc.)	5 (4.6)	3 (1.4)	3.24 (0.75-13.94)	
	Indigenous	272 (7.2)	Did not eat pre-packed sandwiches	56 (69.1)	134 (70.2)	Ref.	0.441
			Greengrocers	0 (0)	1 (0.5)	n/a***	
			Supermarket	22 (27.2)	54 (28.3)	0.97 (0.54-1.75)	
			Other place (e.g. restaurant etc.)	3 (3.7)	2 (1)	3.59 (0.58-22.07)	

Eating lamb	All	305 (11.3)	No	43 (42.6)	100 (49)	Ref.	0.288
			Yes	58 (57.4)	104 (51)	1.30 (0.80-2.10)	
	Indigenous	264 (9.9)	No	32 (41.6)	94 (50.3)	1.00	0.198
			Yes	45 (58.4)	93 (49.7)	1.42 (0.83-2.43)	
Provenience of lamb	All	302 (12.2)	Did not eat lamb	43 (43.9)	100 (49)	Ref.	0.185
			Market stall	0 (0)	4 (2)	n/a***	
			Butcher	14 (14.3)	22 (10.8)	1.48 (0.69-3.16)	
			Supermarket	31 (31.6)	67 (32.8)	1.07 (0.62-1.88)	
			Other place (e.g. restaurant etc.)	4 (4.1)	2 (1)	4.65 (0.82-26.36)	
			Multiple proveniences (at least two different)	6 (6.1)	9 (4.4)	1.55 (0.52-4.62)	
	Indigenous	261 (10.9)	Did not eat lamb	32 (43.2)	94 (50.3)	Ref.	0.453
			Market stall	0 (0)	3 (1.6)	n/a***	
			Butcher	11 (14.9)	20 (10.7)	1.62 (0.70-3.73)	
			Supermarket	27 (36.5)	63 (33.7)	1.26 (0.69-2.30)	
			Other place (e.g. restaurant etc.)	0 (0)	1 (0.5)	n/a***	
			Multiple proveniences (at least two different)	4 (5.4)	6 (3.2)	1.96 (0.52-7.38)	
Eating chicken	All	331 (3.8)	No	5 (4.6)	18 (8.1)	Ref.	0.236
			Yes	104 (95.4)	204 (91.9)	1.83 (0.66-5.08)	
	Indigenous	283 (3.4)	No	4 (5)	17 (8.4)	Ref.	0.329
			Yes	76 (95)	186 (91.6)	1.74 (0.57-5.33)	
Eating fish	All	319 (7.3)	No	33 (31.4)	60 (28)	Ref.	0.531
			Yes	72 (68.6)	154 (72)	0.85 (0.51-1.41)	
	Indigenous	275 (6.1)	No	26 (32.9)	59 (30.1)	Ref.	0.648
			Yes	53 (67.1)	137 (69.9)	0.88 (0.50-1.54)	
Provenience of fish	All	317 (7.8)	Did not eat fish	33 (32)	60 (28)	Ref.	0.325
			Market stall	2 (1.9)	6 (2.8)	0.61 (0.12-3.17)	
			Fishmonger	11 (10.7)	14 (6.5)	1.43 (0.58-3.50)	
			Supermarket	51 (49.5)	116 (54.2)	0.80 (0.47-1.37)	
			Other place (e.g. restaurant etc.)	4 (3.9)	5 (2.3)	1.45 (0.36-5.79)	
			Multiple proveniences (at least two different)	2 (1.9)	13 (6.1)	0.28 (0.06-1.31)	

	Indigenous	273 (6.8)	Did not eat fish	26 (33.8)	59 (30.1)	Ref.	0.613
			Market stall	1 (1.3)	5 (2.6)	0.45 (0.05-4.08)	
			Fishmonger	7 (9.1)	14 (7.1)	1.13 (0.41-3.14)	
			Supermarket	41 (53.2)	104 (53.1)	0.89 (0.50-1.61)	
			Other place (e.g. restaurant etc.)	1 (1.3)	4 (2)	0.57 (0.06-5.32)	
			Multiple proveniences (at least two different)	1 (1.3)	4 (2)	0.23 (0.03-1.87)	
Eating at a pub or a restaurant	All	337 (2)	No	46 (40.4)	81 (36.3)	Ref.	0.470
			Yes	68 (59.6)	142 (63.7)	0.84 (0.53-1.34)	
	Indigenous	289 (1.4)	No	37 (44)	80 (39)	Ref.	0.430
			Yes	47 (56)	125 (61)	0.81 (0.49-1.36)	
No. times eating at a pub or a restaurant	All	326 (5.2)	0	46 (43.4)	81 (36.8)	Ref.	$0.143^{T}/0.437^{C}$
			1-2	44 (41.5)	83 (37.7)	0.93 (0.56-1.56)	
			3-4	5 (4.7)	36 (16.4)	0.24 (0.09-0.67)	
			5+	11 (10.4)	20 (9.1)	0.97 (0.43-2.20)	
	Indigenous	283 (3.4)	0	37 (45.7)	80 (39.6)	1.00	$0.043^{T}/0.066^{C}$
			1-2	38 (46.9)	79 (39.1)	1.04 (0.60-1.80)	
			3-4	3 (3.7)	32 (15.8)	0.20 (0.06-0.70)	
			5+	3 (3.7)	11 (5.4)	0.59 (0.15-2.24)	
Eating at a canteen	All	305 (11.3)	No	86 (85.1)	182 (89.2)	Ref.	0.306
			Yes	15 (14.9)	22 (10.8)	1.44 (0.71-2.92)	
	Indigenous	263 (10.2)	No	65 (84.4)	166 (89.2)	Ref.	0.275
			Yes	12 (15.6)	20 (10.8)	1.53 (0.71-3.31)	
No. times eating at a canteen	All	301 (12.5)	0	86 (86.9)	182 (90.1)	Ref.	$0.559^{T}/0.978^{C}$
			1-2	5 (5.1)	7 (3.5)	1.51 (0.47-4.90)	
			3-4	2 (2)	3 (1.5)	1.41 (0.23-8.60)	
			5+	6 (6.1)	10 (5)	1.27 (0.45-3.61)	
	Indigenous	259 (11.6)	0	65 (86.7)	166 (90.2)	Ref.	$0.477^{T}/0.919^{C}$
			1-2	3 (4)	6 (3.3)	1.28 (0.31-5.26)	
			3-4	2 (2.7)	3 (1.6)	1.70 (0.28-10.42)	
			5+	5 (6.7)	9 (4.9)	1.42 (0.46-4.39)	
Eating at a takeaway or	All	324 (5.8)	No	46 (41.8)	94 (43.9)	Ref.	0.717

fast food			Yes	64 (58.2)	120 (56.1)	1.09 (0.68-1.73)	
	Indigenous	277 (5.5)	No	35 (42.7)	86 (44.1)	Ref.	0.828
			Yes	47 (57.3)	109 (55.9)	1.06 (0.63-1.78)	
No. times eating at a takeaway or fast food	All	307 (10.8)	0	46 (46.5)	94 (45.2)	Ref.	$0.917^{T}/0.670^{C}$
			1-2	32 (32.3)	81 (38.9)	0.81 (0.47-1.39)	
			3-4	18 (18.2)	22 (10.6)	1.67 (0.82-3.42)	
			5+	3 (3)	11 (5.3)	0.56 (0.15-2.09)	
	Indigenous	264 (9.9)	0	35 (46.7)	86 (45.5)	Ref.	$0.714^{T}/0.558^{C}$
			1-2	23 (30.7)	72 (38.1)	0.78 (0.43-1.45)	
			3-4	16 (21.3)	20 (10.6)	1.97 (0.91-4.23)	
			5+	1 (1.3)	11 (5.8)	0.22 (0.03-1.80)	
GENERAL HOUSEHOLD DETAILS							
No. of adults (≥16y) in the house	All	343 (0.3)	1	15 (12.8)	52 (23)	Ref.	$0.419^{T}/0.868^{C}$
			2	80 (68.4)	128 (56.6)	2.17 (1.14-4.10)	
			3	16 (13.7)	32 (14.2)	1.73 (0.75-3.98)	
			4+	6 (5.1)	14 (6.2)	1.49 (0.49-4.53)	
	Indigenous	293	1	13 (15.1)	50 (24.2)	Ref.	$0.963^{T}/0.703^{C}$
			2	60 (69.8)	113 (54.6)	2.04 (1.03-4.05)	
			3	9 (10.5)	31 (15)	1.12 (0.43-2.92)	
			4+	4 (4.7)	13 (6.3)	1.18 (0.33-4.24)	
Any contact with a person with diarrhoea outside the house	All	315 (8.4)	No	94 (87)	187 (90.3)	Ref.	0.370
			Yes	14 (13)	20 (9.7)	1.39 (0.67-2.88)	
	Indigenous	271 (7.5)	No	75 (93.8)	172 (90.1)	Ref.	0.328
			Yes	5 (6.3)	19 (9.9)	0.60 (0.22-1.68)	

*percentages refer to the proportion of participants with missing information for the variable; **percentages refer to the proportion among cases and controls that answered the question; ***Odds ratio with 95% confidence interval; *** Odds ratio not calculated because no variation present in the data; ^{T} Chi-Square for trend; ^{CT} logistic regression on the categorized variable; ^{C} logistic regression on the continuous form of the variable

APPENDIX 4: *Giardia* molecular genotyping results

Table 4A: *Giardia* reference sequences used for sub-assemblage genotyping. Sequences are indicated by their GenBank accession number and host of origin.

		Gene			
Species	**Assemblage**	***bg***	***gdh***	***tpi***	***ssu-rRNA***
G. duodenalis	AI	X85958 human	M84604 human	L02120 human	M54878 human
	AII	AY072723 human	AY178737 human	U57897 human	AF199446 human
	AIII	EU216429 red deer	EU637582 fallow deer	DQ650648 fallow deer	DQ100287 roe deer
	BIII	AY072726 human	EF685684 human	AF069561 human	AF199447 human
	BIV	AY072728 human	AY178738 human	AF069560 human	AF113898 human
	C	AY545646 dog	U60983 dog	AY228641 dog	AF199449 dog
	D	AY545647 dog	U60986 dog	DQ246216 dog	AF199443 dog
	E	AY072729 pig	AY178741 pig	KF891311 water buffalo	AF199448 goat
	F	AY647264 cat	AF069057 cat	AF069558 cat	AF199444 cat
	G	EU769221 rat	AY178746 rat	EU781013 rat	AF199450 rat
G. muris	-	EF455599 mouse	*Not available*	*Not available*	M73682 mouse
G. ardeae	-	*Not available*	AF069060 blue heron	AF069564 blue heron	*Not available*

Table 4B: List of the 157 isolates successfully sequenced at the *bg* gene ordered by sub-assemblage and genotype.

Sub-assemblage	Subtype	Identity (ID) to closest reference or previously described isolate (GenBank Acc. No.) *(no. of differences)*	Host(s) or environmental sample of isolation of matching sequences in GenBank	No. isolates	Isolates ID *(sequenced in only one direction)*
AII	**A2**	100%ID Reference AII (AY072723)	Human, sheep, bison, wastewater treatment plant, reclaimed wastewater	16	13/CT, 13/C161, 13/C133, 13/C123, 12/G, 12/C88, 12/C46, 08/H3, 10/H21, 11/11, 11/12, 11/27, 12/C14, 12/C15, 12/C16, 12/C41
	A2a*	99%ID Reference AII *(1)*		1	13/C146
	A3	100%ID A3 isolate ECUST13495 (JX898210)	Human, dog, bison, oryx, wastewater treatment plant, reclaimed wastewater	23	12/C70, 09/10, 13/C164, 09/22, 13/C144, 13/C122, 13/C117, 12/C105, 12/C101, 12/C91, 12/C74, 12/C65, 12/C56, 12/C31, 11/67, 11/29, 10/H51, 10/30, 08/13 *(10/3, 12/C98, 13/C127, 13/C162)*
AI	**AI**	100%ID Reference AI (X85958)	Human, dog, cat, cattle, goat, horse, moose, planigale, dolphin	1	11/13
	AIa*	99%ID Reference AI *(1)*		1	10/H12
AI/II ?	***AI/2***	*100% ID References AI and AII*		*1*	*(13/DA)*
A heterogeneous**	**n/a**	/		5	10/20, 12/C89, 10/H8 *(10/24, 09/15)*
BIII	**BIII**	100%ID Reference BIII (AY072726)	Human	7	10/28, 11/1, 12/C20, 12/C28, 12/C110, 13/CC, 13/CM
	BIIIa*	99%ID Reference BIII *(1)*		1	12/C75
BIV	**BIV**	100%ID Reference BIV (AY072728)	Human	1	12/C23
	BIVa*	99%ID Reference BIV *(1)*		1	13/C151
	BIVb*	99%ID Reference BIV *(2)*		1	12/C19
	Sweh198	100%ID Sweh198 isolate (HM165226)	Human	1	11/69
	GU417	100%ID GU417 isolate (JQ303245)		1	13/C137
Other B	**B1**	100%ID B1 isolate Nij5 (AY072725)	Human	1	11/28
	B1-1	100% ID B1-1 isolate Sweh042 (HM165214)	Human, macaque, gazelle, tamandua	3	11/71, 12/C95, 11/3

	B1-2	100% ID B1-2 isolate Sweh033 (HM165213)	Human, dog, lemur catta	11	13/C160, 13/CF, 13/C149, 13/C147, 13/C145, 12/D, 12/C80, 12/C11, 12/C9, 12/C7 *(12/C84)*
	B1-2a*	99% ID B1-2 isolate Sweh033 *(1)*		1	13/CO
	B1-3	100%ID B1-3 isolate Sweh001 (HM165208)	Human, cattle	48	13/CH, 13/C118, 13/C126, 12/C106, 12/C102, 12/C93, 12/C85, 12/C83, 12/C79, 12/C77, 12/C73, 12/C72, 12/C61, 12/C59, 12/C45, 12/C44, 12/C43, 12/C37, 12/C35, 12/C26, 12/C21, 12/C18, 12/C5, 12/C1, 12/A, 11/50, 11/24, 11/10, 10/H9, 10/H6, 08/20, 08/8, 13/C130, 13/C132, 13/C140, 13/C155, 13/CI, 13/CP, 12/C2, 12/C3, 13/C131 *(09/11, 10/13, 10/15, 10/21, 12/C25, 13/C113, 13/C115)*
	B1-5	100%ID B1-5 isolate Sweh047 (HM165216)	Human	6	12/C87, 12/C92, 12/C96, 13/C152, 13/C165, 13/CE
	BG-Ber2	100% ID BG-Ber2 isolate (DQ090523)	Human	4	12/C81, 13/C157, 11/58 *(11/36)*
	BG-Ber6	100% ID BG-Ber6 isolate (DQ090527)	Human	1	12/C99
	BG-Ber1	100% ID BG-Ber1 isolate (DQ090522)	Human	1	13/CG
B heterogeneous **	**n/a**	/		19	13/C129, 11/44, 12/C24, 12/C30, 13/CL, 11/35, 12/C17, 12/C6, 12/C27, 13/C124, 13/C125, 13/C119, 13/C148, 12/C34, 10/H43, 12/C39, 11/59, 12/C114 *(10/7)*
Unsure	**A+B**	SNPs at 17 position along the gene sequence compatible with a mixed A and B template		1	12/C71

*sequence with at least one novel nucleotide polymorphism; **sequence with at least one heterogeneous position

Table 4C: Nucleotide polymorphisms at the *bg* locus in the assemblage A isolates successfully sequenced in both directions. The sequence of one representative isolate from each genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	242	245	248	278	335	350	353	364	368	383	391	402	416	419	434	458	461	467	479	497	501	557	563	576	578	610	623	635	647
AI (X85958)	a	g	a	c	g	t	t	t	a	a	g	g	a	c	a	a	t	a	t	a	g	g	g	a	c	g	c	c	t
11/13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10/H12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G*	.	.	.	.	.
AII (AY072723)	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13/CT	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13/C146	.	A	.	.	.	.	.	A*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ECUST13495 (JX898210)	.	A	.	.	.	.	.	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12/C70	.	A	.	.	.	.	.	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AIII (EU216429)	.	.	G	T	.	C	.	.	G	G	.	.	.	.	G	G	.	.	.	.	.	A	.	.	.	.	.	.	.
10/20	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.
12/C89	.	A	.	.	.	.	.	.	.	R	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10/H8	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the AI reference sequence; IUPAC nucleotide ambiguity codes: R=AG

Table 4D: Heterogeneous positions in the isolate 12/C71 *bg* sequence.

Isolate (GenBank Acc.No.)	242	245	248	278	284	287	317	335	350	353	368	374	383	402	416	419	434	458	461	467	479	494	497	500	501	518	548	557	566	578	623	635	647
AI (X85958)	A	G	A	C	C	G	C	G	T	T	A	C	A	G	A	C	A	A	T	A	T	C	A	G	G	A	T	G	G	C	C	C	T
AII (AY072723)	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AIII (EU216429)	.	.	G	T	.	.	.	.	C	.	G	.	G	.	.	.	G	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
BIII (AY072727)	G	.	G	.	G	A	T	A	C	G	G	T	G	.	G	T	G	G	C	G	C	A	G	A	.	G	C	.	A	T	T	T	C
BIV (AY072728)	.	.	G	.	G	.	T	A	C	G	G	T	G	.	G	T	G	G	C	G	C	A	.	A	.	G	C	.	A	T	.	T	C
12/C71	R	R	R	.	.	.	.	R	Y	K	R	.	R	R	R	Y	R	.	Y	R	Y	.	R	.	R	.	.	.	.	Y	Y	Y	Y

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; dots indicate identity to the AI reference sequence; IUPAC nucleotide ambiguity codes: R=AG, Y=CT, K=GT

Table 4E: Nucleotide polymorphisms at the *bg* locus in assemblage B isolates successfully sequenced in both directions and without heterogeneous positions. The sequence of one representative isolate from each identified genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	Nucleotidic position														
	227	242	287	335	380	386	413	482	497	524	527	587	623	641	647
BIII (AY072726)	g	g	a	a	g	g	g	g	g	g	g	g	t	g	c
10/28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12/C75	.	.	.	.	A*	.	.	.	.	.	.	.	.	.	.
BIV (AY072728)	.	A	G	.	.	.	.	.	A	.	.	.	C	.	.
12/C23	.	A	G	.	.	.	.	.	A	.	.	.	C	.	.
13/C151	.	A	G	.	.	A*	.	.	A	.	.	.	C	.	.
12/C19	.	A	G	.	.	.	.	.	.	.	A*	.	C	.	.
Sweh198 (M165226)	A	A	G	.	.	.	.	.	A	.	.	A	C	.	.
11/69	A	A	G	.	.	.	.	.	A	.	.	A	C	.	.
GU417 (JQ303245)	.	A	G	.	.	.	.	.	A	.	.	.	G	.	.
13/C137	.	A	G	.	.	.	.	.	A	.	.	.	G	.	.
B1 (AY072725)	.	.	.	.	.	.	A	.	A	A	.	.	.	A	.
11/28	.	.	.	.	.	.	A	.	A	A	.	.	.	A	.
B1-1 (HM165214)	.	.	.	.	.	.	A	.	A	.	.	.	.	A	.
11/71	.	.	.	.	.	.	A	.	A	.	.	.	.	A	.
B1-2 (HM165213)	.	.	.	.	.	.	.	.	A	.	.	.	.	A	.
13/C160	.	.	.	.	.	.	.	.	A	.	.	.	.	A	.
13/CO	.	.	.	.	.	.	.	.	A	.	.	.	C*	A	.
B1-3 (HM165208)	.	.	.	G	.	.	A	A	A	.	.	.	.	A	T
13/CH	.	.	.	G	.	.	A	A	A	.	.	.	.	A	T
B1-5 (HM165216)	.	.	.	G	.	.	A	A	A	.	.	.	.	A	.
12/C87	.	.	.	G	.	.	A	A	A	.	.	.	.	A	.
BG-Ber6 (DQ090527)	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.
12/C99	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.
BG-Ber1 (DQ090522)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
13/CG	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
BG-Ber2 (DQ090523)	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.
12/C81	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the BIII reference sequence

Table 4F: Nucleotide polymorphisms at the *bg* locus in assemblage B isolates successfully sequenced in both directions and showing heterogeneous positions. The sequence without heterogeneous positions of one representative isolate from each genotype is also shown.

Isolate (GenBank Acc.No.)	242	245	263	287	327	335	343	359	371	374	377	380	381	386	389	405	413	444	458	471	473	479	482	490	497	515	524	527	566	569	587	605	617	623	641	647	650
BIII (AY072726)	g	g	g	a	a	a	t	c	g	t	c	g	a	g	c	t	g	a	g	t	g	c	g	t	g	t	g	g	a	c	g	c	c	t	g	c	g
12/C75	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BIV (AY072728)	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.
13/C151	A	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.
12/C19	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	C	.	.	.
Sweh198 (M165226)	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	.	C	.	.	.
GU417 (JQ303245)	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	G	.	.	.
B1 (AY072725)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	A	.	.
B1-1 (HM165214)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	.
B1-2 (HM165213)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	.
13/CO	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	A	.	.
B1-3 (HM165208)	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	A	T	.
B1-5 (HM165216)	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	A	.	.
BG-Ber6 (DQ090527)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.
BG-Ber1 (DQ090522)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
BG-Ber2 (DQ090523)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
10/H43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	Y	.	.	.	Y	.	.	.	.	.	.	.
11/35	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
11/44	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	A	Y	.
11/59	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	.

12/C6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	.
12/C17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	.	.	R	.	.	.	.	Y	R	.	.
12/C24	.	.	.	.	R	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	A	T	.
12/C27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	R	.	.
12/C30	.	.	.	R	.	R	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	R	.	A	.	.	.	.	.	.	.	.	Y	A	T	.
12/C34	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	Y	Y	.	.	R
12/C39	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13/C114	.	.	.	.	.	.	Y	.	.	C	T	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13/C119	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.
13/C124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	Y	.	Y	.	.	.
13/C125	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	Y	.	.	.
13/C129	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	R	.	.	R	Y	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.
13/C148	.	.	.	.	.	.	.	Y	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.
13/CL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	R	.	.	.	.	.	R	.	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; dots indicate identity to the AI reference sequence; IUPAC nucleotide ambiguity codes: R=AG, Y=CT, K=GT

Table 4G: List of the 146 isolates successfully sequenced at the *gdh* gene ordered by sub-assemblage and genotype.

Sub-assemblage	Subtype	Identity (ID) to closest reference or previously described isolate (GenBank Acc. No.) *(no. of differences)*	Host(s) or environmental sample of isolation of matching sequences in GenBank	No. isolates	Isolates ID *(sequenced in only one direction)*
AII	**AII**	100%ID Reference AII (AY178737)	Human, cattle, dolphin, porpoise, gull	27	12/C91, 11/34, 10/8, 10/H21, 10/H51, 11/12, 11/37, 11/8, 12/C105, 12/C14, 12/C15, 12/C32, 12/C71, 12/C88, 12/C89, 12/C94, 13/C120, 13/C121, 13/C127, 13/C133, 13/C135, 13/C161, 13/CT, 08/22, 09/16 *(11/43, 12/C46)*
	ECUST2196	100%ID ECUST2196 isolate (JX994237) 99%ID Reference AII *(1)*	Human	18	08/H3, 11/27, 11/29, 11/67, 12/C101, 12/C16, 12/C65, 12/C68, 12/C70, 12/C74, 12/G, 13/C117, 13/C122, 13/C123, 13/C134, 13/C162, 13/DA, 09/10
	ECUST2196a*	99%ID ECUST2196 isolate *(1)*		1	12/C41
	ISSGd198	100%ID ISSGd198 isolate (EU278608)	Human	3	10/H8, 12/C31, 13/C164
A heterogeneous**	**n/a**			2	12/C56, 08/11
BIV	**BIV**	100%ID Reference BIV (AY178738)	Human, chinchilla	45	13/CO, 10/24, 10/H43, 10/H6, 11/10, 11/24, 11/58, 11/71, 12/C1, 12/C106, 12/C113, 12/C20, 12/C21, 12/C24, 12/C25, 12/C26, 12/C3, 12/C35, 12/C37, 12/C43, 12/C44, 12/C45, 12/C59, 12/C7, 12/C72, 12/C73, 12/C75, 12/C79, 12/C84, 12/C85, 12/C87, 12/C9, 12/C92, 12/C93, 12/C96, 12/E, 13/C115, 13/C130, 13/C131, 13/C132, 13/C147, 13/CF, 13/CP, 08/20, 08/8
Other B	**ISSGd167**	100%ID ISSGd167 isolate (EU637587)	Human	13	08/25, 11/35, 11/2, 11/28, 11/66, 12/C83, 12/C95, 13/C152, 13/C160,

					13/C165, 13/CE, 13/DH *(12/C80)*
	Ad-158	100% ID Ad-158 isolate (AY178753)	Marmoset	2	13/C145, 13/C149
	Ba*	99%ID Ad-158 isolate *(2)* 99%ID Reference BIV *(3)*		1	13/CM
	Bb*	*99%ID B0 isolate (AY178738) (2)* *99%ID Reference BIV (2)*		*1*	*(13/C155)*
	Bc*	99%ID B4 isolate (EF507682) *(2)* 99%ID Reference BIV *(3)*		1	10/28
	Bd*	99%ID B isolate (AY178756) *(2)* 99%ID Reference BIII *(3)*		1	12/C77
	Be*	99%ID B isolate (EU637588) *(5)* 99%ID Reference BIII *(6)*		1	11/38
	Bf*	99%ID B isolate (EU637588) *(2)* 99%ID Reference BIII *(7)*		1	13/C118
	Bg*	99% B isolate (EF507654) *(4)* 99%ID Reference BIII *(6)*		1	12/C19
	Bh*	99% ID B isolate (EU362955) *(5)* 99%ID Reference BIII *(5)*		1	12/C53
	Bi*	99%ID B isolate (HM134212) *(5)* 99%ID Reference BIV *(6)*		1	12/C11
	Bl*	99%ID B isolate (HM134216) *(7)* 99%ID Reference BIV *(8)*		1	11/36
B heterogeneous**	**n/a**	/		25	13/C154, 13/CL, 13/C148, 13/C140, 13/C125, 13/C124, 12/C81, 12/C39, 12/C30, 12/C6, 12/C2, 11/69, 11/44, 11/3, 12/C17, 12/C99, 13/C151, 13/CG, 12/C18, 13/CN, 13/C126, 11/1, 13/C157, 13/C119, 13/CC

*sequence with at least one novel nucleotide polymorphism; **sequence with at least one heterogeneous position

Table 4H: Nucleotide polymorphisms at the *gdh* locus in the assemblage A isolates successfully sequenced in both directions. The sequence of one representative isolate from each genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	554	557	560	567	575	584	599	602	608	638	662	683	716	746	749	750	770	776	782	784	794	797	833	836	842	848	866	899	905
Reference AI (M84604)	g	c	g	g	g	g	a	a	a	g	g	g	g	g	a	a	a	g	a	c	a	g	a	g	g	g	a	g	t
Reference AIII (EU637582)	A	T	A	.	A	A	G	G	.	.	.	A	.	A	G	.	G	A	G	.	G	A	G	A	.	.	G	A	.
Reference AII (AY178737)	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	G	.	.	.	.	.	.	.	.	A	G	.	.
12/C91	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	G	.	.	.	.	.	.	.	.	A	G	.	.
ECUST2196 (JX994237)	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.
08/H3	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.
12/C41	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	.	.	.	.	.	.	.	.	T*	A	G	.	.
ISSGd198 (EU278608)	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	A	G	.	.
10/H8	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	A	G	.	.
08/11	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	.	G	.	.	.	.	.	.	.	.	A	G	.	Y
12/C56	.	.	.	A	A	.	G	G	G	A	A	.	A	.	.	W	.	.	.	.	.	.	.	.	.	A	G	.	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the AI reference sequence; IUPAC nucleotide ambiguity codes: Y=CT, W=AT

Table 4I: Nucleotide polymorphisms at the *gdh* locus in assemblage B isolates successfully sequenced in both directions and without heterogeneous positions. The sequence of one representative isolate from each identified genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	572	578	602	635	644	662	676	713	731	737	746	764	770	779	794	803	842	845	857	863	872	899	908	917	923	929	935	950	983	989
Reference BIII (EF685684)	g	a	g	g	t	c	t	a	c	c	g	a	a	c	g	a	c	t	c	g	g	g	g	c	g	g	t	g	a	g
Reference BIV (AY178738)	.	.	.	.	C	A	.	.	.	.	A	.	G	.	.	.	.	.	T	.	.	.	A	.	.	A	.	.	.	.
13/CO	.	.	.	.	C	A	.	.	.	.	A	.	G	.	.	.	.	.	T	.	.	.	A	.	.	A	.	.	.	.
ISSGd167 (EU637587)	.	.	.	.	C	A	.	G	.	.	A	G	G	T	.	.	.	.	T	.	A	A	.	.	.	A	.	.	.	.
08/25	.	.	.	.	C	A	.	G	.	.	A	G	G	T	.	.	.	.	T	.	A	A	.	.	.	A	.	.	.	.
Ad-158 (AY178753)	.	.	.	.	C	A	.	G	.	.	A	.	G	.	.	.	.	.	T	.	A	.	.	.	.	A	.	.	.	.
13/C145	.	.	.	.	C	A	.	G	.	.	A	.	G	.	.	.	.	.	T	.	A	.	.	.	.	A	.	.	.	.
13/CM	.	.	.	.	C	A	.	.	.	.	A	.	G	.	.	.	.	C*	T	.	A	.	.	.	.	A	.	.	.	.
10/28	.	.	.	.	C	A	.	.	.	.	A	.	G	.	.	.	.	.	T	A*	.	.	.	.	.	.	.	.	.	.
12/C77	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A*	.	.
11/36	.	G*	.	.	C	A	C*	.	.	.	A	.	G	.	A*	.	.	.	.	.	.	.	.	T*	.	A	.	.	.	A*
11/38	.	.	A*	A*	.	A	.	.	.	.	.	.	G	.	.	.	T*	.	.	.	.	.	.	.	.	A	.	.	.	.
12/C11	A*	.	.	.	C	A	.	.	.	.	.	.	G	.	.	.	.	.	T	.	.	.	A	.	.	.	C*	.	.	.
12/C19	.	.	.	.	.	A	.	.	T*	.	.	.	.	.	.	G*	.	.	.	.	.	.	.	.	A*	.	.	A*	.	.
12/C53	.	.	.	.	.	.	.	.	.	T*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A*	.	.	A*	.	.
13/C118	.	.	.	A*	.	A	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A*	.	.	A*	G*	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the BIII reference sequence.

Table 4L: Nucleotide polymorphisms at the *gdh* locus in assemblage B isolates successfully sequenced in both directions and showing heterogeneous positions. The sequence without heterogeneous positions of one representative isolate from each genotype is also shown.

Isolate (GenBank Acc.No.)	572	578	602	635	641	644	653	662	676	685	701	704	713	731	737	740	746	764	770	779	803	815	827	833	842	845	857	863	872	882	896	899	907	908	917	923	929	950	983	989	1001	1004
BIII (EF685684)	g	a	g	g	c	t	g	c	t	t	g	g	a	c	c	g	g	a	a	c	a	c	g	a	c	t	c	g	g	a	g	g	t	g	c	g	g	g	a	g	c	g
BIV (AY178738)	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.
ISSGd167	.	.	.	.	.	C	.	A	.	.	.	.	G	.	.	.	A	G	G	T	.	.	.	.	.	.	T	.	A	.	.	A	.	.	.	.	A	.	.	.	.	.
Ad-158	.	.	.	.	.	C	.	A	.	.	.	.	G	.	.	.	A	.	G	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	A	.	.	.	.	.
13/CM	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	C	T	.	A	.	.	.	.	.	.	.	A	.	.	.	.	.
10/28	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12/C77	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
11/36	.	G	.	.	.	C	.	A	C	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	A	.	.
11/38	.	.	A	A	.	.	.	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
12/C11	A	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
12/C19	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.
12/C53	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.
13/C118	.	.	.	A	.	.	.	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	G	.	.	.
11/1	.	.	.	.	Y	Y	.	M	.	.	.	.	.	.	.	R	R	.	G	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	A	.	.	R	Y	.
11/3	.	.	.	.	.	C	.	A	.	.	.	K	G	.	.	.	A	G	G	T	.	.	.	.	.	.	T	.	A	.	.	A	.	.	.	.	A	.	.	.	.	.
11/44	.	.	.	.	.	C	.	A	.	.	.	.	R	.	.	.	A	R	G	Y	.	.	.	.	.	.	T	.	R	.	.	R	.	R	.	.	A	.	.	.	.	.
11/69	.	.	.	R	.	Y	.	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
12/C2	.	.	.	.	.	C	.	A	.	.	.	.	R	.	.	.	A	R	G	T	.	.	.	.	.	.	T	.	R	.	.	R	.	R	.	.	A	.	.	.	.	.
12/C6	.	.	.	.	.	C	.	A	.	.	.	.	R	.	.	.	A	R	G	Y	.	.	.	.	.	.	T	.	R	.	.	R	.	R	.	.	A	.	.	.	.	.
12/C17	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	Y	Y	T	R	R	.	.	.	.	R	.	.	R	A	.	.	Y	.

12/C18	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	R	A	.	.	A	.	.	.	.	.
12/C30	.	.	.	.	.	C	.	A	.	.	.	.	R	.	.	.	A	R	G	Y	.	.	.	.	.	.	T	.	R	.	.	R	.	R	.	.	A	.	.	.	.	.
12/C39	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	R	.	.	.	.	.	.	.	.	R	A	.	.	.	.
12/C81	R	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.
12/C99	.	.	R	A	.	.	.	A	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	R	.	.	.
13/C119	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	R	.	.	.	Y	.	R	.	.	.	.	.	.	R	.	R	.	.	.	R
13/C124	.	.	.	.	.	Y	.	M	.	.	.	.	.	.	.	R	R	.	R	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13/C125	.	.	.	R	.	Y	.	A	.	.	.	.	.	.	.	.	R	.	G	.	.	.	.	.	Y	.	Y	.	.	.	.	.	.	.	.	R	R	A	R	.	Y	.
13/C126	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	T	.	.	R	.	.	.	A	.	.	A	.	.	.	.	.
13/C140	.	.	R	A	.	.	.	A	.	.	.	.	.	Y	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	R	.	.	.
13/C148	.	.	.	.	.	Y	.	A	.	.	.	.	.	.	.	.	R	.	G	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	R	R	.	.	.	.
13/C151	.	.	.	A	.	.	.	A	.	.	R	.	.	.	Y	.	R	.	G	.	.	.	.	R	.	.	.	.	.	.	R	.	.	.	.	A	.	A	R	.	.	.
13/C154	.	.	.	R	.	Y	.	A	.	.	.	.	.	.	.	.	R	.	G	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	R	.	R	R	R	G	.	.	.
13/C157	.	R	.	R	.	Y	.	A	.	.	.	.	.	.	.	.	R	.	G	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	Y	.	R	R	.	.	.	.
13/CC	.	.	.	.	.	Y	.	M	.	Y	.	.	.	.	.	.	.	.	R	.	.	Y	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.
13/CG	.	.	.	.	.	C	.	A	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	A	.	.	.	.
13/CL	.	.	.	.	.	.	.	M	.	.	.	.	.	.	.	.	R	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.
13/CN	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.

Nucleotide polymorphisms (in capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the BIII reference sequence. IUPAC nucleotide ambiguity codes: R=AG, Y=CT, K=GT, M=AC

Table 4M: List of the 163 isolates successfully sequenced at the *tpi* gene ordered by sub-assemblage and genotype

Sub-assemblage	Subtype	Identity (ID) to closest reference or previously described isolate (GenBank Acc. No.) *(no. of differences)*	Host(s) or environmental sample of isolation of matching sequences in GenBank	No. isolates	Isolates ID *(sequenced in only one direction)*
AII	**AII**	100%ID Reference AII (U57897)	Human, cattle, sheep, dog, spider monkey, seal, dolphin, gull	51	12/C89, 13/CT, 13/DA, 13/C164, 13/C162, 13/C161, 13/C135, 13/C133, 13/C127, 13/C123, 13/C122, 13/C117, 12/G, 12/C105, 12/C88, 12/C74, 12/C70, 12/C68, 12/C65, 12/C56, 12/C32, 12/C31, 12/C15, 11/67, 11/43, 11/29, 11/8, 10/H51, 10/H21, 10/H19, 10/H8, 08/22, 08/H3, 09/10, 09/16, 09/22, 10/8, 10/20, 12/C16 *(11/12, 11/27, 12/C14, 12/C41, 12/C91, 12/C94)*. Short product: 12/C46, 12/C57, 12/C71, 12/C97 *(11/37, 13/C120)*
AI	**Martha**	100%ID Martha isolate (KJ776589) 99%ID Reference AI (L02120)	Human, dog, cat, sheep, alpaca	1	11/13
BIII	**BIII**	100%ID Reference BIII (AF069561)	Human, dog, macaque, seal, wastewater	11	12/C23, 12/C99, 13/C119, 13/C151, 13/C154, 13/DB, 13/CN *(11/30)*. Short product: 11/70, 12/C19, 13/CB
	HS29	100%ID HS29 isolate (KC632557) 99%ID Reference BIII *(1)*	Human	1	Short product 13/C163
	Sweh060	100%ID Sweh060 isolate (HM140714) 99%ID Reference BIII *(2)*	Human	1	Short product 13/C137
	HS114	100% ID HS114 isolate (KC632558) 99%ID Reference BIII *(1)*	Human	1	Short product 11/15
	BIIIa*	99%ID Reference BIII *(1)*		1	10/H43
	BIIIb*	99%ID Reference BIII *(1)*		1	10/H35
BIV	**BIV**	100%ID Reference BIV (AF069560)	Human, dog, seal, dolphin, gull, eider, urban wastewater	50	13/C160, 13/CF, 13/C155, 13/C147, 13/C131, 13/C130, 13/C126, 13/C115, 12/C113, 12/D, 12/C106, 12/C95, 12/C83, 12/C82, 12/C79, 12/C72, 12/C59, 12/C45, 12/C44, 12/C43, 12/C37, 12/C26, 12/C25,

					12/C20, 12/C7, 12/C6, 12/C3, 11/71, 11/34, 11/24, 11/10, 11/3, 10/H20, 10/H18, 10/H9, 10/H6, 09/11, 08/8, 13/CO, 08/25, 13/CP *(10/H44, 13/CS, 8/20).* Short product: 12/C63, 12/C80, 12/C61, 11/44, 11/28, 10/3
	VB906855	100%ID VB906855 isolate (KF843920) 99%ID Reference BIV *(1)*	Human, macaque, dog, beaver, gazelle, seal, dolphin, gull, tamandua, urban wastewater	13	10/H25, 12/A, 12/C9, 12/C84, 12/C87, 12/C92, 12/C96, 13/C145, 13/C152, 13/C165, 13/DH *(13/CE).* Short product: 10/28
	BIVa*	99%ID Reference BIV *(1)*		1	12/C18
	BIVb*	99%ID Reference BIV *(1)*		1	Short product 11/25
	BIVc*	*99%ID Reference BIV (1)*		*1*	*(12/C50)*
Other B	**Sweh171**	100%ID Sweh171 isolate (HM140722) 99%ID Reference BIII *(3)*	Human	2	10/H14, 13/C149
	Ba*	99%ID RM1 isolate (GU564279) *(2)* 99%ID Reference BIV *(6)*		1	12/C64
	Bb*	99%ID Sweh171 isolate (HM140722) *(3)* 99%ID Ref. BIII *(4)*		1	13/CG
B heterogeneous**	**n/a**			25	13/C125, 13/C140, 13/C124, 12/C77, 12/C35, 12/C30, 12/C27, 12/C21, 12/C1, 11/1, 13/C148, 11/36, 11/69, 13/CL, 13/CM, 12/C17, 13/C157, 11/35 *(12/C11, 12/C2, 13/CC).* Short product: 12/C108, 12/C85, 13/CD, 8/11

*sequence with at least one novel nucleotide polymorphism; **sequence with at least one heterogeneous position

Table 4N: Nucleotide polymorphisms at the *tpi* locus in the assemblage A isolates successfully sequenced in both directions. The sequence of one representative isolate from each genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	658	663	705	820	826	868	883	895	913
AI (L02120)	g	a	g	t	g	t	t	t	c
Martha (KJ776589)	.	.	.	C	.	.	.	.	.
11/13	.	.	.	C	.	.	.	.	.
AII (U57897)	A	.	.	.	.	.	.	.	.
12/C89	A	.	.	.	.	.	.	.	.
AIII (DQ650648)	.	G	A	.	A	C	C	C	T

Nucleotide polymorphisms (capitals) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the AI reference sequence

Table 4O: Nucleotide polymorphisms at the *tpi* locus in assemblage B isolates successfully sequenced in both directions without heterogeneous positions. The sequence of one representative isolate from each genotype is shown in comparison with the relative most similar reference or isolate retrieved in GenBank.

Isolate (GenBank Acc.No.)	574	586	616	628	646	655	664	750	753	777	802	847	859	889	892	895	913
BIII (AF069561)	t	t	c	c	g	t	g	t	t	t	c	c	g	g	g	c	a
12/C23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10/H35	.	.	T*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10/H43	.	.	.	.	.	.	.	.	.	.	T*	.	.	.	.	.	.
HS29 (KC632557)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
13/C163	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Sweh060 (HM140714)	.	.	.	.	.	.	A	.	C	.	.	.	.	.	.	.	.
13/C137	.	.	.	.	.	.	A	.	C	.	.	.	.	.	.	.	.
HS114 (KC632558)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
11/15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
BIV (AF069560)	.	.	.	T	.	.	.	.	.	.	.	T	.	A	A	.	.
13/C160	.	.	.	T	.	.	.	.	.	.	.	T	.	A	A	.	.
VB906855 (KF843920)	.	.	.	.	.	.	.	.	.	.	.	T	.	A	A	.	.
10/H25	.	.	.	.	.	.	.	.	.	.	.	T	.	A	A	.	.
12/C18	.	.	.	T	A*	.	.	.	.	.	.	T	.	A	A	.	.
11/25	.	.	.	T	.	.	.	.	.	.	.	T	.	A	A	.	T*
Sweh171 (HM140722)	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.
10/H14	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.
RM1 (GU564279)	C	.	.	.	.	C	.	.	.	.	.	.	.	A	A	.	.
12/C64	C	G*	.	.	.	C	.	.	.	.	T*	.	.	A	A	.	.
13/CG	.	.	.	.	.	.	.	C*	.	C*	.	.	A	.	A	.	.

Nucleotide polymorphisms (capital letters) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the BIII reference sequence.

Table 4P: Nucleotide polymorphisms at the *tpi* locus in assemblage B isolates successfully sequenced in both directions and showing heterogeneous positions. The sequence without heterogeneous positions of one representative isolate from each genotype is also shown.

Isolate (GenBank Acc.No.)	574	586	615	616	628	629	634	646	655	664	700	714	745	750	753	760	777	790	802	821	843	847	852	859	868	887	889	890	892	895	913
BIII (AF069561)	t	t	g	c	c	a	c	g	t	g	g	t	g	t	t	t	t	c	c	g	t	c	g	g	t	g	g	g	g	c	a
10/H35	.	.	.	T*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10/H43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T*	.	.	.	.	.	.	.	.	.	.	.	.
13/C163	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Sweh060 (HM140714)	.	.	.	.	.	.	.	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HS114 (KC632558)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
BIV (AF069560)	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	A	.	.
VB906855 (KF843920)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	A	.	.
12/C18	.	.	.	.	T	.	.	A*	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	A	.	.
11/25	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	A	.	T*
Sweh171 (HM140722)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	A	.	.
RM1 (GU564279)	C	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.
12/C64	C	G*	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T*	.	.	.	.	.	.	.	A	.	A	.	.
13/CG	.	.	.	.	.	.	.	.	.	.	.	.	.	C*	.	.	C*	.	.	.	.	.	.	A	.	.	.	.	A	.	.
11/1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.
11/35	.	.	T	.	.	.	.	.	.	.	A	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11/36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	Y	.	R	.	R	.	.
11/69	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12/C1	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	R	.	.
12/C17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	Y	.
12/C21	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.
12/C27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	Y	.

12/C30	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.
12/C35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	A	.	A	.	.
12/C77	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.
13/C124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	A	.	.
13/C125	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	R	.	R	Y	.
13/C140	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	R	.	R	.	.
13/C148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	A	.	A	.	.
13/C157	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.
13/CL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	A	.	.
13/CM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	Y	.
08/11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.
12/C85	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	T	.	.	.	.	A	.	A	.	.
12/C108	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	A	A	.	.	A	.	A	.	.
13/CD	.	.	.	.	.	.	.	.	.	.	.	W	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Nucleotide polymorphisms (capital letters) are numbered from the beginning of the AI reference sequence; novel polymorphisms are indicate with an asterisk; dots indicate identity to the BIII reference sequence. IUPAC nucleotide ambiguity codes: R=AG, Y=CT, W=AT, S=GC

APPENDIX 5: Univariable comparisons between assemblage A and B cases

Table 5A: List of exposure variables that returned a $p>0.2$ in the comparison between cases infected with assemblage A (n=39) and B (n=69) that returned the enhanced surveillance study questionnaire. Odds ratios were calculated using assemblage B as the baseline category.

			Assemblage n (%)**			
Variable	**No. valid (% missing)***	**Category**	**A**	**B**	**OR (95% CI)*****	***p*-value**
Travelling abroad (outside the UK)	105 (2.8)	No	25 (65.8)	47 (70.1)	Ref.	0.644
		Yes	13 (34.2)	20 (29.9)	1.22 (0.52-2.86)	
No. glasses of un-boiled tap water per day	94 (13)	0	3 (9.1)	5 (8.2)	Ref.	0.674
		1-2	13 (39.4)	17 (27.9)	1.27 (0.26-6.33)	
		3-4	8 (24.2)	20 (32.8)	0.67 (0.13-3.47)	
		5+	9 (27.3)	19 (31.1)	0.79 (0.15-4.05)	
Drinking from the mains water supply	103 (4.6)	No	7 (18.9)	12 (18.2)	Ref.	0.926
		Yes	30 (81.1)	54 (81.8)	0.95 (0.34-2.68)	
Drinking from a private water supply	103 (4.6)	No	36 (97.3)	66 (100)	Ref.	0.359
		Yes	1 (2.7)	0 (0)	n/a****	
Drinking bottled water	103 (4.6)	No	17 (45.9)	26 (39.4)	Ref.	0.518
		Yes	20 (54.1)	40 (60.6)	0.76 (0.34-1.72)	
Drinking un-boiled water from a lake, river or stream	103 (4.6)	No	37 (100)	64 (97)	Ref.	0.535
		Yes	0 (0)	2 (3)	n/a****	
Swimming or paddling in a swimming pool	108	No	23 (59)	48 (69.6)	Ref.	0.265
		Yes	16 (41)	21 (30.4)	1.59 (0.70-3.60)	
Frequency of swimming or paddling in a swimming pool (no. times)	96 (11.1)	0	23 (71.9)	48 (75)	Ref.	$0.371^{CT}/0.548^{C}$
		1-2	6 (18.8)	9 (14.1)	1.39 (0.44-4.38)	
		3-4	3 (9.4)	3 (4.7)	2.09 (0.39-11.15)	
		5-6	0 (0)	2 (3.1)	n/a****	
		7+	0 (0)	2 (3.1)	n/a****	
Swimming in a lake, pond or river	108	No	35 (89.7)	64 (92.8)	Ref.	0.720
		Yes	4 (10.3)	5 (7.2)	1.46 (0.37-5.80)	
Practising watersports in freshwater	108	No	38 (97.4)	68 (98.6)	Ref.	1.000

		Yes	1 (2.6)	1 (1.4)	1.79 (0.11-29.43)	
Practising watersports in the sea	108	No	35 (89.7)	66 (95.7)	Ref.	0.250
		Yes	4 (10.3)	3 (4.3)	2.51 (0.53-11.87)	
Fishing	108	No	37 (94.9)	68 (98.6)	Ref.	0.295
		Yes	2 (5.1)	1 (1.4)	3.68 (0.32-41.90)	
Walking in the countryside	108	No	29 (74.4)	44 (63.8)	Ref.	0.259
		Yes	10 (25.6)	25 (36.2)	0.61 (0.25-1.45)	
Picnicking	108	No	37 (94.9)	63 (91.3)	Ref.	0.708
		Yes	2 (5.1)	6 (8.7)	0.57 (0.11-2.96)	
Went camping	108	No	38 (97.4)	67 (97.1)	Ref.	1.000
		Yes	1 (2.6)	2 (2.9)	0.88 (0.08-10.05)	
Keeping a pet	102 (5.5)	No	22 (61.1)	32 (48.5)	Ref.	0.300
		Yes	14 (38.9)	34 (51.5)	0.60 (0.26-1.37)	
Keeping dogs	102 (5.5)	No	27 (75)	49 (74.2)	Ref.	0.933
		Yes	9 (25)	17 (25.8)	0.96 (0.38-2.45)	
Keeping cats	102 (5.5)	No	30 (83.3)	50 (75.8)	Ref.	0.374
		Yes	6 (16.7)	16 (24.2)	0.62 (0.22-1.77)	
Keeping rabbits	102 (5.5)	No	35 (97.2)	61 (92.4)	Ref.	0.420
		Yes	1 (2.8)	5 (7.6)	0.35 (0.04-3.10)	
Keeping birds	102 (5.5)	No	34 (94.4)	60 (90.9)	Ref.	0.709
		Yes	2 (5.6)	6 (9.1)	0.59 (0.11-3.08)	
Keeping horses	102 (5.5)	No	36 (100)	65 (98.5)	Ref.	1.000
		Yes	0 (0)	1 (1.5)	n/a****	
No. of different pet species owned	102 (5.5)	0	22 (61.1)	32 (48.5)	Ref.	$0.337^{CT}/0.135^{C}$
		1	10 (27.8)	23 (34.8)	0.63 (0.25-1.59)	
		2	4 (11.1)	7 (10.6)	0.83 (0.22-3.18)	
		3+	0 (0)	4 (6.1)	n/a****	
Visited a premise with animals	92 (14.8)	No	20 (62.5)	45 (75)	Ref.	0.210
		Yes	12 (37.5)	15 (25)	1.80 (0.71-4.53)	
Visited a farm, stable or horse riding schools	92 (14.8)	No	29 (90.6)	54 (90)	Ref.	1.000
		Yes	3 (9.4)	6 (10)	0.93 (0.22-3.40)	
Visited a wildlife park or zoo	92 (14.8)	No	31 (96.9)	59 (98.3)	Ref.	1.000
		Yes	1 (3.1)	1 (1.7)	1.90 (0.11-31.48)	
Touching cattle while on the premise	91 (15.7)	No	29 (90.6)	55 (93.2)	Ref.	0.657

		Yes	3 (9.4)	4 (6.8)	1.42 (0.30-6.79)	
Touching pigs while on the premise	91 (15.7)	No	32 (100)	58 (98.3)	Ref.	1.000
		Yes	0 (0)	1 (1.7)	n/a****	
Touching sheep while on the premise	91 (15.7)	No	31 (96.9)	56 (94.9)	Ref.	1.000
		Yes	1 (3.1)	3 (5.1)	0.60 (0.06-6.04)	
Touching horses while on the premise	91 (15.7)	No	31 (96.9)	53 (89.8)	Ref.	0.414
		Yes	1 (3.1)	6 (10.2)	0.28 (0.03-2.48)	
Touching goats while on the premise	91 (15.7)	No	32 (100)	58 (98.3)	Ref.	1.000
		Yes	0 (0)	1 (1.7)	n/a****	
Touching chickens while on the premise	91 (15.7)	No	31 (96.9)	58 (98.3)	Ref.	1.000
		Yes	1 (3.1)	1 (1.7)	1.87 (0.11-30.95)	
Touching deers while on the premise	91 (15.7)	No	32 (100)	58 (98.3)	Ref.	1.000
		Yes	0 (0)	1 (1.7)	n/a****	
Eating any fruit	97 (10.2)	No	4 (11.8)	3 (4.8)	Ref.	0.236
		Yes	30 (88.2)	60 (95.2)	0.37 (0.08-1.78)	
Eating any salad items	96 (11.1)	No	5 (14.7)	12 (19.4)	Ref.	0.568
		Yes	29 (85.3)	50 (80.6)	1.39 (0.45-4.35)	
No. of adults (≥16y) in the house	104 (3.7)	1	12 (32.4)	14 (20.9)	Ref.	$0.493^{CT}/0.603^{C}$
		2	18 (48.6)	40 (59.7)	0.52 (0.20-1.36)	
		3	3 (8.1)	8 (11.9)	0.44 (0.09-2.03)	
		4+	4 (10.8)	5 (7.5)	0.93 (0.20-4.28)	
Changing nappies	98 (9.2)	No children in the house or children not in nappies	30 (85.7)	50 (79.4)	Ref.	0.522
		Child in nappies but not changing nappies	3 (8.6)	5 (7.9)	1.00 (0.22-4.49)	
		Child in nappies and changing nappies	2 (5.7)	8 (12.7)	0.42 (0.08-2.09)	
Any other person with diarrhoea in the house	99 (8.3)	No	32 (91.4)	56 (87.5)	Ref.	0.742
		Yes	3 (8.6)	8 (12.5)	0.66 (0.16-2.65)	
Any contact with a person with diarrhoea outside the house	92 (14.8)	No	29 (85.3)	53 (91.4)	Ref.	0.490
		Yes	5 (14.7)	5 (8.6)	1.83 (0.49-6.84)	

*percentage refers to the proportion of participants with missing information for the variable; **percentages refer to the proportion among the cases that answered the question; ***Odds ratio with 95% confidence interval; **** Odds ratio not calculated because no variation present in the data; ^{C}Logistic regression on the continuous variable; ^{CT} Logistic regression on the categorized variable.

Table 5B: List of variables that returned a *p*-value>0.2 in univariable analysis of *Giardia* assemblage A cases (n=28), assemblage B cases (n=57) and healthy controls (n=226) that returned the case-control study questionnaire.

Variable	Data subset	No. valid (% missing)*	Category	Cases n (%)**	Controls n (%)**	OR (95% CI)***	*p*-value
SOCIODEMOGRAPHICS AND AREA AND SEASON VARIABLES							
Gender	Assemblage A	254	Female	11 (39.3)	110 (48.7)	Ref.	0.348
			Male	17 (60.7)	116 (51.3)	1.47 (0.66-3.27)	
	Assemblage B	283	Female	25 (43.9)	110 (48.7)	Ref.	0.516
			Male	32 (56.1)	116 (51.3)	1.21 (0.68-2.18)	
Ethnicity	Assemblage A	251 (1.2)	Non-white *(e.g. Asian, Chinese, Black African or Caribbean or other)*	2 (7.1)	21 (9.4)	Ref.	1.000
			White	26 (92.9)	202 (90.6)	1.35 (0.30-6.10)	
	Assemblage B	279 (1.4)	Non-white *(e.g. Asian, Chinese, Black African or Caribbean or other)*	6 (10.7)	21 (9.4)	Ref.	0.769
			White	50 (89.3)	202 (90.6)	0.87 (0.33-2.26)	
Level of neighbourhood deprivation (IMD rank)	Assemblage A	253 (0.4)	Very low (24,363-32,482)	9 (32.1)	58 (25.8)	Ref.	$0.598^{CT}/0.212^{C}$
			Low (16,242-24,362)	9 (32.1)	58 (25.8)	1.00 (0.37-2.70)	
			Moderate (8,121-16,241)	3 (10.7)	41 (18.2)	0.47 (0.12-1.85)	
			Very high (1-8,120)	7 (25)	68 (30.2)	0.66 (0.23-1.89)	
	Assemblage B	282 (1)	Very low (24,363-32,482)	12 (21.1)	58 (25.8)	Ref.	$0.551^{CT}/0.794^{C}$
			Low (16,242-24,362)	17 (29.8)	58 (25.8)	1.42 (0.52-3.23)	
			Moderate (8,121-16,241)	14 (24.6)	41 (18.2)	1.65 (0.69-3.93)	
			Very high (1-8,120)	14 (24.6)	68 (30.2)	0.99 (0.43-2.32)	
Reported area of living	Assemblage A	249 (2)	City	5 (19.2)	42 (18.8)	Ref.	0.998
			Town	12 (46.2)	104 (46.6)	0.97 (0.32-2.92)	
			Village	9 (34.6)	77 (34.5)	0.98 (0.31-3.12)	
	Assemblage B	279 (1.4)	City	11 (19.6)	42 (18.8)	Ref.	0.363
			Town	31 (55.4)	104 (46.6)	1.14 (0.52-2.47)	
			Village	14 (25)	77 (34.5)	0.69 (0.29-1.66)	
Season	Assemblage A	254	Winter (December-February)	4 (14.3)	48 (21.2)	Ref.	0.593
			Spring (March-May)	10 (35.7)	70 (31)	1.71 (0.51-5.78)	
			Summer (June-August)	8 (28.6)	46 (20.4)	2.08 (0.59-7.41)	
			Autumn (September-November)	6 (21.4)	62 (27.4)	1.16 (0.31-4.35)	

	Assemblage B	283	Winter (December-February)	15 (26.3)	48 (21.2)	Ref.	0.790
			Spring (March-May)	18 (31.6)	70 (31)	0.82 (0.38-1.79)	
			Summer (June-August)	9 (15.8)	46 (20.4)	0.63 (0.25-1.57)	
			Autumn (September-November)	15 (26.3)	62 (27.4)	0.77 (0.34-1.74)	
HEALTH DETAILS							
Suffering from irritable bowel syndrome (IBS)	Assemblage A	238 (6.3)	No	23 (92)	199 (93.4)	Ref.	0.678
			Yes	2 (8)	14 (6.6)	0.81 (0.17-3.79)	
	Assemblage B	264 (6.7)	No	47 (92.2)	199 (93.4)	Ref.	0.758
			Yes	4 (7.8)	14 (6.6)	1.21 (0.38-3.84)	
OCCUPATIONAL DETAILS							
Having a potentially at-risk occupation	Assemblage A	253 (0.4)	Not currently working	21 (75)	129 (57.3)	Ref.	0.451
			Working but not with any at-risk occupation	6 (21.4)	83 (36.9)	0.44 (0.17-1.15)	
			Working with animals	0 (0)	1 (0.4)	n/a***	
			Working with manure/slurry	0 (0)	1 (0.4)	n/a***	
			Working with human faeces	0 (0)	7 (3.1)	n/a***	
			Contact with sewage	0 (0)	1 (0.4)	n/a***	
			Contact with freshwaters	0 (0)	1 (0.4)	n/a***	
			More than one at-risk occupation	1 (3.6)	2 (0.9)	3.07 (0.27-35.39)	
	Assemblage B	281 (0.7)	Not currently working	27 (48.2)	129 (57.3)	Ref.	0.652
			Working but not with any at-risk occupation	25 (44.6)	83 (36.9)	1.44 (0.78-2.65)	
			Working with animals	0 (0)	1 (0.4)	n/a***	
			Working with manure/slurry	0 (0)	1 (0.4)	n/a***	
			Working with human faeces	3 (5.4)	7 (3.1)	2.05 (0.50-8.43)	
			Contact with sewage	0 (0)	1 (0.4)	n/a***	
			Contact with freshwaters	1 (1.8)	1 (0.4)	4.78 (0.29-78.78)	
			More than one at-risk occupation	0 (0)	2 (0.9)	n/a***	
ANIMAL CONTACT							
Visiting or working at a farm	Assemblage A	253 (0.4)	No	25 (92.6)	192 (85)	Ref.	0.389
			Yes	2 (7.4)	34 (15)	0.45 (0.10-2.00)	
	Assemblage B	280 (1.1)	No	49 (90.7)	192 (85)	Ref.	0.270
			Yes	5 (9.3)	34 (15)	0.58 (0.21-1.55)	
Touching animals while visiting or working at a farm	Assemblage B	279 (1.4)	Neither visiting nor working at a farm	49 (90.7)	192 (85.3)	Ref.	0.550
			Not touching animals while at a farm	2 (3.7)	12 (5.3)	0.65 (0.14-3.01)	
			Touching animals while at a farm	3 (5.6)	21 (9.3)	0.56 (0.16-1.95)	

Touching animals at a wildlife park or zoo	Assemblage B	280 (1.1)	Not visiting a wildlife park or zoo	50 (90.9)	217 (96.4)	Ref.	0.233
			Visiting but not touching any animal	3 (5.5)	6 (2.7)	2.17 (0.52-8.97)	
			Visiting and touching an animal	2 (3.6)	2 (0.9)	4.34 (0.60-31.56)	
RECREATIONAL ACTIVITIES							
Frequency of swimming or paddling in a swimming pool (no. times)	Assemblage A	242 (4.7)	0	15 (55.6)	177 (81.2)	Ref.	$0.340^{CT}/0.344^{C}$
			1-2	3 (12.5)	17 (7.8)	2.08 (0.55-7.92)	
			3-4	3 (12.5)	12 (5.5)	2.95 (0.75-11.62)	
			5-6	2 (8.3)	6 (2.8)	3.93 (0.73-21.21)	
			7+	1 (4.2)	6 (2.8)	1.97 (0.22-17.43)	
	Assemblage B	268 (5.3)	0	34 (68)	177 (81.2)	Ref.	$0.256^{CT}/0.669^{C}$
			1-2	9 (18)	17 (7.8)	2.76 (1.13-6.69)	
			3-4	4 (8)	12 (5.5)	1.73 (0.53-5.70)	
			5-6	1 (2)	6 (2.8)	0.87 (0.10-7.44)	
			7+	2 (4)	6 (2.8)	1.73 (0.34-8.96)	
Frequency of using a Jacuzzi or a hot tub (no. times)	Assemblage B	253 (10.6)	0	37 (90.2)	203 (95.8)	Ref.	$0.387^{CT}/0.258^{C}$
			1-2	3 (7.3)	6 (2.8)	2.74 (0.66-11.46)	
			3+	1 (2.4)	3 (1.4)	1.83 (0.18-18.06)	
Immersing the head underwater while using a Jacuzzi or hot tub	Assemblage B	250 (11.7)	Not using a Jacuzzi or a hot tub	37 (94.9)	203 (96.2)	Ref.	0.562
			Using a Jacuzzi or a hot tub without immersing the head	2 (5.1)	6 (2.8)	1.83 (0.35-9.41)	
			Using Jacuzzi or a hot tub immersing the head	0 (0)	2 (0.9)	n/a***	
Swimming in a lake, pond or river	Assemblage A	238 (6.3)	No	23 (95.8)	209 (97.7)	Ref.	0.475
			Yes	1 (4.2)	5 (2.3)	1.82 (0.20-16.23)	
	Assemblage B	266 (6)	No	52 (100)	209 (97.7)	Ref.	0.587
			Yes	0 (0)	5 (2.3)	n/a***	
Swimming in the sea	Assemblage A	243 (4.3)	No	22 (91.7)	206 (94.1)	Ref.	0.649
			Yes	2 (8.3)	13 (5.9)	1.44 (0.30-6.80)	
	Assemblage B	272 (3.9)	No	50 (94.3)	206 (94.1)	Ref.	1.000
			Yes	3 (5.7)	13 (5.9)	0.95 (0.26-3.46)	
Practising watersports in freshwater	Assemblage A	244 (3.9)	No	24 (100)	219 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
	Assemblage B	273 (3.5)	No	53 (100)	219 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	

Practising watersports in the sea	Assemblage A	244 (3.9)	No	24 (100)	219 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
	Assemblage B	273 (3.5)	No	53 (100)	219 (99.5)	Ref.	1.000
			Yes	0 (0)	1 (0.5)	n/a***	
Going fishing	Assemblage A	235 (7.5)	No	26 (100)	206 (98.6)	Ref.	1.000
			Yes	0 (0)	3 (1.4)	n/a***	
	Assemblage B	263 (7.1)	No	54 (100)	206 (98.6)	Ref.	1.000
			Yes	0 (0)	3 (1.4)	n/a***	
Going hunting		Exposure not reported in either cases or controls					
Going picnicking	Assemblage A	235 (7.5)	No	25 (96.2)	188 (90)	Ref.	0.482
			Yes	1 (3.8)	21 (10)	0.36 (0.05-2.78)	
	Assemblage B	262 (7.4)	No	49 (92.5)	188 (90)	Ref.	0.580
			Yes	4 (7.5)	21 (10)	0.73 (0.24-2.23)	
Going camping	Assemblage A	235 (7.5)	No	26 (100)	206 (98.6)	Ref.	1.000
			Yes	0 (0)	3 (1.4)	n/a***	
	Assemblage B	262 (7.4)	No	52 (98.1)	206 (98.6)	Ref.	1.000
			Yes	1 (1.9)	3 (1.4)	1.32 (0.13-12.96)	
Going caravanning	Assemblage A	235 (7.5)	No	24 (92.3)	199 (94.3)	Ref.	0.656
			Yes	2 (7.7)	12 (5.7)	1.38 (0.29-6.55)	
	Assemblage B	264 (6.7)	No	52 (98.1)	199 (94.3)	Ref.	0.476
			Yes	1 (1.9)	12 (5.7)	0.32 (0.04-2.51)	
WATER CONSUMPTION							
Drinking un-boiled water straight from the tap	Assemblage A	251 (1.2)	No	1 (3.8)	23 (10.2)	Ref.	0.485
			Yes	25 (96.2)	202 (89.8)	2.85 (0.37-22)	
	Assemblage B	281 (0.7)	No	4 (7.1)	23 (10.2)	Ref.	0.484
			Yes	52 (92.9)	202 (89.8)	1.48 (0.49-4.46)	
No. glasses of un-boiled tap water per day	Assemblage A	241 (5.1)	0	1 (4.2)	22 (10.1)	Ref.	$0.532^{CT}/0.195^{C}$
			1-2	12 (50)	98 (45.2)	2.69 (0.33-21.82)	
			3-4	5 (20.8)	60 (27.6)	1.83 (0.20-16.58)	
			5+	6 (25)	37 (17.1)	3.57 (0.40-31.62)	
	Assemblage B	267 (5.7)	0	4 (8)	22 (10.1)	Ref.	$0.500^{CT}/0.738^{C}$
			1-2	21 (42)	98 (45.2)	1.18 (0.37-3.78)	
			3-4	19 (38)	60 (27.6)	1.74 (0.53-5.69)	
			5+	6 (12)	37 (17.1)	0.89 (0.23-3.51)	

Drinking un-boiled water from a lake, river or stream	Assemblage A	226 (11)	No	24 (96)	199 (99)	Ref.	0.298
			Yes	1 (4)	2 (1)	4.15 (0.36-47.45)	
	Assemblage B	252 (11)	No	50 (98)	199 (99)	Ref.	0.494
			Yes	1 (2)	2 (1)	1.99 (0.18-22.39)	
Water from the tap reported having an unusual smell	Assemblage A	246 (3.1)	No	26 (100)	215 (97.7)	Ref.	1.000
			Yes	0 (0)	5 (2.3)	n/a***	
	Assemblage B	275 (2.8)	No	53 (96.4)	215 (97.7)	Ref.	0.630
			Yes	2 (3.6)	5 (2.3)	1.62 (0.31-8.59)	
FOOD CONSUMPTION							
FOOD CONSUMPTION HABITS							
No. times per week eating salads or raw vegetables	Assemblage A	244 (3.9)	0	7 (26.9)	34 (15.6)	Ref.	$0.392^{CT}/0.106^{C}$
			1-2	9 (34.6)	71 (32.6)	0.62 (0.21-1.79)	
			3-4	6 (23.1)	54 (24.8)	0.54 (0.17-1.74)	
			5+	4 (15.4)	59 (27.1)	0.33 (0.09-1.21)	
	Assemblage B	274 (3.2)	0	9 (16.1)	34 (15.6)	Ref.	$0.704^{CT}/0.192^{C}$
			1-2	20 (35.7)	71 (32.6)	1.06 (0.44-2.58)	
			3-4	16 (28.6)	54 (24.8)	1.12 (0.44-2.82)	
			5+	11 (19.6)	59 (27.1)	0.70 (0.26-1.87)	
No. times per week eating cooked vegetables	Assemblage A	246 (3.1)	0	0 (0)	6 (2.7)	Ref.	$0.392^{CT}/0.095^{C}$
			1-2	5 (18.5)	26 (11.9)	1.12 (0.21-1.79)	
			3-4	10 (37)	63 (28.8)	1.54 (0.17-1.90)	
			5+	12 (44.4)	124 (56.6)	0.33 (0.09-1.21)	
	Assemblage B	274 (3.2)	0	2 (3.6)	6 (2.7)	Ref.	$0.669^{CT}/0.271^{C}$
			1-2	4 (7.3)	26 (11.9)	0.46 (0.07-3.13)	
			3-4	19 (34.5)	63 (28.8)	0.90 (0.17-4.86)	
			5+	30 (54.5)	124 (56.6)	0.73 (0.14-3.78)	
Peeling raw fruit before eating	Assemblage A	248 (2.4)	No	18 (69.2)	173 (77.9)	Ref.	0.319
			Yes	8 (30.8)	49 (22.1)	1.57 (0.64-3.83)	
	Assemblage B	277 (2.1)	No	43 (78.2)	173 (77.9)	Ref.	0.968
			Yes	12 (21.8)	49 (22.1)	0.98 (0.48-2.01)	
Washing raw fruit before eating	Assemblage A	246 (3.1)	No	8 (33.3)	70 (31.5)	Ref.	0.857
			Yes	16 (66.7)	152 (68.5)	0.92 (0.38-2.25)	

	Assemblage B	276 (2.5)	No	22 (40.7)	70 (31.5)	Ref.	0.198
			Yes	32 (59.3)	152 (68.5)	0.67 (0.36-1.24)	
FOOD CONSUMPTION DURING THE EXPOSURE WINDOW							
Eating cooked vegetables	Assemblage A	248 (2.4)	No	1 (3.7)	14 (6.3)	Ref.	1.000
			Yes	26 (96.3)	207 (93.7)	1.76 (0.22-13.93)	
	Assemblage B	276 (2.5)	No	4 (7.3)	14 (6.3)	Ref.	0.764
			Yes	51 (92.7)	207 (93.7)	0.86 (0.27-2.73)	
Provenience of cooked vegetables	Assemblage A	242 (4.7)	Did not eat cooked vegetables	1 (3.7)	14 (6.5)	Ref.	0.877
			Market	2 (7.4)	11 (5.1)	2.54 (0.20-31.86)	
			Greengrocers	1 (3.7)	8 (3.7)	1.75 (0.10-31.96)	
			Supermarket	20 (74.1)	142 (66)	1.97 (0.25-15.82)	
			Homegrown	0 (0)	3 (1.4)	n/a***	
			Other place (e.g. restaurant etc.)	0 (0)	2 (0.9)	n/a***	
			Multiple proveniences (at least two different)	3 (11.1)	35 (16.3)	1.20 (0.11-12.54)	
	Assemblage B	264 (6.7)	Did not eat cooked vegetables	4 (8.2)	14 (6.5)	Ref.	0.117
			Market	3 (6.1)	11 (5.1)	0.95 (0.18-5.19)	
			Greengrocers	2 (4.1)	8 (3.7)	0.87 (0.13-5.90)	
			Supermarket	32 (65.3)	142 (66)	0.79 (0.24-2.56)	
			Homegrown	0 (0)	3 (1.4)	n/a***	
			Other place (e.g. restaurant etc.)	4 (8.2)	2 (0.9)	7.00 (092-53.23)	
			Multiple proveniences (at least two different)	4 (8.2)	35 (16.3)	0.40 (0.09-1.83)	
Drinking fruit juice	Assemblage A	246 (3.1)	No	8 (29.6)	58 (26.5)	Ref.	0.728
			Yes	19 (70.4)	161 (73.5)	0.86 (0.35-2.06)	
	Assemblage B	274 (3.2)	No	14 (25.5)	58 (26.5)	Ref.	0.877
			Yes	41 (74.5)	161 (73.5)	1.05 (0.54-2.08)	
Eating pre-packed sandwiches	Assemblage A	235 (7.5)	No	18 (72)	144 (68.6)	Ref.	0.726
			Yes	7 (28)	66 (31.4)	0.85 (0.34-2.13)	
	Assemblage B	266 (6)	No	36 (64.3)	144 (68.6)	Ref.	0.542
			Yes	20 (35.7)	66 (31.4)	1.21 (0.65-2.25)	
Provenience of pre-packed	Assemblage A	233 (8.3)	Did not eat pre-packed sandwiches	18 (72)	144 (69.2)	Ref.	0.816
			Greengrocers	0 (0)	1 (0.5)	n/a***	

sandwiches			Supermarket	7 (28)	60 (28.8)	0.93 (0.37-2.35)	
			Other place (e.g. restaurant etc.)	0 (0)	3 (1.4)	n/a***	
	Assemblage B	262 (7.4)	Did not eat pre-packed sandwiches	36 (66.7)	144 (69.2)	Ref.	0.686
			Greengrocers	0 (0)	1 (0.5)	n/a***	
			Supermarket	16 (29.6)	60 (28.8)	1.07 (0.55-2.07)	
			Other place (e.g. restaurant etc.)	2 (3.7)	3 (1.4)	2.67 (0.43-16.56)	
Provenience of beef	Assemblage A	245 (3.5)	Did not eat beef	3 (12)	56 (25.5)	Ref.	0.542
			Market stall	0 (0)	4 (1.8)	n/a***	
			Butcher	3 (12)	28 (12.7)	2.00 (0.38-10.55)	
			Supermarket	15 (60)	104 (47.3)	2.69 (0.75-9.70)	
			Other place (e.g. restaurant etc.)	1 (4)	4 (1.8)	4.67 (0.39-55.73)	
			Multiple proveniences (at least two different)	3 (12)	24 (10.9)	2.33 (0.44-12.40)	
	Assemblage B	272 (3.9)	Did not eat beef	9 (17.3)	56 (25.5)	Ref.	0.215
			Market stall	0 (0)	4 (1.8)	n/a***	
			Butcher	7 (13.5)	28 (12.7)	1.56 (0.52-4.61)	
			Supermarket	27 (51.9)	104 (47.3)	1.61 (0.71-3.67)	
			Other place (e.g. restaurant etc.)	4 (7.7)	4 (1.8)	6.22 (1.31-29.44)	
			Multiple proveniences (at least two different)	5 (9.6)	24 (10.9)	1.30 (0.39-4.27)	
Eating lamb	Assemblage A	226 (11)	No	11 (50)	100 (49)	Ref.	1.000
			Yes	11 (50)	104 (51)	0.96 (0.40-2.32)	
	Assemblage B	255 (9.9)	No	23 (45.1)	100 (49)	Ref.	0.616
			Yes	28 (54.9)	104 (51)	1.17 (0.63-2.17)	
Provenience of lamb	Assemblage A	225 (11.4)	Did not eat lamb	11 (49)	100 (49)	Ref.	0.910
			Market stall	0 (0)	4 (2)	n/a***	
			Butcher	3 (14.3)	22 (10.8)	1.24 (0.32-4.82)	
			Supermarket	6 (28.6)	67 (32.8)	0.81 (0.29-2.31)	
			Other place (e.g. restaurant etc.)	0 (0)	2 (1)	n/a***	
			Multiple proveniences (at least two different)	1 (4.8)	9 (4.4)	1.01 (0.12-8.74)	
	Assemblage B	253 (10.6)	Did not eat lamb	23 (46.9)	100 (49)	Ref.	0.225
			Market stall	0 (0)	4 (2)	n/a***	
			Butcher	5 (10.2)	22 (10.8)	0.99 (0.34-2.88)	
			Supermarket	14 (28.6)	67 (32.8)	0.91 (0.44-1.89)	
			Other place (e.g. restaurant etc.)	3 (6.1)	2 (1)	6.52 (1.03-41.3)	
			Multiple proveniences (at least two different)	4 (8.2)	9 (4.4)	1.93 (0.55-6.83)	

Eating pork	Assemblage A	231 (9.1)	No	9 (39.1)	84 (40.4)	Ref.	1.000
			Yes	14 (60.9)	124 (59.6)	1.05 (0.44-2.55)	
	Assemblage B	259 (8.5)	No	16 (31.4)	84 (40.4)	Ref.	0.236
			Yes	35 (68.6)	124 (59.6)	1.48 (0.77-2.85)	
Provenience of pork	Assemblage A	231 (9.1)	Did not eat pork	9 (39.1)	84 (40.4)	Ref.	0.901
			Market stall	0 (0)	4 (1.9)	n/a***	
			Butcher	2 (8.7)	15 (7.2)	1.24 (0.24-6.34)	
			Supermarket	11 (47.8)	92 (44.2)	1.12 (0.44-2.83)	
			Other place (e.g. restaurant etc.)	0 (0)	3 (1.4)	n/a***	
			Multiple proveniences (at least two different)	1 (4.3)	10 (4.8)	0.93 (0.11-8.15)	
	Assemblage B	256 (9.5)	Did not eat pork	16 (33.3)	84 (40.4)	Ref.	0.842
			Market stall	1 (2.1)	4 (1.9)	1.31 (0.14-12.52)	
			Butcher	3 (6.3)	15 (7.2)	1.05 (0.27-4.05)	
			Supermarket	23 (47.9)	92 (44.2)	1.31 (0.65-2.65)	
			Other place (e.g. restaurant etc.)	2 (4.2)	3 (1.4)	3.50 (0.54-22.65)	
			Multiple proveniences (at least two different)	3 (6.3)	10 (4.8)	1.57 (0.39-6.36)	
Eating chicken	Assemblage A	249 (2)	No	2 (7.4)	18 (8.1)	Ref.	1.000
			Yes	25 (92.6)	204 (91.9)	1.10 (024-5.04)	
	Assemblage B	274 (3.2)	No	2 (3.8)	18 (8.1)	Ref.	0.288
			Yes	50 (96.2)	204 (91.9)	2.21 (0.50-9.82)	
Eating fish	Assemblage A	240 (5.5)	No	6 (23.1)	60 (28)	Ref.	0.593
			Yes	20 (76.9)	154 (72)	1.30 (0.50-3.39)	
	Assemblage B	263 (7.1)	No	17 (34.7)	60 (28)	Ref.	0.356
			Yes	32 (65.3)	154 (72)	0.73 (0.38-1.42)	
Provenience of fish	Assemblage A	240 (5.5)	Did not eat fish	6 (23.1)	60 (28)	Ref.	0.412
			Market stall	1 (3.8)	6 (2.8)	1.67 (0.17-16.25)	
			Fishmonger	2 (7.7)	14 (6.5)	1.43 (0.26-7.84)	
			Supermarket	15 (57.7)	116 (54.2)	1.29 (0.48-3.50)	
			Other place (e.g. restaurant etc.)	2 (7.7)	5 (2.3)	4.00 (0.63-25.24)	
			Multiple proveniences (at least two different)	0 (0)	13 (6.1)	n/a***	
	Assemblage B	261 (7.8)	Did not eat fish	17 (36.2)	60 (28)	Ref.	0.516
			Market stall	1 (2.1)	6 (2.8)	0.59 (0.07-5.23)	

			Fishmonger	5 (10.6)	14 (6.5)	1.26 (0.40-3.40)	
			Supermarket	21 (44.7)	116 (54.2)	0.64 (0.31-1.30)	
			Other place (e.g. restaurant etc.)	2 (4.3)	5 (2.3)	1.41 (0.25-7.93)	
			Multiple proveniences (at least two different)	1 (2.1)	13 (6.1)	0.27 (0.03-2.23)	
Eating at a pub or a restaurant	Assemblage A	251 (1.2)	No	13 (46.4)	81 (36.3)	Ref.	0.298
			Yes	15 (53.6)	142 (63.7)	0.66 (0.30-1.45)	
	Assemblage B	279 (1.4)	No	23 (41.1)	81 (36.3)	Ref.	0.511
			Yes	33 (58.9)	142 (63.7)	0.82 (0.45-1.49)	
No. times eating at a pub or a restaurant	Assemblage A	246 (3.1)	0	13 (50)	81 (36.8)	Ref.	$0.224^{CT}/0.116^{C}$
			1-2	10 (38.5)	83 (37.7)	0.75 (031-1.81)	
			3-4	1 (3.8)	36 (16.4)	0.17 (0.02-1.37)	
			5+	2 (7.7)	20 (9.1)	0.62 (0.13-2.99)	
	Assemblage B	272 (3.9)	0	23 (44.2)	81 (36.8)	Ref.	$0.353^{CT}/0.313^{C}$
			1-2	21 (40.4)	83 (37.7)	0.89 (0.46-1.73)	
			3-4	4 (7.7)	36 (16.4)	0.39 (0.13-1.21)	
			5+	4 (7.7)	20 (9.1)	0.70 (0.22-2.27)	
Eating at a canteen	Assemblage A	226 (11)	No	20 (90.9)	182 (89.2)	Ref.	1.000
			Yes	2 (9.1)	22 (10.8)	0.83 (0.18-3.78)	
	Assemblage B	255 (9.9)	No	42 (82.4)	182 (89.2)	Ref.	0.180
			Yes	9 (17.6)	22 (10.8)	1.77 (0.76-4.13)	
No. times eating at a canteen	Assemblage A	223 (12.2)	0	20 (95.2)	182 (90.1)	Ref.	$0.237/0.464^{C}$
			1-2	0 (0)	7 (3.5)	n/a***	
			3-4	1 (4.8)	3 (1.5)	3.03 (0.30-30.5)	
			5+	0 (0)	10 (5)	n/a***	
	Assemblage B	253 (10.6)	0	42 (82.4)	182 (90.1)	Ref.	$0.486^{CT}/0.583^{C}$
			1-2	4 (7.8)	7 (3.5)	2.48 (0.69-8.85)	
			3-4	1 (2)	3 (1.5)	1.44 (0.15-14.23)	
			5+	4 (7.8)	10 (5)	1.73 (0.52-5.80)	
Eating at a takeaway or fast food	Assemblage A	240 (5.5)	No	10 (38.5)	94 (43.9)	Ref.	0.595
			Yes	16 (61.5)	120 (56.1)	1.25 (0.54-2.89)	
	Assemblage B	268 (5.3)	No	19 (35.2)	94 (43.9)	Ref.	0.245
			Yes	35 (64.8)	120 (56.1)	1.44 (0.78-2.68)	

No. times eating at a takeaway or fast food	Assemblage A	231 (9.1)	0	10 (43.5)	94 (45.2)	Ref.	0.381
			1-2	9 (39.1)	81 (38.9)	1.04 (0.40-2.70)	
			3-4	4 (17.4)	22 (10.6)	1.71 (0.49-5.96)	
			5+	0 (0)	11 (5.3)	n/a***	
	Assemblage B	255 (9.9)	0	19 (40.4)	94 (45.2)	Ref.	$0.491^{CT}/0.626^{C}$
			1-2	17 (36.2)	81 (38.9)	1.04 (0.51-2.13)	
			3-4	9 (19.1)	22 (10.6)	2.02 (0.81-5.07)	
			5+	2 (4.3)	11 (5.3)	0.90 (0.18-4.39)	
Eating at a barbecue	Assemblage A	226 (11)	No	21 (95.5)	180 (88.2)	Ref.	0.481
			Yes	1 (4.5)	24 (11.8)	0.36 (0.05-2.78)	
	Assemblage B	255 (9.9)	No	45 (88.2)	180 (88.2)	Ref.	1.000
			Yes	6 (11.8)	24 (11.8)	1.00 (0.39-2.59)	
No. times eating at a barbecue	Assemblage A	225 (11.4)	0	21 (95.5)	180 (88.7)	Ref.	0.439
			1-2	1 (5.3)	18 (8.9)	0.48 (0.06-3.75)	
			3-4	0 (0)	5 (2.5)	n/a***	
	Assemblage B	253 (10.6)	0	45 (90)	180 (88.7)	Ref.	$0.322^{CT}/0.568^{C}$
			1-2	5 (10)	18 (8.9)	1.11 (0.39-3.15)	
			3-4	0 (0)	5 (2.5)	n/a***	
GENERAL HOUSEHOLD DETAILS							
No. of adults (≥16y) in the house	Assemblage A	254	1	6 (21.4)	52 (23)	Ref.	$0.581^{CT}/0.501^{C}$
			2	19 (67.9)	128 (56.6)	1.29 (0.49-3.40)	
			3	2 (7.1)	32 (14.2)	0.54 (0.10-2.85)	
			4+	1 (3.6)	14 (6.2)	0.62 (0.07-5.57)	
	Assemblage B	283	1	6 (10.5)	52 (23)	Ref.	$0.100^{CT}/0.793^{C}$
			2	41 (71.9)	128 (56.6)	2.78 (1.11-6.93)	
			3	6 (10.5	32 (14.2)	1.62 (0.48-5.47)	
			4+	4 (7)	14 (6.2)	2.47 (0.61-10.00)	
Any contact with a person with diarrhoea outside the house	Assemblage A	231 (9.1)	No	22 (91.7)	187 (90.3)	Ref.	1.000
			Yes	2 (8.3)	20 (9.7)	0.85 (0.19-3.88)	
	Assemblage B	260 (8.1)	No	45 (84.9)	187 (90.3)	Ref.	0.255
			Yes	8 (15.1)	20 (9.7)	1.66 (0.69-4.02)	

*percentages refer to the proportion of participants with missing information for the variable in the specific dataset; **percentages refer to the proportion among cases and controls that answered the question; ***Odds ratio with 95% confidence interval; *** Odds ratio not calculated because no variation present in the data; ^{CT} logistic regression on the categorized variable; ^{C} logistic regression on the continuous variable

REFERENCES

Abdolrasouli, A., McMillan, A. and Ackers, J. P. (2009). Sexual transmission of intestinal parasites in men who have sex with men. *Sexual Health,* **6**, 185-194.

Abou-Shady, O., El Raziky, M. S., Zaki, M. M. and Mohamed, R. K. (2011). Impact of *Giardia lamblia* on growth, serum levels of zinc, copper, and iron in Egyptian children. *Biological trace element research,* **140**, 1-6.

Adam, R. D. (2000). The *Giardia lamblia* genome. *International Journal for Parasitology,* **30**, 475-484.

Adam, R. D. (2001). Biology of *Giardia lamblia*. *Clinical Microbiology Reviews,* **14**, 447-475.

Adams, P. J., Monis, P. T., Elliot, A. D. and Andrew Thompson, R. C. (2004). Cyst morphology and sequence analysis of the small subunit rDNA and ef1α identifies a novel *Giardia* genotype in a quenda (*Isoodon obesulus*) from Western Australia. *Infection, Genetics & Evolution,* **4**, 365-370.

Al-Mohammed, H. I. (2011). Genotypes of *Giardia intestinalis* clinical isolates of gastrointestinal symptomatic and asymptomatic Saudi children. *Parasitology Research,* **108**, 1375-1381.

Aldeen, W. E., Carroll, K., Robison, A., Morrison, M. and Hale, D. (1998). Comparison of nine commercially available enzyme-linked immunosorbent assays for detection of *Giardia lamblia* in fecal specimens. *Journal of clinical microbiology,* **36**, 1338-1340.

Alexander, C., Jones, B., Inverarity, D. and Pollock, K. G. J. (2014). Genotyping of *Giardia* isolates in Scotland: A descriptive epidemiological study. *Epidemiology And Infection,* **142**, 1636-1639.

Almeida, A., Delgado, M.L., Soares, S.C., Castro, A.O., Moreira, M.J., Mendonca, C.M., Canada, N.B. and Correia da Costa, J.M. (2006). Genotype analysis of *Giardia* isolated from asymptomatic children in Northern Portugal. *Journal of Eukariotic Microbiology*, **53** (S1) 177-178.

Almeida, A., Pozio, E. and Caccio, S. M. (2010). Genotyping of *Giardia duodenalis* Cysts by New Real-Time PCR Assays for Detection of Mixed Infections in Human Samples. *Applied & Environmental Microbiology*, **76**, 1895-1901.

Aloisio, F., Filippini, G., Antenucci, P., Lepri, E., Pezzotti, G., Cacciò, S. M. and Pozio, E. (2006). Severe weight loss in lambs infected with *Giardia duodenalis* assemblage B. *Veterinary Parasitology,* **142**, 154-158.

Amar, C. F. L., Dear, P. H., Pedraza-Díaz, S., Looker, N., Linnane, E. and McLauchlin, J. (2002). Sensitive PCR-restriction fragment length polymorphism assay for detection and genotyping of *Giardia duodenalis* in human feces. *Journal of clinical microbiology*, **40**, 446-452.

Amar, C. F. L., East, C. L., Gray, J., Iturriza-Gomara, M., Maclure, E. A. and McLauchlin, J. (2007). Detection by PCR of eight groups of enteric pathogens in 4,627 faecal samples: Re-examination of the English case-control Infectious Intestinal Disease Study (1993-1996). *European Journal of Clinical Microbiology and Infectious Diseases*, **26**, 311-323.

Andrews, R. H., Adams, M., Boreham, P. F. L., Mayrhofer, G. and Meloni, B. P. (1989). *Giardia intestinalis*: electrophoretic evidence for a species complex. *International Journal for Parasitology*, **19**, 183-190.

Ankarklev, J., Jerlström-Hultqvist, J., Ringqvist, E., Troell, K. and Svärd, S. G. (2010). Behind the smile: cell biology and disease mechanisms of *Giardia* species. *Nature Reviews. Microbiology*, **8**, 413-422.

Ankarklev, J., Svärd, S. G. and Lebbad, M. (2012a). Allelic sequence heterozygosity in single *Giardia* parasites. *BMC Microbiology*, **12**: 65.

Ankarklev, J., Hestvik, E., Lebbad, M., Lindh, J., Kaddu-Mulindwa, D. H., Andersson, J. O., Tylleskär, T., Tumwine, J. K. and Svärd, S. G. (2012b). Common Coinfections of *Giardia intestinalis* and *Helicobacter pylori* in Non-Symptomatic Ugandan Children. *PLoS Neglected Tropical Diseases*, **6**(8): e1780.

Anuar, T. S., Nor Azreen, S., Salleh, F. M. and Moktar, N. (2014). Molecular epidemiology of giardiasis among Orang Asli in Malaysia: Application of the triosephosphate isomerase gene. *BMC Infectious Diseases*, **14**: 78.

Armson, A., Reid, S., Ryan, U. M., Johnson, J., Yang, R. and Thompson, J. (2009). *Giardia* genotypes in pigs in Western Australia: Prevalence and association with diarrhoea. *Experimental Parasitology*, **121**, 381-383.

Atwill, E. R., Perea, L. and McDougald, N. K. (2000). Cross-sectional study of faecal shedding of *Giardia duodenalis* and *Cryptosporidium parvum* among packstock in the Sierra Nevada Range. *Equine veterinary journal*, **32**, 247-252.

Aydin, A. F., Besirbellioglu, B. A., Avci, I. Y., Tanyuksel, M., Araz, E. and Pahsa, A. (2004). Classification of *Giardia duodenalis* parasites in Turkey into groups a and B using restriction fragment length polymorphism. *Diagnostic microbiology and infectious disease*, **50**, 147-151.

Ballweber, L. R., Xiao, L., Bowman, D. D., Kahn, G. and Cama, V. A. (2010). Giardiasis in dogs and cats: update on epidemiology and public health significance. *Trends In Parasitology*, **26**, 180-189.

Barigye, R., Dyer, N. W., Newell, T. K., Khaitsa, M. L., Trout, J. M., Santin, M. and Fayer, R. (2008). Molecular and immunohistochemical detection of assemblage E, *Giardia duodenalis* in scouring North Dakota calves. *Veterinary Parasitology,* **157**, 196-202.

Barsky, A. J., Peekna, H. M. and Borus, J. F. (2001). Somatic symptom reporting in women and men. *Journal of General Internal Medicine,* **16**, 266-275.

Batchelor, D. J., Tzannes, S., Graham, P. A., Wastling, J. M., Pinchbeck, G. L. and German, A. J. (2008). Detection of Endoparasites with Zoonotic Potential in Dogs with Gastrointestinal Disease in the UK. *Transboundary & Emerging Diseases,* **55**, 99-104.

Beck, R., Sprong, H., Pozio, E. and Cacciò, S. M. (2012). Genotyping of *Giardia duodenalis* isolates from dogs: Lessons from a multilocus sequence typing study. *Vector-Borne and Zoonotic Diseases,* **12**, 206-213.

Bello, J., Núñez, F. A., González, O. M., Fernández, R., Almirall, P. and Escobedo, A. A. (2011). Risk factors for *Giardia* infection among hospitalized children in Cuba. *Annals of Tropical Medicine and Parasitology,* **105**, 57-64.

Bénéré, E., Van Assche, T., Van Ginneken, C., Peulen, O., Cos, P. and Maes, L. (2012a). Intestinal growth and pathology of *Giardia duodenalis* assemblage subtype AI, AII, B and e in the gerbil model. *Parasitology*, **139**, 424-433.

Bénéré, E., Van Assche, T., Cos, P. and Maes, L. (2012b). Intrinsic susceptibility of *Giardia duodenalis* assemblage subtypes AI, AII, B and EIII for nitric oxide under axenic culture conditions. *Parasitology Research*, **110**, 1315-1319.

Betancourt, W. Q. and Rose, J. B. (2004). Drinking water treatment processes for removal of *Cryptosporidium* and *Giardia*. *Veterinary Parasitology*, **126**, 219-234.

Birky, C. W. (2010). *Giardia* Sex? Yes, but how and how much? *Trends In Parasitology,* **26**, 70-74.

Boone, J. H., Wilkins, T. D., Nash, T. E., Brandon, J. E., Macias, E. A., Jerris, R. C. and Lyerly, D. M. (1999). TechLab and alexon *Giardia* enzyme-linked immunosorbent assay kits detect cyst wall protein 1. *Journal of clinical microbiology,* **37**, 611-614.

Boontanom, P., Mungthin, M., Tan-ariya, P., Naaglor, T. and Leelayoova, S. (2011). Epidemiology of giardiasis and genotypic characterization of *Giardia duodenalis* in preschool children of a rural community, central Thailand. *Tropical Biomedicine,* **28**, 32-39.

Breathnach, A. S., McHugh, T. D. and Butcher, P. D. (2010). Prevalence and clinical correlations of genetic subtypes of *Giardia lamblia* in an urban setting. *Epidemiology & Infection,* **138**, 1459-1467.

Broglia, A., Weitzel, T., Harms, G., Cacció, S. M. and Nöckler, K. (2013). Molecular typing of *Giardia duodenalis* isolates from German travellers. *Parasitology Research,* **112**, 3449-3456.

Bundy, D. A. P. (1988). Gender-dependent patterns of infections and disease. *Parasitology Today,* **4**, 189-193.

Cacciò, S. M., De Giacomo, M. and Pozio, E. (2002). Sequence analysis of the β-giardin gene and development of a polymerase chain reaction-restriction fragment length polymorphism assay to genotype *Giardia duodenalis* cysts from human faecal samples. *International Journal for Parasitology,* **32**, 1023-1030.

Cacciò, S. M., Beck, R., Lalle, M., Marinculic, A. and Pozio, E. (2008). Multilocus genotyping of *Giardia duodenalis* reveals striking differences between assemblages A and B. *International Journal for Parasitology,* **38**, 1523-1531.

Cacciò, S. M. and Sprong, H. (2010). *Giardia duodenalis*: Genetic recombination and its implications for taxonomy and molecular epidemiology. *Experimental Parasitology,* **124**, 107-112.

Calderaro, A., Gorrini, C., Montecchini, S., Peruzzi, S., Piccolo, G., Rossi, S., Gargiulo, F., Manca, N., Dettori, G. and Chezzi, C. (2010). Evaluation of a real-time polymerase chain reaction assay for the laboratory diagnosis of giardiasis. *Diagnostic microbiology and infectious disease,* **66**, 261-267.

Cantey, P. T., Roy, S., Lee, B., Cronquist, A., Smith, K., Liang, J. and Beach, M. J. (2012). Study of Nonoutbreak Giardiasis: Novel Findings and Implications for Research. *American Journal of Medicine*.

Carlin, E. P., Garrett, J., Lorentzen, L., Bowman, D. D. and Scarlett, J. M. (2006). Prevalence of *Giardia* in symptomatic dogs and cats throughout the United States as determined by the IDEXX SNAP Giardia test. *Veterinary therapeutics : research in applied veterinary medicine,* **7**, 199-206.

Casemore, D. (2006). Towards a US national estimate of the risk of endemic waterborne disease - Sero-epidemiologic studies. *Journal of Water and Health,* **4**, 121-164.

Castro-Hermida, J. A., Delafosse, A., Pors, I., Ares-Mazas, E. and Chartier, C. (2005). *Giardia duodenalis* and *Cryptosporidium parvum* infections in adult goats and their implications for neonatal kids. *Veterinary Record,* **157**, 623-627.

Castro-Hermida, J. A., Almeida, A., González-Warleta, M., Correia Da Costa, J. M., Rumbo-Lorenzo, C. and Mezo, M. (2007a). Occurrence of *Cryptosporidium parvum* and *Giardia duodenalis* in healthy adult domestic ruminants. *Parasitology Research,* **101**, 1443-1448.

Castro-Hermida, J. A., Mezo, M. and González-Warleta, M. (2007b). Natural infection by *Cryptosporidium parvum* and *Giardia duodenalis* in sheep and goats in Galicia (NW Spain). *Small Ruminant Research,* **72**, 96-100.

Castro-Hermida, J. A., Gonza´lez-Warleta, M. and Mezo, M. (2014). *Cryptosporidium* spp. and *Giardia duodenalis* as pathogenic contaminants of water in Galicia, Spain: The need for safe drinking water. *International Journal of Hygiene and Environmental Health*. Published ahead of print: doi:10.1016/j.ijheh.2014.09.001

Cavalier-Smith, T. (2003). The excavate protozoan phyla Metamonada Grassé emend. (Anaeromonadea, Parabasalia, Carpediemonas, Eopharyngia) and Loukozoa emend. (Jakobea, Malawimonas): Their evolutionary affinities and new higher taxa. *International Journal of Systematic and Evolutionary Microbiology,* **53**, 1741-1758.

Chalmers, R. M., Smith, R., Elwin, K., Clifton-Hadley, F. A. and Giles, M. (2011). Epidemiology of anthroponotic and zoonotic human cryptosporidiosis in England and Wales, 2004-2006. *Epidemiology And Infection,* **139**, 700-712.

Claerebout, E., Casaert, S., Dalemans, A. C., De Wilde, N., Levecke, B., Vercruysse, J. and Geurden, T. (2009). *Giardia* and other intestinal parasites in different dog populations in Northern Belgium. *Veterinary Parasitology,* **161**, 41-46.

Coles, C. L., Levy, A., Dagan, R., Deckelbaum, R. J. and Fraser, D. (2009). Risk factors for the initial symptomatic *Giardia* infection in a cohort of young Arab-Bedouin children. *Annals of Tropical Paediatrics,* **29**, 291-300.

Conroy, D. A. (1960). A note on the occurrence of *Giardia* sp. in a Christmas pudding. *Revista iberica de parasitologia,* **20**, 567-571.

Cook, N., Nichols, R. A. B., Wilkinson, N., Paton, C. A., Barker, K. and Smith, H. V. (2007). Development of a method for detection of *Giardia duodenalis* cysts on lettuce and for simultaneous analysis of salad products for the presence of *Giardia* cysts and *Cryptosporidium* oocysts. *Applied And Environmental Microbiology,* **73**, 7388-7391.

Cotton, J. A., Beatty, J. K. and Buret, A. G. (2011). Host parasite interactions and pathophysiology in *Giardia* infections. *International Journal for Parasitology,* **41**, 925-933.

CPMP (1990). Good Clinical Practice for Trials on Medicinal Products in the European Community: CPMP Working Party on Efficacy of Medicinal Products. *Pharmacology & Toxicology,* **67**, 361-373.

Daly, E. R., Roy, S. J., Blaney, D. D., Manning, J. S., Hill, V. R., Xiao, L. and Stull, J. W. (2010). Outbreak of giardiasis associated with a community drinking-water source. *Epidemiology & Infection,* **138**, 491-500.

Daniels, C. W. and Belosevic, M. (1995). Comparison of the course of infection with *Giardia* muris in male and female mice. *International Journal for Parasitology,* **25**, 131-135.

Davies, A. P., Campbell, B., Evans, M. R., Bone, A., Roche, A. and Chalmers, R. M. (2009). Asymptomatic carriage of protozoan parasites in children in day care centers in the united kingdom. *Pediatric Infectious Disease Journal,* **28**, 838-840.

DeRegnier, D. P., Cole, L., Schupp, D. G. and Erlandsen, S. L. (1989). Viability of *Giardia* cysts suspended in lake, river, and tap water. *Applied And Environmental Microbiology,* **55**, 1223-1229.

de Wit, M.A.S., Koopmans, M.P.G, Kortbeek, L.M., van Leeuwen, N.J, Bartelds, A.I.M and van Duynhoven, Y.T.H.P. (2001). Gastroenteritis in sentinel general practices, the Netherlands. *Emerging infectious Diseases,* **7** (1), 82-91.

Dykes, A. C., Juranek, D. D., Lorenz, R. A., Sinclair, S., Jakubowski, W. and Davies, R. (1980). Municipal waterborne giardiasis. Epidemiologic investigation - beavers implicated as a possible reservoir. *Annals of Internal Medicine,* **92**, 165-170.

European Center for Disease Control and Prevention (2011). Annual epidemiological report 2011 - Reporting on 2009 surveillance data and 2010 epidemic intelligence data. ECDC, Stockholm. Available at: http://www.ecdc.europa.eu/en/publications/_layouts/forms/Publication_DispForm.aspx?List=4f55ad51-4aed-4d32-b960-af70113dbb90&ID=496

Ekdahl, K. and Andersson, Y. (2005). Imported giardiasis: impact of international travel, immigration, and adoption. *American Journal of Tropical Medicine and Hygiene,* **72**, 825-830.

Ellam, H., Verlander, N. Q., Lamden, K., Cheesbrough, J. S., Durband, C. A. and James, S. (2008). Surveillance of giardiasis in Northwest England 1996-2006: impact of an enzyme immunoassay test. *Eurosurveillance,* **13**, 18977-18977.

Elwin, K., Fairclough, H. V., Hadfield, S. J. and Chalmers, R. M. (2013). *Giardia duodenalis* typing from stools: A comparison of three approaches to extracting DNA, and validation of a probe-based real-time PCR typing assay. *Journal of Medical Microbiology,* **63**, 38-44.

Epe, C., Rehkter, G., Schnieder, T., Lorentzen, L. and Kreienbrock, L. (2010). *Giardia* in symptomatic dogs and cats in Europe—Results of a European study. *Veterinary Parasitology*, **173**, 32-38.

Espelage, W., an der Heiden, M., Stark, K. and Alpers, K. (2010). Characteristics and risk factors for symptomatic *Giardia lamblia* infections in Germany. *BMC Public Health*, **10**, 41-41.

Faustini, A., Marinacci, C., Fabrizi, E., Marangi, M., Recchia, O., Pica, R., Giustini, F., Marca, A. l., Nacci, A., Panichi, G. and Perucci, C. A. (2006). The impact of the Catholic Jubilee in 2000 on infectious diseases. A case-control study of giardiasis, Rome, Italy 2000-2001. *Epidemiology And Infection*, **134**, 649-658.

Fayer, R., Dubey, J. P. and Lindsay, D. S. (2004). Zoonotic protozoa: from land to sea. *Trends In Parasitology*, **20**, 531-536.

Feng, Y. and Xiao, L. (2011). Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clinical Microbiology Reviews*, **24**, 110-140.

Foronda, P., Bargues, M. D., Abreu-Acosta, N., Periago, M. V., Valero, M. A., Valladares, B. and Mas-Coma, S. (2008). Identification of genotypes of *Giardia intestinalis* of human isolates in Egypt. *Parasitology Research*, **103**, 1177-1181.

Franzén, O., Jerlström-Hultqvist, J., Castro, E., Sherwood, E., Ankarklev, J., Reiner, D. S., Palm, D., Andersson, J. O., Andersson, B. and Svärd, S. G. (2009). Draft Genome Sequencing of *Giardia intestinalis* Assemblage B Isolate GS: Is Human Giardiasis Caused by Two Different Species? *PLoS Pathogens*, **6**, 1-14.

Gagnon, F., Duchesne, J. F., Lévesque, B., Gingras, S. and Chartrand, J. (2006). Risk of giardiasis associated with water supply in an endemic context. *International Journal of Environmental Health Research*, **16**, 349-359.

Garcia, L. S., Shum, A. C. and Bruckner, D. A. (1992). Evaluation of a new monoclonal antibody combination reagent for direct fluorescence detection of *Giardia* cysts and *Cryptosporidium* oocysts in human fecal specimens. *Journal of clinical microbiology*, **30**, 3255-3257.

Gelanew, T., Lalle, M., Hailu, A., Pozio, E. and Cacciò, S. M. (2007). Molecular characterization of human isolates of *Giardia duodenalis* from Ethiopia. *Acta Tropica*, **102**, 92-99.

Geurden, T., Casaert, S., Vercruysse, J., Claerebout, E., Berkvens, D., Geldhof, P., Levecke, B. and Martens, C. (2008). Mixed *Giardia duodenalis* assemblage A and E infections in calves. *International Journal for Parasitology*, **38**, 259-264.

Geurden, T., Levecke, B., Cacció, S. M., Visser, A., de Groote, G., Casaert, S., Vercruysse, J. and Claerebout, E. (2009). Multilocus genotyping of *Cryptosporidium* and *Giardia* in non-outbreak related cases of diarrhoea in human patients in Belgium. *Parasitology*, **136**, 1161-1168.

Geurden, T., De Wilde, N., Vercruysse, J., Claerebout, E., Vandenhoute, E., Pohle, H. and Casaert, S. (2010). The effect of a fenbendazole treatment on cyst excretion and weight gain in calves experimentally infected with *Giardia duodenalis*. *Veterinary Parasitology*, **169**, 18-23.

Ghoneim, N. H., Abdel-Moein, K. A. and Saeed, H. (2012). Fish as a possible reservoir for zoonotic *Giardia duodenalis* assemblages. *Parasitology Research*, **110**, 2193-2196.

Gillespie, I. A., O'Brien, S. J., Frost, J.A., Adak, J.K., Horby, P., Swan, A.V., Painter, M.J., and Neil, K.R. (2002). A Case-Case Comparison of *Campylobacter coli* and *Campylobacter jejuni* Infection: A Tool for Generating Hypotheses. *Emerging Infectious Diseases*, **8** (9), 937-942.

Gillespie, I. A., O'Brien, S. J., Penman, C., Tompkins, D., Cowden, J. and Humphrey, T. J. (2008). Demographic determinants for *Campylobacter* infection in England and Wales: Implications for future epidemiological studies. *Epidemiology And Infection*, **136**, 1717-1725.

Gillespie, I. A., Mook, P., Adak, G.K., O'Brien, S. J., and McCarthy, N.D. (2012). The "Case-Chaos Study" as an Adjunct or Alternative to Conventional Case-Control Study Methodology. *American Journal of Epidemiology*, **176** (6), 497-505.

Graczyk, T. K., Girouard, A. S., Tamang, L., Nappier, S. P. and Schwab, K. J. (2006). Recovery, bioaccumulation, and inactivation of human waterborne pathogens by the Chesapeake Bay nonnative oyster, *Crassostrea ariakensis*. *Applied And Environmental Microbiology*, **72**, 3390-3395.

Gray, S. F., Gunnell, D. J. and Peters, T. J. (1994). Risk factors for giardiasis: a case-control study in Avon and Somerset. *Epidemiology And Infection*, **113**, 95-102.

Grazioli, B., Matera, G., Laratta, C., Schipani, G., Guarnieri, G., Spniniello, E., Imeneo, M., Amorosi, A., Focà, A. and Luzza, F. (2006). *Giardia lamblia* infection in patients with irritable bowel syndrome and dyspepsia: A prospective study. *World Journal of Gastroenterology*, **12**, 1941-1944.

Grit, G. H., Bénéré, E., Ehsan, A., De Wilde, N., Claerebout, E., Vercruysse, J., Maes, L. and Geurden, T. (2012). *Giardia duodenalis* cyst survival in cattle slurry. *Veterinary Parasitology*, **184**, 330-334.

Hanevik, K., Kristoffersen, E., Svard, S., Bruserud, O., Ringqvist, E., Sørnes, S. and Langeland, N. (2011). Human cellular immune response against

Giardia lamblia 5 years after acute giardiasis. *Journal of Infectious Diseases,* **204**, 1779-1786.

Haque, R., Roy, S., Kabir, M., Stroup, S. E., Mondal, D. and Houpt, E. R. (2005). *Giardia* assemblage A infection and diarrhea in Bangladesh. *The Journal Of Infectious Diseases,* **192**, 2171-2173.

Harter, L., Frost, F., Grunenfelder, G., Perkins-Jones, K. and Libby, J. (1984). Giardiasis in an Infant and Toddler Swim Class. *American Journal of Public Health,* **74**, 155-156.

Health Protection Agency (2011a). *Giardia lamblia* laboratory reports. Available at:http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Giardia/EpidemiologicalData/.

Health Protection Agency (2011b). *Campylobacter* infections by age and sex, England and Wales, 2000-2010. Available at: http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/GastrointestinalDisease/EpidemiologicalData/.

Health Protection Agency (2013). *Giardia lamblia* - 2010 update. Available at: http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Giardia/EpidemiologicalData/.

Hellard, M.E., Sinclair, M.I., Hogg, G.G. and Fairley, C.K. (2000). Prevalence of enteric pathogens among community based asymptomatic individuals. *Journal of Gastroenterology and Hepatology*, **15** (3), 290-293.

Helmy, M. M. F., Abdel-Fattah, H. S. and Rashed, L. (2009). Real-time Pcr/rflp assay to detect *Giardia intestinalis* genotypes in human isolates with diarrhea in Egypt. *Journal of Parasitology,* **95**, 1000-1004.

Hill, D. R. (2001). *Giardia lamblia*. In *Principles and practice of clinical parasitology* (eds. Gillespie, S. H., and Pearson, R. D.), John Wiley & Sons Ltd, Chichester; UK.

Hippisley-Cox, J. and Vinogradova, Y. (2009). Trends in Consultation Rates in General Practice 1995/1996 to 2008/2009: Analysis of the QResearch® database. Available at: http://www.hscic.gov.uk/catalogue/PUB01077/trencons-rate-gene-prac-95-09-95-09-rep.pdf.

Homan, W. L. and Mank, T. G. (2001). Human giardiasis: Genotype linked differences in clinical symptomatology. *International Journal for Parasitology,* **31**, 822-826.

Hopkins, R. M., Meloni, B. P., Groth, D. M., Wetherall, J. D., Reynoldson, J. A. and Thompson, R. C. A. (1997). Ribosomal RNA sequencing reveals differences between the genotypes of *Giardia isolates* recovered from humans and dogs living in the same locality. *Journal of Parasitology,* **83**, 44-51.

Hoque, M. E., Hope, V. T., Scragg, R., Kjellstrom, T. and Lay-Yee, R. (2001). Nappy handling and risk of giardiasis. *Lancet,* **357**, 1017.

Hoque, M. E., Hope, V. T. and Scragg, R. (2002a). *Giardia* infection in Auckland and New Zealand: Trends and international comparison. *New Zealand Medical Journal,* **115**, 121-123.

Hoque, M. E., Hope, V. T., Kjellström, T., Scragg, R. and Lay-Yee, R. (2002b). Risk of giardiasis in Aucklanders: a case-control study. *International Journal of Infectious Diseases,* **6**, 191-197.

Hoque, M. E., Hope, V. T., Scragg, R. and Kjellström, T. (2003). Children at risk of giardiasis in Auckland: a case-control analysis. *Epidemiology And Infection,* **131**, 655-662.

Hoque, M. E., Hope, V. T., Scragg, R. and Graham, J. (2005). Under-notification of giardiasis in Auckland, New Zealand: A capture-recapture estimation. *Epidemiology And Infection,* **133**, 71-79.

Hörman, A., Korpela, H., Sutinen, J., Wedel, H. and Hänninen, M. L. (2004). Meta-analysis in assessment of the prevalence and annual incidence of *Giardia* spp. and *Cryptosporidium* spp. infections in humans in the Nordic countries. *International Journal for Parasitology,* **34**, 1337-1346.

Hunter, P. R., Hughes, S., Woodhouse, S., Syed, Q., Verlander, N. Q., Chalmers, R. M., Morgan, K., Nichols, G., Beeching, N. and Osborn, K. (2004). Sporadic cryptosporidiosis case-control study with genotyping. *Emerging Infectious Diseases,* **10**, 1241-1249.

Istre, G. R., Dunlop, T. S., Gaspard, G. B. and Hopkins, R. S. (1984). Waterborne giardiasis at a mountain resort: Evidence for acquired immunity. *American Journal of Public Health,* **74**, 602-604.

Jephcott, A. E., Begg, N. T. and Baker, I. A. (1986). Outbreak of giardiasis associated with mains water in the United Kingdom. *Lancet,* **1**, 730-732.

Jerlström-Hultqvist, J., Franzén, O., Ankarklev, J., Xu, F., Nohýnková, E., Andersson, J. O., Svärd, S. G. and Andersson, B. (2010a). Genome analysis and comparative genomics of a *Giardia intestinalis* assemblage E isolate. *BMC Genomics,* **11**, 543-543.

Jerlström-Hultqvist, J., Ankarklev, J. and Svärd, S. G. (2010b). Is human giardiasis caused by two different giardia species? *Gut Microbes,* **1**.

John, D. T. and Petri, W. A. (2006). *Markell and Voge's Medical Parasitology,* Philadelphia: Saunders. 9th ed.

Johnson, D. C., Reynolds, K. A., Gerba, C. P., Pepper, I. L. and Rose, J. B. (1995). Detection of *Giardia* and *Cryptosporidium* in marine waters. *Water Science and Technology,* **31**, 439-442.

Johnson, D. C., Enriquez, C. E., Pepper, I. L., Davis, T. L., Gerba, C. P. and Roset, J. B. (1997). Survival of *Giardia*, *Cryptosporidium*, *Poliovirus* and *Salmonella* in marine waters. *Water Science & Technology,* **35**, 261-268.

Johnston, A.R., Gillespie, T.R., Rwego, I.B., Tranby McLachlan, T.L., Kent, A.D. and Goldberg, T.L. (2010). Molecular epidemiology of cross-species *Giardia duodenalis* transmission in Western Uganda. *PLoS Neglected Tropical Diseases,* **4**(5), e683.

Júlio, C., Vilares, A., Oleastro, M., Ferreira, I., Gomes, S., Monteiro, L., Nunes, B., Tenreiro, R. and Ngelo, H. (2012). Prevalence and risk factors for *Giardia duodenalis* infection among children: A case study in Portugal. *Parasites and Vectors,* **5**: 22.

Kaboré, H., Levallois, P., Michel, P., Payment, P., Déry, P. and Gingras, S. (2010). Association between potential zoonotic enteric infections in children and environmental risk factors in Quebec, 1999-2006. *Zoonoses And Public Health,* **57**, 195-205.

Karabiber, N. and Aktas, F. (1991). Foodborne giardiasis. *Lancet (British edition),* **7**, 376-377.

Karanis, P., Kourenti, C. and Smith, H. (2007). Waterborne transmission of protozoan parasites: A worldwide review of outbreaks and lessons learnt. *Journal of Water & Health,* **5**, 1-38.

Katz, D. E., Heisey-Grove, D., Beach, M., Dicker, R. C. and Matyas, B. T. (2006). Prolonged outbreak of giardiasis with two modes of transmission. *Epidemiology And Infection,* **134**, 935.

Keystone, J. S., Keystone, D. L. and Proctor, E. M. (1980). Intestinal parasitic infections in homosexual men: prevalence, symptoms and factors in transmission. *Canadian Medical Association Journal,* **123**, 512-514.

Khan, S. M., Debnath, C., Pramanik, A. K., Xiao, L., Nozaki, T. and Ganguly, S. (2011). Molecular evidence for zoonotic transmission of *Giardia duodenalis* among dairy farm workers in West Bengal, India. *Veterinary Parasitology,* **178**, 342-345.

Klein, S. L. (2004). Hormonal and immunological mechanisms mediating sex differences in parasite infection. *Parasite immunology,* **26**, 247-264.

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., Holden, M. T. G., Dougan, G., Bentley, S. D., Parkhill, J. and Peacock, S. J. (2012). Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. *PLoS Pathogens,* **8**(8): e1002824.

Kramer, M. H., Herwaldt, B. L., Craun, G. F., Calderon, R. L. and Juranek, D. D. (1996). Waterborne disease: 1993 and 1994. *Journal American Water Works Association*, **88**, 66-80.

Laishram, S., Kang, G. and Ajjampur, S. S. R. (2012). Giardiasis: A review on assemblage distribution and epidemiology in India. *Indian Journal of Gastroenterology,* **31**, 3-12.

Lalle, M., Pozio, E., Capelli, G., Bruschi, F., Crotti, D. and Cacciò, S. M. (2005). Genetic heterogeneity at the β-giardin locus among human and animal isolates of *Giardia duodenalis* and identification of potentially zoonotic subgenotypes. *International Journal for Parasitology,* **35**, 207-213.

Lasek-Nesselquist, E., Welch, D. M. and Sogin, M. L. (2010). The identification of a new *Giardia duodenalis* assemblage in marine vertebrates and a preliminary analysis of *G. duodenalis* population biology in marine systems. *International Journal for Parasitology*, **40**, 1063-1074.

Laupland, K. B. and Church, D. L. (2005). Population-based laboratory surveillance for *Giardia* sp. and *Cryptosporidium* sp. infections in a large Canadian health region. *BMC Infectious Diseases,* **5**: 72.

Lauwaet, T., Davids, B. J., Reiner, D. S. and Gillin, F. D. (2007). Encystation of *Giardia lamblia*: a model for other parasites. *Current Opinion in Microbiology*, **10**, 554-559.

Lebbad, M., Ankarklev, J., Tellez, A., Leiva, B., Andersson, J. O. and Svärd, S. (2008). Dominance of *Giardia* assemblage B in León, Nicaragua. *Acta Tropica,* **106**, 44-53.

Lebbad, M., Mattsson, J. G., Christensson, B., Ljungström, B., Backhans, A., Andersson, J. O. and Svärd, S. G. (2010). From mouse to moose: Multilocus genotyping of *Giardia* isolates from various animal species. *Veterinary Parasitology,* **168**, 231-239.

Lebbad, M., Petersson, I., Karlsson, L., Botero-Kleiven, S., Andersson, J. O., Svenungsson, B. and Svärd, S. G. (2011). Multilocus genotyping of human *Giardia* isolates suggests limited zoonotic transmission and association between assemblage B and flatulence in children. *PLoS Neglected Tropical Diseases,* **5**(8): e1262.

Levecke, B., Meulemans, L., Dalemans, T., Casaert, S., Claerebout, E. and Geurden, T. (2011). Mixed *Giardia duodenalis* assemblage A, B, C and E infections in pet chinchillas (*Chinchilla lanigera*) in Flanders (Belgium). *Veterinary Parasitology,* **177**, 166-170.

Levy, D. A., Bens, M. S., Craun, G. F., Calderon, R. L. and Herwaldt, B. L. (1998). Surveillance for waterborne-disease outbreaks--United States, 1995-1996. *MMWR. CDC Surveillance Summaries: Morbidity And Mortality*

Weekly Report. CDC Surveillance Summaries / Centers For Disease Control, **47**, 1-34.

Lobo, M. L., Xiao, L., Antunes, F. and Matos, O. (2009). Occurrence of *Cryptosporidium* and *Giardia* genotypes and subtypes in raw and treated water in Portugal. *Letters in Applied Microbiology*, **48**, 732-737.

Maddox-Hyttel, C., Langkjær, R. B., Enemark, H. L. and Vigre, H. (2006). *Cryptosporidium* and *Giardia* in different age groups of Danish cattle and pigs—Occurrence and management associated risk factors. *Veterinary Parasitology*, **141**, 48-59.

Maia-Brigagão, C., Morgado-Díaz, J. A. and De Souza, W. (2012). *Giardia* disrupts the arrangement of tight, adherens and desmosomal junction proteins of intestinal cells. *Parasitology International*, **61**, 280-287.

Majewska, A. C. (1994). Successful experimental infections of a human volunteer and Mongolian gerbils with *Giardia* of animal origin. *Transactions Of The Royal Society Of Tropical Medicine And Hygiene*, **88**, 360-362.

Mallon, M. E., MacLeod, A., Wastling, J. M., Smith, H. and Tait, A. (2003). Multilocus genotyping of *Cryptosporidium parvum* Type 2: population genetics and sub-structuring. *Infection, Genetics And Evolution*, **3**, 207-218.

Mank, T. G., Zaat, J. O. M., Deelder, A. M., Van Eijk, J. T. M. and Polderman, A. M. (1997). Sensitivity of microscopy versus enzyme immunoassay in the laboratory diagnosis of giardiasis. *European Journal of Clinical Microbiology and Infectious Diseases*, **16**, 615-619.

Manser, M., Granlund, M., Edwards, H., Saez, A., Petersen, E., Evengard, B. and Chiodini, P. (2014). Detection of *Cryptosporidium* and *Giardia* in clinical laboratories in Europe-a comparative study. *Clinical Microbiology and Infection*, **20**, 65-71.

Mayrhofer, G., Andrews, R. H., Ey, P. L. and Chilton, N. B. (1995). Division of *Giardia* isolates from humans into two genetically distinct assemblages by electrophoretic analysis of enzymes encoded at 27 loci and comparison with *Giardia muris*. *Parasitology*, **111**, 11-17.

McClelland, E. E. and Smith, J. M. (2011). Gender specific differences in the immune response to infection. *Archivum Immunologiae et Therapiae Experimentalis*, **59**, 203-213.

Minetti, C., Taweenan, W., Hogg, R., Featherstone, C., Randle, N., Latham, S. M. and Wastling, J. M. (2014). Occurrence and Diversity of *Giardia duodenalis* Assemblages in Livestock in the UK. *Transboundary and Emerging Diseases*, **61**(6): e60-e67.

Mintz, E. D., Hudson-Wragg, M., Mshar, P., Cartter, M. L. and Hadler, J. L. (1993). Foodborne giardiasis in a corporate office setting. *Journal of Infectious Diseases,* **167**, 250-253.

Mohammed Mahdy, A. K., Lim, Y. A. L., Johari, S., Wan, K. L. and Hesham Al-Mekhlafi, M. S. (2008). Risk factors for endemic giardiasis: highlighting the possible association of contaminated water and food. *Transactions Of The Royal Society Of Tropical Medicine And Hygiene,* **102**, 465-470.

Monis, P. T. and Andrews, R. H. (1998). Molecular epidemiology: assumptions and limitations of commonly applied methods. *International Journal for Parasitology,* **28**, 981-987.

Monis, P. T., Andrews, R. H., Mayrhofer, G. and Ey, P. L. (2003). Genetic diversity within the morphological species *Giardia intestinalis* and its relationship to host origin. *Infection, Genetics & Evolution,* **3**, 29-38.

Monis, P. T., Caccio, S. M. and Thompson, R. C. A. (2009). Variation in *Giardia*: towards a taxonomic revision of the genus. *Trends In Parasitology,* **25**, 93-100.

Morrison, H. G., McArthur, A. G., Gillin, F. D., Aley, S. B., Adam, R. D., Olsen, G. J., Best, A. A., Cande, W. Z., Chen, F., Cipriano, M. J., Davids, B. J., Dawson, S. C., Elmendorf, H. G., Hehl, A. B., Holder, M. E., Huse, S. M., Kim, U. U., Lasek-Nesselquist, E., Manning, G., Nigam, A., Nixon, J. E. J., Palm, D., Passamaneck, N. E., Prabhu, A., Reich, C. I., Reiner, D. S., Samuelson, J., Svard, S. G. and Sogin, M. L. (2007). Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science,* **317**, 1921-1926.

Naiman, H. L., Sekla, L. and Albritton, W. L. (1980). Giardiasis and other intestinal parasitic infections in a Manitoba residential school for the mentally retarded. *Canadian Medical Association Journal,* **122**, 185-189.

Nash, T. E., Herrington, D. A., Losonsky, G. A. and Levine, M. M. (1987). Experimental human infections with *Giardia lamblia*. *The Journal Of Infectious Diseases,* **156**, 974-984.

Nasser, A. M., Vaizel-Ohayon, D., Aharoni, A. and Revhun, M. (2012). Prevalence and fate of *Giardia* cysts in wastewater treatment plants. *Journal of applied microbiology,* **113**, 477-484.

Naumova, E. N., Jagai, J. S., Matyas, B., DeMaria, A., MacNeill, I. B. and Griffiths, J. K. (2007). Seasonality in six enterically transmitted diseases and ambient temperature. *Epidemiology And Infection,* **135**, 281-292.

Nygård, K., Schimmer, B., Søbstad, Ø., Walde, A., Tveit, I., Langeland, N., Hausken, T. and Aavitsland, P. (2006). A large community outbreak of

waterborne giardiasis-delayed detection in a non-endemic urban area. *BMC Public Health,* **6**, 141-141.

O'Handley, R. M. and Olson, M. E. (2006). Giardiasis and cryptosporidiosis in ruminants. *The Veterinary Clinics Of North America. Food Animal Practice,* **22**, 623-643.

O'Reilly, C. E., Bowen, A. B., Perez, N. E., Sarisky, J. P., Shepherd, C. A., Miller, M. D., Hubbard, B. C., Herring, M., Buchanan, S. D., Fitzgerald, C. C., Hill, V., Arrowood, M. J., Xiao, L. X., Hoekstra, R. M., Mintz, E. D. and Lynch, M. F. (2007). A waterborne outbreak of gastroenteritis with multiple etiologies among resort island visitors and residents: Ohio, 2004. *Clinical Infectious Diseases,* **44**, 506-512.

Office for National Statistics (2013a). Official Labour Market Statistics: Mid-year population estimates, 2011. Available at: http://www.nomisweb.co.uk/.

Office for National Statistics (2013b). Neighbourhood Statistics. Available at: http://www.neighbourhood.statistics.gov.uk/dissemination/.

Olson, M. E., Thorlakson, C. L., Deselliers, L., Morck, D. W. and McAllister, T. A. (1997). *Giardia* and *Cryptosporidium* in Canadian farm animals. *Veterinary Parasitology,* **68**, 375-381.

Olson, M. E., Guselle, N., McAllister, T. A., Goh, J. and Phillips, M. (1999). *Giardia* cyst and *Cryptosporidium* oocyst survival in water, soil, and cattle feces. *Journal of environmental quality,* **28**, 1991-1996.

Olson, M. E., Ceri, H. and Morck, D. W. (2000). *Giardia* vaccination. *Parasitology Today,* **16**, 213-217.

Ouattara, M., N'Guéssan, N. A., Yapi, A. and N'Goran, E. K. (2010). Prevalence and spatial distribution of *Entamoeba histolytica/dispar* and *Giardia lamblia* among schoolchildren in Agboville area (Côte d'ivoire). *PLoS Neglected Tropical Diseases,* **4** (1): e574.

Palmer, C. S., Rees, R., Robertson, I. D., Thompson, R. C. A. and Traub, R. J. (2008). National study of the gastrointestinal parasites of dogs and cats in Australia. *Veterinary Parasitology,* **151**, 181-190.

Panaro, M. A., Cianciulli, A., Mitolo, V., Mitolo, C. I., Acquafredda, A., Brandonisio, O. and Cavallo, P. (2007). Caspase-dependent apoptosis of the HCT-8 epithelial cell line induced by the parasite *Giardia intestinalis*. *FEMS Immunology and Medical Microbiology,* **51**, 302-309.

Pet Food Manufacturers Association (2013). 2011 Pet population and ownership. Available at: http://www.pfma.org.uk/pet-population-2008-2011/.

Public Health Agency Canada (2009). C-EnterNet 2009 Annual Report. National Integrated Enteric Pathogen Surveillance Program. Available at: http://www.phac-aspc.gc.ca/c-enternet/pubs/2009/index-eng.php.

Plutzer, J., Ongerth, J. and Karanis, P. (2010). *Giardia* taxonomy, phylogeny and epidemiology: Facts and open questions. *International Journal of Hygiene and Environmental Health*, **213**, 321-333.

Ponce-Macotela, M., Martínez-Gordillo, M. N., Bermúdez-Cruz, R. M., Salazar-Schettino, P. M., Ortega-Pierres, G. and Ey, P. L. (2002). Unusual prevalence of the *Giardia intestinalis* A-II subtype amongst isolates from humans and domestic animals in Mexico. *International Journal for Parasitology*, **32**, 1201-1202.

Popruk, S., Thima, K., Udonsom, R., Rattaprasert, P. and Sukthana, Y. (2011). Does silent *Giardia* infection need any attention? *Open Tropical Medicine Journal*, **4**, 26-32.

Porter, J. D., Ragazzoni, h. P., Buchanon, J. D., Waskin, H. A., Juranek, D. D. and Parkin, W. E. (1988). *Giardia* Transmission in a Swimming Pool. *American Journal of Public Health*, **78**, 659-662.

Prucca, C. G., Rivero, F. D. and Lujn, H. D. (2011). Regulation of antigenic variation in *Giardia lamblia*. Vol. 65 pp. 611-630.

Puebla, L. J., Núñez, F. A., Fernández, Y. A., Fraga, J., Rivero, L. R., Millán, I. A., Valdés, L. A. and Silva, I. M. (2014). Correlation of *Giardia duodenalis* assemblages with clinical and epidemiological data in Cuban children. *Infection, Genetics And Evolution*, **23**, 7-12.

Rauch, A. M., Van, R., Bartlett, A. V. and Pickering, L. K. (1990). Longitudinal study of *Giardia lamblia* infection in a day care center population. *Pediatric Infectious Disease Journal*, **9**, 186-189.

Read, C., Walters, J., Robertson, I. D. and Thompson, R. C. A. (2002). Correlation between genotype of *Giardia duodenalis* and diarrhoea. *International Journal for Parasitology*, **32**, 229-231.

Read, C. M., Monis, P. T. and Andrew Thompson, R. C. (2004). Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infection, Genetics & Evolution*, **4**, 125-130.

Rendtorff, R. C. (1954). The Experimental Transmission of Human Intestinal Protozoan Parasites. II. *Giardia lamblia* Cysts given in Capsules. *American Journal of Hygiene*, **59**, 209-220.

Rimhanen-Finne, R., Hänninen, M., Vuento, R., Laine, J., Jokiranta, T. S., Snellman, M., Pitkänen, T., Miettinen, I. and Kuusi, M. (2010). Contaminated water caused the first outbreak of giardiasis in Finland, 2007: a descriptive study. *Scandinavian Journal of Infectious Diseases*, **42**, 613-619.

Rimhanen-Finne, R., Sakari Jokiranta, T., Virtanen, M. J. and Kuusi, M. (2011). *Giardia* and *Cryptosporidium* infection in Finland: A registry-based study of their demographic determinants. *APMIS,* **119**, 735-740.

Ringqvist, E., Palm, J. E. D., Skarin, H., Hehl, A. B., Weiland, M., Davids, B. J., Reiner, D. S., Griffiths, W. J., Eckmann, L., Gillin, F. D. and Svärd, S. G. (2008). Release of metabolic enzymes by *Giardia* in response to interaction with intestinal epithelial cells. *Molecular and biochemical parasitology,* **159**, 85-91.

Roberts, L. and Zeibig, E. (2013). Specimen Collection and Processing. In *Clinical Parasitology - a practical approach* (ed. Zeibig, E.), Elsevier, St Louis, Missouri.

Roberts, L. S., Schmidt, G. D. and Janovy, J. (2009). Other Flagellated Protozoa. In *Gerald D. Schmidt & Larry S. Roberts' Foundations of parasitology* Boston, MA ; McGraw-Hill Higher Education, 2009. 8th ed.

Robertson, B., Sinclair, M. I., Forbes, A. B., Veitch, M., Kirk, M., Cunliffe, D., Willis, J. and Fairley, C. K. (2002). Case-control studies of sporadic crytosporidiosis in Melbourne and Adelaide, Australia. *Epidemiology And Infection,* **128**, 419-431.

Robertson, I. D., Thompson, R. C. A., Lymbery, A. J. and Irvin, P. J. (2000). The role of companion animals in the emergence of parasitic zoonoses. *International Journal for Parasitology,* **30**, 1369-1377.

Robertson, L. J., Forberg, T., Hermansen, L., Gjerde, B. K. and Langeland, N. (2007a). Molecular characterisation of *Giardia* isolates from clinical infections following a waterborne outbreak. *Journal of Infection,* **55**, 79-88.

Robertson, L. J. (2007b). The potential for marine bivalve shellfish to act as transmission vehicles for outbreaks of protozoan infections in humans: A review. *International Journal of Food Microbiology,* **120**, 201-216.

Robertson, L. J., Hanevik, K., Escobedo, A. A., Mørch, K. and Langeland, N. (2010). Giardiasis--why do the symptoms sometimes never stop? *Trends In Parasitology,* **26**, 75-82.

Robertson, L. J., Sprong, H., Ortega, Y. R., van der Giessen, J. W. B. and Fayer, R. (2014). Impacts of globalisation on foodborne parasites. *Trends In Parasitology,* **30**, 37-52.

Rodríguez-Hernández, J., Canut-Blasco, A. and Martín-Sánchez, A. M. (1996). Seasonal prevalences of *Cryptosporidium* and *Giardia* infections in children attending day care centres in Salamanca (Spain) studied for a period of 15 months. *European Journal of Epidemiology,* **12**, 291-295.

Roxström-Lindquist, K., Palm, D., Reiner, D., Ringqvist, E. and Svärd, S. G. (2006). *Giardia* immunity – an update. *Trends In Parasitology*, **22**, 26-31.

Ryan, U. and Cacciò, S. M. (2013). Zoonotic potential of *Giardia*. *International Journal for Parasitology*, **43**, 943-956.

Sahagún, J., Clavel, A., Goñi, P., Seral, C., Llorente, M. T., Castillo, F. J., Capilla, S., Arias, A. and Gómez-Lus, R. (2008). Correlation between the presence of symptoms and the *Giardia duodenalis* genotype. *European Journal of Clinical Microbiology and Infectious Diseases*, **27**, 81-83.

Santín, M., Cortés Vecino, J. A. and Fayer, R. (2013). A large scale molecular study of *Giardia duodenalis* in horses from Colombia. *Veterinary Parasitology*, **196**, 31-36.

Savioli, L., Smith, H. and Thompson, A. (2006). *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. *Trends In Parasitology*, **22**, 203-208.

Scaramozzino, P., Mazzanti, S., Scholl, F., De Liberato, C., Spaziani, A., Di Cave, D., Berrilli, F. and D'Orazi, C. (2009). A study of the prevalence and genotypes of *Giardia duodenalis* infecting kennelled dogs. *Veterinary Journal*, **182**, 231-234.

Singer, S. M., Elmendorf, H. G., Conrad, J. T. and Nash, T. E. (2001). Biological selection of variant-specific surface proteins in *Giardia lamblia*. *Journal of Infectious Diseases*, **183**, 119-124.

Smith, H. V., Cacciò, S. M., Tait, A., McLauchlin, J. and Thompson, R. C. A. (2006a). Tools for investigating the environmental transmission of *Cryptosporidium* and *Giardia* infections in humans. *Trends In Parasitology*, **22**, 160-167.

Smith, A., Reacher, M., Smerdon, W., Adak, G. K., Nichols, G. and Chalmers, R. M. (2006b). Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992-2003. *Epidemiology And Infection*, **134**, 1141-1149.

Smith, H. V., Cacciò, S. M., Cook, N., Nichols, R. A. B. and Tait, A. (2007). *Cryptosporidium* and *Giardia* as foodborne zoonoses. *Veterinary Parasitology*, **149**, 29-40.

Snel, S. J. B. M. G. V. K. (2009a). The epidemiology of giardiasis in New Zealand, 1997–2006. *The New Zealand Medical Journal*, **122**, 62-75.

Snel, S. J., Baker, M. G., Kamalesh, V., French, N. and Learmonth, J. (2009b). A tale of two parasites: the comparative epidemiology of cryptosporidiosis and giardiasis. *Epidemiology & Infection*, **137**, 1641-1650.

Soares, R. M., de Souza, S. L. P., Silveira, L. H., Funada, M. R., Richtzenhain, L. J. and Gennari, S. M. (2011). Genotyping of potentially zoonotic *Giardia*

duodenalis from exotic and wild animals kept in captivity in Brazil. *Veterinary Parasitology*, **180**, 344-348.

Solaymani-Mohammadi, S. and Singer, S. M. (2010). *Giardia duodenalis*: The double-edged sword of immune responses in giardiasis. *Experimental Parasitology*, **126**, 292-297.

Soliman, R. H., Fuentes, I. and Rubio, J. M. (2011). Identification of a novel Assemblage B subgenotype and a zoonotic Assemblage C in human isolates of *Giardia intestinalis* in Egypt. *Parasitology International*, **60**, 507-511.

Sousa, M. C., Morais, J. B., Machado, J. E. and Poiares-Da-Silva, J. (2006). Genotyping of *Giardia lamblia* human isolates from Portugal by PCR-RFLP and sequencing. *Journal of Eukaryotic Microbiology*, **53**, S174-S176.

Sprong, H., Cacciò, S. M. and van der Giessen, J. W. B. (2009). Identification of zoonotic genotypes of *Giardia duodenalis*. *PLoS Neglected Tropical Diseases*, **3**, 1-12.

Stark, D., van Hal, S., Marriott, D., Ellis, J. and Harkness, J. (2007). Irritable bowel syndrome: A review on the role of intestinal protozoa and the importance of their detection and diagnosis. *International Journal for Parasitology*, **37**, 11-20.

Steketee, R. W., Reid, S., Cheng, T., Stoebig, J. S., Harrington, R. G. and Davis, J. P. (1989). Recurrent Outbreaks of Giardiasis in a Child Day Care Center, Wisconsin. *American Journal of Public Health*, **79**, 485-490.

Stuart, J. M., Orr, H. J., Warburton, F. G., Jeyakanth, S., Pugh, C., Morris, I., Sarangi, J. and Nichols, G. (2003). Risk Factors for Sporadic Giardiasis: A Case-Control Study in Southwestern England. *Emerging Infectious Diseases*, **9**, 229.

Sulaiman, I. M., Fayer, R., Bern, C., Gilman, R. H., Trout, J. M., Schantz, P. M., Das, P., Lal, A. A. and Lihua, X. (2003). Triosephosphate Isomerase Gene Characterization and Potential Zoonotic Transmission of *Giardia duodenalis*. *Emerging Infectious Diseases*, **9**, 1444-1452.

Swaminathan, A., Torresi, J., Schlagenhauf, P., Thursky, K., Wilder-Smith, A., Connor, B. A., Schwartz, E., Vonsonnenberg, F., Keystone, J. and O'Brien, D. P. (2009). A global study of pathogens and host risk factors associated with infectious gastrointestinal disease in returned international travellers. *The Journal Of Infection*, **59**, 19-27.

Takumi, K., Swart, A., Mank, T., Lasek-Nesselquist, E., Lebbad, M., Cacciò, S. M. and Sprong, H. (2012). Population-based analyses of Giardia duodenalis is consistent with the clonal assemblage structure. *Parasites & Vectors*, 168.

Tam, C. C., Rodrigues, L. C. and O'Brien, S. J. (2003). The study of infectious intestinal disease in England: What risk factors for presentation to general

practice tell us about potential for selection bias in case-control studies of reported cases of diarrhoea. *International journal of epidemiology,* **32**, 99-105.

Tam, C. C., Rodrigues, L. C., Viviani, L., Dodds, J. P., Evans, M. R., Hunter, P. R., Gray, J. J., Letley, L. H., Rait, G., Tompkins, D. S. and O'Brien, S. J. (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.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology And Evolution,* **30**, 2725-2729.

Taylor, M. A., Coop, R. L. and Wall, R. L. (2007). *Veterinary parasitology,* Oxford : Blackwell Pub., 2007. 3rd ed.

Taweenan, Weerapol (2010). Molecular epidemiology of *Giardia duodenalis*. PhD thesis. University of Liverpool, United Kingdom.

Thompson, R. C. A. and Monis, P. T. (2004). Variation in *Giardia*: Implications for Taxonomy and Epidemiology. *Advances in Parasitology,* **58**, 69-137.

Thompson, R. C. A., O'Handley, R. and Palmer, C. S. (2008). The public health and clinical significance of *Giardia* and *Cryptosporidium* in domestic animals. *Veterinary Journal,* **177**, 18-25.

Thompson, R. C. A., Lymbery, A. J. and Smith, A. (2010). Parasites, emerging disease and wildlife conservation. *International Journal for Parasitology,* **40**, 1163-1170.

Thompson, R. C. A. and Monis, P. (2012). *Giardia*-From Genome to Proteome. *Advances in Parasitology,* **78**, 57-95.

Traub, R. J., Monis, P. T. and Robertson, I. D. (2005a). Molecular epidemiology: A multidisciplinary approach to understanding parasitic zoonoses. *International Journal for Parasitology,* **35**, 1295-1307.

Traub, R., Wade, S., Read, C., Thompson, A. and Mohammed, H. (2005b). Molecular characterization of potentially zoonotic isolates of *Giardia duodenalis* in horses. *Veterinary Parasitology,* **130**, 317-321.

Traub, R. J., Monis, P. T., Robertson, I., Irwin, P. and Mencke, N. (2004). Epidemiological and molecular evidence supports the zoonotic transmission of *Giardia* among humans and dogs living in the same community. *Parasitology,* **128**, 253.

Traub, R. J., Inpankaew, T., Reid, S. A., Sutthikornchai, C., Sukthana, Y., Robertson, I. D. and Thompson, R. C. A. (2009). Transmission cycles of *Giardia duodenalis* in dogs and humans in Temple communities in Bangkok-

a critical evaluation of its prevalence using three diagnostic tests in the field in the absence of a gold standard. *Acta Tropica,* **111**, 125-132.

Trout, J. M., Santín, M., Greiner, E. and Fayer, R. (2005). Prevalence and genotypes of *Giardia duodenalis* in post-weaned dairy calves. *Veterinary Parasitology,* **130**, 177-183.

Trout, J. M., Santín, M., Greiner, E. and Fayer, R. (2004). Prevalence of *Giardia duodenalis* genotypes in pre-weaned dairy calves. *Veterinary Parasitology,* **124**, 179-186.

Trout, J. M., Fayer, R. and Santin, M. (2007). Prevalence of *Giardia duodenalis* genotypes in adult dairy cows. *Veterinary Parasitology,* **147**, 205-209.

Tzannes, S., Wastling, J., German, A. J., Pinchbeck, G. L., Batchelor, D. J. and Graham, P. A. (2008). Prevalence of *Cryptosporidium*, *Giardia* and *Isospora* species infections in pet cats with clinical signs of gastrointestinal disease. *Journal of Feline Medicine and Surgery,* **10**, 1-8.

Veronesi, F., Passamonti, F., Cacciò, S., Diaferia, M. and Piergili Fioretti, D. (2010). Epidemiological survey on equine *Cryptosporidium* and *Giardia* infections in Italy and molecular characterization of isolates. *Zoonoses And Public Health,* **57**, 510-517.

Verweij, J. J., Schinkel, J., Laeijendecker, D., Van Rooyen, M. A. A., Van Lieshout, L. and Polderman, A. M. (2003). Real-time PCR for the detection of *Giardia lamblia*. *Molecular and Cellular Probes,* **17**, 223-225.

Verweij, J. J. (2014). Application of PCR-based methods for diagnosis of intestinal parasitic infections in the clinical laboratory. *Parasitology*. Published ahead of print. Available at http://dx.doi.org/10.1017/S0031182014000419.

Warburton, A. R., Jones, P. H. and Bruce, J. (1994). Zoonotic transmission of giardiasis: a case control study. *Communicable Disease Report. CDR Review,* **4**, R32-R36.

Wensaas, K. A., Langeland, N., Hanfvevik, K., Mørch, K., Eide, G. E. and Rortveit, G. (2012). Irritable bowel syndrome and chronic fatigue 3 years after acute giardiasis: Historic cohort study. *Gut,* **61**, 214-219.

World Health Organisation (1979). Parasitic Zoonoses. Report of a WHO Expert Committee with the participation of FAO. Technical Report Series No. 637. *World Health Organization, Geneva*. Available at: http://apps.who.int/iris/handle/10665/41353.

World Health Organisation (2008). Guidelines for Drinking-water Quality, incorporating first and second addenda to third edition, Volume 1, Recommendations. *World Health Organization, Geneva*. Available at: http://www.who.int/water_sanitation_health/dwq/gdwq3rev/en/.

Wielinga, C. M. and Thompson, R. C. A. (2007). Comparative evaluation of Giardia duodenalis sequence data. *Parasitology*, **134**, 1795-1821.

Winkworth, C. L., Learmonth, J. J., Matthaei, C. D. and Townsend, C. R. (2008). Molecular characterization of *Giardia* isolates from calves and humans in a region in which dairy farming has recently intensified. *Applied And Environmental Microbiology*, **74**, 5100-5105.

Xiao, L. and Fayer, R. (2008). Molecular characterisation of species and genotypes of *Cryptosporidium* and *Giardia* and assessment of zoonotic transmission. *International Journal for Parasitology*, **38**, 1239-1255.

Xiao, L. and Herd, R. P. (1994). Epidemiology of equine *Cryptosporidium* and *Giardia* infections. *Equine veterinary journal*, **26**, 14-17.

Xiao, L., Herd, R. P. and Bowman, G. L. (1994). Prevalence of *Cryptosporidium* and *Giardia* infections on two Ohio pig farms with different management systems. *Veterinary Parasitology*, **52**, 331-336.

Yang, R., Reid, A., Lymbery, A. and Ryan, U. (2010). Identification of zoonotic *Giardia* genotypes in fish. *International Journal for Parasitology*, **40**, 779-785.

Yoder, J. S., Harral, C. and Beach, M. J. (2010). Giardiasis surveillance - United States, 2006-2008. *MMWR. Surveillance Summaries: Morbidity And Mortality Weekly Report. Surveillance Summaries / CDC*, **59**, 15-25.

Younas, M., Shah, S. and Talaat, A. (2008). Frequency of *Giardia lamblia* infection in children with recurrent abdominal pain. *JPMA. The Journal Of The Pakistan Medical Association*, **58**, 171-174.