Science for Communities

Quantitative Microbial Risk Assessment Pilot Study

September 2020

Authors:

Margaret Leonard (ESR)
Brent Gilpin (ESR)
Beverley Horn (ESR)
Sarah Coxon (ESR)
Bridget Armstrong (ESR)
Joanne Hewitt (ESR)
Megan Devane (ESR)
Pradip Gyawali (ESR)
Paula Scholes (ESR)
Kirstin Thom (ESR)
Angela Cornelius (ESR)
Ashley Orton (ESR)
David Hayman (Massey University)
Anthony Pita (Massey University)

Prepared for: **Ministry for the Environment**

Client report No: **FW20007**

Reviewers: **Dr Richard Muirhead, AgResearch**

Professor Gillian Lewis, University of Auckland

Acknowledgements

We would like to thank Juliet Milne, NIWA, who liaised with the Councils to select suitable sites for the Pilot study and prepared the contact list.

We would like to acknowledge and thank the council staff who undertook the sampling, braving some unpleasant weather and frustrations with courier services and the Council managers who have supported this project in principle and made the resources available.

We would like to thank staff from the following Regional, District and City councils, for their time and efforts in undertaking sample collection for this study.

Auckland Council

Bay of Plenty Regional Council

Environment Canterbury

Environment Southland

Gisborne District Council

Greater Wellington Regional Council

Horizons Regional Council

Marlborough District Council

Nelson City Council

Northland Regional Council

Taranaki Regional Council

Waikato Regional Council

West Coast Regional Council

We would also like to thank iwi and hapū who generously gave their time and knowledge of their awa.

We would also like to acknowledge our reviewers who provided excellent feedback.

Manager

Wim Nijhof
Group Leader
Human & Ecological Health

Author

Dr Margaret Leonard

Senior Scientist
Water and Wastewater

DISCLAIMER

The Institute of Environmental Science and Research Limited (ESR) has used all reasonable endeavours to ensure that the information contained in this client report is accurate. However, ESR does not give any express or implied warranty as to the completeness of the information contained in this client report or that it will be suitable for any purposes other than those specifically contemplated during the Project or agreed by ESR and the Client.

Contents

List of Tables

List of Figures

1 Executive summary

The current Freshwater Recreational Guidelines (MfE, 2003) and the National Policy Statement for Freshwater Management (NPS-FM 2020) (New Zealand Government 2020) are based on data collected during the 1998-2000 Freshwater Microbiology Research Programme (FMRP), a Quantitative Microbial Risk Assessment[1] (QMRA) and an association found between the pathogen *Campylobacter* and faecal indicator bacteria *Escherichia coli* (McBride et al, 2002). There have been significant changes in land use and waste management practices in the 20 years since the FMRP, raising some uncertainty over the relevance of the FMRP-derived relationships between pathogens and indicators. In view of the fundamental importance of these relationships, and the need for further confidence in the guidelines they underpin, an understanding of the current prevalence of pathogens, and relevance to human health risk is required.

This report describes a pilot study undertaken to inform the options for the design, number of samples, costs, and logistics for a large-scale replacement study for the 1998-2000 FMRP and QMRA. The pilot also enabled a selection of new methodologies to be trialled. Due to the limited number of samples collected, the aims of the pilot did not include establishing the current state of concentrations of pathogens and faecal indicators in the rivers sampled.

There has also been a shift to acknowledge Māori values, their intergenerational perspectives and to incorporate these in science/ policy responses. The NPS-FM 2020 is centred on giving effect to Te Mana o Te Wai, which needs to be taken into account in the formulation of the next stage of this project.

1.1 Pilot study

1.1.1 Indicators and pathogens in New Zealand rivers in 2020

Between February and March 2020, 52 water samples were collected from 16 rivers (initially characterised as six urban, five dairy farming and five sheep & beef farming) from around New Zealand. They were analysed for the levels of *Escherichia coli*, enterococci, *Campylobacter*, *Salmonella*, Shiga toxin-producing *E. coli* (STEC), *Cryptosporidium*, *Giardia*, norovirus, enterovirus, adenovirus and a set of faecal source tracking (FST) markers (human, ruminant and wildfowl). The river sampling sites were selected on the basis the site regularly had elevated *E. coli* concentrations recorded during previous monitoring.

Key results of the pilot study were:

- *E. coli* were detected in all samples, with 17 samples from eight different rivers having ≥1,000 *E. coli* MPN/100 mL.
- All the rivers contained wildfowl markers, and were the only source identified in 10 of the samples. Four of these samples were taken from two of the rivers, which based on observed land use, would have been characterised as sheep & beef or dairy farming.

[1] A QMRA is a framework to combine information on the particular pathogen(s) and the potential dose or exposure (a function of the concentration of pathogens in the water and the volume of water that might be ingested during recreation), to estimate the risk of infection and illness

- All six urban rivers contained human FST markers, in all the samples tested. However, five of the samples from two of the urban rivers also contained significant concentrations of ruminant FST markers.
- Conversely, while the ruminant FST marker confirmed the observed land use of four dairy farming, and four sheep & beef farming rivers, two of these rivers also contained human FST markers in at least some of the samples tested.
- Using a combination of culture and quantitative PCR (qPCR)-based methodologies, *Campylobacter* were detected in 39 samples from 14 rivers, *Salmonella* in 12 samples from eight rivers, and STEC in 13 samples from 11 rivers. Using culture methods, the maximum level of *Campylobacter* was 92 MPN /100ml, with 85% of detections ≤4.1 MPN/100mL. *Salmonella* and STEC detections by MPN culture were all ≤0.25 MPN/100ml.
- Whole genome sequencing confirmed likely pathogenic strains of *Campylobacter, Salmonella* and STEC.
- Using traditional microscopy techniques *Giardia* was detected in 42 samples from 15 rivers, with 40 samples having ≤ 24 cysts/100 L, and a maximum of 250 cysts/100 L. C*ryptosporidium* was detected in 22 samples from 12 rivers with a maximum of 31 oocysts/100 L. Quantitative PCR only detected *Giardia* in one sample. *Cryptosporidium parvum* and *C. hominis*, were not detected in any sample by qPCR.
- Human adenovirus (HAdV) was not detected in any sample, while noroviruses were detected in five rivers, and enterovirus in two rivers. However, the concentration of viruses in the samples were too low to quantify and only one sample contained both viruses.

Direct comparisons with the 1998-2000 FMRP survey (McBride et al, 2002) are limited by differences in detection levels, methodology, pilot study sample size and targeting of more contaminated rivers in 2020. With those caveats in mind:

- The pilot study samples had higher median concentrations of *E. coli*, being more like the samples from FMRP with higher concentrations of *E. coli*.
- The pilot study had a higher prevalence of *Giardia* and *Cryptosporidium*, however the concentrations detected were similar to FMRP.
- Adenovirus were detected in a third of FMRP samples but were not detected in the pilot study.
- In the FMRP 9% of samples had *Campylobacter* >110 MPN/100mL, while the maximum level in the pilot study was 92 MPN/100mL. Adjusting for differences in detection limits, the frequency of detection of *C. jejuni* was higher in the pilot study (40%), than the FMRP (30%).

The key conclusions from these observations are:

- Potentially pathogenic micro-organisms were detected in 94% of the samples, but generally at very low concentrations. Above water quality criterion for *E. coli* of 540 MPN/100 mL, the prevalence of pathogens increased compared to below the criterion.
- Half the samples with viruses detected were associated with human contaminated samples, but the other half were from rivers without obvious human sources of contamination.
- Faecal source tracking markers were readily detected in almost all the samples from all the rivers, allowing a source identification.

- Land use has been shown to influence variations in pathogens and provides information on potential sources important for management. Faecal source tracking was a tool that was unavailable in 1998-2000, and in this pilot study confirmed the frequent occurrence of more than one source of contamination, and also that observed land use does not always match the sources of contamination. This is important information for effective management of water quality. Future studies should include both visual confirmation of land use and FST.

1.1.2 qPCR methodology as an alternative to traditional microbiological analyses

An aim of the project was to determine if qPCR could give similar or better sensitivity for enumerating bacteria and protozoa against traditional culture and microscope-based techniques.

The correlations between *E. coli* culture and qPCR (R^2= 0.85) were strong and calibration curves are presented. A reasonable correlation was determined with enterococci and qPCR for samples which had high concentrations of human FST only. Where there were high concentrations of ruminant and human FST, the correlation was not as strong. Correlations between bacterial and protozoan pathogens were not achieved, which is likely because most were only detected at low concentrations. The combination of qPCR and culture increased the number of samples where *Campylobacter*, *Salmonella* and STEC were detected and therefore provides a better understanding of the pathogens present.

Quantitative PCR is likely to detect pathogens when present at higher concentrations, and allows estimation of pathogen presence at lower concentrations. The lack of an isolate however means whole genome sequencing can not be used to confirm pathogenic potential. Further calibration work is required to convert qPCR results to a concentration for input into the QMRA.

1.1.3 Iwi engagement

The pilot study initiated engagement with iwi and hapū for each site resulting in kōrero with 21 iwi and hapū about the project and their awa. This kōrero needs to be developed and extended to co-design the larger study and a new framework which encompasses mātauranga Māori.

1.2 Full study for QMRA of recreational freshwater

A large-scale study plan akin to the 1998-2000 FMRP has been detailed and costings provided to the Ministry for the Environment. It is proposed to include *E. coli,* enterococci, with *Campylobacter* as the target pathogen as well as *Salmonella* and STEC. Protozoa are included as desirable, but challenging due to sampling logistics. There is limited evidence from this study or others that virus concentrations will correlate with other indicator organisms, except at very high concentrations (Korajkic et al, 2018). At those concentrations (where faecal source tracking markers and indicators are high and easily detected) the risks from viruses and protozoa can be predicted much more readily based on assessment of faecal indicators and sources using risk assessment approaches. Therefore it is proposed that viruses are excluded, which will reduce costs considerably.

Iwi should be engaged in site selection to ensure that sites of cultural significance are included. These sites may not have a history of testing as Council selection criteria for monitoring may differ.

1.3 Alternative considerations

To make a significant impact in the near future on improving recreational water quality management, we propose consideration of complimentary work streams to strengthen existing

freshwater recreational water quality guidelines to develop a framework. The three step framework would support evidence based decisions on how to investigate water quality when guidelines are exceeded, supported by worked examples and could be undertaken in phases.

As illustrated in Figure 1 there are three key components to this.

- Step 1: Detection of faecal contamination using indicators such as *E. coli* (possibly in conjunction with enterococci) and other chemical or environmental measures to compare against guidelines.
- Step 2: Explicit guidance on how to investigate water quality to determine sources of contamination, and therefore appropriate mitigations or interventions.
- Step 3: If the source analysis suggests low risk sources of pollution, or sources for which mitigations or interventions are politically, socially or economically challenging, then site specific health risk assessment using tools such as QMRA and scenario modelling may be required to reclassify risk, or to support necessary interventions when the guideline expectations are not met.

This needs direct linkage with industry and research programmes such as Our Land & Water to guide intervention options, and ongoing monitoring programmes for trend analysis of water quality. Most importantly this needs to give effect to Te Mana o Te Wai and incorporate Māori values and approaches such as a Cultural Health Index (CHI).

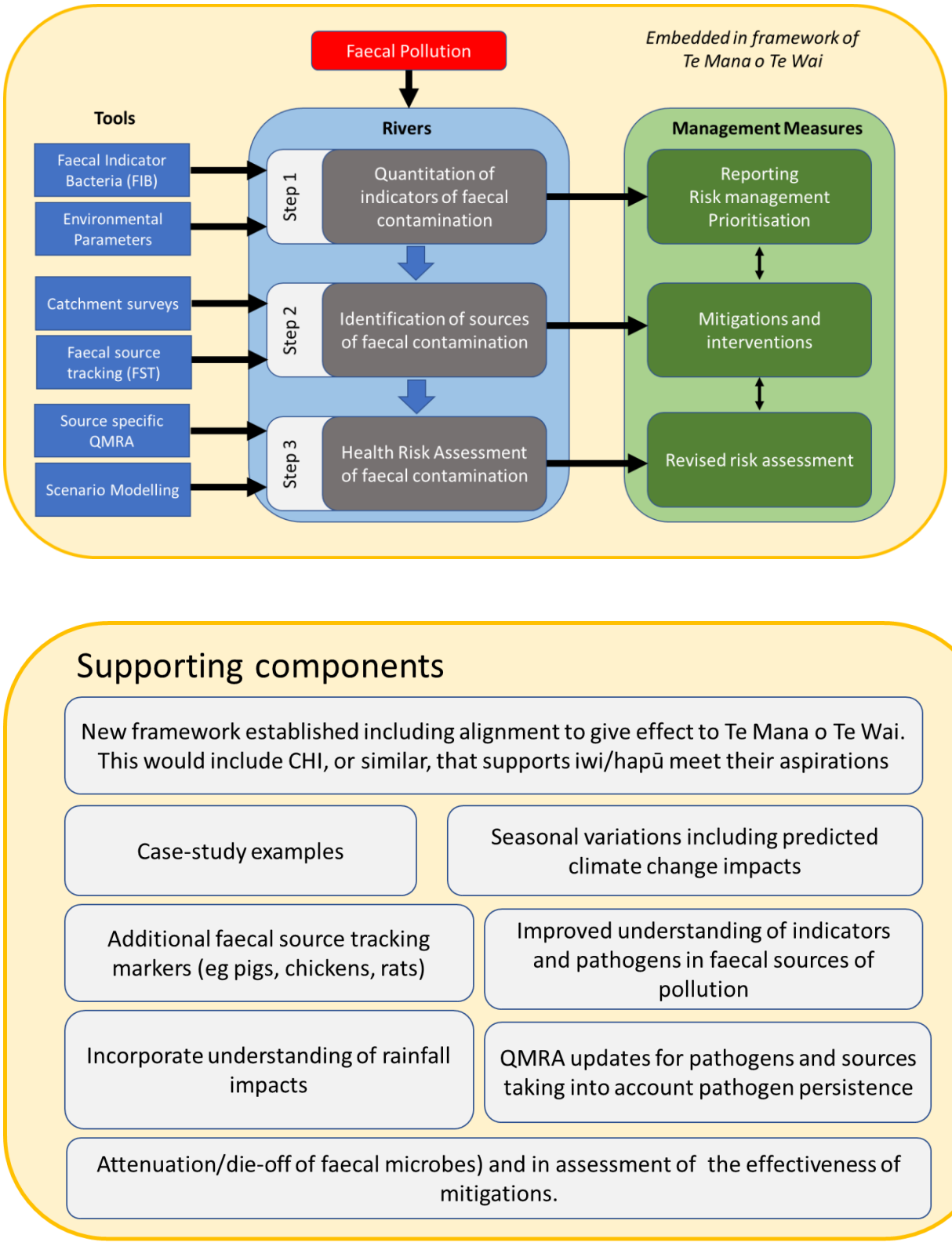

Figure 1: **Conceptual framework for water quality assessment using quantitative microbial risk assessment with supporting components. Wire diagram adapted from Savio et al, 2018.**

2 Introduction

2.1 Pilot study aims

The aims of the pilot study were to:

- generate new data on the concentrations of pathogens present in New Zealand rivers
- determine if qPCR methods can achieve equivalent or better sensitivity for enumeration of bacteria and protozoa against traditional culture-based and microscope analysis methods
- initiate engagement with local iwi for each river where there is a sampling site
- refine logistics for sampling and analysis
- accurately cost the full study, including a full study plan.

This pilot study would allow determination of the likely statistical power of the full study including whether a smaller number of samples would achieve necessary outcomes for a Quantitative Microbial Risk Assessment[2] (QMRA).

2.2 Pathogens in recreational water

The presence of microbial pathogens (bacteria, viruses and protozoa) in recreational waters poses a health risk to those using the water for swimming, food gathering, and other primary, or other contact activities, such as kayaking or waka ama. The majority of waterborne pathogens that cause human illness, including *Campylobacter*, *Salmonella*, enteric viruses, *Giardia* and *Cryptosporidium*, are associated with human and/or animal faeces. Contamination of waterways with faecal material may result from the discharge of inadequately treated sewage, leaking sewage pipes, combined sewage-stormwater discharges, septic tank discharges or leaks, run-off from urban and/or agricultural land, and direct deposition from farm or wild animals. Health effects resulting from contact with contaminated water include gastrointestinal illness, respiratory illness and skin infections.

In 2016, a total of 16,305 notifications were reported through New Zealand's notifiable disease database, EpiSurv (ESR, 2017). Over 5,400 of these cases were asked whether they had recreational water contact leading up to their illness, with 21% of them answering in the affirmative. Recreational water includes rivers, lakes, ocean, swimming pools and spas. The highest rates of recreational water contact were for cases caused by *Giardia* (33% of respondents), *Cryptosporidium* (26%) and STEC (26%), while reported rates were lower for *Salmonella* (20%) and *Campylobacter* (16%). Recreational water contact is one of a number of possible risk factors that cases can be exposed to and commonly cases will have more than one possible risk factor indicated in the notification. The source of the pathogen may be difficult to confirm and the indication of a risk factor on a notification does not confirm the risk factor was associated with the case becoming ill.

The pathogens selected for analysis for this pilot study, which were determined in a previous report (Gilpin et al, 2018) are:

- Bacterial pathogens
  - *Campylobacter*

[2] A QMRA is a framework to combine information on the particular pathogen(s) and the potential dose or exposure (a function of the concentration of pathogens in the water and the volume of water that might be ingested during recreation), to estimate the risk of infection and illness.

  - *Salmonella*
  - Shiga toxin-producing *E. coli* (STEC)
- Enteric viruses
  - adenovirus
  - norovirus types GI and GII
  - enterovirus
- Protozoa
  - *Giardia*
  - *Cryptosporidium.*

A brief summary of the individual pathogens, their role in human illness and presence in water is given in Appendix A.

2.3 Indicator organisms

The microbial quality of recreational waters is monitored to protect public health. Direct monitoring for the presence of pathogens in water is impractical, as pathogens tend to be present in a population intermittently. Once excreted pathogen concentrations decrease from die off, attenuation and dilution in the river and so they are likely to occur at low concentrations. Analyses are complex and expensive. Rather, 'indicator organisms' are used to monitor microbial water quality. Indicator organisms are not usually pathogenic themselves. However, as they are typically found in the intestinal tract of warm-blooded animals, they are indicative of faecal contamination and the potential presence of pathogens. The most commonly used indicators of faecal contamination are *Escherichia coli* (*E. coli*), faecal coliforms and enterococci, whose presence is quick, cheap and easy to test compared to pathogen analysis, and there are robust, standard methods to provide consistency of results allowing data comparison.

An assumption in using indicator organisms as a proxy for health risk is that the presence and concentration of microbial indicators varies consistently with that of pathogens (Harwood et al, 2014). In the literature review by Korajkic et al (2018), 23 studies examined the indicator-pathogen relationship in freshwaters, only 13 reported a statistically significant relationship between at least one indicator and at least one pathogen. *E. coli* was the indicator that had the greatest number of significant pathogen relationships, while *Cryptosporidium* and *Giardia* were the pathogens most commonly reported to correlate with indicators, followed by STEC, *Salmonella* and *Campylobacter*. Significant correlations between indicator organisms and viral pathogens were reported less frequently.

Faecal indicator organisms, including *E. coli* and enterococci, have been used to assess the quality of recreational waters for over a century (Korajkic et al, 2018), and despite the limitations involved in using indicator organisms as a proxy for health risk, they remain an important tool in monitoring the suitability of water for recreational use.

2.4 Derivation of current New Zealand guidelines

The 2003 Ministry for the Environment and Ministry of Health Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas "The Guidelines" (MfE, 2003) were developed to help water managers monitor, report on and control the public health risk posed by microbiological contamination of recreational waters. The freshwater component of the Guidelines uses the indicator organism *E. coli* to assess water quality, with numeric guideline values developed from the findings of the 1998-2000 FMRP (McBride et al, 2002).

The FMRP was designed to include concerns at the time about zoonotic pathogens from point and diffuse sources of animal waste from the agricultural industry as well as contributions from feral mammals, birds and human sources. The 25 sites in the nationwide survey were chosen to represent different land uses and associated faecal impacts (dairy farming, beef and sheep farming, municipal, wildfowl and forested/undeveloped). Microbial water quality was determined from water samples collected fortnightly for 15 months and analysed for 10 pathogens and indicators.

The Guidelines were developed using the FMRP study data (MfE, 2003). The Guideline bands were determined by matching the percentiles of the risk of campylobacteriosis illness derived from a QMRA using the *Campylobacter* data recorded by the study, with the percentiles of *E.coli* concentrations recorded by the study. *Campylobacter* was chosen because this pathogen was most frequently detected in the FMRP study and a medium correlation was observed with the indicator *E. coli* using Spearman's rank correlation coefficient. The interpretation of faecal indicator bacteria concentrations with respect to risks associated with recreation in New Zealand freshwaters are likely to be unique compared to other locations around the world, due to the strong influence of mixes of human and rural contamination sources present in New Zealand relative to those in other countries. Waterways are graded A-D, according to the 95th percentile concentration of *E. coli* over a five year period and a sanitary survey of the catchment. In line with the World Health Organisation (WHO) Annapolis approach (WHO, 1999), grading was complimented by surveillance and the actions were required where elevated concentrations of *E. coli* were measured in a single sample. For example when a single sample was above 550 MPN/100 mL daily sampling was required and the public advised of a health risk.

In 2014, the National Policy Statement Freshwater Management (NPS-FM 2014) (New Zealand Government 2014) introduced freshwater management objectives for water quality including recreational water. Suitability for recreational use was described using *E. coli* as the "attribute" - a measurable characteristic of fresh water, which supports particular values (ie recreational use). The "attribute state" is the level to which an attribute is to be managed and again there were four states A (best) to D (worst). The risk of *Campylobacter* infection was estimated for each grade. The numeric attribute states specified the 95th percentile *E. coli* concentrations. These risk assessments and numeric values were based on assessment of data collected in the previous study from McBride et al (2002). The attribute states were recalculated in 2017 National Policy Statement Freshwater Management (NPS-FM 2017) (New Zealand Government 2017) and defined five colour bands specified by the median *E. coli*, the percentage of exceedances above 260/100 mL and above 540/100 mL, and the 95th percentile of the *E. coli* concentration (60 samples from a maximum of 5 years). As well as a 95th percentile value the attribute states were defined as Blue (best), Green, Yellow, Orange and Red (worst) and the risk profile is described for each attribute state.

Bathing season surveillance sampling, responding to the results from individual samples, which was not included in 2014, was reintroduced in 2017 with requirements for action to investigate *E. coli* results greater than 260/100 mL, and to inform the public that the site is unsuitable for recreation if individual *E. coli* results are greater than 540/100 mL.

McBride and Soller (2017) discuss the differences between the different values used to classify recreational water in The Guidelines 2003, NPS-FM 2014 and NPS-FM 2017. The 2017 grade Blue is similar to 2003 Guideline Grade C and NPS-FM 2014 Grade B, while the Orange grade is more restrictive than the NPS-FM 2014 Grade C. The minimum grade acceptable for swimming is the bottom of the yellow grade. By specifying that the grading is determined from more than 60 samples collected over five years, irrespective of weather conditions, it is noted that the concentrations of *E. coli* may be elevated.

These attribute states have not changed in the new NPS-FM 2020, but attribute states (Excellent, Good, Fair, Poor) have been added during the bathing season. These attributes are assigned according to the 95th percentile *E. coli* concentrations. A national bottom line is specified which is 540 *E. coli* /100 mL, below this value, sites are classified as "Poor". If a desired attribute state is not met then Council must prepare an Action plan to achieve the attribute state within a specified timeframe.

2.5 Overseas guidelines

The United States of America (USA) and Canadian Recreational Water Guidelines for indicator bacteria are associated with illness rates derived from epidemiological/microbiological studies at lake swimming sites (Health Canada, 2012; USEPA, 2012). None of these studies were for rivers. The guidelines are also based on human pollution sources which are consistently higher risk than animal sources (Soller et al, 2010a, Soller et al, 2010b).

A review of the risk assessments underpinning the new attribute states for NPS-FM 2017 was undertaken, discussing the implications of the different gradings and an assessment of how many rivers would meet the NPS-FM 2017 criteria (McBride and Soller, 2017). The report includes a table comparing the different New Zealand gradings and those of the USA, European Union (UN) and WHO. Table 1 shows the numeric values for 2003 Guidelines and the 2014, and 2017 iterations of the NPS_FM, against the risk profile for the number of *Campylobacter* infections. The actual risk varies slightly to descriptions in the NPS-FM Attribute table, as it includes longer term water quality data in terms of exceedances.

The USA RWQC set at 36 illness/1000, the EU Excellent Grade and recommendations by WHO are similar to the Blue Grade, while the minimum grade for swimming in New Zealand, Yellow, is slightly less restrictive than the EU Good Grade. It is important to note this is a simple overview as each jurisdiction will have different rules and methods for calculating suitability for swimming. The USA and WHO guidelines are based on 3-5% risk of illness in recreational users, whereas the New Zealand guidelines are based on infection, which may not lead to illness. While faecal indicator concentrations are similar, the risk may be lower in New Zealand as the risk assessment includes animal sources. Overseas, the risk is based on human sewage contamination only, which would pose a consistently higher risk to health.

Table 1: Comparisons between New Zealand and overseas approaches to setting of guidelines for freshwater recreation from McBride and Soller 2017

													0.1% Campy Infection		1% Campy Infection		5% Campy Infection					
Percentile of Distribution	2.5	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	97.5	>97.5
Predicted Campylobacter infection / 1000	0	0	0	0	0	0	0	0	0	0	0	0	1	3	10	18	26	72	131	329	435	>435
E. coli MPN/100mL	-	4	9	14	29	32	40	51	66	91	110	131	154	191	260	332	461	613	980	1986		>1986
2003 MfE/MoH Attribute State A												130										
2003 MfE/MoH Attribute State B															260							
2003 MfE/MoH Attribute State C																	550					
2003 MfE/MoH Attribute State D																			>550			
2014 NPS Attribute State A															260							
2014 NPSAttribute State B																	540					
2014 NPS Attribute State C																			1000			
2014 NPS Attribute State D																						>1000
2017 NPS Clean Water Blue												130					540					
2017 NPS Clean Water Green												130							800			
2017 NPS Clean Water Yellow												130								2100		
2017 NPS Clean Water Orange													>130 but <260								2580	
2017 NPS Clean Water Red															>260							>2580
USA RWQC - Set as 32/1000 ILLNESSES, effective CV=1.14											100						454					
USA RWQC - Set as 36/1000 ILLNESSES, effective CV=1.15												126						570				
WHO Guidelines - NOT in terms of EC, but the ENT values closely corrspond to USA ENT values which are equivalently protective to EC values via the EC derivation - estimated to be 50/1000 ILLNESS											Estimated						Estimated					
EU - Excellent Quality Bathing Water											Estimated						500					
EU - Good Quality Bathing Water													Estimated						1000			

Median value
95th %ile

***E. coli* swimmability thresholds: International comparison**

NZ	USA	EU	
A/B: 260*	126$ & 410*	Excellent:	500*
MAS: 540*	100$ & 320*	Good:	1,000*
		Sufficient:	900*

Units: #/100mL. *95%ile, ^90%ile. $geometric mean
NZ max Campy infection risks are 1% (A/B) and 5% (MAS)
USA criteria are for max illness risks of 3.6% or 3.2%
EU freshwater guidellines are derived starting with WHO guidelines but are not specifically risk-based

MPN =most probable number RWQC = Recreational Water Quality Criteria MfE = Ministry for the Environment MoH = Ministry of Health

2.6 Why a fresh look at science underpinning New Zealand guideline is important

An essential underpinning of Guidelines is a robust understanding of the relationship between the pathogens that may be present in the water and the indicator organisms used to monitor water quality. It is important in the New Zealand context, to have confidence in the science underpinning the relationship between indicators and pathogens in the water. Specifically, to understand which are the best indicators of microbial water quality, and what methods should be used for their assessment. The sections below describe the significant changes over the last 20 years raising some uncertainty over the relevance of the FMRP derived relationships between pathogens and indicators for future water quality guidelines.

2.6.1 New analysis methods

Since 2003, the development and routine use of molecular techniques, such as quantitative polymerase chain reaction (qPCR) and whole genome sequencing (WGS) allows for more sensitive and specific detection of micro-organisms. This may allow for improved characterisation of the relationship between the concentrations of various pathogens and concentrations of indicators. Further, these techniques allow for the characterisation of indicators and pathogens based on their DNA, including information on the serotype, genotype and/or potential virulence of pathogens. For example, not all strains of a species may cause illness, or may cause illness of varying severity (eg most *E. coli* are harmless, however a small percentage are pathogenic, such as STEC, which carry the virulent *stx1* and/or *stx*2 genes).

Molecular analysis can also assist in providing information on the likely source of contamination (eg human, cattle, sheep, wildfowl), by investigating the presence of particular micro-organisms or particular genotypes of widely dispersed micro-organisms that are specific to a certain host animal. Source attribution is important, as scientists have a growing appreciation of the different health risks that might be posed by faecal contamination from different sources.

2.6.2 Land use changes

In addition to improved analytical methodologies, there have also been significant changes in land use patterns within New Zealand, intensification in some regions, as well as changes in land management practices (eg stock exclusion from waterways, irrigation of agricultural effluent to land). These may affect how micro-organisms are introduced to, and survive in waterways, and hence the relationship between indicators and pathogens in the water.

2.6.3 Te Mana o Te Wai

The NPS-FM 2020 gives effect to Te Mana o Te Wai, requiring the incorporation of Maori values and approaches. While microbial health risk addresses a component of iwi health and well-being, it is just one aspect. Cultural Health Index (CHI) are recognised as an important tool for local iwi and hapū to meet their aspirations, that consider aspects such as mauri, tikanga and cultural practices, and an understanding of what is important for monitoring cultural health in the local context. Although the pilot study was focused on logistics and confirming the prevalence of pathogens, the initial engagement initiated with iwi and hapū at each site needs to be developed and extended to co-design a new framework which encompasses Mātauranga Māori.

2.6.4 Updating underpinning Guideline data

Following a review of relevant scientific literature published since the FMRP, as well as records of disease incidence in New Zealand, recommendations were made through five reports, on which

pathogenic and indicator micro-organisms should be evaluated in such a study, the experimental design and methodology for a QMRA study suitable for updating the FMRP (Gilpin et al, 2018, Moriarty et al, 2018, Lake et al, 2018, Milne et al, 2018, Horn et al, 2018). These recommendations included:

- the number of rivers to be sampled to give geographical coverage (n = 30)
- determinants for selection of rivers, including history of contamination, existing datasets, adjacent land uses
- frequency of sampling (fortnightly) and sampling duration (18 months) for 1200 samples
- pathogens and indicators to be analysed.

Further to receiving the report on the Survey Design and QMRA Project Plan (Moriarty et al, 2018) the Ministry for the Environment identified that a pilot study would be beneficial to confirm size and scale of a national project. Sites with a history of *E. coli* exceedances were chosen and while that might represent a worst-case scenario for New Zealand rivers, if there is a low level of positive detection of pathogens, then a larger study size would be required. Conversely, if there were higher concentrations of pathogens detected, then a smaller study may achieve the desired outcome.

3 Methodology

3.1 Selection of sites and sampling plan

Sites were selected from the full list determined in the Freshwater Microbiological Sciences Review Stage One (Milne et al, 2018). The 16 sites were selected based on land uses likely to lead to the contamination of freshwater with pathogens and which had a history of *E. coli* exceedances. The dominant land uses targeted were:

- six urban sites
- five sites where the major influence is likely to be sheep & beef farming
- five sites where the major influence is likely to be dairy farming.

The selected sites were geographically distributed across New Zealand, and provided the opportunity to commence engagement with 13 councils and local iwi. The sites and their observed land use are given in Appendix B.

For the pilot study, it was decided that 16 rivers would sampled five times each, giving 80 samples. This was largely driven by logistical and financial constraints. It was estimated that 50% of samples would have *Campylobacter*, but the prevalence of *Salmonella* and STEC could be much lower, perhaps as low as 5-10%. While these might represent a worst-case scenario of rivers, if there is a low level of positive detection of pathogens, then a larger study size will be required. Conversely, if there were higher concentrations of pathogens detected, then a smaller study may achieve the desired outcome.

3.2 Analytical methods

Each sample collection event involved the collection of a 250 mL, 6 L and a 10 L water sample, as well as the filtration of up to 100 L of water through a Filta-Max cartridge. Figure 2 provides an overview of the analytical approach, with the actual volumes used in the analysis. At least one isolate each of *Campylobacter, Salmonella* and STEC from each positive sample were analysed by whole genome sequencing (WGS). Faecal Source Tracker (FST) markers for ruminants, human and avian were analysed to identify contamination sources. Details of the methodologies are given in Appendix C.

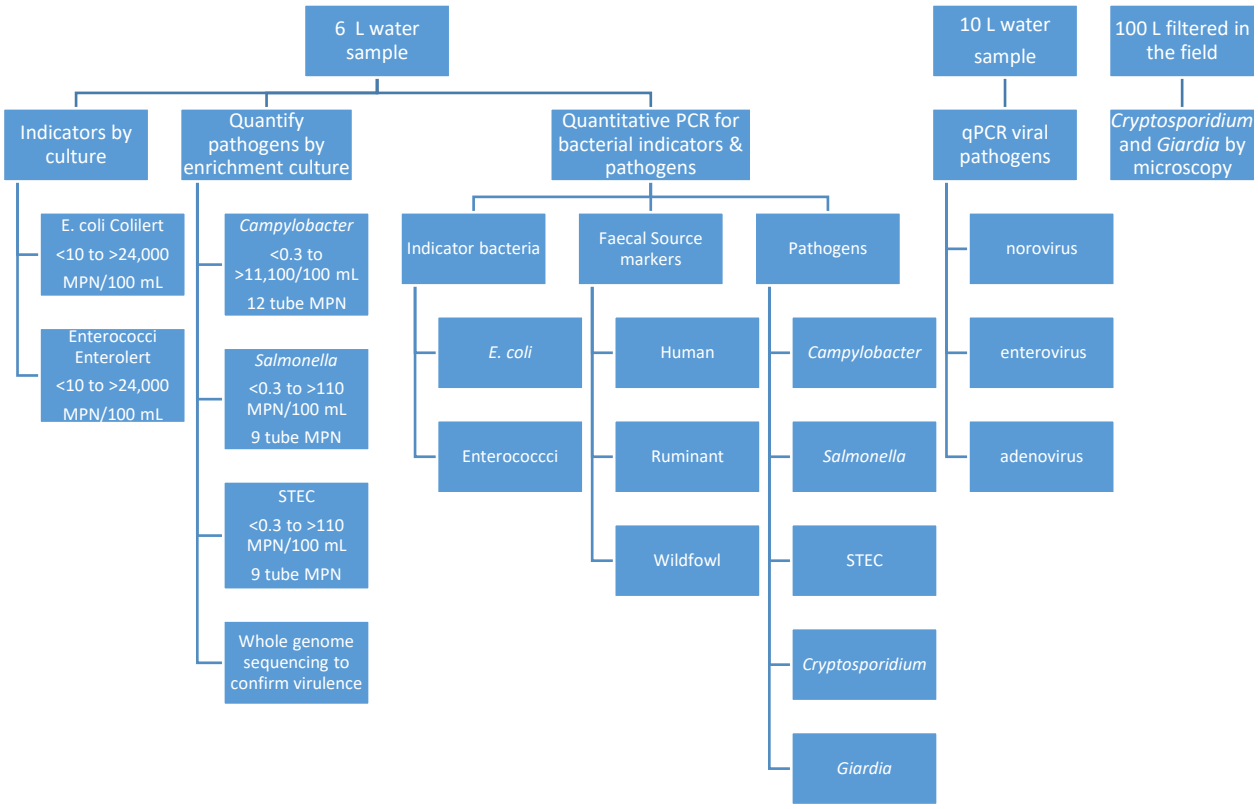

Figure 2: Overview of sample analysis during pilot study

3.3 Statistical analysis

The pilot study was designed to collect a small amount of data to support the design of a larger study. As such, most of the statistical descriptions of the data are of a high-level, descriptive nature. The graphs have been generated using the R statistical software (R Core Team, 2018) or Microsoft Excel. Pearson correlation coefficients and Spearman's rank correlation coefficients have been calculated in Excel.

The power calculations for detecting a given change in proportions for a given sample size were calculated using the GPower 3 software (Faul et al, 2007) using the option for a single sample binomial test.

3.4 Environmental data

Environmental and chemical measurements were also taken in the field during sampling or from data records. They are summarised in Table 2.

Table 2: Field measurements

Parameter	Source
Photo of river at sampling time	Cell phone
Flow	Gauged site (or taken from flow recorded, or modelled)
Turbidity	Council equipment if equipment available Laboratory test
Clarity	Black disk or clarity tube
Conductivity	Council meter if equipment available Laboratory test
pH	Council equipment if equipment available Laboratory test
Dissolved Oxygen	Council meter
Water temperature	Council meter
Time sampling started and finished	Watch/cell phone
Volume filtered	Flow metre of filtration unit
Presence of animals	Observation
Rainfall at time sampling	Observation
Wind direction and strength	Observation
Sunlight	Observation
Rainfall previous 24, 48 and 72 hours	Records from nearest weather station

3.5 Engagement with iwi

Local iwi and hapū were contacted to inform them of this study, and invite them to the initial site visit or to kōrero and discuss the project in the context of a particular site. During these kōrero we were able to explain that for the pilot study, the sites needed to have a high potential for pathogen contamination and realised that the sites may not be of the most cultural significance for iwi and hapū. Feedback was requested to gather information about how a cultural assessment process might work for individual iwi and hapū.

4 Results

4.1 River sampling

The sites are presented in Figure 3 labelled A-P, with detailed location and land use provided in Appendix B. In presenting the results, the rivers labelled A-P in Figure 3, are re-labelled as rivers 1-16 to provide anonymity. In the plots, data is presented for each river (labelled 1-16), and each sampling event labelled a-d.

Council staff, iwi and landowners were invited to site visits in November or December 2019, where sampling techniques were demonstrated, and the study team confirmed the site characteristics. Feedback from staff showed that the site visit and training was very much appreciated. Assembling the filtering equipment for protozoan sampling was not a procedure with which staff were familiar. The site visits also made subsequent communication over issues and interpretation of results easier. During the sampling period, the distribution of sampling equipment and containers went smoothly and the level of documentation and instructions were generally considered to be appropriate. Issues that arose were addressed by emails and further instructions. The introduction of the ice bath for sampling, for example, was very useful in rapidly reducing the temperature of the samples. The courier system worked well, delivering samples to the laboratory early in the morning except for two locations, where it became necessary to change the courier company to achieve required delivery times.

Sampling occurred fortnightly from each of the 16 rivers, from 3 February 2020 until 23 March (7 weeks), resulting in 63 samples being collected by council staff. Eight sets of samples were discarded due to the laboratories being closed by Level 4 Covid19 lockdown and three sets of samples didn't arrive within the 24 hours timeframe required, due to inability of the courier to deliver overnight. Changing the courier company resolved this issue. Three or four sampling rounds were completed for each river, depending on the initial sampling dates.

Fifty samples were analysed using culture methods (one sample was analysed within 48 hours and the data is available but not used for analysis), 52 samples were analysed for viruses and protozoa, as they are not as time sensitive, and 51 samples were analysed by qPCR for FST markers and pathogens. The raw analytical data is presented in Appendix D.

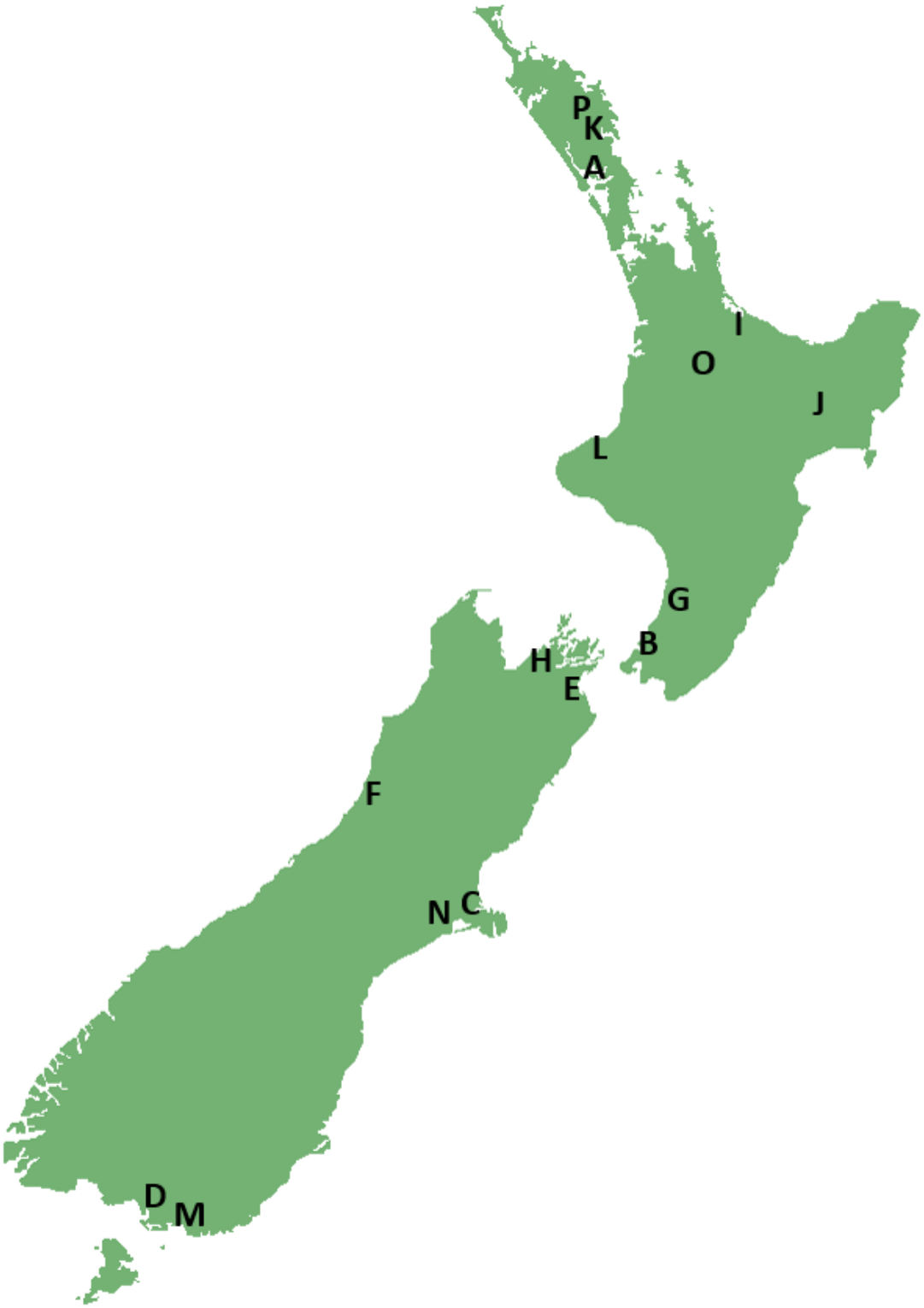

Figure 3: **Approximate location of 16 sampling sites A-P**

4.2 Faecal Source Tracking (FST) markers

Sites were initially selected based on observed land use and/or previous FST data to provide examples of rivers likely to be contaminated from either urban, dairy farming or beef & sheep farming sources. In this study, markers which are specific for humans, ruminants and birds have been used, with FST results for 51 samples presented in Table 16, Appendix D and their distribution in each river shown and discussed in Appendix E.

These results are summarised in Figure 4, with the labels at the top representing the dominant faecal source for each river. In most samples more than one source was present. Wildfowl markers were detected in all of the rivers and in 11 of the samples, were the only source identified. This included all three samples from river 15 and river 16 which, based on observed land use assessment, were selected as sheep & beef farming and dairy farming sites, but reclassified as wildfowl only. All six urban rivers (rivers 1 -6) contained human FST markers, in all the samples tested. However, five of the samples from two of the urban rivers (rivers 1 and 2) also contained significant concentrations of ruminant FST markers. Conversely, while the ruminant FST marker confirmed the observed land use of four dairy farming and four sheep & beef farming rivers, two of the samples (river 7 and river 11) also contained human markers.

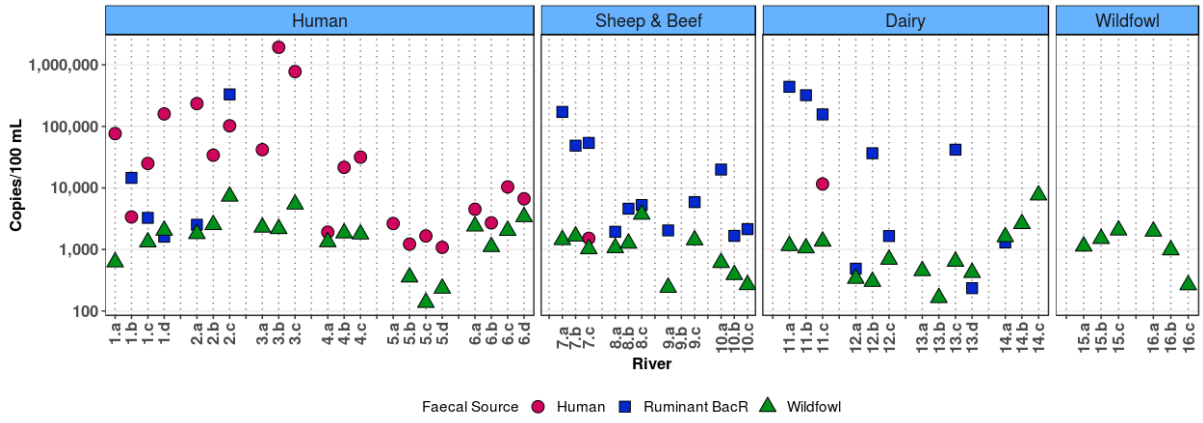

Figure 4: Faecal source tracking summary for the rivers sampled

Note: No FST data is available for 9b.

4.3 Faecal Indicator Bacteria (FIB)

4.3.1 E. coli

E. coli were detected in all samples using Colilert assays, with one sample from river 1 being above the upper detection limit (24,000 MPN/100 mL). There was a range in concentrations of at least one order of magnitude within all the rivers (Figure 5). Seventeen samples from eight different rivers had ≥1,000 *E. coli*/100 mL, including the sample which arrived late. The six rivers categorised as human-impacted consistently had the highest concentrations of *E. coli* (>1,000 MPN/100 mL). Another six samples had >540 *E. coli* MPN/100 mL, 12 samples between 260 and 540 *E. coli* MPN/100 mL, and 16 samples <260 *E. coli* MPN/100 mL.

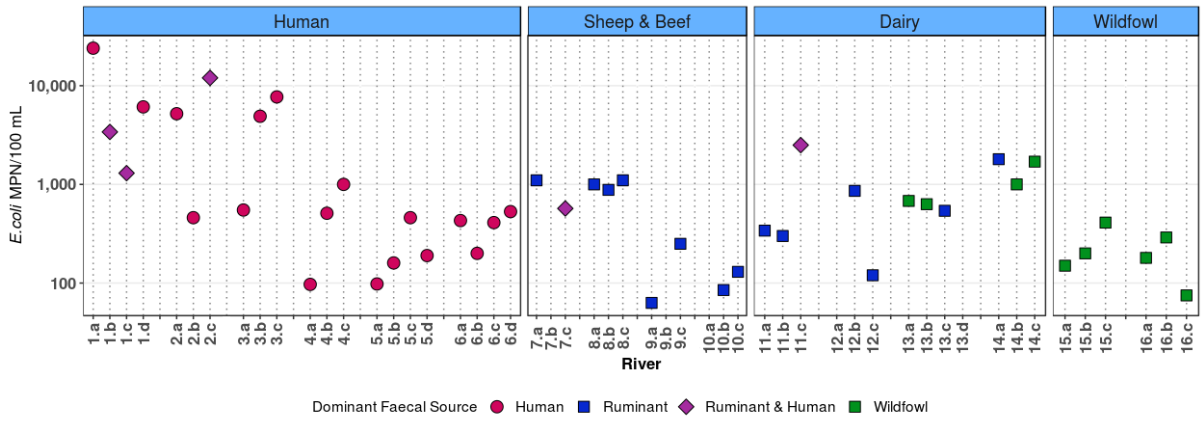

Figure 5: ***E. coli* concentrations, coloured by the dominant faecal sources in each sample**

Note: Excludes 12a which is at the detection limit (10MPN/100 mL), no *E. coli* data available for 7b, or 10a, no FST data available for 9b, above 24,000 MPN/100 mL is reported as 24,000 MPN/100 mL.

E. coli were also quantified by qPCR. A strong correlation (R^2 = 0.85) was observed between *E. coli* by the traditional Colilert method and qPCR as shown in (Figure 6).

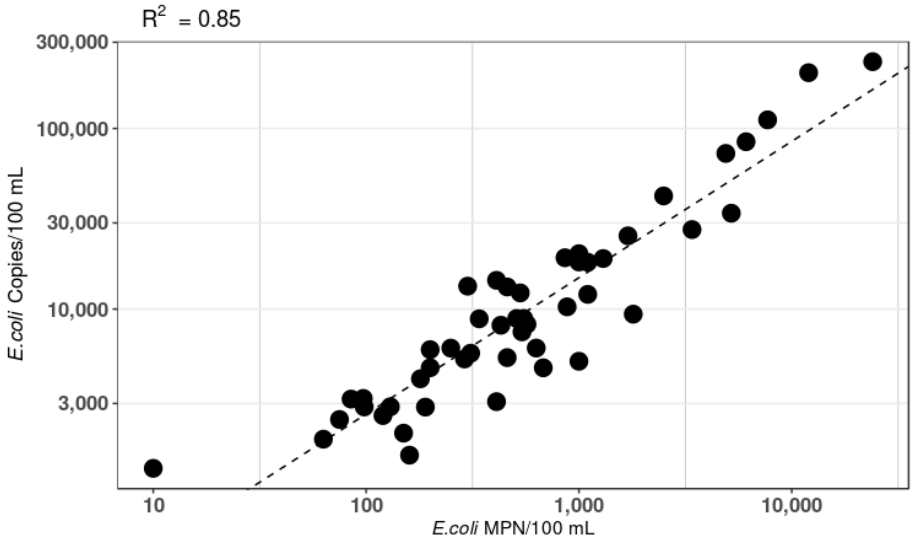

Figure 6: **Correlation of culture and qPCR methods for *E. coli* across all data**

4.3.2 Enterococci

Enterococci were detected in 44 of samples (86%), including the sample which arrived late, with seven below the limit of detection (10 MPN/100 mL), six at the level of detection and none above the upper limit of detection. Concentrations of enterococci ranged from <10-990 MPN/100 mL. The data for each river is shown and against the dominant faecal sources in Figure 7.

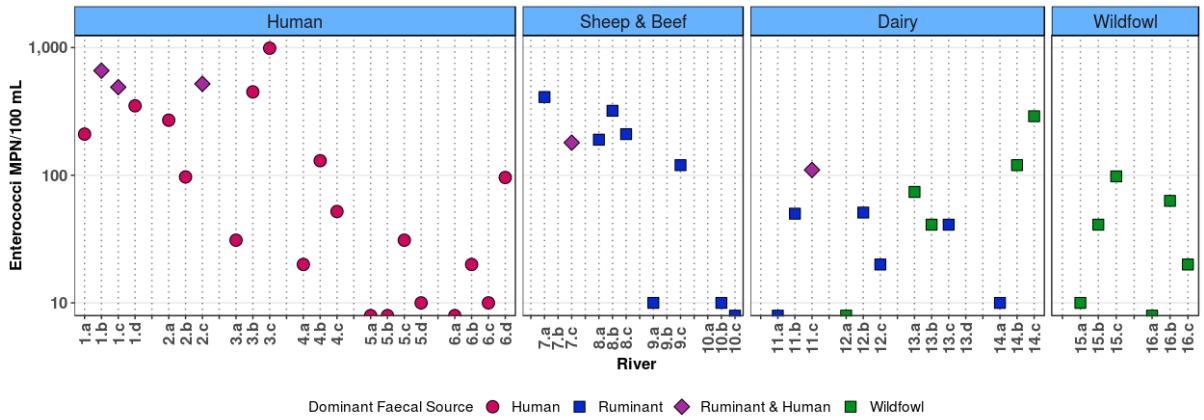

Figure 7: Enterococci concentrations, coloured by the dominant faecal sources in each sample

Note no enterococci data available for 7b, or 10a, no FST data available for 9b

Studies from the USA have found a strong association with enterococci concentrations when measured by qPCR and culture for both marine and freshwaters and have developed and tested two standard methods for *Enterococcus* spp. (USEPA, 2018). A strong correlation was not found during this study, $R^2 = 0.187$ (Figure 8).

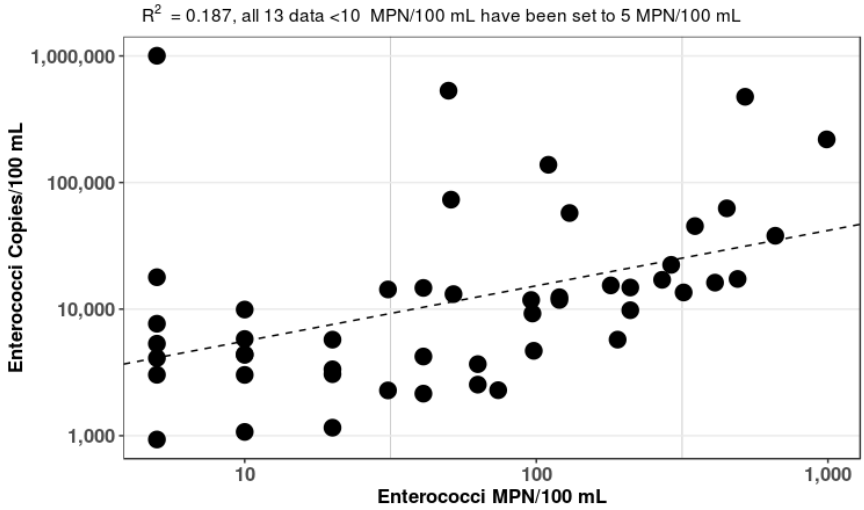

Figure 8: Correlation between enterococci culture method (MPN/100 mL) and qPCR (copies/100 mL) across all data

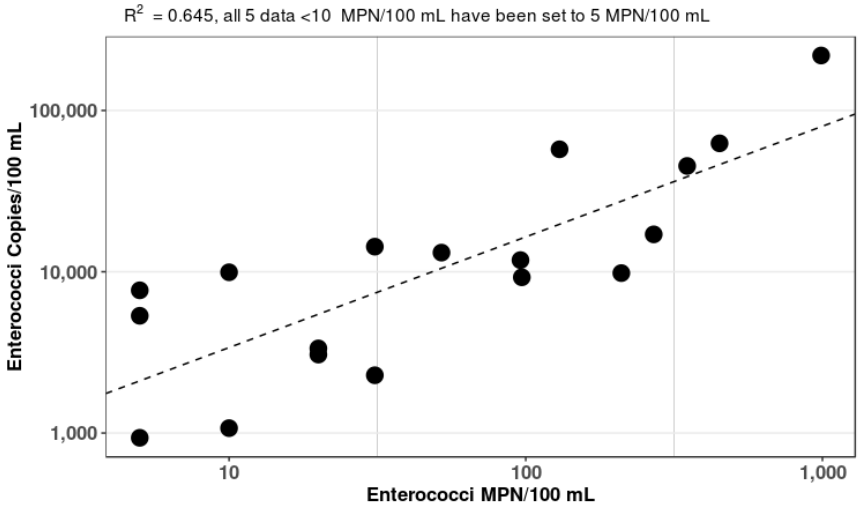

Figure 9: Correlation between enterococci culture method (MPN/100 mL) and qPCR sample (copies/100 mL) impacted by human sources

When this analysis is repeated using only samples where human sources of contamination dominated, the correlation between culture and qPCR is much stronger with $R^2 = 0.645$ (Figure 9). The strength of the correlation (R^2 value) reduces when the sources are a mix of high concentrations of human and ruminant FST ($R^2 = 0.588$).

4.4 Pathogenic bacteria

Using MPN culture enrichment, *Campylobacter* spp. were detected in 34 samples from all but two of the rivers (Figure 10). *Salmonella* spp. were detected in nine of the samples, and STEC in only one sample. In only one sample was *Salmonella* or STEC detected in the absence of *Campylobacter* (river 5).

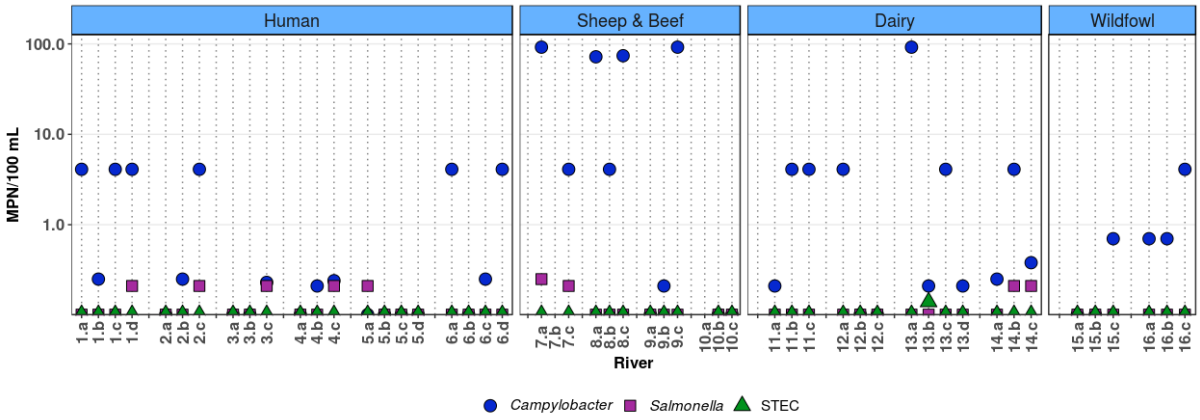

Figure 10: Bacterial pathogen summary for the rivers sampled

Note: When pathogens were not detected, the data overlay each other on the X axis. No data for 7b or 10a.

4.4.1 *Campylobacter*

Of the 34 samples with *Campylobacter* detected by MPN culture (68%), five samples had between 72 and 92 MPN/100 mL, 14 samples had 4.1 MPN/100 mL, and the other 15 were below 1 MPN/100 mL.

Detailed results for *Campylobacter* are described in Appendix F, including the WGS results from sequencing isolates from 32 river samples. The WGS data indicates that the *Campylobacter* were a genetically diverse group of isolates, with the exception of two of the isolates from river 7. The isolates were mainly *C. jejuni* (30 isolates), with two *C. lari.* Seven isolates of two new *Campylobacter* species were also recorded (six *Campylobacter* spp. 1 isolates, and one *Campylobacter* spp. 2 isolate). There is no evidence that these new *Campylobacter* spp. 1 and 2 cause disease in human, so the health risk is unknown.

Quantitative PCR detected *C. jejuni* in 33 samples and *C. coli* in 13 samples (Figure 11 and Table 18). There were four samples where *Campylobacter* were detected by MPN culture, but not by qPCR. Three of these had MPN/100 mL of 4.1 (two were from the same river), and the fourth was 0.21 MPN/100 mL. *C. jejuni* was detected by qPCR in 27 of the 34 samples which had *C. jejuni* detected by MPN culture. Of the samples with more than 10 MPN/100 mL, all were detected by qPCR. There were five samples where *Campylobacter* were detected by qPCR, but not by MPN culture. These qPCR concentrations were low (maximum of 72 copies/100 mL). *C. coli* was not identified in any of the MPN culture samples, but was detected in 13 samples from 10 rivers by qPCR. Ten of these samples had other species of *Campylobacter* detected by MPN culture, which may suggest preferential isolation of *C. jejuni* using MPN culture.

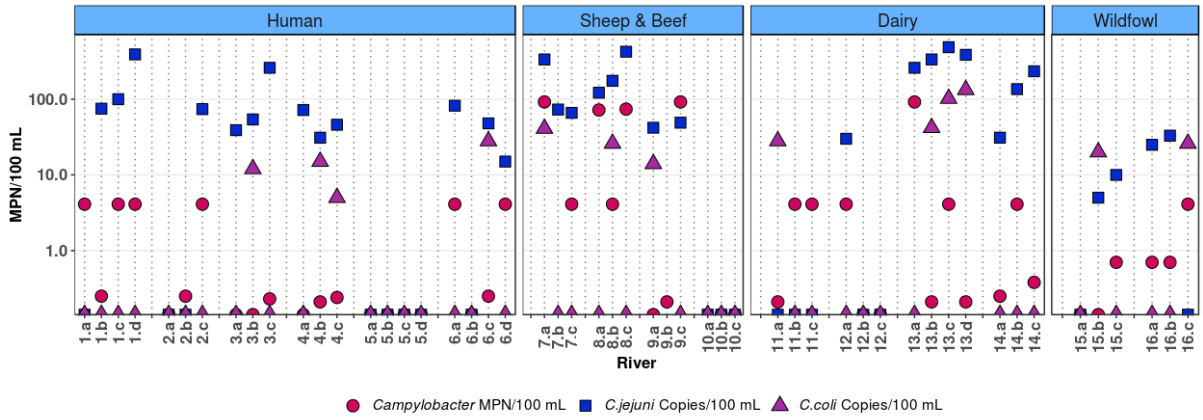

Figure 11: *Campylobacter* MPN culture and qPCR results for the rivers sampled

Note: When pathogens were not detected, the data overlay each other on the X axis. No MPN data for 7b or 10a

A comparison of the traditional culture methods and qPCR showed that the linear correlation between qPCR and traditional method for *Campylobacter* was not strong (Figure 12). Reasons for this include differences in detection limits and the non-linear resolution of the MPN dilutions used. Essentially the MPN analysis is a statistical method and only gives results in bands depending on how many tubes show a positive result at each of the dilutions. Appendix C gives the MPN table (Table 14) which shows that if the actual level is between 1 and 10, the only possible results are 1.4 and 4.1. Comparison of *Campylobacter* detection by culture versus qPCR for *C. jejuni/C. coli*, found 81.6% agreement between the methods with Cohen's k of 0.58 (moderate agreement).

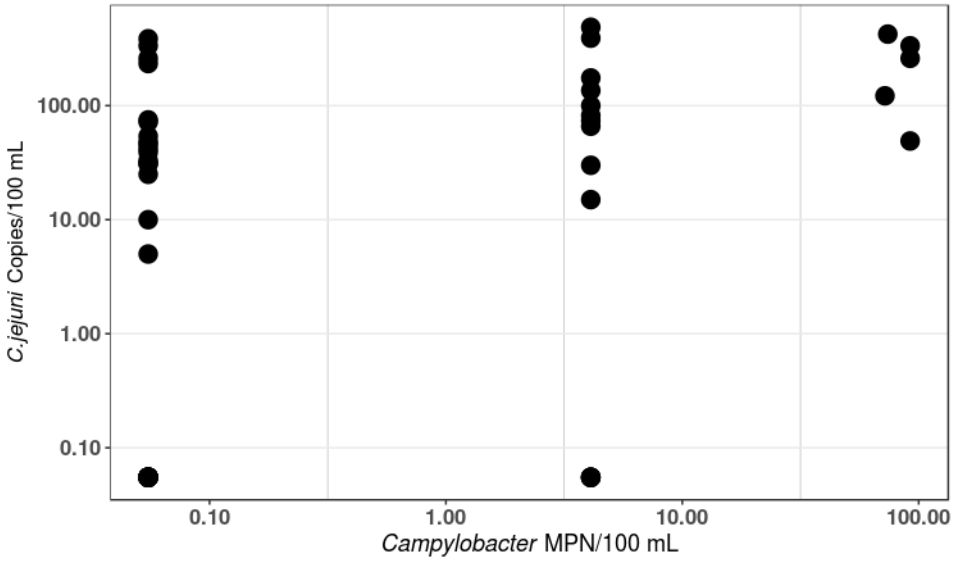

Figure 12: Correlation of culture and qPCR methods for *Campylobacter* (MPN) and *C. jejuni* (qPCR) across all data

4.4.2 Salmonella and STEC

Salmonella was detected in eight rivers (Figure 13). By MPN culture it was detected in nine (18%) of 50 samples with concentrations of 0.21 MPN/100 mL (eight samples) and one sample at 0.25 MPN/100 mL. Of the five river samples with PCR detection of *Salmonella*, two samples had no *Salmonella* isolated from MPN culture, while the other two had 0.21 and 0.25 MPN/100 mL. One sample was delivered after 24 hours, so the MPN result was not available for comparison. WGS was used to confirm isolates were S*almonella* and the four *Salmonella* serovars that were identified are discussed in Appendix G.

Salmonella samples were tested by qPCR for the presence of two virulence genes – *inv*A and *ttr*. The *inv*A gene was detected in four samples, and *ttr* in two samples with one in the river 3 sample containing both *inv*A and *ttr*. The levels were very low, at or below the level of quantitation. The *inv*A PCR assay is more sensitive than *ttr* when tested on pure strains in our laboratory which may explain why at the low levels of *Salmonella* identified there were more detections of *inv*A.

STEC

STEC was detected in 11 rivers, in 13 samples by qPCR, but by MPN culture it was detected in only one sample at a concentration of 0.14 MPN/100 mL (Figure 13).

Direct qPCR analysis on DNA extracts from water samples, detected the *stx*1 gene in seven samples at concentrations of between 7 and 172 copies/100 mL. Four of these also contained the *stx*2 gene at concentrations of 42-280 copies/100mL, and another six *stx*2 gene only at concentrations of 2-33 copies/100mL. The sample with the E. coli O177:H25 isolate with *stx*2 gene present (0.14 MPN/100 mL), was qPCR positive for *stx*1, at a level of 7 copies/100 mL. These seemingly incongruent results probably reflect the low levels present, and that the sample may have also contained bacteria with the *stx*1 gene but enrichment and isolation didn't find them. Another sample from this river was positive by qPCR for *stx*2, which would suggest bacteria with both *stx*1 and *stx*2 can be present. Figure 13 summarises the detection of *Salmonella* and STEC by culture and qPCR.

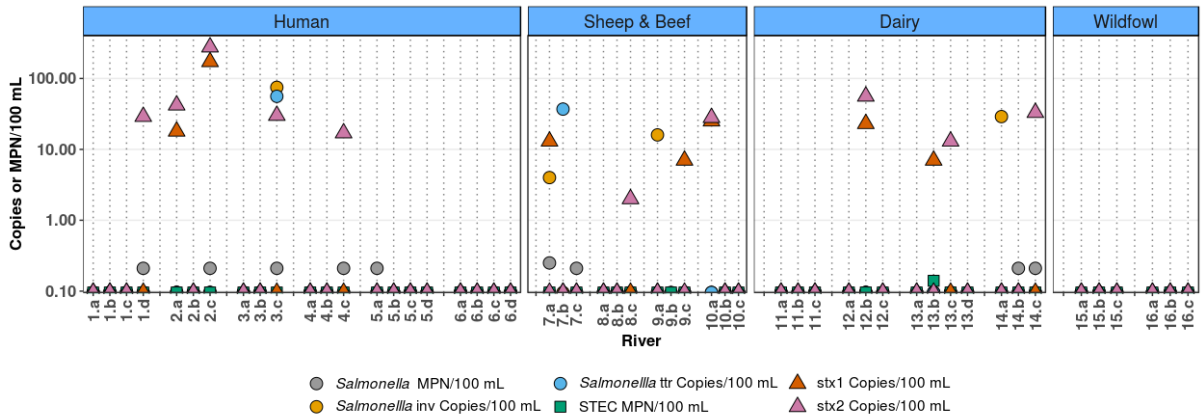

Figure 13: *Salmonella* and STEC summary for the rivers sampled

Note: When pathogens were not detected, the data overlay each other on the X axis. No MPN data for 7b or 10a

Salmonella whole genome sequencing

Between one and four isolates from each of the positive nine samples were confirmed as *Salmonella* by whole genome sequencing. One presumptive *Salmonella* result was excluded based on whole genome sequencing. Isolates were all identified as *Salmonella enterica* with the serovars Enteritidis, Typhimurium, Emek or Saintpaul. Details of WGS and information on the serovars are given in Appendix G.

Escherichia whole genome sequencing

The MPN culture for STEC produced a single positive plate culture detection from which an isolate was characterised as a *stx*2 strain from river 13b. In addition, following positive detection by PCR directly on the MPN enrichment broths, there were two other isolates from two samples that were only *eae* positive. Details on WGS are discussed in Appendix G.

4.5 Protozoa

Protozoa are detected by filtering a large volume of water (ideally 100 litres) in the field. In the laboratory the protozoa are eluted from the filter and concentrated before a subsample is placed on to a slide for visual detection. Owing to suspended solids and turbidity in a river it is not possible to always filter 100 L, in which case filtering was stopped after an hour. The data and volumes filtered are presented in Appendix D, Table 20. The average volume filtered in this study was 73.2 litres, with a minimum of 2 litres before the filter blocked. The recovery rate through filtration, elution off the filter and microscopy is estimated to be between 15 and 55%. Results presented here are not adjusted by recovery rates. Results are presented in Figure 14 as counts ((oo)cysts/100 L) and as qPCR detections (copies/100 mL).

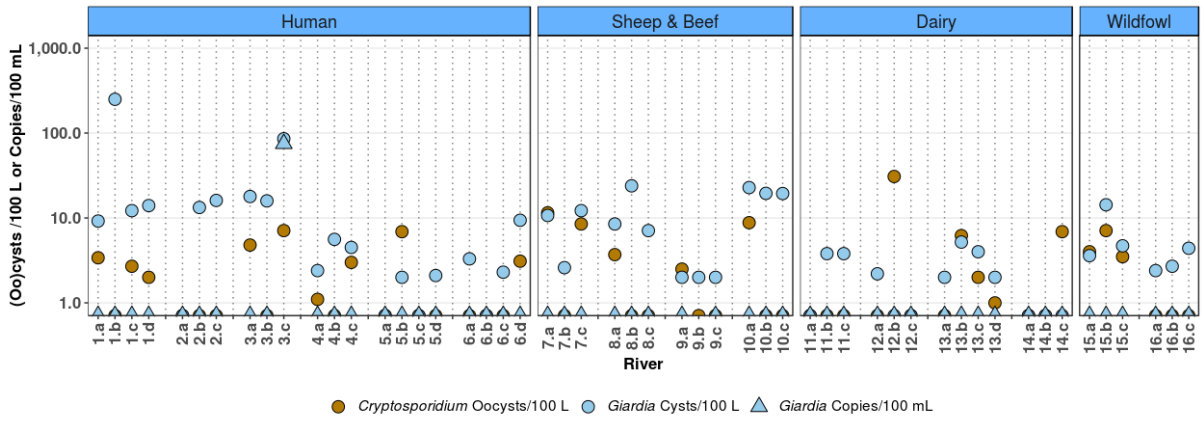

Figure 14: Protozoa summary for the rivers sampled

Note: When pathogens were not detected, the data overlay each other on the X axis

4.5.1 Cryptosporidium

Cryptosporidium was detected in 22 (42%) of samples at concentrations from 1 to 30.8 oocysts/100 L. At the highest concentration this is only 0.3 oocyst/L. Quantitative PCR of the *Cryptosporidium* species commonly found associated with illness in humans, *C. parvum* and *C. hominis,* were not detected in any sample.

4.5.2 Giardia

Giardia were detected in 42 samples (81%) ranging in concentration from 2 up to 250 cysts/100 L. There were 27 samples with between 2 and 9 cysts/100 L, 13 samples were between 11 and 24 cysts/100 L, one sample at 86 cysts/100 L and one at 250 cysts/100 L.

Giardia was only detected by qPCR in one sample (75 copies/100 mL), which was the sample with the 86 cysts/100 L. For this sample 900 mL of river water was filtered for qPCR which would correspond to less than 1 cyst in the 900 mL. The microscopy method has an estimated recovery of 15-55% which would correspond to between 573 and 156 cysts/100 L, and therefore, between 2 and 6 cysts in 900 mL tested by qPCR.

Giardia was not detected by qPCR in the samples with the highest concentration of 250 cysts/100 L. This was a turbid sample. For the traditional microscopy analysis, only 2 L of water could be filtered in the field before the filter blocked. The concentration may therefore be overestimated. For the qPCR only 500 mL of river water could be filtered. Assuming a concentration of 250 cysts/100 L this equates to 2.5 cysts/L, which would be only 1.25 cysts in the 500 mL tested by qPCR.

4.6 Viruses

Out of 52 samples, 21 samples contained potential RT-PCR inhibition of the viral analyses. The murine norovirus RT-qPCR indicated that the recovery of viruses from water samples ranged between 3-12%.

Norovirus GI was detected in six samples, norovirus GII in seven samples, and enterovirus in two samples. For norovirus GI and GII there were only two samples (one sample each) where all three replicates were positive (+ve), while in five samples, two of the replicates were positive, and in eight samples only one of the replicates was positive. All were below the limit of quantitation and at the limit of detection, and so only the CT value is given where a result was positive, otherwise the result

is reported as “not detected” (ND) giving a negative (-ve) result. The results with the three replicates are given in Table 3.

Table 3: Virus PCR results with three replicate analyses

River	HAdV (CT)	+ve /-ve	NoV GI (CT)	+ve /-ve	NoV GII (CT)	+ve /-ve	Enterovirus (CT)	+ve /-ve
2.a	ND/ND/ND	-ve	38.3/ND/39.1	+ve	ND/ND/ND	-ve	ND/ND/ND	-ve
2.c	ND/ND/ND	-ve	39.7/38.4/39.9	+ve	ND/ND/39.3	+ve	ND/ND/ND	-ve
3.a	ND/ND/ND	-ve	ND/ND/ND	-ve	ND/ND/39.7	+ve	ND/ND/ND	-ve
3.b	ND/ND/ND	-ve	ND/ND/38.8	+ve	38.2/ND/38.5	+ve	ND/39.9/ND	+ve
3.c	ND/ND/ND	-ve	ND/ND/ND	-ve	38.2/38.3/37.4	+ve	ND/ND/ND	-ve
9.c	ND/ND/ND	-ve	ND/ND/ND	-ve	ND/ND/ND	-ve	ND/39.2/ND	+ve
10.c	ND/ND/ND	-ve	ND/ND/ND	-ve	37.8/37.1/ND	+ve	ND/ND/ND	-ve
14.b	ND/ND/ND	-ve	37.9/ND/ND	+ve	40.1/ND/ND	+ve	ND/ND/ND	-ve
14.b	ND/ND/ND	-ve	ND/ND/ND	-ve	ND/39.7/ND	+ve	ND/ND/ND	-ve
14.c	ND/ND/ND	-ve	41.3/40.5/ND	+ve	ND/ND/ND	-ve	ND/ND/ND	-ve
15.b	ND/ND/ND	-ve	ND/40.5/38.9	+ve	ND/ND/ND	-ve	ND/ND/ND	-ve

The presence of viruses in each river is presented in Figure 15. Human adenovirus (HAdV) was not detected in any sample. One sample (from river 3) contained all three viruses and two samples (from river 2 and river 14) had both norovirus GI and GII. River 3 and river 14 had all three samples positive for at least one virus.

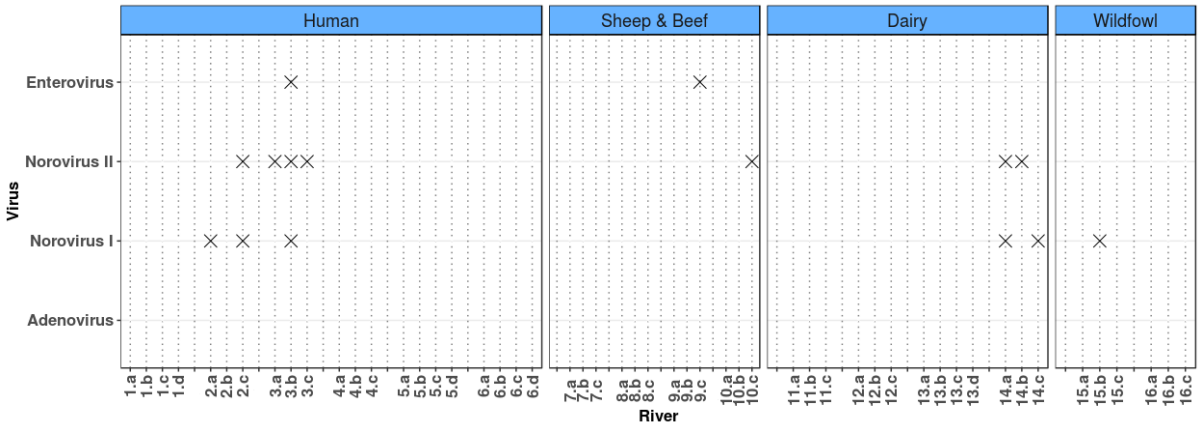

Figure 15: Virus detection by RT-qPCR for the rivers sampled

4.7 Field data

Field and laboratory physiochemical data are presented in Table 21 and discussed in Appendix H. Not all councils have the same equipment, nor all the equipment to measure the range of water quality parameters requested, so some parameters were measured in the field and some in the laboratory. Temperature, dissolved oxygen (DO) and pH were measured in the field, while turbidity and conductivity were measured in the laboratory (Table 21).

Figure 16 shows the range of data across all sites for DO (48 measurements), field pH (37 measurements) and water temperature (52 measurements). The median DO was 9 mg/L, but four samples had concentrations below 5.0 mg/L, which would adversely affect aquatic life. The average

field pH was 7.55, which is within the 6.5-8.5 range expected for rivers. The lowest pH was measured after an extreme rain event and the three other readings below pH 6.5 are from the same river. Low turbidity was measured in most samples, with a median of 2.25 NTU, indicating good clarity, but there were three events when turbidity was greater than 10 NTU; two of the three measurements were taken when there had been a significant rainfall event within the previous 72 hours. The water temperature reflects the late summer–early autumn sampling period. The average temperature in South Island rivers was 15.9°C and 18.6°C in the North Island rivers.

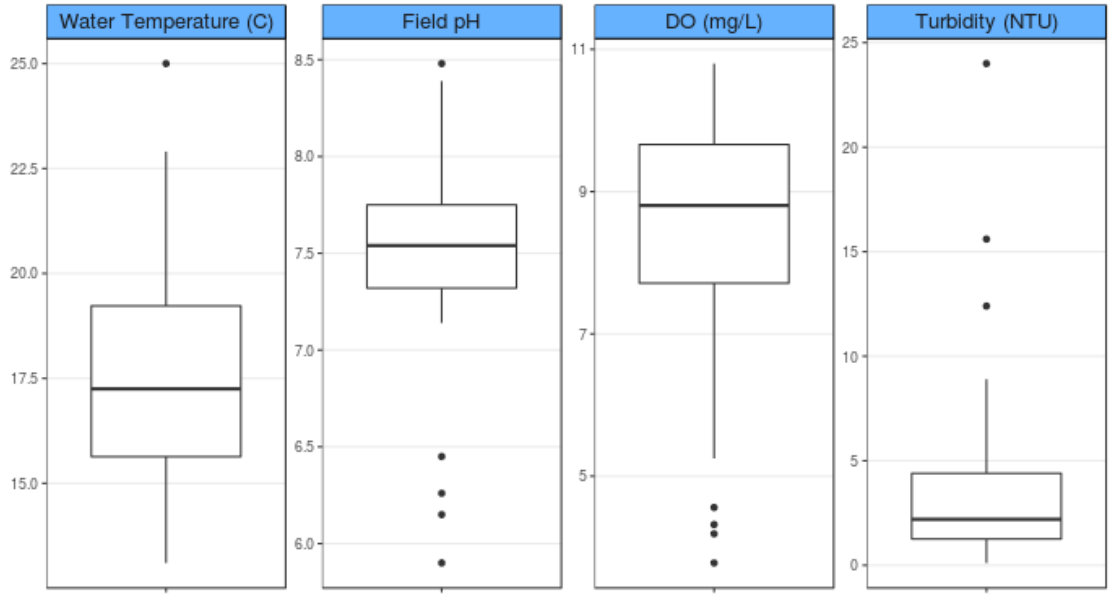

Figure 16: Box plots of water temperature, field pH, dissolved oxygen (DO), and turbidity

Note: The box plots represent the interquartile range 25th-75th with the median shown as the heavy line within the box. The whiskers extend 1.5 times the interquartile range. Outliers are plotted as individual points

Conductivity was low at each river, except for the river with tidal influence, river 6. Excluding river 6 the median conductivity was 230 µS/cm with a range of 88- 362 µS/cm. Salt water intrusion from the tidal influence was low with conductivity measurements of 1711-3410 µS/cm (salinity of <0.2-1.2 ppt). Freshwater is usually 0-1500 µS/cm and typical sea water has a conductivity value of about 50,000 µS/cm (salinity of 33-35ppt).

The individual results for each river for pH, DO, flow, turbidity and rainfall are discussed in Appendix H and illustrate that individual rivers had different characteristics. Water clarity, turbidity and flow appear to be characteristic of individual rivers.

5 Discussion

5.1 Overview

The purpose of this pilot study was to collect data to inform the design of a larger study for a QMRA for recreational water quality. A key aim of the pilot study was to determine the size of the proposed full study of 30 rivers (Phase 2), so that it would be statistically robust, to inform an update of the QMRA undertaken in the 1998-2000 FRMP (McBride et al, 2002).

This pilot study addresses the three of the key question necessary for a QMRA (Figure 17).

a) What microbiological pathogens are present?

b) What are their concentrations?

c) Which pathogens are most likely to cause infection (based on a and b above)?

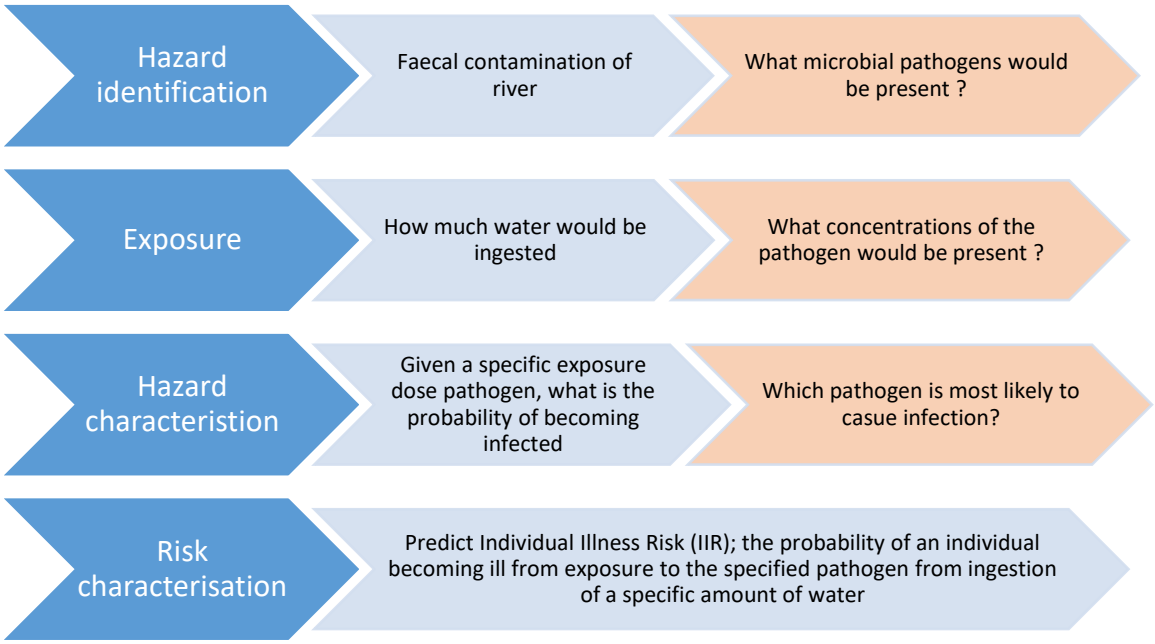

Figure 17: Key steps of a QMRA

Therefore, this pilot study of 16 geographically diverse rivers, commissioned by the Ministry for the Environment, generated data on the current detection rates and concentrations of FIB and pathogens (bacteria, protozoa and viruses) present in New Zealand rivers, to also guide assessments of likely significant changes since the FMRP 1998-2000. Between February and March 2020 samples were collected from 16 rivers, which was a COVID-19 truncated/impacted version of the intended 80 samples from 16 rivers.

The FMRP had identified land use as an explanatory variable for correlation between *Campylobacter* and *E. coli* (McBride et al, 2002) and this was taken into account in the selection of the 16 sites. The sites represented land use impacts dominated either by human, dairy farming, or sheep & beef farming and were selected because they had a history of elevated concentrations of *E. coli.* Faecal source tracking qPCR methods, which were not available for the original FRMP study, were used to

indicate faecal pollution and more accurately identify faecal sources. FIB have been measured, but because of the limited size of the pilot study, statistically robust correlations between FIB and pathogens were not expected.

In addition, the pilot study evaluated whether new, more cost-effective methodologies such as qPCR methods, which allow identification of human relevant pathogens, could be incorporated into Phase 2 of the study.

Water quality parameters are discussed as they have been found to be associated with the presence and concentration of pathogens in studies by McBride et al (2002) and Bradshaw et al (2016).

5.2 Concentrations of FIB and pathogens in selected New Zealand rivers in 2020

The pilot study has identified that there are frequent detections of pathogens by culture in New Zealand rivers, however, in general, they are identified at low concentrations as summarised in Table 4.

Table 4: Summary of quartiles, median and range of FIB, *Campylobacter*, *Cryptosporidium* and *Giardia*

Micro-organism	Number of samples	Concentration of micro-organisms				
		Minimum	First Quartile	Median	Third Quartile	Maximum
***E. coli* MPN/100 mL**	50	10	193	485	1075	>24,000
Enterococci MPN/100 mL	50	<10	13	52	188	990
***Campylobacter* MPN/100 mL**	50	<0.11	<0.11	0.3	4.1	92
***Cryptosporidium* oocysts/100 L**	52	ND	ND	<0.11	3.4	30.8
***Giardia* cysts /100 L**	52	ND	2	3.8	12.2	250

Note: *Salmonella* and STEC are not included due to low number of detections. ND= not detected

5.2.1 Campylobacter

The pilot study has identified *Campylobacter* spp. concentrations in freshwater ranging from not detected (<0.11 MPN/100mL) to three samples of 92 MPN/100mL. These three highest concentrations were sampled from three different rivers, with ruminant FST dominating in two of the samples and wildfowl in the third sample. *Campylobacter* was detected in 39/50 samples by qPCR and/or MPN. Comparison of the C*. jejuni* in this study with other New Zealand isolates shows the species isolated in this pilot study have similar genotypes to isolates from water sources and wildfowl (Appendix F).

5.2.2 Salmonella and STEC

Salmonella were detected in eight rivers, in nine samples by MPN culture and two other samples by qPCR. STEC was detected in 11 rivers in 13 samples. It was detected in one sample by MPN culture at a concentration of 0.14 MPN/100 mL and in 16 other samples by qPCR.

The majority of *Salmonella* isolates are associated with human illness, with one type which is rarely found in humans, associated with cattle. The STEC identified is uncommon in humans in New Zealand but is associated with human and animal waste. WGS detected another uncommon *Escherichia* isolate *E. albertii*. Details on the *Salmonella* and STEC isolates are given in Appendix G.

5.2.3 Viruses

Unlike the FMRP study, where adenovirus was detected in 32% of samples across all land use categories, it was not detected in any samples during this Pilot study. The norovirus and enteroviruses detected in this study were at the limits of detection. A hypothesis of this pilot study was that FST analysis could be used to identify samples for subsequent viral analysis. While this was partially true, the detection of viruses in half the samples without significant human FST markers negates this approach. However, the virus detections by qPCR without infectivity studies may not indicate infectious virus, and persistence of the viruses in the water environment may mean that the source of the low-level viruses is very distant – both temporally and/or spatially.

5.2.4 Protozoa

Cryptosporidium and *Giardia* were detected in the rivers at low concentrations ranging from 1- 30.8 oocyts/100 L and 2 to 250 cysts/100 L, respectively, with *Giardia* being the most frequently detected pathogen in the study (81% of samples) using traditional microscopy. No *Cryptosporidium* were the types which are associated with human illness.

Protozoa are very hardy with *Giardia* remaining infective for 11 weeks in water at 4°C and *Cryptosporidium* for more than 12 weeks, but infectivity will rapidly degrade at 25°C (Olson et al, 1999), so their presence may not be indicative of recent pollution. The high detection rate of *Giardia* in these rivers were selected based on sites with historically high *E. coli* concentrations. This association of *Giardia* with high *E. coli* concentrations highlights the usefulness of FIB as indicators of a faecal contamination input to a river and the requirement for further assessments. An association between *E. coli* and the prevalence of protozoa has been reported for New Zealand drinking source water (Phiri et al, 2020).

5.3 What concentrations of pathogens are a health risk?

The concentrations of pathogens found in the river water samples in this pilot study can be translated to an estimated health risk via the dose response relationship (Hazard Characterisation in Figure 17). The dose response relationship models the amount of water (and consequently pathogens) ingested and the likelihood of illness. There are many different events that can lead to ingestion of river water during recreation. The calculations presented in Table 5 gives the single point dose response estimates for the probability of infection given the concentrations of pathogens found in the water samples of the pilot study. These estimates use the maximum amount of water that would be expected to be swallowed during a swimming exposure event, and therefore, take a worst-case approach. Most swimming sessions would result in much less water being consumed. These estimates do not take into account the variability of consumption amount or the range of concentrations that would be taken into account in a full QMRA.

Table 5: Infection risk from doses of non-viral pathogenic micro-organisms found in the water samples of the pilot study

Micro-organism	Concentrations recorded in study	Average dose based on 280 mL water consumed	Single point estimate of probability of infection given dose from consuming 280 ml of water[b]
Campylobacter	4 MPN/100mL 92 MPN/100mL	12 258	0.12 (120 from 1000 exposures) 0.40 (400 from 1000 exposures)
Salmonella	0.25 MPN/100mL	6	< 0.0001 (<1 from 10,000)
STEC	0.14 MPN/100mL	< 1	< 0.005 (< 5 from 1000)
Cryptosporidium	31 cysts/100L	< 1	< 0.008 (< 8 from 1000)
Giardia	85 cysts/100L 250 cysts/100L[a]	< 1 < 1	< 0.005 (<5 from 1000) < 0.015 (<15 from 1000)

[a] Estimate based on sampling of 2 L of water, before filter clogged.

[b] Dose response calculated using the liberal dose response relationships (Lake et al, 2018)

The 52 samples in the pilot study were taken from sites associated with high *E. coli* monitoring data, and therefore, associated with a higher risk of faecal contamination. In these samples, the highest infection risk for non-virus micro-organisms remains from *Campylobacter*, which mirrors the findings of the FMRP study.

Norovirus was detected in 11 samples at levels below the limit of quantitation. While there is still some debate over the number of virons or viron clusters to cause illness at low doses, it is possible that single digit doses of virons or viron clusters are likely to cause infection in ~70% of the population (Messner et al, 2014).

While *Giardia* was the most prevalent pathogen in the pilot study, Table 5 shows that only at very high concentrations would there be likely to be >1% risk of illness. This estimated risk of infection is equivalent to Blue (A) grade water quality in Table 9, or a "Good" classification described NPS-FM 2020. None of the 725 samples analysed in the FMRP were at this concentration (McBride et al, 2002).

5.4 How the pilot study results compare to the 1998-2000 study

A key driver for the pilot was to understand whether the concentrations of pathogens in rivers had significantly changed since 1998-2000, as this would inform planning for Phase 2. This section compares the results from the two studies to see if there is any evidence of changes to the types and concentrations of pathogens and indicators identified in the rivers between the two studies.

The prevalence of detection of micro-organisms in the two studies is compared in Table 6.

Table 6: Comparison of the prevalence of FIB and pathogens in the FMRP 1998- 2000 and 2020 studies

Micro-organism	FMRP 1998-2000 Study		2020 Pilot Study	
	% detection	Detection limit	% detection	Detection limit
Total coliforms	Not measured	-	100	≥ 10 / 100mL
E. coli	99	≥ 1 /100mL	100	≥ 10 / 100mL
Enterococci	Not measured	-	86	≥ 10 / 100mL
***Campylobacter* spp.**	60	≥ 0.3 /100mL	68 46[b]	≥ 0.11 /100mL ≥ 0.3 /100mL
***Salmonella* spp.**	10	≥ 0.12 /100mL	18	≥ 0.11 /100mL
STEC	Not measured	-	2	≥ 0.11 /100mL
***Cryptosporidium* oocysts**	5	in up to 100L[a]	42	in up to 100L[a]
***Giardia* cysts**	8	in up to 100L[a]	81	in up to 100L[a]
Enterovirus	33	in 10L	4	in 10L
Adenovirus	32	in 10L	Not detected	in 10L
Norovirus GI	Not measured	-	12	in 10L
Norovirus GII	Not measured	-	12	in 10L

[a] Volume of water filtered, depended on how long it took for the filter to be clogged.
[b] For *Campylobacter* a comparison using same detection limits as FMRP.

Observations on the prevalence of detection and concentrations of micro-organisms between the two studies in Table 6 are given below.

- *E. coli* was present in all samples at concentrations at, or greater than, 10/100 mL, which is similar to the FMRP study
- *Salmonella*, *Cryptosporidium* oocysts and *Giardia* cysts were identified in a greater percentage of pilot study samples, but at low concentrations
- The limit of detection for *Campylobacter* was 0.3 MPN/100 mL in the FMRP and in the pilot study was 0.11 MPN/100 mL. Applying this higher detection limit to the results of the pilot study gives the *Campylobacter* detection rate of 46% which is lower than in the FMRP.

The data on land use concentrations of *E.coli* and *Giardia* can also be compared with FMRP. The samples from the sheep & beef and the dairy farming sites that only had wildfowl FST (n=8) have been assigned to wildfowl, while the other samples are in the observed land use (Figure 18), which includes the five samples that had a mix of high concentrations of human and ruminant FST.

In 2020, *E. coli* median concentrations were similar for all land uses (Figure 18), while in the FMRP, *E. coli* concentrations were spatially distributed with concentrations higher in “Birds” and lower in “Municipal” (Figure 19). The distribution of *E. coli* in 2020 is mostly in the upper quartile, compared to FMRP and the medians for all categories are higher than in FMRP, except for wildfowl, which only had six samples in the pilot study. This is not unexpected as the sites were selected on the basis on a history of elevated *E. coli* concentrations.

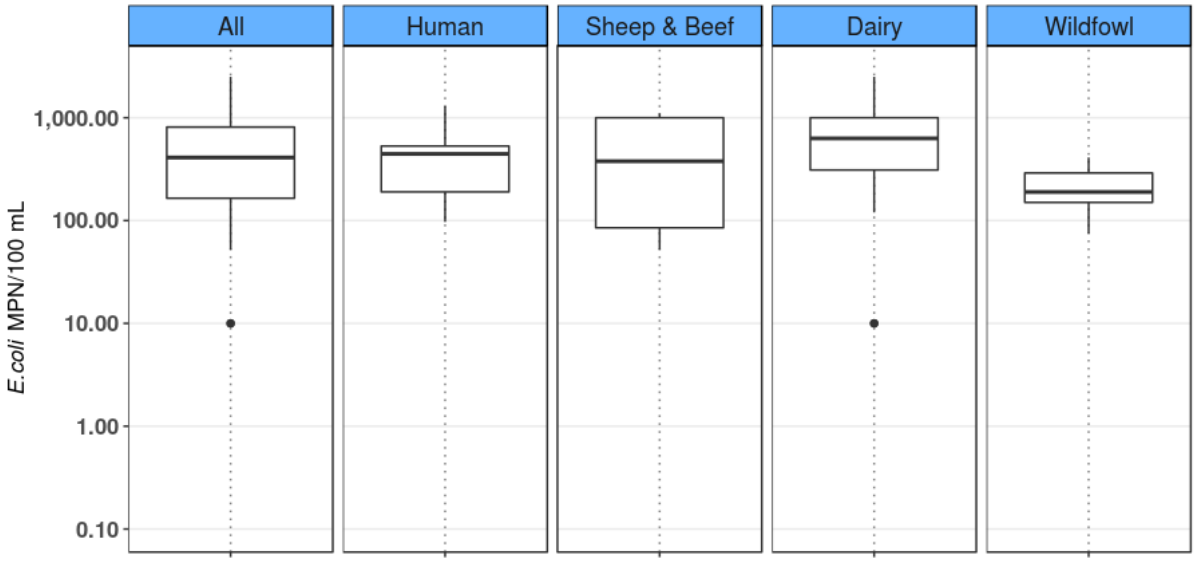

Figure 18: Spatial variability of all *E. coli* in 2020 across land use, determined from FST

Note: there are only six wildfowl samples

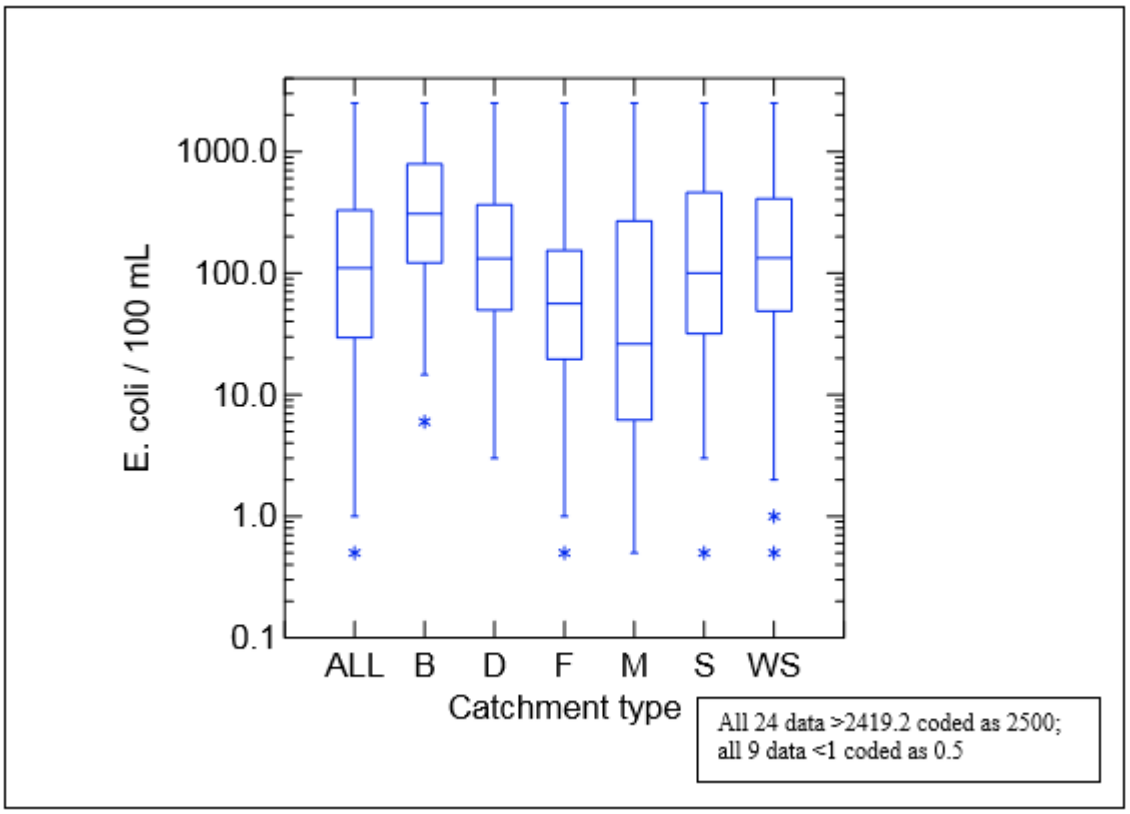

Figure 19: Spatial variability of *E. coli* across land use, from McBride et al, 2002

The concentrations of *Giardia* in the pilot study water samples were within the ranges seen for the FMRP study. However, there were some differences in the ranges of countable values when the data was split by land use. Some of this may be due to the small number of samples in the pilot study, once stratified by land use.

Figure 20 presents box plots for Giardia in the pilot study and Figure 21 data from the FMRP. The upper range of the *Giardia* concentrations are higher for Human (pilot study) compared to Municipal (M: FMRP) and for sheep & beef farming (pilot study) compared to Sheep (S: FMRP).

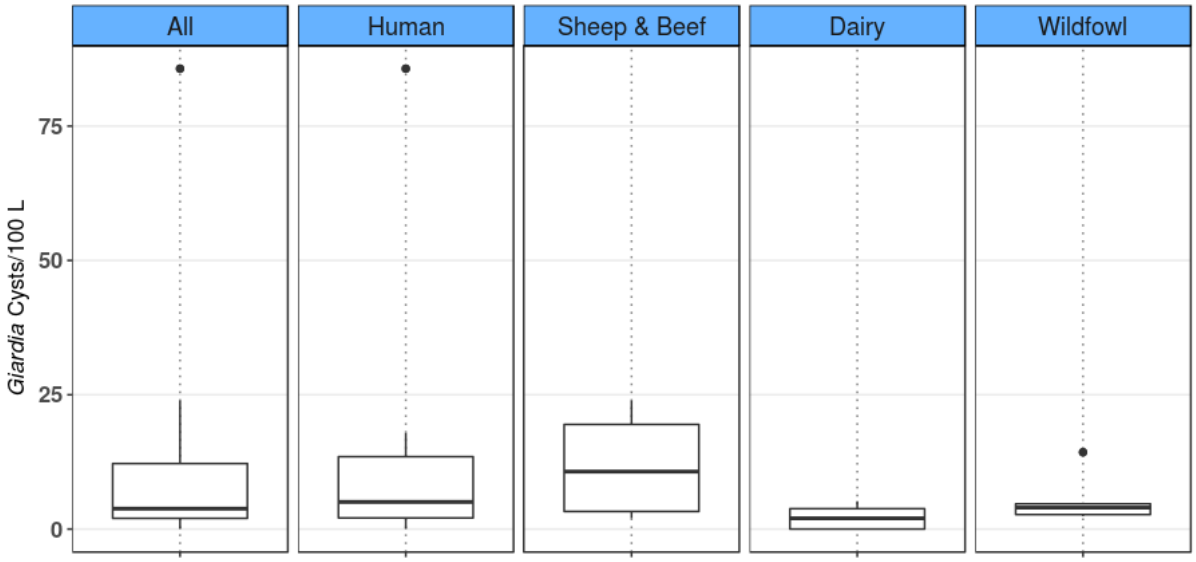

Figure 20: Spatial variability of *Giardia* in 2020 across land use, determined by FST in 2020 (excludes 250 cycts/100 L)

Note: there are only six wildfowl samples

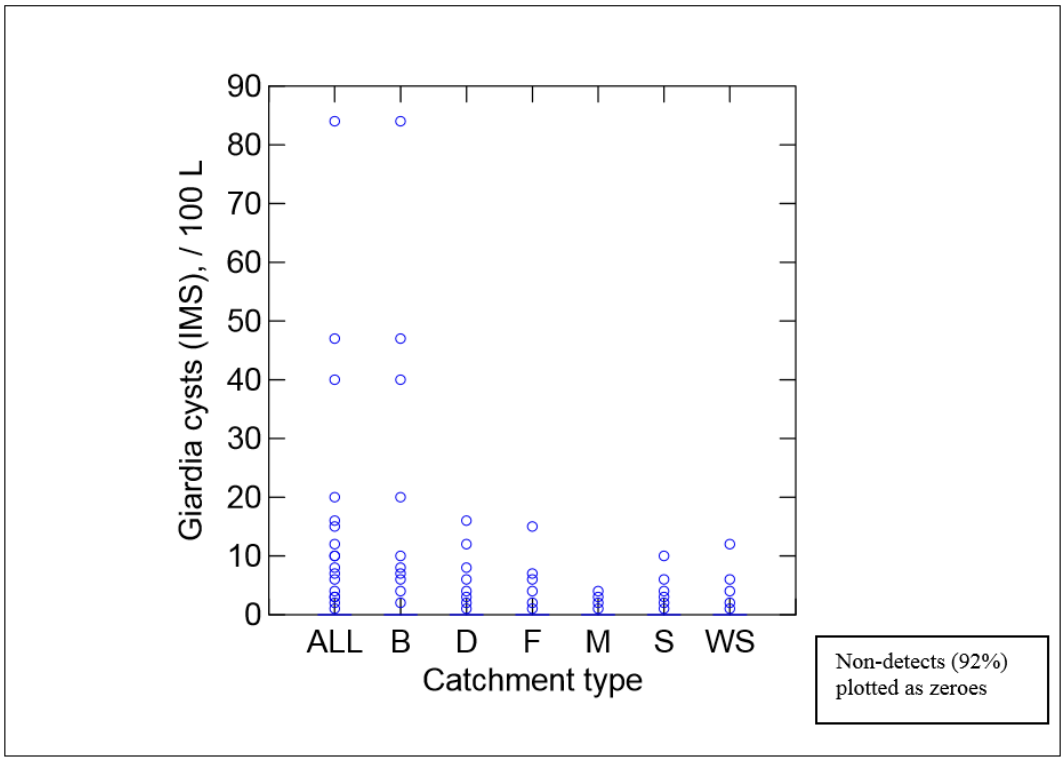

Figure 21: Spatial variability of *Giardia* across land uses in FMRP, from McBride et al, 2002

From the high concentrations of *E. coli* measured it would be expected that high concentrations of pathogens were also present. However, the number of *Campylobacter* concentrations at or near the upper detection limit is lower. McBride et al (2002) recorded 64 out of 726 (9%) of samples with *Campylobacter* concentrations greater than 100 MPN/100 mL. The upper limit of quantification for the FMRP study was 110 MPN/100mL. Other sampling conducted by ESR in Southland found 8% of samples contained *Campylobacter* concentrations in the range 100 to 999 MPN/100mL and 2% of samples >1,000 MPN/100mL. No samples in the 2020 dataset exceeded the upper detection limit or were greater than 100 MPN/100 mL.

- Taking into account that the maximum concentration of 250 cysts/100 L *Giardia* may be an overestimate as only 2 L was sampled (not 100 L), then the next highest concentration (84 cycts/100 L) is similar to that recorded in 1998-2000 (McBride et al, 2002).
- Where detected, all *Salmonella* concentrations were less than 1 MPN/100 mL. In the 1998-2000 study, four data were recorded above the detection limit of 110 MPN/100 mL.
- The range of concentrations for *Cryptosporidium* is lower than those observed in the 1998-2000 study, with a maximum of 30.8 oocysts/100 L in 2020 study and about 120 oocysts/100 L in the FMRP (McBride et al, 2002). Most *Cryptosporidium* oocyst concentrations were below 10/100 L and most *Giardia* cyst concentrations below 20/100 L.
- The *Campylobacter* species most likely to be associated with human illness are *C. jejuni* and *C. coli*. In this study, a higher proportion of samples contained *C. jejuni* (40%) than in the FMRP study, where only half of the samples positive for *Campylobacter* contained *C. jejuni* ie 30%.

Enterovirus was detected less frequently compared to the FRMP study. Unlike the FMRP study, adenovirus was not detected in this study. Because the virus results are presence/absence it is not possible to comment if the overall distribution of counts likely in the water is different between the two studies.

Some of the differences in the prevalence observed in the pilot study compared to the 1998-2000 study may be due to the temporal nature of the sources of pathogens. In the previous study, enterovirus detection was greatest in the winter months and dipped in late summer, while *Salmonella* peaked in winter and *Campylobacter* peaked over the first summer period. Site selection may also play a part for the higher frequency of detection and concentrations, as the 16 sites in the pilot study had a history of high concentrations of *E. coli.*

5.5 Association between indicators and pathogens

The NPS-FM 2020 uses a range of *E. coli* concentrations in freshwater to group freshwater sites into risk-based attribute bands. A concentration of 540 *E. coli*/100 mL is considered to present an unacceptable risk to human health for recreational use. The USA uses a criterion of 35 MPN/100 mL enterococci or 120 MPN/100 mL *E. coli* as their criteria for Freshwater Contact Recreation, which was based on studies where sites were impacted by human sewage. In this study, 28 samples were below the New Zealand criterion of *E. coli* 540 MPN/100 mL and 19 samples were below the USA criteria for acceptable freshwater recreational activity (enterococci 35 MPN/100mL). Only seven samples were below 120 MPN/100 mL in this pilot study. The higher concentrations of *E. coli* than enterococci measured in this pilot study are consistent with other studies on FIB (Korajkic et al, 2018).

Table 7 and Table 8 tabulate the number of samples where different pathogens were detected above and below the criteria discussed for:

- *E. coli* - greater than and less than, or equal to, 540 MPN/100 mL (Table 7) and less than 130 MPN/100mL.
- Enterococci -greater than and less than, or equal to, 35 MPN/100 mL (Table 8).

As two samples arrived too late for analysis by culture, this analysis is based on 50 samples. STEC is not included as it was only found in one sample. Note that many of the detections in this pilot study, relate to very low concentrations of micro-organisms, so the table does not inform the possible health risk, it illustrates the possible occurrence (or prevalence) of the pathogen in the water above and below the FIB criteria.

The tables show detections of pathogenic micro-organisms occur more frequently above the criteria for both *E. coli* and enterococci except for *Giardia* (Tables 7 - 9). *Giardia* are present in similar proportions, when split by *E. coli* concentrations, but they are proportionally lower when split by enterococci concentrations.

Table 7: Number of samples with pathogens detected stratified by *E. coli* concentration

	Total Samples	*E. coli* ≤ 540 MPN / 100mL (n=28)		*E. coli* > 540 MPN/ 100mL (n=22)	
		Number of samples with detected result	Percentage of samples with detected result	Number of samples with detected result	Percentage of samples with detected result
Campylobacter	50	16	57	18	82
Salmonella	50	1	4	8	36
Cryptosporidium	50	9	32	12	55
Giardia	50	23	82	17	77
Norovirus GI, GII and enterovirus	50	3	11	8	36

Table 8: Number of samples with pathogens detected stratified by enterococci concentration

	Total Samples	Enterococci ≤ 35 / 100mL (n=19)		Enterococci > 35 / 100mL (n=31)	
		Number of samples with detected result	Samples with detected result (%)	Number of samples with detected result	Samples with detected result (%)
Campylobacter	50	7	37	27	87
Salmonella	50	1	5	8	26
Cryptosporidium	50	5	26	16	52
Giardia	50	13	68	27	87
Norovirus GI, GII and enterovirus	50	3	16	8	26

The range of pathogen concentrations varied between water samples from not detected (ND) to higher concentrations across the two *E. coli* and enterococci stratifications, with a higher upper range generally observed in the indicator criteria with higher concentrations (Table 9).

Table 9: Range of micro-organism concentrations observed when stratified by *E. coli* or enterococci concentration.

Category	*Campylobacter* (MPN/100 mL)	*Salmonella* (MPN/100 mL)	*Cryptosporidium* (cysts/100 L)	*Giardia* (cysts/100 L)
> 540 *E. coli* /100mL	0.21 - 92	0.21 - 0.25	1 – 30.8	2 – 250[b]
≤ 540 *E. coli* /100mL	0.21 – 92[a]	0.21	1 – 7.1	2 – 19.5
>35 enterococci/100mL	0.21 - 92	0.21 - 0.25	1 – 30.8	2 – 250[b]
≤35 enterococci/100mL	0.21- 4.1	0.21	1.1 - 6.9	2 - 19.5

a: Single value of 92 MPN/100 ml associated with a rainfall event (*E. coli* was 250 MPN/100 mL)
b: 250 MPN/100 L based on 2 L sample.

The diversity of the types of pathogens organisms observed in samples was also observed to increase above the *E. coli* criterion. Above 540 MPN/100mL samples with four or five types of pathogen were detected (viruses are grouped as one type of pathogen, as their detection was very low). When *E. coli* concentration was below 540 MPN/100 mL, no samples had four or more types of pathogens. All three samples with no detected pathogens had *E. coli* concentrations <540 MPN/100 mL.

In the FMRP, a Spearman's rank correlation coefficient[3] was used to assess the strength of the association between FIB and pathogens. *Campylobacter* data is only in limited bands (i.e. discontinuous) due to the statistical nature of the MPN method and there is insufficient data for this type of analysis. *Giardia* was detected more frequently than *Campylobacter* and had higher concentrations. Spearman's rank correlation coefficient was applied to *E. coli* and enterococci concentrations ranked against *Giardia* (41 and 36 paired samples), respectively. The correlation coefficients were significant and moderate at 0.534 and 0.519, respectively ($p \leq 0.05$). There are too few *Salmonella*, STEC and *Cryptosporidium* concentrations to determine a statistically robust Spearman's rank correlation coefficient.

5.6 Determining the sources of pollution

5.6.1 FST

Understanding the source of faecal contamination assists in making decision about management and mitigation measures. It was an explanatory factor for the variation in concentrations of pathogens in FMRP (McBride et al, 2002). FST is a tool that wasn't available in 1998-2000 and the FMRP study relied on a catchment assessment. The land use classification "Birds", was found to have the highest rate of detection of *Giardia*, *Cryptosporidium*, and adenovirus (McBride et al, 2002). These adenovirus results, in particular, highlighted that many of the sites in that study were likely to be impacted by multiple sources rather than solely the single, observed classification.

[3] A coefficient near 1, or -1 means that the parameters are strongly correlated (either both increasing, or one increasing as the other decreases, respectively. A coefficient of 0 means no correlation)

In the current study, many of the sites selected had been tested previously using FST which was one of the selection criteria for inclusion. The FST analysis in this pilot study matched the observed classification in most rivers, but not all, and not all samples within a river had the same dominant FST marker. For two of the rivers selected to represent farmed sources, only wildfowl markers were detected, and in some of the other farming-impacted rivers, only some samples contained ruminant markers. In contrast, all six urban rivers contained human faecal source markers, in all the samples tested. Two farming-impacted rivers also contained significant concentrations of human markers which may reflect the presence of on-site wastewater treatment and discharge systems. The ubiquitous presence of wildfowl markers highlighted that wildfowl will be a source of both indicators and pathogens in most rivers. This may confound assessment of urban and farmed sources. FST provides clarity about known and unknown sources which allows for evidence based decision making on management and mitigation.

Using FST shows that there were more categories of faecal contamination sources detected than the observed land use. Consequently, FST provides valuable information on what activities may need to be better managed. There is insufficient data to provide robust statistical analysis of the land uses (as determined by FST) and pathogens, but comparisons of observed land use categories and FST data (Table 17) are given below.

- *Campylobacter* was always detected in samples with high concentrations of ruminant and human markers.
- The pathogenic *E. coli* was detected only in one sample, with the contamination source attributed to wildfowl by FST markers.
- *Giardia* was always detected in samples with high concentrations of ruminant and human markers.
- *Cryptosporidium* was more frequently detected in samples dominated by wildfowl markers (50%).
- *Salmonella* was detected in samples with high concentrations of human FST markers (human and ruminant-human).
- Viruses, norovirus GI and GII and enterovirus were most often found in the samples with high concentrations of human FST markers (human and ruminant–human), but they were also present in the samples with high concentrations of the other FST markers (wildfowl and ruminant markers) and low human FST markers.

Birds may act as vectors, picking up material from one area and depositing it in another area, so it is not surprising that those markers are associated with viruses. Rural sites also have dwellings, lifestyle blocks or toilet facilities for people using the recreational areas, so all of these sites may be a source of human viruses.

Despite mixed sources of pollution, FST was able to clearly demonstrate human, ruminant and wildfowl sources of pollution in the rivers. While faecal source identification does not directly relate to health risk, assessment of the health risk from the identified source, when calibrated by concentrations of indicators could be used to estimate the overall potential health risk at a river. FST qPCR methods, which were not available for the original FRMP study, were confirmed as valuable tools in alerting authorities to potential sources of pathogens associated with specific faecal sources. Greater refinement of the health risk assessment could be achieved with additional knowledge of the average prevalence and concentration of pathogens in the faeces of various animal species.

5.6.2 Faecal sources

As these pathogens are mostly at low concentrations in the river water, an alternative approach would be to sample faecal material from the potential sources. Concentrations of pathogens in faecal sources from wildfowl, cows, sheep, pigs, dogs and rats would be higher and more readily detected. The approach of characterising the pathogens present in the faeces of New Zealand animal/avian species would generate a useful pathogen data set. The New Zealand-specific data already determined and the knowledge gaps are presented in Table 24, Appendix I. Faecal source data could be used to understand the potential health risk associated with a particular animal/avian source when that source is identified as a contributor to faecal contamination in a waterway. Complementing the data with river water measurement of *Campylobacter* concentrations would take into account attenuation of pathogen viability where faecal material is not directly deposited into waterways.

5.7 Quantitative PCR as an alternative to traditional microbiological analyses

One of the aims of this study was to assess via direct comparison with the traditional methods (culture or microscopy), the feasibility of detecting pathogens directly in filtered water samples by quantitative PCR.

5.7.1 Correlation between bacterial qPCR and culture method

MPN analysis of pathogens is a resource intense activity that is not feasible to undertake on a routine basis. Quantitative PCR offers the advantage that, provided a sample is collected and filtered, decisions on what pathogens to test can be made later and undertaken on a sequential basis. MPN analysis requires immediate analysis of the samples but does provide the potential for isolation of pathogens for characterisation by biochemical and/or whole genome analysis. In contrast, quantitative PCR does not, currently, allow cost-effective characterisation of pathogens, although improvements in metagenomic analyses of water DNA extracts may provide this in future studies. A useful compromise is to perform single enrichments of 1 L water samples to provide pathogen isolates, and to concurrently, filter water samples for DNA extraction for quantification of pathogens by qPCR.

A strong correlation was observed between the *E. coli* qPCR and the MPN culture method. For enterococci this was only observed in samples where strong human sources were detected. The Enterolert MPN culture method detects both faecal enterococci and the environmental species of enterococci that have no association with faecal sources (Devane et al, 2020). Therefore, the strong association of enterococci with human faecal sources may reflect that the Enterolert was measuring faecally–derived *Enterococcus* species not environmental sources. For these indicator organisms, high concentrations of gene copy equivalents were detected, making for robust, reliable detection.

In contrast qPCR bacterial pathogen concentrations were very low. The highest level of *C. jejuni* detected was 490 copies/100 mL, with a *C. coli* maximum of 130 copies. Except for one sample *E. coli* with *stx*1 and *stx*2 concentrations of 170 and 280 copies/100 mL, all other pathogen detections were less than 100 copies/100 mL. *Campylobacter* was detected in 39/50 samples by qPCR and/or MPN and qPCR for *Salmonella* and STEC detected more positives by qPCR than by MPN culture, but again detection was at very low levels of the pathogen target. Combining qPCR and culture MPN increased the number of samples in which bacterial pathogens were detected. This pilot study provides some support for a single enrichment of 1 L water samples to provide pathogen isolates, and to concurrently, filter water samples for DNA extraction for quantification of pathogens by qPCR.

The low concentrations of bacterial pathogens and banded enumeration results for the MPN method meant that there was no strong statistical correlation between MPN and qPCR methods for any of these bacterial pathogens. Generating this type of calibration curve between the two methodologies for detection of pathogens in water samples is going to be difficult and compounded by the different organisms detected.

Having demonstrated a useful relationship between detection of FIB by culture and by qPCR in this study, a calibration of qPCR using serially diluted cultured isolates or samples of known concentrations of target microbes would be a better way to determine qPCR relationship to the actual number of microbial species in a water sample.

5.7.2 Correlation between protozoa qPCR and microscopy method

The correlation between traditional methods of microscopy for determining *Cryptosporidium* and *Giardia* and qPCR was not representative because of the level of the lower detection limit, which is dependent on the volumes of water filtered. A maximum of 2 L was sampled for qPCR detection of protozoa, compared with 2-100 L volumes for detection by microscopy methods. At present, there is insufficient data for a correlation between the two methods. In addition, qPCR is likely to only be useful where *Giardia* or *Cryptosporidium* concentrations are greater than 100 cysts/100 L.

5.8 Impact of rainfall

As rainfall saturates the ground, land runoff occurs, which may transport pathogens present in the faecal material deposited on the ground to rivers. Surface flow has been shown to be responsible for 68% of the flux of *E. coli* from poorly drained pasture, and is most significant in the early spring period (Monaghan et al, 2016).

In this study, rainfall data for the rivers was taken from the nearest rainfall station, with only one site within a kilometre, the remaining rainfall measurement sites were 2-18 km from the river. It was noted by Councils that the rainfall data may not accurately reflect the rainfall in the immediate catchment. Conditions that would result in surface runoff flow were not able to be determined during the study, so a minimum of 2.5 mm of rain in the past 72 hours has been used as a criterion for a rainfall event.

This study did not specifically target rainfall events, but thirteen samples were collected following rainfall events of 2.5 mm or more in the previous 48 hours and another nine samples collected when there had been rainfall in the previous 72 hours (Table 21). *Giardia* was the most frequently detected pathogen with the highest concentrations and there was a good number and range of FIB concentrations, so these three micro-organisms were used to assess rainfall impacts, within the limited nature of a pilot study. The concentrations of *E. coli*, enterococci and *Giardia* data were normally distributed so a two tailed t-test could be used to determine if rainfall affected the mean concentrations of *E. coli*, enterococci and *Giardia* collected when there had been prior rainfall and without recent rainfall (i.e. no rainfall within 72 hours).The t-test showed that there was a significant difference in the means for enterococci data collected when there had been rainfall within 48 hours prior to collection ($p = 0.05$), but not for *E. coli* or *Giardia*.

The following observations are made on the prevalence of samples with pathogens detected.

- *Campylobacter* was detected in 10/11 samples with rainfall within 48 hours of sampling and 16/22 samples where there was rainfall within 72 hours prior to sampling.
- Five of the seven samples with NoV GII and both samples with enterovirus occurred when there had been rainfall over the past 72 hours.

- Three of the five highest concentrations of *Campylobacter* were samples collected following rainfall.
- Of the five samples with significant concentrations of both ruminant and human contamination, four were collected within 48 hours of rain events of 28-43 mm.

As most of New Zealand was in drought during the current study, there were less rainfall events in many areas. Sampling over periods where rainfall is expected to be more frequent would provide data on the risks from land runoff following rain events.

5.9 Water quality parameters

Turbidity can be used as a surrogate for rainfall. In the FMRP, it was identified as an important explanatory variable for the concentrations of pathogens and indicators and was proposed as a key parameter, rather than flow, as it was easier to measure (McBride et al, 2002).

Turbidity data depends on the instrument used and so it was measured in the laboratory to achieve consistency, although it would be at the 24 hour time limit for measurement. It was noted that the laboratory data and field data shows significant differences (R^2=0.77). It would be better if all field data was recorded with the same equipment.

While McBride et al (2002) found that turbidity and rainfall were correlated, there was a poor linear correlation across all rivers (R^2 values of 0.186 and 0.157 for 48 and 72 hours cumulative rainfall, respectively) and Spearman's rank correlation coefficient (0.15). The turbidity dataset was dominated by a few very high results with nine results greater than 5 NTU. Of these nine samples, five were from sampling occasions when there had been significant rainfall, between 11mm and 43 mm, in the previous 72 hours. Turbidity and water clarity did not have a good correlation either.

Correlations between water quality parameters and pathogens and been found for DO, pH, conductivity, total suspended solids and temperature (Bradshaw et al, 2016). They reported a negative correlation between temperature and *Campylobacter* and a positive correlation between temperature and *Salmonella* using Spearman's rank correlation coefficient. In the pilot study a moderate negative correlation was found with temperature for enterococci and *E. coli* (-0.46 and -0.37, respectively, $p<0.05$). Weaker correlations were found with DO and *Giardia* with a correlation coefficient of 0.31 ($p\leq0.05$). No other statistically significant correlations were found with conductivity, turbidity or pH (Table 10).

Table 10: Spearman's rank correlation coefficients for water quality parameters, FIB and pathogens. Statistically significant results are shaded

	pH	DO	Temperature	Conductivity	Turbidity
E. coli	-0.29	0.17	-0.37	-0.09	0.24
Enterococci	0.09	0.23	-0.46	-0.26	0.13
Giardia	0.30	0.31	-0.27	0.00	-0.09

Continued collection of temperature, DO, conductivity, pH and turbidity should continue in a larger study to provide a larger dataset to assess their potential as explanatory factors and to provide information on the general water quality of the site.

5.10 Iwi engagement

Pilot study sites were selected on criteria which did not include the values, aspirations and perspectives of iwi and hapū, although some sites have particular relevance as they were identified in Iwi Management Plans eg Kaiate Falls at Kaiate River, and Hatea River at Whangarei Falls. Engagement with iwi and hapū about this project was initiated through the regional and unitary councils, including contact with iwi liaison officers, where available.

When the project team visited each site, iwi and hapū were invited to meet to kōrero about the project or to kōrero by phone. After initial contact, 21 representatives from iwi and hapū attended a site visit or engaged in kōrero. Feedback was received on the value of the site visit and/or kōrero to introduce the project. These iwi and hapū representatives were contacted again after sampling was finished to inform them of progress, the impact of Covid19 on the project, to identify preferred methods of communication and, as participants, what information they would like as feedback.

5.10.1 Feedback

General feedback on the pilot study was requested.

- What worked well?
- What could be done better?
- What should we be doing for the future?

Key themes from the feedback from iwi and hapū were:

Health - ensuring that recreational water was of suitable quality

- people still swim at the sites as “it is part of who we are”.
- mahinga kai and recreational water quality are interlinked.

Choice of site - the lack of consultation with iwi and hapū in the choice of pilot study sites[4].

- Some recreational spots are very popular and there is an ongoing potential health risk, as swimming, or collection of mahinga kai, still occurs despite notices.
- It is a priority to retain a relationship with the awa.
- An interest was expressed in iwi and hapū identifying different freshwater sites.

Information feedback options proposed included

- hui
- technical paper such as a conference paper
- newsletter type pamphlet with a narrative rather than a numeric approach, highlighting “what it means” and “so what”.

Barriers to engaging

While a Cultural Health Index (CHI) approach had been proposed in the initial full study, there was only a small opportunity for contact with iwi and hapū in the pilot study. It has been identified that

[4] It was explained that the pilot sites had been chosen because they were highly impacted so that the methods could be tested and to ensure that it would be likely that there were pathogens present

resourcing was an issue for iwi and hapū, in that there are a high number of requests for engagement and only limited resources to respond.

Another barrier is identifying the most appropriate iwi or hapū person to contact, as processes vary around the country. Some Councils had close contacts with iwi or iwi liaison staff and in other areas alternative ways of identifying contacts were used where possible, which included personal connections and recommendations. As the project progresses, it is expected that contacting people who are mandated to speak on behalf of their iwi and hapū will become easier.

Process

Comments echoed those made previously in the development stage hui, that the process should have more comprehensive engagement from the beginning for co-design. There were a variety of views on other information which could be collected with some embracing a CHI approach (Rainforth and Harmsworth, 2019), while others proposed more individual approaches. Enthusiasm was expressed for using both the CHI and western science approaches which support iwi and hapū requests that sites of cultural significance be included to explore both systems of knowledge. All approaches require someone with spare time to undertake them.

5.10.2 General comments

The pilot study afforded the opportunity for valuable engagement with local iwi and hapū. Feedback from iwi and hapū has been collated to inform the Phase 2 design of the freshwater study. Of particular note were iwi concerns about the lack of incorporation of sites that were of significance to Māori and that reflect iwi and hapū values, aspirations and perspectives of tangata whenua. Selection of such sites in Phase 2 would enable iwi and hapū to gain a better understanding of water quality of those river sites, beyond the usual monitoring information. This information would then facilitate management of their awa.

Relationships initiated during this current work can be used to invite input into phase 2 selection of sites and incorporate at least one of cultural significance in each region. As a minimum CHI, or similar should be developed, with financial support to acknowledge the intellectual property which iwi and hapū bring, and to facilitate fuller participation.

5.11 Summary

The goal of this study was to inform the implementation of a large-scale study to replace the 1998-2000 FMRP study. This includes understanding the role of new methodologies and the current concentrations and prevalence of FIB and pathogens in a selection of New Zealand rivers. The faecal indicator *E. coli* was detected in all samples in this pilot study, with 32% equal to, or having a concentration greater than, 1000 MPN/100 mL. All target pathogens, *Campylobacter*, *Salmonella*, STEC, *Cryptosporidium* and *Giardia* were detected by traditional methods, and/or qPCR, mostly at low concentrations.

Compared to the FMRP, the *E. coli* data shows that the sites were more contaminated than in the FMRP study. There was an increase in the prevalence of protozoan pathogens, but concentrations were similar. More of the *Campylobacter* spp. were identified as being of the types that cause illness in humans (*C. jejuni* and *C. coli*) than in the FMRP.

Where FIB exceed water quality criteria*, E. coli* 540 MPN/100 mL (New Zealand) or enterococci 35 MPN/100 mL (USA), pathogens were more prevalent, and more types of pathogens were also observed. This pilot study, with its limited dataset, supports the continued use of FIB as useful indicators of faecal contamination, and the potential for pathogens to be present in a water body.

Faecal source tracking qPCR methods, which were not available for the original FRMP study, were confirmed as valuable tools in alerting authorities to potential sources of pathogens associated with specific faecal sources. This enables better evidence based decision making on management and mitigation.

E. coli qPCR methods can provide equivalent detection with traditional culture-based methods. This is also true for enterococci when the source is dominated by human only sources. The value of qPCR methods as a supplement to culturing methods was illustrated by the detection and quantification of *Campylobacter coli*, STEC and *Salmonella* by qPCR.

Environmental factors and water quality were assessed for their potential impact on pathogen concentrations. Even though there had been very low rainfall at most sites during the sampling period and consequently a small dataset, rainfall had a significant impact on between the mean concentrations of enterococci in samples collected with rainfall within 48 hours of sampling and the rest of the data enterococci.

While turbidity had been identified as an explanatory variable in FMRP (McBride et al, 2002), it was not evident in this study in terms of correlation with flow, rainfall or water clarity. There was a moderate association between water temperature and FIB and a weaker association with protozoa and DO. Continued collection of temperature, DO, conductivity, pH and turbidity should continue in a larger study to provide a larger dataset to assess their potential as explanatory factors and to provide information on the general water quality of the site.

The pilot study afforded the opportunity for valuable engagement with local iwi and hapū at each river location. Feedback from iwi and hapū have been collated and will inform the Phase 2 design of the freshwater survey. Of particular note were iwi concerns about incorporation of river sites that were of significance to Māori. Selection of such sites in Phase 2 would enable iwi and hapū to gain a better understanding of those river sites, beyond what the normal resources of iwi and hapū could deliver. This information would then facilitate their implementation of impact mitigations.

It is recommended that the iwi and hapū contacts that have been made are further developed and extended to allow iwi and hapū to assist in co-design of the project which would include nominating sites that are important to them and to reflect iwi and hapū values, aspirations and the perspectives of tangata whenua. As a minimum CHI, or similar should be developed for the sites in the full study and supported financially to acknowledge the intellectual property which iwi and hapū bring, and to facilitate fuller participation.

6 Next steps and recommendations

6.1 Full study for QMRA for freshwater recreational guidelines

6.1.1 Study plan based on feedback from pilot study

The collection of samples was carried out by staff from the Regional and Unitary Councils. Feedback from Councils indicated that the sampling required extra resources, as the time taken for sampling was longer than routine water quality sampling, mostly due to the time required for filtering samples on-site for protozoa, which normally took an hour. Some sampling sites required two staff to assist with measurements. In most cases, doing more than two sites in a day and delivering the samples to the courier on time, would not be possible. Sites closer to the Council offices and/or courier drop off were an advantage as travelling could take an hour or more for one sample.

However, if the sampling was more similar to the council's routine sampling then more samples could be collected in one day. Most councils had specific recreational water quality monitoring programmes over the summer and sampling could fit into this allocation of resources. Only one council confirmed it would have resources to do extensive sampling throughout the year, although other councils mentioned needing a lead in time of up to six months to budget and organise resources. At some sites with ongoing water quality issues, council projects had been initiated which could provide some additional resources. Unless additional funding was provided, a simpler and quicker sampling and field measurement programme is recommended to fit in with Council resources.

To co-design the project with iwi and hapū, there will be the need to be support for hui to include sites of importance to iwi and hapū and to gather knowledge about the awa from a more holistic Te Mana o Te Wai perspective. This might require different resources and assessments. These sites may not have a history of testing as Council selection criteria for monitoring may differ.

Using the information from the pilot study, a full study plan, akin to the FMRP 1998-2000, has been prepared with costings and provided to the Ministry for the Environment. *E. coli* and enterococci are the indicator organisms most readily applicable to routine, widespread surveillance and they have proven to be useful indicators of contaminated water internationally and in New Zealand. *Campylobacter* is the important target pathogen, as the basis of the previous standards, its frequent detection in this pilot study and its ongoing high rates of human illness in New Zealand. *Salmonella* and STEC are also included. FST analysis was shown to be effective in determining the likely faecal risk and potential sources of the contamination.

Although the concentrations were low, the frequent detection of protozoa in this pilot study highlights the importance of this group of pathogens. This may reflect the choice of sites known to have elevated *E. coli* concentrations. We have included them in the full study proposal, but recognise that many councils would need additional resourcing to undertake this sampling. Separation of the protozoa sampling into a dedicated study may be a more practical option.

There is limited evidence from this study or others that virus concentrations will correlate with other indicator organisms, except at very high concentrations (Korajkic et al, 2018). Therefore it is proposed to that viruses are excluded, which will reduce costs considerably. At those concentrations (where faecal source tracking markers and indicators are high and easily detected) the risks from viral and protozoa can be predicted much more readily based on assessment of faecal indicators and sources using risk assessment approaches.

6.1.2 How effective would a full study be to underpin the QMRA?

The driver for the proposed full study is to provide robust reasoning for both the setting of water quality human health risk attributes for freshwater recreational sites, and to inform the structure and parameters of the QMRA, which underpin the attribute bands and potentially be a tool to help councils comply with NPS-FM requirements. The use of data in the QMRA has a number of stages as shown in Figure 22.

Figure 22: Data flow for the freshwater QMRA

The full study could inform the following:

- which indicator micro-organisms and environmental variables should be included as inputs into the risk model
- how to convert from indicator variables to pathogen doses
- how to evaluate the public health risk.

The dose-response relationship will not be informed by a full study and is unlikely to be updated in the foreseeable future due to ethical reasons. There is not sufficient historical notified case typing data for campylobacteriosis available to consider differential infectivity rates in the dose response relationship.

Table 11 gives a summary of how the data produced by a full study will contribute to the study goals along with the associated strengths and risks of using this format of data collection.

Table 11: Summary of the goals for a full study and the associated data-related strengths and risks to underpin a QMRA revision of the freshwater microbial water quality guidelines

Improved understanding of the relationship between pathogens potentially in the water and indicators used for Water Quality monitoring.	
How	Collect up-to-date pathogen, indicator prevalence and concentration data from a wide range of rivers across New Zealand. Use data to understand the relationships between pathogen and indicator variables.
Strengths	Data will be spatially and temporally diverse, not focusing on specific scenarios, so potentially more applicable to national targets. Data collected in a standardised way. Data mainly chosen from sites where we expect to have higher *E. coli* concentrations. Data collected over 20 months (two bathing seasons) to account for variations in seasonal weather patterns.
Risks	Unexpected weather patterns, may produce biased dataset, eg if a very dry year. The variability of the data across the diverse range of river sites may make it difficult to improve on the conclusions of the 1998-2000 study.

	Highly contaminated rivers targeted in the study may not answer the queries/concerns of iwi and hapū as outlined in their feedback on the chosen river sites not being representative of sites of cultural significance.
Improved understanding of the impact of land use and environmental factors on water quality and health risk.	
How	Collect data on land use, environmental and river data alongside the micro-organism data. Use data to conduct multivariate modelling to find relationships between variables.
Strengths	Data covering a wide range of locations and land use types
Risks	There will be less flexibility of choosing when to sample in a large-scale study, which means some environmental data may not be well represented in the final dataset. To understand the effects of specific factors on health risk, enough samples need to be taken relating to the different states of these factors. This is unlikely to happen with sampling at regular time intervals. Actual risks may not be that strongly associated with the dominant land use in large catchments.
Improved human health risk estimates	
How	Improved understanding of the relationship between pathogens, indicators and environmental factors would improve human health risk estimates for different scenarios.
Strengths	Including FST and genotyping information may help to inform health risk when combined with analysis of the potential loading of faeces and methods of transfer to the river.
Risks	Source attribution studies require prior knowledge of the potential sources of the genotypes, it is possible to fall into the trap of only seeing what you are looking for. *Campylobacter* notification isolates are not routinely typed, so will be difficult to match types to potential health risk on a national scale. Health risk estimates can only be improved if robust relationships can be established between proposed monitoring variables and the presence and potential doses of pathogens consumed by recreational water users.

6.2 Alternative Approaches

Updating the data on the state of water quality in New Zealand rivers and consequent risk of infection provides essential evidence to underpin the current Guidelines. However, an ongoing concern with the current Guidelines is the lack of a formalised approach when water quality criteria are exceeded. As well as the overriding requirement to increase the number of rivers in New Zealand that are suitable for primary contact, and to improve water quality, the NPS-FM 2020 requires Councils to prepare Action Plans where Attribute numeric values are exceeded in a catchment (Table 22, NPS-FM, 2020). To do so, councils need to better understand the causes of poor water quality.

The potential causes of faecal pollution are not necessarily those which are most obvious. A key advance in this pilot study compared to the FMRP of 1998-2000 (McBride et al, 2002) has been the ability to use FST to more accurately identify the sources of faecal pollution. A consistent, nationwide, less complex framework with standardised tools to determine the cause of faecal contamination, and site-specific QMRA to assess risk to human health, supported by case studies would assist Councils to prepare more effective Action Plans to improve the water quality.

There are two main components to water quality assessment as illustrated in Figure 23. The source of the pathogens will dictate what pathogens might be present. The concentrations of pathogens may be reduced during the transport to the water body, where the health risk is expressed based on the source of the pathogens and pathways. As demonstrated in this pilot project, there are logistical and analytical constraints around the routine direct detection of pathogens, nor is the detection of target pathogens sufficient to provide protection from all pathogens. Instead, indicator organisms provide a cost-effective pathway to estimate health risk, because they signal the occurrence of a faecal contamination event, and therefore, the likely presence of pathogens.

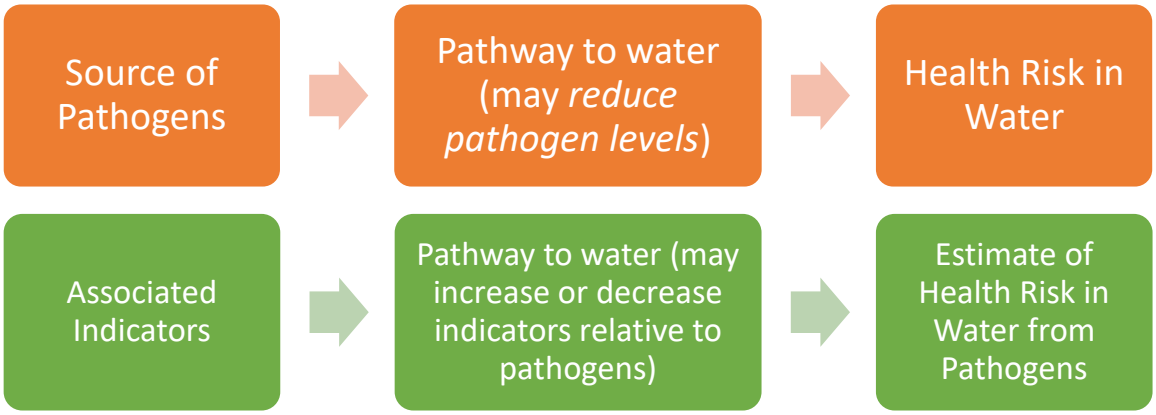

Figure 23: Summary of process for assessing health risk from contaminated water

Andreas Farnleitner and colleagues at Karl Landsteiner University of Health Sciences in Austria have recently proposed a "framework for integrated faecal pollution analysis and management" for drinking water source protection, based on the World Health Organization approach (Savio et al, 2018, Farnleitner et al, 2018). There are three interacting levels ("three-step approach") which characterise the approach:

a) is there a problem with faecal pollution?

b) if yes, what is the reason for it? and

c) what is the actual health risk related to the faecal source(s) that contributes to the observed pollution?

This drinking water-based source assessment approach is equally applicable to recreational water assessment. To make a transformational impact on recreational water quality management in the near future, we propose that we should progress directly to recreational water quality guidelines which crucially include specific guidance on what should be done when guidelines are exceeded and how to do this. In Figure 24 we present a conceptual framework for improvement of water quality. Many of the components of this framework are in existence: NPS-FM 2020, Freshwater Management Plans, State of the Environment monitoring and reporting, Our Land and Water programmes, Recreational Water Quality Guidelines, FMRP. The drivers would be that the framework simplifies assessment of health risk and expedites action. This framework has the advantage of being similar to source water protection and therefore will be mutually supportive. Table 12 illustrates the alignment with the WHO framework for drinking water.

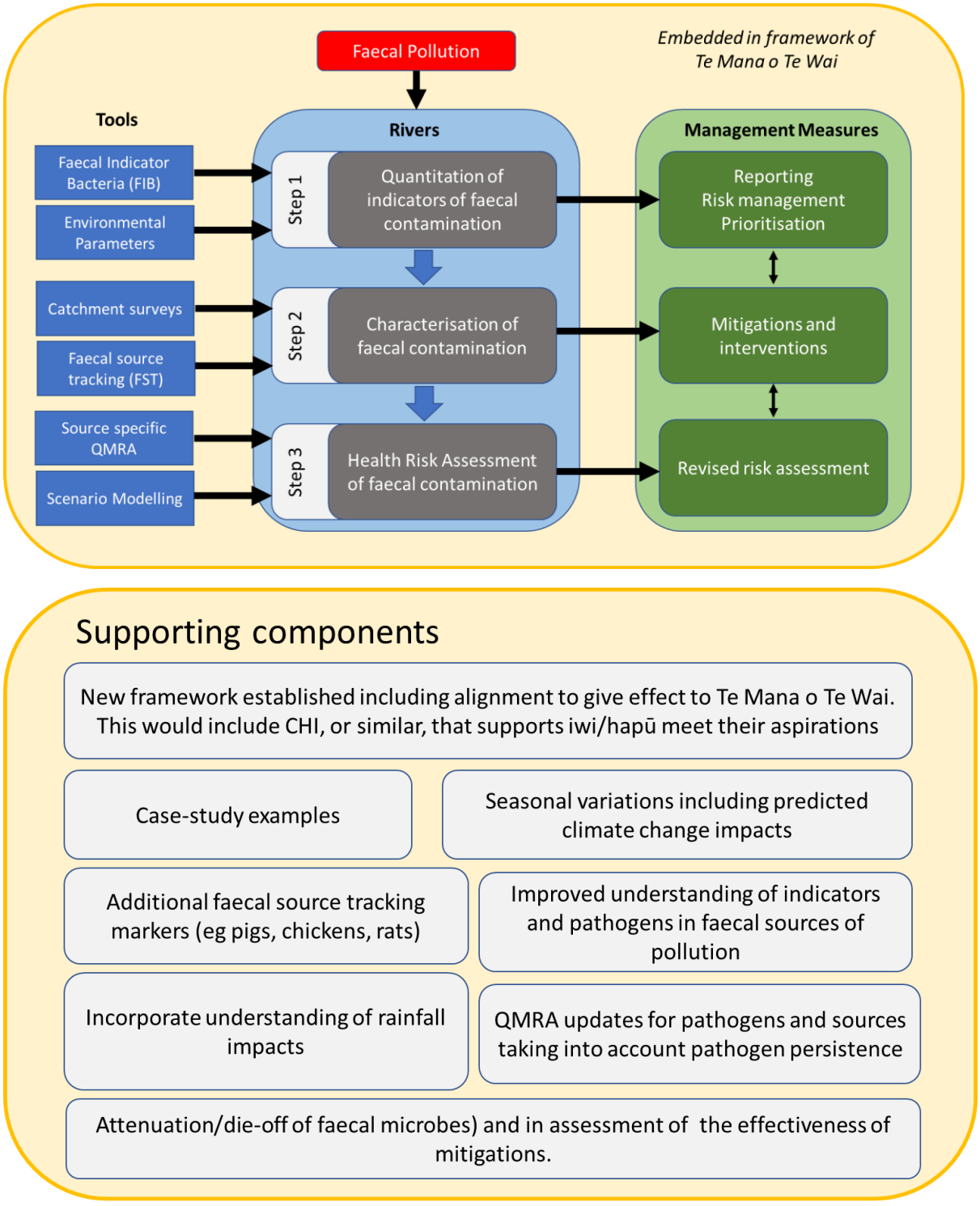

Figure 24: Conceptual framework for water quality assessment using quantitative microbial-risk assessment with supporting components. Wire diagram adapted from Savio et al, 2018.

Table 12: Alignment with WHO framework

WHO Approach to Drinking Water	Management Measures	Framework
Is there a problem with faecal pollution at this site?	Report	• Monitoring of faecal contamination using indicators such as *E. coli* (possibly in conjunction with enterococci) and other water quality physiochemical or environmental measures and CHI • Do we have enough data to make an assessment – how many samples, over what time period eg Tables 9 and 22 of NPS-FM 2020
If there is, what is the reason?	Catchment management and pollution protection	• Explicit guidance on how to investigate water quality to determine sources of contamination, and therefore appropriate mitigations or interventions (direct linkage with industry and research programmes such as Our Land & Water to guide intervention options) • What is the source of the indicators • Catchment assessment with FST (Genetic/PCR MST markers) to characterise the type of faecal contamination and variability of faecal sources
What is the actual health risk related to the faecal source(s)?	Mitigation	• In the absence of other information, we should assume that contamination is from human and or direct deposition from the ruminant source. This is a worst case scenario but provides maximum public health protection. This would be the current, default QMRA which is based on this assumption with default numeric water quality values. • If source of faecal contamination is not human or direct deposition of cattle faecal material, use case studies QMRA to assess health risk for other sources (eg bird, pigs, dogs) • Provide guidance to community on where it is safe (or unsafe) to swim • Guide regulators and community on where changes to practices or interventions are needed, and importantly what should be done and how. • If mitigations or interventions are not immediately able to be implemented, or faecal source analysis suggests low risk sources of pollution, then site-specific health risk assessment using tools such as QMRA and scenario modelling may be required to support interventions or justify modification to health risk.

6.2.1 Te Mana o Te Wai

The NPS-FM 2020 requires Freshwater Management that gives effect to Te Mana o Te Wai in managing freshwater in New Zealand, recognising that the first obligation is to the health of the water. Long term visions for water bodies are to be informed by iwi aspirations. The health of an awa can be assessed using models such as a CHI, eg Mauri compass or Mauri model. Different tools for assessing the health of a water body are used all across New Zealand (Rainforth and Harmsworth, 2019). They noted that Waikato and Bay of Plenty use a wide range of the tools. This was reflected in discussions with iwi, who didn't support a standardised tool. Assessment of each awa will be specific to each iwi and hapū. Key elements may include mauri, tikanga and cultural practices, significant sites, the absence or abundance of species and their health, as well as access to sites, landscape, land use, riparian habitat and water quality. Many iwi and hapū were very enthusiastic about jointly using their model CHI and western science, bringing the two knowledge systems together. There is also the juxtaposition of a holistic view of an awa, alongside the site specific nature of the recreational water quality QMRA project. However, any project on freshwater quality can also inform other values, such as mahinga kai, as it will provide an understanding of pathogens within the awa.

This project provided the opportunity to engage with 21 iwi and hapū and interested parties either at the site, or kōrero while visiting each region. Feedback included the need for action to be taken to actually improve water quality (the "so what" factor). By developing a tiered framework, site specific QMRA and tools to assist Councils to investigate and manage sources of faecal contamination, this project assists iwi and Council with the long-term vision of improving water quality.

There is high interest in the pilot stage of this project and iwi have requested feedback as a technical report, hui and summary of the information for general distribution. Having engaged interest in the project, it is important to continue to provide the opportunity to co-design case studies. The intellectual property which iwi bring to the project needs to be supported through allocation of funding for hui and cultural assessment or equivalent work.

6.2.2 Justification

Councils are required to have action plans where the target numeric attribute states are not being met. The data from the most recent assessment of New Zealand's recreational water quality in Our Freshwater New Zealand 2020, had 375 sites with more than 60 data points over five years, which therefore, could be classified according to the Attribute State Bands in NPS-FM 2017 (Larned et al, 2018). In general, urban sites had the poorest quality water, followed by pastoral sites. There were 264 sites with *E. coli* 95th percentile > 540/100 mL (Poor) and 36 sites with 95th percentile between 260 and 540/100 mL (Fair) according to the NPR-FM 2020. The raw data classified 135 sites as Band E, 114 as Band D, 6 as Band C, 41 as Band B and 79 as Band A. Appendix 3 in the NPS-FM 2020 shows that 29% of freshwater currently does not meet the target for primary contact recreation.

During the bathing surveillance season councils are required to investigate potential sources of microbial contamination when a sample is above 260 *E. coli*/100 mL and at concentrations above 540/100 mL, advise the public that the site is unsuitable for contract recreation until further sampling shows the *E. coli* concentration is less than 540/100 mL. A sanitary inspection may highlight the obvious potential sources of faecal contamination but as discussed above, often the observed catchment use is not the dominant source of faecal contamination. With the advent of FST for better identification of potential sources of faecal contamination, further guidance can support decision making in remedying the situation and complying with the NPS-FM by preparing an evidence-based Action Plan, using scenario-based QMRA. It is proposed that we utilise this new knowledge in FST to advance the Guidelines to support Councils to improve recreational water quality.

6.3 Work streams

The alternative approach outlined above needs refining and testing with researchers in New Zealand and overseas, and with stakeholder across the New Zealand landscape. In particular, the need to give effect to Te Mana o Te Wai and incorporate Māori values and approaches. Furthermore, this approach needs to fit into a higher level assessment framework of Te Mana o Te Wai. While microbial health risk addresses a component of iwi health and well-being, it is just one aspect. A form of CHI is needed that would include the range of matters important for local iwi and hapū to meet their aspirations, that takes into account aspects such as mauri, tikanga and cultural practices, and an understanding of what is important for monitoring cultural health in the local context.

The recommended areas of research and implementation required are listed below.

1. A new framework established including incorporation with Te Mana o Te Wai.
2. While the framework is nationally applicable, it needs to function on specific rivers and catchments. A series of case-studies, building on the pilot study, would provide worked examples of the approach, identify exceptions to the rule (and how to identify them), and validate effectiveness of interventions and mitigations.
3. Improved understanding of indicators and pathogens in sources of pollution. While data exists, it needs to be collated and gaps filled both on specific sources, and on pathogens, by measuring concentrations of pathogens in faecal sources.
4. Faecal source tracking is a key component of this framework. While markers for human, ruminant and wildfowl are well established, tools for the assessment of other sources such as pigs, chickens, rats and other sources may be needed.
5. Incorporate understanding of the impact of rainfall on water quality assessment.
6. Consider seasonal variations in water quality assessment, including predicted climate change impacts.
7. Establish what the impacts of pathways from faecal source to the river are, both in terms of understanding health risk and in assessing the effectiveness of mitigations.
8. QMRA updates for pathogens and sources taking into account pathogen persistence.

Each of these topics has a range of levels of existing information available and varying levels of importance and need. A number could be accomplished in small projects by a range of researchers. For example, Table 24 (Appendix I) summarises knowledge gaps in pathogen prevalence and concentration in animal faeces. QMRA modelling can proceed initially without this information for all sources (using modelled data), but a gradual filling of knowledge gaps would strengthen the efficacy of QMRA.

APPENDIX A: Pathogens selected

A.1 Pathogenic bacteria

A.1.1 *Campylobacter*

Campylobacteriosis has the highest rate of notified bacterial gastroenteric infections in New Zealand with 126.1 cases/100,000 people in 2019 (Pattis et al, 2020). *Campylobacter* has been frequently found in rivers in New Zealand (Devane et al, 2005, Devane et al, 2014, Eyles et al, 2003, Garrett et al, 2007) and overseas (Bradshaw et al, 2016, Vereen Jr et al, 2013).

McBride et al (2002) estimated that 4% of campylobacteroisis could be attributed to recreational water, with Gilpin et al (2013) also identifying recreational water contact as a minor source of illness. A relatively small proportion of notified campylobacteriosis from recreational water contact may still reflect a significant disease burden, due to the large number of cases in New Zealand.

A.1.2 STEC/VTEC/EHEC

STEC/verocytotoxin toxin-producing *E. coli* (VTEC) /enterohemorrhagic *E. coli* (EHEC) have been found in rivers around the world. A study in Taiwan recovered these organisms at rates between 3.6% for STEC to 17.2% for enteropathogenic *E. coli (*EPEC) (Huang et al, 2016).

A.1.3 *Salmonella*

Salmonella is present in sewage effluents or animal faeces which contaminate pasture, soil and water. They do not usually multiply in soil and waters but may survive for long periods (Bell and Kyriakides, 2005). It was detected in New Zealand freshwater at a rate of 10% in the FMRP (McBride et al, 2002). Contamination in the environment can be spread by rodents or wild bird populations (Davies et al, 2013).

A.2 Viruses

A.2.4 Norovirus

Human norovirus genogroups I and II (GI and GII) are major causes of viral gastroenteritis and have been the cause of outbreaks associated with consumption of faecally-contaminated water and food (Kukkula et al, 1997, Hewitt et al, 2007). As norovirus outbreaks often show seasonal tendencies, their presence in wastewater may be more sporadic than other enteric viruses such as human adenoviruses and human polyomaviruses.

There are no recorded outbreaks from exposure of recreational water contaminated with norovirus in New Zealand, but they have been associated with outbreaks internationally (Graciaa et al, 2018). The incidence of norovirus is likely to be under-reported in New Zealand as although gastroenteritis outbreaks are notifiable, norovirus infections are not, although they may be identified as part of an investigation into an outbreak that specifically identifies norovirus.

A.2.5 Adenovirus

Adenoviruses cause infections in all mammals, but are very host-specific - bovine adenoviruses infect cows, human adenoviruses infect humans etc.

For human adenoviruses, there are seven identified species (A-G) and over 50 types. Human adenovirus species F is a major cause of gastroenteritis in young children. Human adenoviruses are ubiquitous in the human population with most humans infected by age 20. Human adenoviruses are

transmitted person to person by direct contact or aerosols through the mouth, nasopharynx or eye or commonly via the faecal-oral route. The transmission of human adenovirus species F is predominately through the faecal-oral route. Non-species F adenoviruses cause a range of symptoms including pneumonia, hepatitis, conjunctivitis and cystitis.

Data on human adenovirus prevalence in the New Zealand population are limited as adenovirus cases, and particularly of diarrhoea, would be unlikely to be reported in the surveillance figures due to the testing algorithms used and reporting requirements in New Zealand. As yet, no reported foodborne or waterborne outbreaks (except those associated with aerosolisation or conjunctivitis) have been associated with human adenovirus species F in New Zealand or overseas.

A.2.6 Enterovirus

Enteroviruses cause a wide range of diseases following exposure by the faecal-oral or airborne transmission routes. Asymptomatic/subclinical infections are common. Enteroviruses have been often used to assess human health risk due originally to the relative ease of culture and frequent detection and have been used as a faecal indicator (Puig et al, 1994, Tani et al, 1995), with standard EPA methods developed (Cashdollar et al, 2013).

A.3 Protozoa

Cryptosporidium and *Giardia* cause diarrhoea and abdominal pain. The (oo)cysts are present in the gut of humans and animals such as cattle, sheep, cats, dogs, rats and possums and can be transferred by infected faecal material.

Several waterborne outbreaks of cryptosporidiosis and giardiasis have been described in New Zealand attributed to contaminated drinking water, both from surface water and groundwater sources, and exposure to recreational water and swimming pools (ESR - Public Health Surveillance: https://surv.esr.cri.nz/surveillance/surveillance.php). Recreational contact was commonly reported as an activity in notified cases of *Giardia* and *Cryptosporidium* (ESR, 2017), although this does not necessarily mean contact recreation was causative.

Cryptosporidium and *Giardia* were detected rather infrequently (5% and 8% respectively) and at low concentrations in all catchment types in rivers in New Zealand during the FMRP (McBride et al, 2002), and typically at concentrations of 1-20 (oo)cysts/100 L. Cyst and oocyst detections were found to be strongly dissociated (McBride et al, 2002). (Oo)cysts can survive adverse conditions in the environment for months until ingested by a new suitable host. *Cryptosporidium* oocysts were found in 42% of the 114 lowland river samples, but in only about 1% of samples collected from intermediate rivers and bush catchments (Ministry of Health, 2019). Most *Cryptosporidium* oocysts in the lowland river samples were found in spring (Phiri et al, 2020). *Giardia* cysts were found in 58% of the 114 lowland river samples, but in only about 8% of samples collected from intermediate rivers and bush catchments (Ministry of Health, 2019). Prevalence peaked in September and December (Phiri et al, 2020).

Recent epidemiological studies have found for giardiasis there is no seasonal pattern in outbreak probability and an inverse association with density of dairy farming cattle, whereas in dairy farming areas cryptosporidiosis outbreaks were observed in spring and positively associated with dairy cattle density and negatively with temperature (Lal et al, 2018). Different protozoa species (eg *C. parvum* vs *C. hominis*) and sub/genotypes have different host associations (Garcia–R et al, 2017).

APPENDIX B: Sampling sites and observed land uses

Table 13: Sampling sites for pilot study

Site No.	Region	Site Name	Median *E. coli/ 100 mL*	Impact Type
A	Auckland	Oteha Stream at Days Bridge	1310	Urban
B	Wellington	Porirua Stream at Town Centre	1000	
C	Canterbury	Heathcote River at Catherine St	326	
D	Southland	Otepuni Creek at Nith St	2050	
E	Marlborough	Taylor River at Riverside Park	246	
F	West Coast	Sawyers Creek at Dixon Park	1457	
G	Manawatu-Wanganui	Manakau at State Highway 1 Bridge	606	Sheep & Beef
H	Nelson	Wakapuaka at Paremata Flats Reserve	288	
I	Bay of Plenty	Kaiate River at Kaiate Falls Rd	240	
J	Gisborne	Wharekopae River at Rere Rockslide	170	
K	Northland	Hatea River at Whangarei Falls	290	
L	Taranaki	Waitara River at Bertrand Rd	1096	Dairy
M	Southland	Moffat Creek at Moffat Road	305	
N	Canterbury	Selwyn River at Coes Ford	206	
O	Waikato	Piako River at Paeroa-Tahuna Rd	340	
P	Northland	Waiotu River at Stage Highway 1	350	

APPENDIX C: Microbiological methods

C.1 Bacterial Methods

C.1.1 *E. coli* and enterococci

River water (10 mL) was analysed for *E. coli* and enterococci using IDEXX Colilert and Enterolert assays respectively (APHA, 2017a; APHA, 2017b). This provides a detection range of <10 MPN/100 mL up to >24,000 MPN/100 mL.

C.1.2 *Campylobacter*

A 12 tube MPN analysis was undertaken to isolate *Campylobacter* spp. River water volumes of 1 x 700 mL, 2 x 100 mL, 3 x 1 mL, 3 x 0.1 mL, 3 x 0.01 mL were analysed (ISO, 2017a; APHA, 2017c; Wong et al, 2004; HC, 2014; MIMM, 2008).

Larger volumes were filtered through 0.22 µm filters, and placed into 25 mL of Bolton broth. Analysis of volumes 1 mL or less were directly added to broth. Broths were incubated at 41.5 +/- 1˚C for 24 h, and then a loopful plated onto modified charcoal-cefoperazone-deoxycholate agar (mCCDA) plates. Putative positive colonies were restreaked onto Columbia Blood Agar (CBA) plates. Purified colonies (1-2) were added into 500 µL of 2% Chelex 100 Sodium form (Sigma Cat#C7901) in sterile Milli-Q (Millipore, Merck) water. The Chelex solution was heated at 95°C for 10 min, centrifuged 16,000 g for 5 min and the supernatant tested by conventional PCR using the *Campylobacter* multiplex assay of Wong et al (2004). Visualisation of the PCR amplicons was carried out on the MultiNA Microchip Electrophoresis System (Shimadzu) using fluorescence detection of the DNA products (Figure 25). This provides a detection range of <0.11 MPN/100 mL up to >11,100 MPN/100 mL. As the isolates were randomly selected for confirmation, the culture sample data is grouped as *Campylobacter* species rather than *C. jejuni* and *C. coli.*

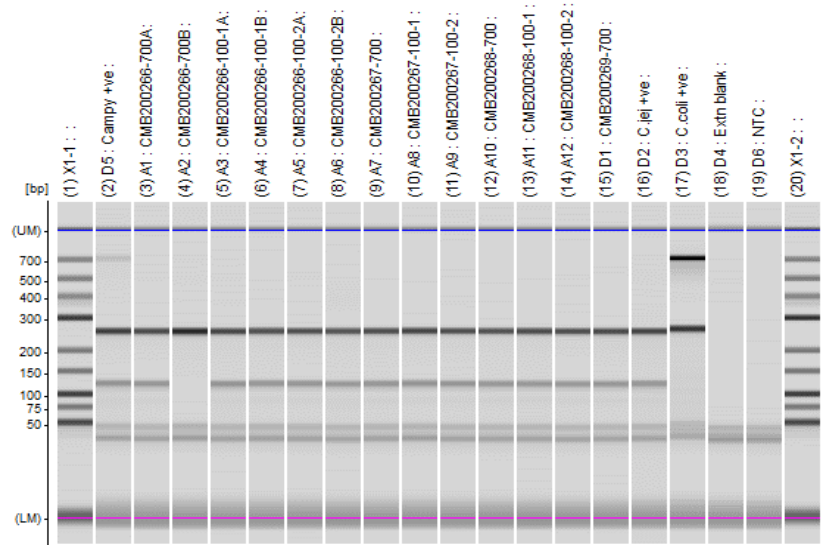

Figure 25: An example of a *Campylobacter* PCR detection using the MultiNA electrophoretic system.
Note: From the top, the gene targets for *C. coli* (695 bp), Thermotolerant *Campylobacter* (246 bp) and *C. jejuni* (99 bp). There is a slight bp offset on the image

Five tube MPN table and confidence limits are presented in Table 14.

Table 14: Five tube MPN and confidence intervals

Tubes in dilution series							
1	2	3	3	3			
700 mL	100 mL	1 mL	0.1 mL	0.01 mL	MPN/ 100 ML (confidence intervals)		
1	0	0	0	0	0.21	(0.01,	1.60)
0	1	0	0	1	0.23	(0.04,	0.75)
0	1	1	0	0	0.24	(0.04,	0.75)
0	2	0	0	0	0.25	(0.04,	0.81)
0	2	1	0	0	0.38	(0.09,	1.01)
1	1	0	0	0	0.70	NC	
1	1	1	0	0	1.41	(0.22,	5.03)
1	2	0	0	0	4.11	NC	
1	2	1	0	0	35.71	NC	
1	2	1	0	1	72.33	(10.66,	238.83)
1	2	1	1	0	73.57	(10.83,	243.27)
1	2	2	0	0	91.78	(12.57,	328.03)
1	2	2	1	0	146.89	(31.68,	439.51)
1	2	3	0	0	231.16	(38.20,	1081.17)
1	2	3	1	0	427.29	(84.50,	1833.59)
1	2	3	2	0	932.80	(177.88,	3340.31)
1	2	3	3	0	2397.90	(408.65,	11500.77)
1	2	3	3	1	4621.83	(1145.84,	16992.44)
1	2	3	3	2	10989.50	NC	

Where there are NC for the confidence interval, the confidence interval was unable to be estimated to due non-convergence of the McBride estimations algorithm (McBride, 2003).

C.1.3 *Salmonella*

A 9 tube MPN analysis was undertaken to isolate *Salmonella* spp.. River water volumes of 1 x 700 mL, 2 x 100 mL, 3 x 1 mL, 3 x 0.1 mL were analysed (ISO, 2017b; PHE, 2015; APHA, 2017d).

Larger volumes were filtered through 0.45 µm filters, and placed into 25 mL of buffered peptone water (BPW) broth. Analysis of volumes 1 mL or less were directly added to the broths. Broths were incubated at 37˚C for 18 h and then 1 mL of BPW was transferred into 10 mL of Muller-Kauffmann Tetrathionate Novobiocin Supplement (MKTTn broth, and 0.1 mL of BPW transferred into 10 mL of Rappaport-Vassiliadis Soya peptone broth (RVS Broth), which were then incubated at 37˚C and 41.5˚C respectively for 24 h. A loopful of each broth was plated onto Xylose Lysine Deoxycholate (XLD) and Hektoen Enteric agar and incubated for 24 h at 37˚C. Putative *Salmonella* colonies were plated on MacConkey agar, tryptic soy agar (TSA), inoculated into tryptone broth, urease broth and onto triple sugar iron agar (TSI) or Lysine Iron Agar (LIA) slopes. Isolates which were consistent with *Salmonella* then had polyO and polyH serology undertaken. If still indicative of *Salmonella*, Microgen

biochemical testing was undertaken. *Salmonella* isolates were then whole genome sequenced. This provides a detection range of <0.11 MPN/100 mL up to >1,110 MPN/100 mL.

C.1.4 STEC

A 9 tube MPN analysis was undertaken to isolate STEC. River water volumes of 1 x 700 mL, 2 x 100 mL, 3 x 1 mL, 3 x 0.1 mL were analysed (ISO/TS, 2012).

Larger volumes were filtered through 0.45 µm filters, and placed into 25 mL of modified Trypticase Soy Broth (TSB) broth with novobiocin. Analysis of volumes 1 mL or less were directly added to the broths. Broths were incubated at 37˚C for 24 h and then the broths tested by STEC multiplex PCR which detected *eae*, *hly*A, *stx*1 and *stx*2 genes (Paton and Paton, 1998). Broths that tested positive were plated on MacConkey Agar with Sorbitol, Cefixime, and Tellurite (CT-SMAC) and MacConkey STEC chrome agar, and incubated for 24 h at 37˚C. Twenty colonies from each plate were streak isolated and purified on CBA plates, and then multiplex PCR performed on pools of 5 isolates (Figure 26). If a pool was positive, then individual isolates were tested by multiplex PCR. This provides a detection range of <0.11 MPN/100 mL up to >1,110 MPN/100 mL.

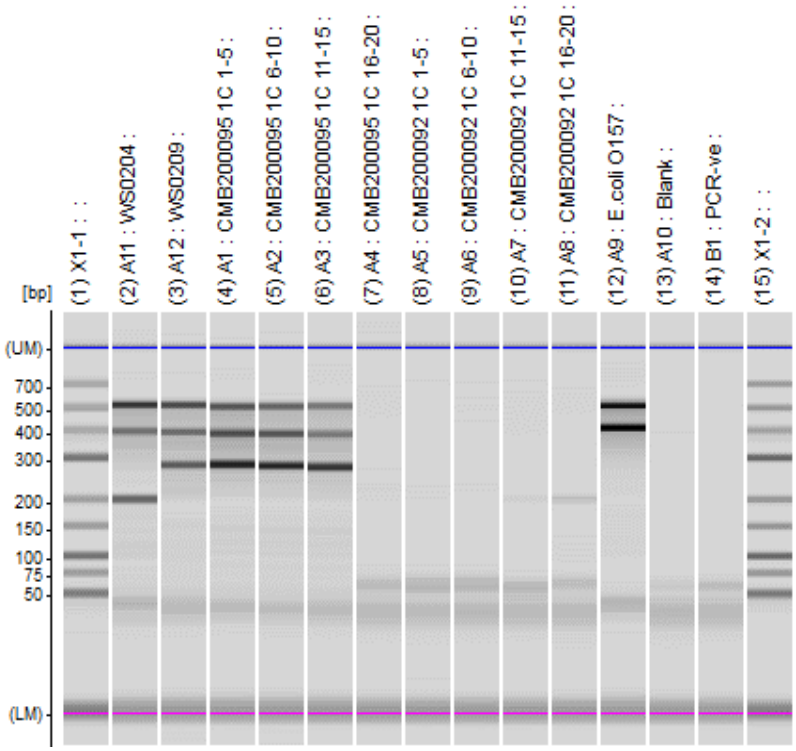

Figure 26: An example of a STEC PCR MultiNA electrophoretic system

Note: Lanes 1 & 15 are size standards. Lanes 2, 3, 12 are positive controls, lanes 13 and 14 are negative controls, remaining lanes pools of isolates being tested. Bands detected from top *hly*A (534 bp), *eae* (384 bp), *stx*2 (255 bp) and *stx*1 (180 bp). Note there is a slight off set of bp on the image

C.2 Protozoa analysis

Protozoa were analysed using USEPA Method 1623: *Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA (EPA, 2013). Briefly, protozoa in the river water are captured by a filter during filtration. The filter setup consists of a filter housing, containing the IDEXX Filta-Max® filter, attached to a submersible pump on the inlet side and a flow meter on the outlet side which records the actual volume of water filtered. Up to 100 L of river water is filtered in the field, or the filter is run for up to

one hour, whichever is achieved first. The filtration time and the start and final readings from the flow meter are recorded. The filter housing was sent to ESR overnight courier on ice (<10°C) where the filter was recovered from the filter housing and sent to Massey University laboratory by overnight courier on ice (<10°C) for analysis. In the laboratory, *Cryptosporidium* oocysts and *Giardia* cysts retained on the filter are eluted and the eluate centrifuged to pellet the (oo)cysts which are isolated using anti-*Cryptosporidium* and anti-*Giardia* immunomagnetic beads and separated from the other material. The (oo)cysts and cysts are stained on well slides with a fluorescent label and DAPI[5]. The stained sampled is examined using UV fluorescence and differential interference contrast microscopy. The number of objects on the slide that meet the size, shape and fluorescence characteristics of *Cryptosporidium* and *Giardia* (oo)cysts are reported. Massey University have determined that the recovery rate for river water varies from 15-55%. The data is reported as the actual count.

C.3 Quantitative PCR analysis of viruses

River water samples (10 L) were collected and sent by courier for overnight delivery to the ESR Environmental Virology laboratory in Porirua. Each water sample was seeded with 1 mL murine norovirus (10^6 plaque forming units (PFU)/mL) as a process control to calculate virus recovery rate and to evaluate potential RT-PCR inhibition.

Each 10 L water samples were first concentrated to 700-900 mL by hollow-fibre ultrafiltration (HFUF) method (Hill et al, 2005) with modifications. In brief, 1% w/v sodium polyphosphate was added to each water sample and filtered through the HFUF membrane at the flow rate of 150-300/min until a final volume of approximately 500 mL was achieved. To further concentrate viruses, polyethylene glycol 6000 precipitation (with a preceding centrifugation step and beef extract elution step where necessary) was performed as described elsewhere (Hewitt et al, 2007). Finally, 2-20 mL concentrates were obtained from each 10 L water sample and stored at -80℃ until viral extraction. The volume of concentrates varied with turbidity of water sample. In the beginning of the study, this was 2 mL for small pellet and 5 mL for larger pellets. With the exception of two samples, from 23 Feb 2020, all samples were suspended in 10 mL. Viral nucleic acid was extracted from 200 µL concentrates using Presto DNA/RNA Extraction Kit (Geneaid Biotech Ltd, Taiwan) with a minor modification (i.e. 20 µL poly A solution was added after PCR inhibition removal step).

All nucleic acid samples were tested for potential RT-PCR inhibition. Those RT-PCR inhibitory samples were further processed using Onestep™ PCR Inhibition Removal Kit (Zymo, USA) and re-tested for inhibition. Virus recovery from each water sample was determined by comparing the concentration of murine norovirus to control nucleic acid sample (nucleic acid extracted murine norovirus spiked into similar volume of distilled water). Controls were included in each RNA extraction and RT-qPCR assay including 1000, 100, and 10 RTPCR units of norovirus GI and GII.

Previously published qPCR/ two-step RT-qPCR assays were used to detect adenovirus (Hernroth et al, 2002), enterovirus (Donaldson et al, 2002), norovirus genogroup I (GI) (Wolf et al, 2010), norovirus GII (Kageyama et al, 2003) and murine norovirus (Hewitt et al, 2011).

C.4 Quantitative PCR analysis of bacteria

Quantitative PCR methodology is described in more detail in Gilpin et al (2018). Two litres of water were collected from each river and up to 1 L filtered in duplicate through a 0.45 µm Millipore filter (range of 400 mL to 2 L total). CD1 buffer from PowerSoil Pro kit (Qiagen), spiked with Salmon Sperm DNA (Sigma) was added to the filter and vortexed. Filters were stored at -20˚C. Sterile beads were

[5] DAPI 4',6-diamidino-2-phenylindole

added to the filter, and tubes placed in a Biospec Minibead beater, where they were beaten for 3 min at 2,500 rpm, centrifuged at 3,500 g for 5 min and supernatant transferred to a new tube. The volume was adjusted to 650 µL using CD1 buffer, if required. Samples were extracted using the PowerSoil Pro protocol on the QiaCube extraction robot. Filters were eluted in 80 µL of elution buffer and then duplicate filters were combined to a final volume of 160 µL for each water sample.

Quantitative PCR analysis was undertaken on a LightCycler 480 (Roche), with each amplification performed in duplicate using 2 µL of DNA extract. The PCR targets are given in Table 15. Each run included negative and positive controls, and gBlock or Ultramer standard curves. All samples were initially tested using primers specific for Sketa salmon sperm as a process and inhibition control. No significant inhibition was detected, so no further adjustment of samples was undertaken. Assays were confirmed to have an efficiency of >90%. To enable relative comparisons between assays, a common slope (1.926) and y intercept (Cycle threshold (CT) 40) was used in calculations to convert CT to copy number equivalents/100 mL of filtered water.

Copies/100 mL = (1.9263626137313^(40-CT))*((160/2)*(100/Volume filtered))

Table 15: Target bacterial genes and methods for qPCR

Micro-organism	Target	Reference
qPCR Bacteria		
E. coli	23S rDNA	Chern et al, 2011
Enterococci	23S rDNA	Haugland et al, 2012; EPA, 2013
C. jejuni	mapA	Best et al, 2003
C. coli	ceuE	Best et al, 2003
Salmonella spp.	invA and ttr	Hoorfar et al, 2000; Malorny et al, 2004
STEC	eae, stx-1, stx-2	Ibekwe et al, 2004; Derzelle et al, 2011
Cryptosporidium parvum, C. hominis	18S rDNA	Mary et al, 2013
Giardia	beta-giardin gene	Baque et al, 2011
Faecal Source Markers		
General	Bacteriodales 16S rRNA	Shanks et al, 2009; Shanks et al, 2010
Human	*Bacteroides* HF183 crAssphage CPQ_056 (crAss) *Bifidobacterium adolescentis* (BiADO)	Ahmed et al, 2019
Ruminant	Bacteroidales 16S rRNA (BacR)	Reischer et al, 2006
Wildfowl	GFD - Unclassified *Helicobacter* spp. 16S rRNA gene E2- *Desulfovibrio*-like organism	Green et al, 2012 Devane et al, 2007

C.5 Categorisation of dominant FST markers

The FST data can be used to categorise the rivers at the sampling points according to the source of faecal material present in the water, at that time, rather than observed surrounding land use. This means that the category may change between sampling events.

To categorise the dominant faecal sources from the concentrations of FST, the following rules were applied.

- Two wildfowl FST markers were used: GFD, and the E2 marker, which is more specific for "duck". For this study GFD and E2 results were combined, with significant concentrations of wildfowl markers considered to be >1,000 copies, but results >100 copies/100 mL reported.

- The BacR marker indicates the presence of ruminant sources of faecal pollution and samples with >1,000 copies/100 mL were considered significant sources. In addition, samples with >100 copies/100 mL BacR and where the ratio of BacR copies to the total GenBac copies was >0.3% were reported.

- Samples with more than a 1,000 copies/100 mL of any of the three human FST markers were used: HF183, CrAssphage and BiADO each indicate a significant human source. The low concentrations of one or more of the human indicative markers in the other samples may be the result of low level human faecal sources, or may the consequence of cross reaction with non-human sources. Since these can't be distinguished, in this study, samples with all three markers and a combined total of 1,000 copies or more were characterised as containing a significant human source of faecal pollution.

C.6 Whole Genome Sequencing (WGS)

Whole genome sequencing allows a more accurate and efficient determination of serotype, virulence potential, and potentially, source attribution of bacterial isolates.

All bacterial pathogenic isolates were recovered from -80°C storage by streak plating onto Columbia blood agar (CBA) incubating the plates at 37°C, 24-48 h. A single colony was inoculated into 10 mL tryptone soya broth and incubated at 37°C, for 18 h prior to DNA extraction. One milliliter of broth culture was used for DNA extraction using the Qiagen DNeasy Blood and Tissue Kit QiaCube (Qiagen, Hilden, Germany). DNA quality and concentration was assessed using Nanodrop (Thermo Fisher Scientific, Waltham, MA. USA), Qubit™ and PicoGreen® (Quant-iT; Thermo Fisher Scientific). Sequencing libraries containing 1 ng of DNA were prepared using Nextera XT chemistry (Illumina, San Diego, CA, USA) for 150 bp pair-end sequencing run on an Illumina NextSeq sequencer according to the manufacturer's recommendations (Illumina).

Sequence quality and species identification was determined using the Nullarbor pipeline (Seemann et al, accessed 2020). Sequence quality was evaluated on a per genome basis using BioNumerics version 7.6.3 (Applied Maths, NV, Belgium). All genomes passed the basic quality metrics for raw sequence data from Illumina sequencers of average Q-score >30 in both reads and at least 40 X average coverage with expected genome sizes (Lindsey et al, 2016). Genetic variability within each taxa was evaluated using Multi-Locus Sequence Typing (MLST), and whole genome MLST (wgMLST). The wgMLST schema was assessed within BioNumerics. Phylogenetic cluster analysis of the isolates was investigated using wgMLST (categorical data values) and analysed using single-linkage algorithm.

Multiplex ligation-dependent probe amplification-binary typing (MBiT) (Cornelius et al, 2014) types were inferred for each genome using BioNumerics 7.6. The MBiT types were compared to those in MBiT library and the sources of previous isolates that clustered with the study isolates were evaluated. Ribosomal MLST (rMLST), average nucleotide identity (ANI) and genome BLAST distance phylogeny (GBDP) were used to identify new isolates.

APPENDIX D: Raw data

Table 16: Faecal indicator bacteria and pathogen raw data for each site

Site	Land use	Dominant Faecal Source	Total Coliforms MPN / 100 mL	*E.coli* MPN / 100 mL	Enterococci MPN / 100 mL	*Campylobacter* MPN / 100 mL	*Salmonella* MPN / 100 mL	*stx* MPN / 100 ml	*Cryptosporidium* oocysts/100 L	*Giardia* cysts/ 100 L	HAdV	NoV GI	NoV GII	Enterovirus	Rainfall (mm) - 24h	Cumulative rainfall (mm) - 48h	Cumulative rainfall (mm) - 72h
1	Urban	Human	>24,000	>24,00	210	4.1	<0.11	<0.11	3.4	9.2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.5	1.0	1.0
1	Urban	Ruminant & Human	24,000	3,400	660	0.25	<0.11	<0.11	ND[6]	250	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	38.5	38.5	38.5
1	Urban	Ruminant & Human	7,700	1,300	490	4.1	<0.11	<0.11	2.7	12.2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	28.0	66.4
1	Urban	Human	20,000	6,100	350	4.1	0.21	<0.11	2	14	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	5.5	5.5	5.5
2	Urban	Human	11,000	5,200	270	<0.11	<0.11	<0.11	ND	ND	ND	38.3/ND/39.1	ND/ND/ND	ND/ND/ND	0.0	1.5	11.0
2	Urban	Human	3,400	460	97	0.25	<0.11	<0.11	ND	13.3	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	4.5
2	Urban	Ruminant & Human	>24,000	12,000	520	4.1	0.21	<0.11	ND	16.1	ND	39.7/38.4/39.9	ND/ND/39.3	ND/ND/ND	0.0	38.0	38.0
3	Urban	Human	4,600	550	31	<0.11	<0.11	<0.11	4.8	17.9	ND	ND/ND/ND	ND/ND/39.7	ND/ND/ND	0.0	0.0	0.0
3	Urban	Human	24,000	4,900	450	<0.11	<0.11	<0.11	ND	15.9	ND	ND/ND/38.8	38.2/ND/38.5	ND/39.9/ND	0.0	18.2	22.4
3	Urban	Human	>24,000	7,700	990	0.23	0.21	<0.11	7.1	85.7	ND	ND/ND/ND	38.2/38.3/37.4	ND/ND/ND	3.2	3.2	3.2
4	Urban	Human	3,700	97	20	<0.11	<0.11	<0.11	1.1	2.4	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
4	Urban	Human	8,200	510	130	0.21	<0.11	<0.11	ND	5.6	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	5.5	5.5

[6] ND not detected

Site	Land use	Dominant Faecal Source	Total Coliforms MPN / 100 mL	*E.coli* MPN / 100 mL	Enterococci MPN / 100 mL	*Campylobacter* MPN / 100 mL	*Salmonella* MPN / 100 mL	*stx* MPN / 100 ml	*Cryptosporidium* oocysts/100 L	*Giardia* cysts/ 100 L	HAdV	NoV GI	NoV GII	Enterovirus	Rainfall (mm) - 24h	Cumulative rainfall (mm) - 48h	Cumulative rainfall (mm) - 72h
4	Urban	Human	13,000	1,000	52	0.24	0.21	<0.11	3	4.5	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
5	Urban	Human	2,900	98	<10	<0.11	0.21	<0.11	ND	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
5	Urban	Human	9,800	160	<10	<0.11	<0.11	<0.11	6.9	2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
5	Urban	Human	8,200	460	31	<0.11	<0.11	<0.11	ND	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
5	Urban	Human	17,000	190	10	<0.11	<0.11	<0.11	ND	2.1	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
6	Urban	Human	5,300	430	<10	4.1	<0.11	<0.11	ND	3.3	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
6	Urban	Human	9,800	200	20	<0.11	<0.11	<0.11	ND	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.2	0.2
6	Urban	Human	5,500	410	10	0.25	<0.11	<0.11	ND	2.3	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	1.0
6	Urban	Human	13,000	530	96	4.1	<0.11	<0.11	3.1	9.4	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
7	S&B[7]	Ruminant	20,000	1,100	410	92	0.25	<0.11	10.7	10.7	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	33.5
7	S&B	Ruminant	*(8,200)*	*(1,800)*	*(150)*	-	-	-	ND	2.6	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	14.5	14.5	14.5
7	S&B	Ruminant & Human	4,600	570	180	4.1	0.21	<0.11	8.5	12.2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
8	S&B	Ruminant	6,900	1,000	190	72	<0.11	<0.11	3.7	8.5	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	9.0
8	S&B	Ruminant	6,500	880	320	4.1	<0.11	<0.11	ND	23.9	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	9.5	9.5
8	S&B	Ruminant	10,000	1,100	210	74	<0.11	<0.11	ND	7.1	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
9	S&B	Ruminant	1,100	63	10	<0.11	<0.11	<0.11	2	2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
9	S&B		3,400	52	52	0.21	<0.11	<0.11	ND	2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.5	4.1

[7] S&B observed sheep and beef land use

Site	Land use	Dominant Faecal Source	Total Coliforms MPN / 100 mL	*E.coli* MPN / 100 mL	Enterococci MPN / 100 mL	*Campylobacter* MPN / 100 mL	*Salmonella* MPN / 100 mL	*stx* MPN / 100 ml	*Cryptosporidium* oocysts/100 L	*Giardia* cysts/ 100 L	HAdV	NoV GI	NoV GII	Enterovirus	Rainfall (mm) - 24h	Cumulative rainfall (mm) - 48h	Cumulative rainfall (mm) - 72h
9	S&B	Ruminant	4,100	250	120	92	<0.11	<0.11	ND	2	ND	ND/ND/ND	ND/ND/ND	ND/39.2/ND	15.2	15.2	15.2
10	S&B	Ruminant	-	-	-	-	-	-	8.8	22.8	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	8.5	10.0	10.0
10	S&B	Ruminant	4,600	85	10	<0.11	<0.11	<0.11	ND	19.5	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
10	S&B	Ruminant	2,400	130	<10	<0.11	<0.11	<0.11	ND	19.4	ND	ND/ND/ND	37.8/37.1/ND	ND/ND/ND	0.0	0.0	3.5
11	Dairy	Ruminant	1,700	340	<10	0.21	<0.11	<0.11	ND	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	11.5
11	Dairy	Ruminant	11,000	300	50	4.1	<0.11	<0.11	ND	3.8	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	2.5
11	Dairy	Ruminant & Human	17,000	2,500	110	4.1	<0.11	<0.11	ND	3.8	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	43.5	43.5
12	Dairy	Wildfowl	790	10	<10	4.1	<0.11	<0.11	ND	2.2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
12	Dairy	Ruminant	17,000	860	51	<0.11	<0.11	<0.11	30.8	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	2.4	24.6
12	Dairy	Ruminant	1,400	120	20	<0.11	<0.11	<0.11	ND	ND	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.2	0.2
13	Dairy	Wildfowl	4,900	680	74	92	<0.11	<0.11	ND	2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
13	Dairy	Wildfowl	4,400	630	41	0.21	<0.11	0.14	6.2	5.2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
13	Dairy	Ruminant	4,400	540	41	4.1	<0.11	<0.11	2	4	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
13	Dairy	Wildfowl	2,200	310	63	0.21	<0.11	<0.11	1	2	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
14	Dairy	Ruminant	3,700	1,800	10	0.25	<0.11	<0.11	ND	ND	ND	37.9/ND/ND	40.1/ND/ND	ND/ND/ND	0.0	0.0	0.0
14	Dairy	Wildfowl	5,200	1,000	120	4.1	0.21	<0.11	ND	ND	ND	ND/ND/ND	ND/39.7/ND	ND/ND/ND	0.5	20.5	20.5
14	Dairy	Wildfowl	6,100	1,700	290	0.38	0.21	<0.11	6.9	ND	ND	41.3/40.5/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
15	S&B	Wildfowl	17,000	150	10	<0.11	<0.11	<0.11	3.6	3.6	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
15	S&B	Wildfowl	14,000	200	41	<0.11	<0.11	<0.11	7.1	14.3	ND	ND/40.5/38.9	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0

Site	Land use	Dominant Faecal Source	Total Coliforms MPN / 100 mL	*E.coli* MPN / 100 mL	Enterococci MPN / 100 mL	*Campylobacter* MPN / 100 mL	*Salmonella* MPN / 100 mL	*stx* MPN / 100 ml	*Cryptosporidium* oocysts/100 L	*Giardia* cysts/ 100 L	HAdV	NoV GI	NoV GII	Enterovirus	Rainfall (mm) - 24h	Cumulative rainfall (mm) - 48h	Cumulative rainfall (mm) - 72h
15	S&B	Wildfowl	16,000	410	98	0.7	<0.11	<0.11	3.5	4.7	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.2	0.2	0.6
16	Dairy	Wildfowl	3,100	180	<10	0.7	<0.11	<0.11	ND	2.4	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
16	Dairy	Wildfowl	9,800	290	63	0.7	<0.11	<0.11	ND	2.7	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0
16	Dairy	Wildfowl	2,900	75	20	4.1	<0.11	<0.11	ND	4.4	ND	ND/ND/ND	ND/ND/ND	ND/ND/ND	0.0	0.0	0.0

Table 17: Comparison of FST and observed land use

Site	Land use	Dominant Faecal Source	Volume filtered ml	General GenBac / 100 ml	Human HF183 / 100 ml	Human CrAssphage / 100 ml	Human BiADO / 100 ml	Human Total / 100 ml	Ruminant BacR / 100 ml	Avian GFD / 100 ml	Avian E2 / 100 ml	Bird Total / 100 ml	Obser-vations of animals
1	Urban	Human	1200	770,000	65,000	3,100	8,600	77,000	190	440	190	620	Birds
1	Urban	Ruminant & Human	500	390,000	1,400	510	1,400	3,300	15,000	49	16	65	Birds
1	Urban	Ruminant & Human	750	380,000	19,000	2,600	3,800	25,000	3,300	1,300	47	1,300	Bird
1	Urban	Human	800	1,600,000	130,000	9,400	25,000	160,000	1,600	1,900	110	2,000	Birds
2	Urban	Human	800	1,800,000	66,000	65,000	100,000	230,000	2,500	1,600	200	1,800	
2	Urban	Human	400	1,300,000	8,400	11,000	14,000	33,000	300	2,400	100	2,500	Birds
2	Urban	Ruminant & Human	400	5,100,000	26,000	31,000	45,000	100,000	330,000	6,800	450	7,300	Birds
3	Urban	Human	2000	850,000	11,000	8,400	22,000	41,000	110	640	1,600	2,300	
3	Urban	Human	2000	13,000,000	490,000	160,000	1,300,000	2,000,000	600	1,300	890	2,200	
3	Urban	Human	900	6,000,000	150,000	110,000	510,000	770,000	930	3,900	1,500	5,400	
4	Urban	Human	2000	310,000	590	240	1,100	1,900	-	750	570	1,300	Birds
4	Urban	Human	2000	610,000	14,000	3,100	4,300	21,000	-	1,500	300	1,800	Birds
4	Urban	Human	1600	770,000	9,000	8,400	14,000	31,000	-	1,300	510	1,800	Bird
5	Urban	Human	1200	89,000	310	2,300	76	2,700	-	17	-	17	
5	Urban	Human	2000	46,000	140	1,100	6	1,200	2	360	-	360	Birds
5	Urban	Human	1000	48,000	110	1,400	100	1,600	-	140	-	140	
5	Urban	Human	1000	58,000	100	960	20	1,100	-	220	12	230	
6	Urban	Human	1200	550,000	3,100	900	470	4,500	-	980	1,400	2,400	Birds
6	Urban	Human	1200	270,000	710	1,200	850	2,800	1	770	340	1,100	Birds

Site	Land use	Dominant Faecal Source	Volume filtered ml	General GenBac / 100 ml	Human HF183 / 100 ml	Human CrAssphage / 100 ml	Human BiADO / 100 ml	Human Total / 100 ml	Ruminant BacR / 100 ml	Avian GFD / 100 ml	Avian E2 / 100 ml	Bird Total / 100 ml	Obser-vations of animals
6	Urban	Human	1000	470,000	3,500	6,300	640	10,000	-	1,400	630	2,000	Birds
6	Urban	Human	1000	390,000	1,200	1,600	3,800	6,600	36	2,100	1,300	3,400	Birds
7	S & B	Ruminant	1200	930,000	-	50	27	77	170,000	1,400	71	1,400	
7	S & B	Ruminant	1000	400,000	170	97	160	430	49,000	1,600	19	1,600	
7	S & B	Ruminant & Human	1000	800,000	220	75	1,200	1,500	54,000	860	160	1,000	
8	*S & B*	*Ruminant*	*1600*	*130,000*	*620*	*44*	*44*	*710*	*1,900*	*720*	*340*	*1,100*	
8	S & B	Ruminant	1600	300,000	280	-	-	280	4,600	1,000	210	1,300	Dog
8	S & B	Ruminant	1200	440,000	390	-	-	390	5,300	3,600	130	3,700	
9	S & B	Ruminant	2000	76,000	20	15	-	35	2,000	160	77	240	Birds
9	S & B		1600	-	-	-	-	-	-	-	-	-	Not fenced
9	S & B	Ruminant	1200	380,000	66	30	-	96	5,900	1,200	180	1,400	
10	S & B	Ruminant	1300	380,000	500	130	2	630	20,000	540	66	610	
10	S & B	Ruminant	1200	290,000	160	-	2	160	1,700	300	89	390	
10	S & B	Ruminant	1200	150,000	170	24	46	240	2,100	200	71	270	
11	Dairy	Ruminant	350	2,000,000	280	470	74	820	440,000	770	380	1,100	
11	Dairy	Ruminant	300	2,100,000	320	260	1	580	320,000	1,000	46	1,100	
11	Dairy	Ruminant & Human	300	1,900,000	2,800	1,400	7,400	12,000	160,000	1,200	110	1,400	
12	Dairy	Wildfowl	1600	150,000	210	-	-	210	490	300	41	340	Birds
12	Dairy	Ruminant	700	340,000	740	-	-	740	37,000	260	38	300	
12	Dairy	Ruminant	1000	150,000	210	-	2	210	1,700	670	23	690	

Site	Land use	Dominant Faecal Source	Volume filtered ml	General GenBac / 100 ml	Human HF183 / 100 ml	Human CrAssphage / 100 ml	Human BiADO / 100 ml	Human Total / 100 ml	Ruminant BacR / 100 ml	Avian GFD / 100 ml	Avian E2 / 100 ml	Bird Total / 100 ml	Obser-vations of animals
13	Dairy	Wildfowl	2000	80,000	55	-	4	59	180	380	65	450	
13	Dairy	Wildfowl	2000	82,000	29	17	32	78	110	160	7	160	
13	Dairy	Ruminant	2000	230,000	82	-	52	130	42,000	630	6	640	
13	Dairy	Wildfowl	2000	70,000	-	-	2	2	240	420	5	420	Dog
14	Dairy	Ruminant	1200	160,000	42	31	2	75	1,500	680	830	1,500	
14	Dairy	Wildfowl	1200	480,000	74	120	2	200	680	1,600	1,000	2,600	
14	Dairy	Wildfowl	800	1,500,000	120	230	-	350	10	5,000	2,700	7,600	
15	S & B	Wildfowl	1600	190,000	-	-	-	-	-	710	420	1,100	Birds
15	S & B	Wildfowl	1600	370,000	33	22	-	55	-	1,300	230	1,500	
15	S & B	Wildfowl	1200	390,000	-	-	-	-	-	1,100	920	2,100	Birds
16	Dairy	Wildfowl	900	220,000	-	-	9	9	-	1,800	190	2,000	Birds
16	Dairy	Wildfowl	900	150,000	17	-	-	17	34	930	51	980	Bird faeces
16	Dairy	Wildfowl	1000	57,000	-	-	-	-	6	260	8	270	Birds on bank

Table 18: Comparison of PCR and culture methodologies for *E. coli*, enterococci and *Campylobacter* with WGS data

Site	*E.coli* MPN / 100 mL	*E.coli* copies / 100 mL	Enterococci MPN / 100 mL	Enterococci copies / 100 mL	*Campylo-bacter* MPN / 100 mL	*C. jejuni* copies / 100 mL	*C. coli* copies /100 mL	*Campylo-bacter* WGS Species (MLST)
1	24,001	230,000	210	9,800	4.1	ND	ND	
1	3,400	28,000	660	38,000	0.25	75	ND	*C. jejuni* (3640), *Campylobacter* 1
1	1,300	19,000	490	17,000	4.10	100	ND	*C. jejuni* (2381), *Campylobacter* 2
1	6,100	84,000	350	45,000	4.10	390	ND	*C. jejuni* (45), *Campylobacter* 1
2	5,200	34,000	270	17,000	<0.11	ND	ND	
2	460	13,000	97	9,200	0.25	ND	ND	*C. jejuni* (995)
2	12,000	200,000	520	480,000	4.10	74	ND	*C. jejuni* (45), *C. lari*
3	550	8,900	31	14,000	<0.11	39	ND	
3	4,900	73,000	450	63,000	<0.11	54	12	
3	7,700	110,000	990	220,000	0.23	260	ND	*C. jejuni* (3640)
4	97	3,200	20	3,100	<0.11	72	ND	
4	510	8,900	130	57,000	0.21	31	15	*C. jejuni* (991)
4	1,000	20,000	52	13,000	0.24	46	5	*C. jejuni* (2381)
5	98	2,900	<10	5,300	<0.11	ND	ND	
5	160	1,600	<10	930	<0.11	ND	ND	
5	460	5,400	31	2,300	<0.11	ND	ND	
5	190	2,900	10	1,100	<0.11	ND	ND	
6	430	8,100	<10	7,700	4.1	82	ND	*C. jejuni* (9820)
6	200	6,000	20	3,300	<0.11	ND	ND	
6	410	14,000	10	9,900	0.25	48	28	*C. jejuni* (177)
6	530	12,000	96	12,000	4.10	15	ND	*C. jejuni* (2381)
7	1,100	12,000	410	16,000	92	340	41	*C. jejuni* (3640), *Campylobacter* 1
7	*(1,800)*	16,000	*(150)*	14,000	-	73	ND	

Site	*E.coli* MPN / 100 mL	*E.coli* copies / 100 mL	Enterococci MPN / 100 mL	Enterococci copies / 100 mL	*Campylo-bacter* MPN / 100 mL	*C. jejuni* copies / 100 mL	*C. coli* copies /100 mL	*Campylo-bacter* WGS Species (MLST)
7	570	8,200	180	15,000	4.10	66	ND	*C. jejuni* (3640)
8	1,000	5,100	190	5,700	72	1220	ND	*C. jejuni* unusual
8	880	10,000	320	14,000	4.1	180	26	*C. jejuni* (45)
8	1,100	18,000	210	15,000	74.00	420	ND	*C. jejuni* (2381)
9	63	1,900	10	4,400	<0.11	42	14	
9	52		52	-	0.21	ND	ND	
9	250	6,100	120	12,000	92.00	49	ND	*C. jejuni* (new)
10	-	30,000	-	59,000	-	ND	ND	
10	85	3,200	10	4,400	<0.11	ND	ND	
10	130	2,900	<10	18,000	<0.11	ND	ND	
11	340	8,800	<10	1,000,000	0.21	ND	28	*C. jejuni* (2389)
11	300	13,000	50	530,000	4.1	ND	ND	*C. jejuni* (45)
11	2,500	42,000	110	140,000	4.10	ND	ND	*C. jejuni* (677), *C. lari*
12	10	1,300	<10	4,100	4.1	30	ND	*C. jejuni* (991)
12	860	19,000	51	73,000	<0.11	ND	ND	
12	120	2,600	20	5,700	<0.11	ND	ND	
13	680	4,700	74	2,300	92	260	ND	*C. jejuni* (9820)
13	630	6,100	41	2,200	0.21	340	42	*C. jejuni* unusual
13	540	7,500	41	4,200	4.10	490	100	*C. jejuni* unusual
13	310	5,700	63	2,500	0.21	390	130	*C. jejuni* unusual
14	1,800	9,400	10	3,000	0.25	31	ND	*C. jejuni* (1956)
14	1,000	18,000	120	12,000	4.1	140	ND	*C. jejuni* (699)
14	1,700	25,000	290	22,000	0.38	230	ND	*C. jejuni* (45)
15	150	2,100	10	5,800	<0.11	ND	ND	
15	200	4,700	41	15,000	<0.11	5	20	
15	410	3,100	98	4,700	0.70	10	ND	*C. jejuni* unusual
16	180	4,100	<10	3,000	0.7	25	ND	*Campylobacter* 1
16	290	5,300	63	3,700	0.7	33	ND	*Campylobacter* 1

Site	*E.coli* MPN / 100 mL	*E.coli* copies / 100 mL	Enterococci MPN / 100 mL	Enterococci copies / 100 mL	*Campylo-bacter* MPN / 100 mL	*C. jejuni* copies / 100 mL	*C. coli* copies /100 mL	*Campylo-bacter* WGS Species (MLST)
16	75	2,400	20	1,200	4.10	ND	26	*C. jejuni* (new), *Campylobacter* 1

Table 19: Comparison of results by culture and qPCR for *Salmonella* and *E. coli* virulence genes

Site	*Salmonella* MPN / 100 ml	*Salmonella* invA copies/ 100 ml	*Salmonella* ttr copies / 100 ml	*stx* MPN / 100 ml	*stx1* PCR	*stx* 1 copies / 100 ml	*stx2* PCR	*stx* 2 copies / 100 ml	*eae*A PCR	*hly*A PCR
1	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
1	<0.11	ND	ND	<0.11	yes	ND	ND	ND	yes	ND
1	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
1	0.21	ND	ND	<0.11	ND	ND	ND	29	yes	ND
2	<0.11	ND	ND	<0.11	ND	18	ND	42	yes	ND
2	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
2	0.21	ND	ND	<0.11	(yes)	172	(yes)	277	yes	(yes)
3	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
3	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
3	0.21	75	56	<0.11	ND	ND	ND	30	yes	ND
4	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
4	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
4	0.21	ND	ND	<0.11	ND	ND	ND	17	yes	ND
5	0.21	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
5	<0.11	ND	ND	<0.11	ND	ND	ND	ND	ND	ND
5	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
5	<0.11	ND	ND	<0.11	ND	ND	ND	ND	ND	ND
6	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
6	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
6	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND

Site	*Salmonella* MPN / 100 ml	*Salmonella* invA copies/ 100 ml	*Salmonella* ttr copies / 100 ml	*stx* MPN / 100 ml	*stx1* PCR	*stx* 1 copies / 100 ml	*stx2* PCR	*stx* 2 copies / 100 ml	*eae*A PCR	*hly*A PCR
6	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
7	0.25	4	ND	<0.11	ND	13	ND	ND	yes	ND
7	-	ND	37	-	-	ND	-	ND	-	-
7	0.21	ND	ND	<0.11	ND	ND	ND	ND	yes	(yes)
8	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
8	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
8	<0.11	ND	ND	<0.11	ND	ND	ND	2	(yes)	ND
9	<0.11	16	ND	<0.11	ND	ND	ND	ND	(yes)	ND
9	<0.11		ND	<0.11	ND		ND		yes	ND
9	<0.11	ND	ND	<0.11	ND	7	ND	ND	yes	ND
10	-	ND	ND	-	-	25	-	28	-	-
10	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
10	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
11	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
11	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	yes
11	<0.11	ND	ND	<0.11	(yes)	ND	(yes)	ND	yes	(yes)
12	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
12	<0.11	ND	ND	<0.11	ND	23	(yes)	56	yes	ND
12	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
13	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
13	<0.11	ND	ND	0.14	ND	7	yes	ND	yes	yes

Site	*Salmonella* MPN / 100 ml	*Salmonella* invA copies/ 100 ml	*Salmonella* ttr copies / 100 ml	*stx* MPN / 100 ml	*stx1* PCR	*stx* 1 copies / 100 ml	*stx2* PCR	*stx* 2 copies / 100 ml	*eae*A PCR	*hly*A PCR
13	<0.11	ND	ND	<0.11	ND	ND	ND	13	yes	ND
13	<0.11	ND	ND	<0.11	(yes)	ND	(yes)	ND	yes	(yes)
14	<0.11	29	ND	<0.11	ND	ND	ND	ND	ND	ND
14	0.21	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
14	0.21	ND	ND	<0.11	ND	ND	ND	33	yes	ND
15	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
15	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
15	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
16	<0.11	ND	ND	<0.11	ND	ND	ND	ND	(yes)	ND
16	<0.11	ND	ND	<0.11	ND	ND	ND	ND	yes	ND
16	<0.11	ND	ND	<0.11	ND	ND	ND	ND	ND	ND

Table 20: Comparison of traditional and qPCR results for protozoa

Site	*Cryptosporidium* oocysts/100 L	*Giardia* cysts / 100 L	*Giardia* copies / 100 mL	Volume filtered (L)
1	3.4	9.2	ND	87
1	ND	250	ND	2
1	2.7	12.2	ND	74
1	2	14	ND	100
2	ND	ND	ND	84
2	ND	13.3	ND	60
2	ND	16.1	ND	56
3	4.8	17.9	ND	84
3	ND	15.9	ND	88
3	7.1	85.7	75	56
4	1.1	2.4	ND	84
4	ND	5.6	ND	36
4	3	4.5	ND	67
5	ND	ND	ND	84
5	6.9	2	ND	102
5	ND	ND	ND	91
5	ND	2.1	ND	95
6	ND	3.3	ND	61
6	ND	ND	ND	60
6	ND	2.3	ND	86
6	3.1	9.4	ND	64
7	10.7	10.7	ND	84
7	ND	2.6	ND	78
7	8.5	12.2	ND	82
8	3.7	8.5	ND	82
8	ND	23.9	ND	88
8	ND	7.1	ND	84
9	2	2	ND	100
9	ND	2	ND	100
9	ND	2	ND	102
10	8.8	22.8	ND	57
10	ND	19.5	ND	87
10	ND	19.4	ND	98
11	ND	ND	ND	57
11	ND	3.8	ND	52
11	ND	3.8	ND	53
12	ND	2.2	ND	91

Site	*Cryptosporidium* oocysts/100 L	*Giardia* cysts / 100 L	*Giardia* copies / 100 mL	Volume filtered (L)
12	30.8	ND	ND	6.5
12	ND	ND	ND	69
13	ND	2	ND	100
13	6.2	5.2	ND	97
13	2	4	ND	100
13	1	2	ND	100
14	ND	ND	ND	57
14	ND	ND	ND	53
14	6.9	ND	ND	29
15	3.6	3.6	ND	83
15	7.1	14.3	ND	84
15	3.5	4.7	ND	86
16	ND	2.4	ND	84
16	ND	2.7	ND	74
16	ND	4.4	ND	68

Table 21: Environmental information and physio-chemical measurements

Site	Water temperature (°C)	Dissolved oxygen (mg/L)	Field pH	Laboratory Turbidity (NTU)	Laboratory Conductivity (µS/cm)	excl 95th %Flow (m3/s) - 24h ave	Flow (m3/s) - 24h ave	Rainfall (mm) - 24h	Cumulative Rainfall (mm) - 48h	Cumulative Rainfall (mm) - 72h	Water clarity (cm)	Water clarity method
1	19.1	7.80	7.37	2.20	203.5	-	-	0.5	1.0	1.0	184	black disc
1	13.1	10.50	7.39	24	166.8	-	-	38.5	38.5	38.5	19	clarity tube
1	14.6	10.10	7.54	3.8	143.7	-	-	0.0	28.0	66.4	144	black disc
1	15.1	10.30	7.32	4.1	196.0	-	-	5.5	5.5	5.5	127	black disc
2	17.1	-	-	4.4	-	0.608	0.608	0.0	1.5	11.0	-	NM
2	14.6	9.77	-	4.8	272.0	0.104	0.104	0.0	0.0	4.5	77	black disc
2	16.0	7.69	8.39	6.4	268.5	0.276	0.276	0.0	38.0	38.0	84	black disc
3	18.2	10.80	-	0.73	258.5	0.160	0.160	0.0	0.0	0.0	310	black disc
3	15.0	10.18	-	1.33	240.3	0.400	0.400	0.0	18.2	22.4	200	black disc
3	17.4	9.59	-	4.9	200.2	0.224	0.224	3.2	3.2	3.2	99	black disc
4	14.7	-	-	0.84	119.7	0.029	0.029	0.0	0.0	0.0	-	NM
4	14.4	-	-	1.57	119.2	0.029	0.029	0.0	5.5	5.5	-	NM
4	15.1	-	-	0.67	115.9	0.034	0.034	0.0	0.0	0.0	-	NM
5	20.6	7.01	7.29	1.89	350.0	0.022	0.022	0.0	0.0	0.0	158	clarity tube
5	19.2	7.10	7.40	1.88	362.0	0.013	0.013	0.0	0.0	0.0	193	clarity tube
5	19.3	3.78	7.14	1.91	300.0	0.011	0.011	0.0	0.0	0.0	208	clarity tube
5	16.9	4.19	7.22	2.4	236.0	0.011	0.011	0.0	0.0	0.0	161	clarity tube
6	18.0	8.40	7.48	5.80	1693.0	0.630	0.630	0.0	0.0	0.0	43	clarity tube
6	17.3	9.65	7.45	5.1	1657.0	0.607	0.607	0.0	0.2	0.2	53	clarity tube
6	16.6	9.08	7.58	4.2	1616.0	0.601	0.601	0.0	0.0	1.0	66	clarity tube
6	16.0	8.46	7.48	4.6	3272.0	0.432	0.432	0.0	0.0	0.0	49	clarity tube
7	19.6	9.02	7.45	2.8	153.3	0.065	0.065	0.0	0.0	33.5	-	NM

Site	Water temperature (°C)	Dissolved oxygen (mg/L)	Field pH	Laboratory Turbidity (NTU)	Laboratory Conductivity (µS/cm)	excl 95th %Flow (m3/s) - 24h ave	Flow (m3/s) - 24h ave	Rainfall (mm) - 24h	Cumulative Rainfall (mm) - 48h	Cumulative Rainfall (mm) - 72h	Water clarity (cm)	Water clarity method
7	18.3	9.19	7.69	-	153.7	0.060	0.060	14.5	14.5	14.5	110	black disc
7	17.2	9.33	7.87	1.93	160.8	0.041	0.041	0.0	0.0	0.0	90	black disc
8	15.0	8.33	-	1	88.5	0.031	0.031	0.0	0.0	9.0	170	black disc
8	14.8	9.21	-	1.28	-	0.031	0.031	0.0	9.5	9.5	130	black disc
8	14.5	9.44	-	0.11	-	0.023	0.023	0.0	0.0	0.0	131	black disc
9	16.7	10.56	8.19	0.46	234.0	0.354	0.354	0.0	0.0	0.0	630	black disc
9	15.7	9.36	7.30	0.42	230.6	0.319	0.319	0.0	0.5	4.1	-	NM
9	15.8	8.32	7.63	-	228.8	0.310	0.310	15.2	15.2	15.2	520	black disc
10	16.9	8.58	8.15	-	313.6	0.129	0.129	8.5	10.0	10.0	71	clarity tube
10	19.9	7.92	8.21	1.5	274.2	0.111	0.111	0.0	0.0	0.0	161	clarity tube
10	17.7	8.60	8.48	1.1	249.7	0.236	0.236	0.0	0.0	3.5	216	clarity tube
11	15.2	9.70	5.90	8.9	203.4	0.250	0.250	0.0	0.0	11.5	42	black disc
11	15.4	10.59		15.6	220.3	0.030	0.030	0.0	0.0	2.5	31	black disc
11	15.8	8.55	7.81	7	251.1	0.268	0.268	0.0	43.5	43.5	42	black disc
12	21.9	8.12	7.75	1.26	125.5	4.430	4.430	0.0	0.0	0.0	213	black disc
12	18.4	9.53	8.07	12.4	102.6	-	22.536	0.0	2.4	24.6	45	black disc
12	21.2	9.01	7.88	6.7	106.7	-	6.609	0.0	0.2	0.2	-	NM
13	18.4	7.41	7.20	0.21	264.9	0.425	0.425	0.0	0.0	0.0	800	black disc
13	19.5	10.40	7.63	0.25	261.0	0.373	0.373	0.0	0.0	0.0	800	black disc
13	17.1	9.95	7.60	0.24	253.7	0.328	0.328	0.0	0.0	0.0	800	black disc
13	17.0	9.90	7.72	0.26	262.2	0.341	0.341	0.0	0.0	0.0	324	clarity tube
14	17.7	5.91	-	2.3	-	0.383	0.383	0.0	0.0	0.0	188	black disc
14	17.9	6.03	-	3.2	-	0.405	0.405	0.5	20.5	20.5	131	black disc

Site	Water temperature (°C)	Dissolved oxygen (mg/L)	Field pH	Laboratory Turbidity (NTU)	Laboratory Conductivity (µS/cm)	excl 95th %Flow (m3/s) - 24h ave	Flow (m3/s) - 24h ave	Rainfall (mm) - 24h	Cumulative Rainfall (mm) - 48h	Cumulative Rainfall (mm) - 72h	Water clarity (cm)	Water clarity method
14	18.8	5.51	-	3.9	-	0.368	0.368	0.0	0.0	0.0	153	black disc
15	25.0	9.51	7.53	2.10	305.8	0.090	0.090	0.0	0.0	0.0	170	black disc
15	21.3	8.53	7.57	1.8	272.2	0.084	0.084	0.0	0.0	0.0	-	NM
15	21.2	7.72	7.72	1.64	279.0	0.068	0.068	0.2	0.2	0.6	160	black disc
16	22.9	5.25	6.45	3.00	97.5	0.132	0.132	0.0	0.0	0.0	180	black disc
16	20.8	4.32	6.15	2.8	97.1	0.088	0.088	0.0	0.0	0.0	-	NM
16	21.5	4.56	6.26	2.8	97.5	0.075	0.075	0.0	0.0	0.0	168	black disc

APPENDIX E: FST marker results

E.1 Wildfowl FST markers

Two wildfowl FST markers were used: GFD, and the E2 marker which is more specific for "duck". Wildfowl FST markers were ubiquitous, with GFD detected in all samples and 23 samples with GFD concentrations greater than 1,000 copies/100 mL (Figure 27).

The E2 marker was detected in 48 samples, although only six of these were greater than 1,000 copies/100 mL. For this study GFD and E2 results were combined, to give "Wildfowl Total", with combined concentrations of >1,000 copies in 32 samples. Only two samples had concentrations less than 100 copies indicating that wildfowl faecal pollution was likely to be a contributor to most of the samples and all of the rivers tested.

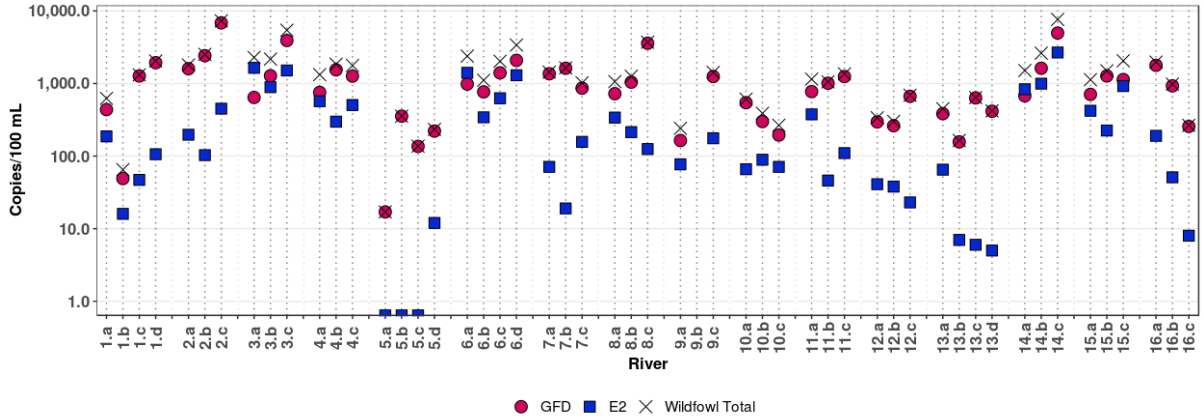

Figure 27: Wildfowl FST markers

No data is available for 9b

E.2 Ruminant FST markers

The BacR marker indicates the presence of ruminant sources of faecal pollution and was detected in 35 samples, of which 23 samples had >1,000 copies/100 mL, 11 were >10,000 copies/100 mL, and a maximum of 440,000 copies was observed. For each positive sample, the BacR concentration was expressed as a ratio of the total GenBac concentration measured in the sample. Samples with less than 1,000 copies of BacR AND where that level was less than 0.3% of the GenBac were considered insignificant levels. The samples in which ruminant faecal markers are potentially a significant contributor are shown in red in Figure 28.

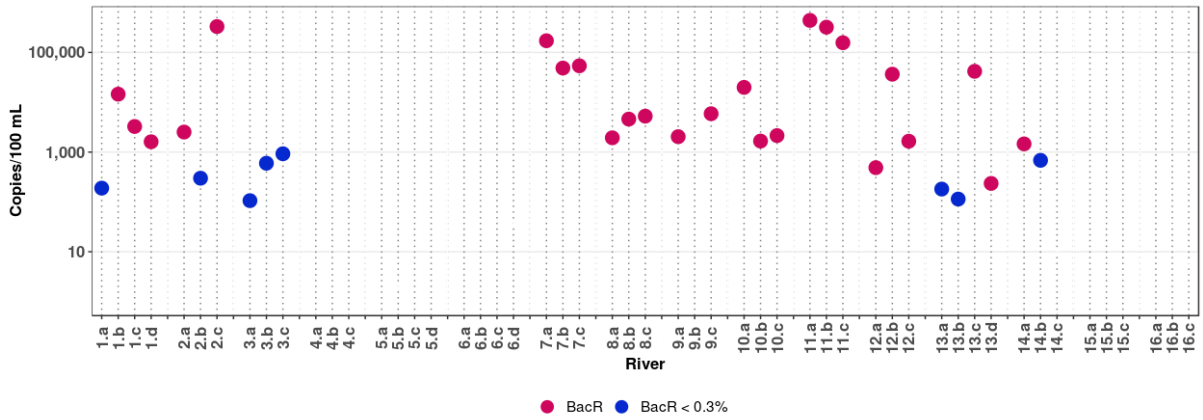

Figure 28: Ruminant FST markers

Note: Only data above 100 copies/100mL are shown. No data is available for 9b

E.3 Human FST markers

Three human FST markers were used: HF183, CrAss phage and BiADO. While high concentrations of these markers are indicative of human sources of pollution, low concentrations of these markers can be found in other sources. Samples with more than a 1,000 copies/100 mL of any of these markers indicate a significant human source. The low concentrations of one or more of the human indicative markers in the other samples may be the result of low level human faecal sources or may be the consequence of non-human sources. Since these can't be distinguished, in this study, samples with all three markers and a combined total of 1,000 copies or more were characterised as containing a significant human source of faecal pollution Figure 29.

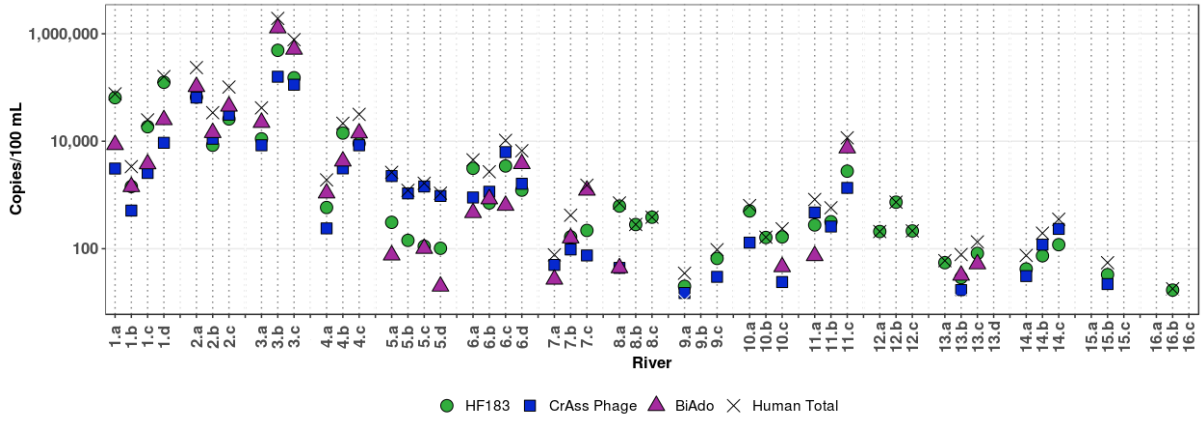

Figure 29: Human indicative FST markers

Note: No data is available for 9b

APPENDIX F: *Campylobacter* MPN and WGS results

F.1 MPN and WGS results

The results for each tube used in the 5 tube MPN are given in Table 22. There are some unexpected results, also noted in the McBride et al (2002) study. An example of this is where 700 mL sample is negative, but the lower volume samples are positive. This occurred in only seven samples. Those all had at least one of the 100 mL samples positive, and MPN values of 0.38 or lower. All of the samples with higher concentrations followed the expected pattern of results.

Table 22: *Campylobacter* MPN and WGS results

	MPN VOLUME (ML)					MPN/			MBIT
RIVER	700	100	1	0.1	0.01	100 ML	SPECIES	MLST ST	PREVIOUS SOURCES
1A	1	2	0	0	0	4.1	*(Campylobacter)***		
1B	0	2	0	0	0	0.25	*C. jejuni*	ST3640	water/wildfowl
1C	1	2	0	0	0	4.1	*C. jejuni*	ST2381	water/wildfowl
							Campylobacter spp. 2		
1D	1	2	0	0	0	4.1	*C. jejuni*	ST45	human
							Campylobacter spp. 1		
2A	0	0	0	0	0	<0.11			
2B	0	2	0	0	0	0.25	*C. jejuni*	ST995	poultry/water
2C	1	2	0	0	0	4.1	*C. jejuni*	ST45	human
							C. lari		wildfowl/human
3A	0	0	0	0	0	<0.11			
3B	0	0	0	0	0	<0.11			
3C	0	1	0	0	1	0.23	*C. jejuni*	ST3640	water/wildfowl
4A	0	0	0	0	0	<0.11			
4B	1	0	0	0	0	0.21	*C. jejuni*	ST991	water/wildfowl
4C	0	1	1	0	0	0.24	*C. jejuni*	ST2381	wildfowl/human
5A	0	0	0	0	0	<0.11			
5B	0	0	0	0	0	<0.11			
5C	0	0	0	0	0	<0.11			
5D	0	0	0	0	0	<0.11			
6A	1	2	0	0	0	4.1	*C. jejuni*	ST9820	water/wildfowl
6B	0	0	0	0	0	<0.11			
6C	0	2	0	0	0	0.25	*C. jejuni*	ST177	water/wildfowl
6D	1	2	0	0	0	4.1	*C. jejuni*	ST2381*	water/wildfowl
7A	1	2	2	0	0	92	*C. jejuni*	ST3640	water/wildfowl
							Campylobacter spp. 1		
7C	1	2	0	0	0	4.1	*C. jejuni*	ST3640	water/wildfowl
8A	1	2	1	0	1	72	*C. jejuni*	(~ST2381)	water/wildfowl

	MPN VOLUME (ML)					MPN/			MBIT
RIVER	700	100	1	0.1	0.01	100 ML	SPECIES	MLST ST	PREVIOUS SOURCES
8B	1	2	0	0	0	4.1	*C. jejuni*	ST45	human/poultry/ ruminant
8C	1	2	1	1	0	74	*C. jejuni*	ST2381	water/wildfowl
9A	0	0	0	0	0	<0.11			
9B	1	0	0	0	0	0.21	*(Campylobacter)***		
9C	1	2	2	0	0	92	*C. jejuni*	NewST*	water/wildfowl
10B	0	0	0	0	0	<0.11			
10C	0	0	0	0	0	<0.11			
11A	1	0	0	0	0	0.21	*C. jejuni*	ST2389	water/wildfowl
11B	1	2	0	0	0	4.1	*C. jejuni*	ST45	human/poultry/ ruminant /wildfowl
11C	1	2	0	0	0	4.1	*C. jejuni*	ST677	wildfowl/human
							C. lari		water/wildfowl
12A	1	2	0	0	0	4.1	*C. jejuni*	ST991	water/wildfowl
12B	0	0	0	0	0	<0.11			
12C	0	0	0	0	0	<0.11			
13A	1	2	2	0	0	92	*C. jejuni*	ST9820	water/wildfowl
13B	1	0	0	0	0	0.21	*C. jejuni*	(~ST2381)	water/wildfowl
13C	1	2	0	0	0	4.1	*C. jejuni*	(~ST2381)	water/wildfowl
13D	1	0	0	0	0	0.21	*C. jejuni*	(~ST2381)	water/wildfowl
14A	0	2	0	0	0	0.25	*C. jejuni*	ST1956	human
14B	1	2	0	0	0	4.1	*C. jejuni*	ST699	water/wildfowl
14C	0	2	1	0	0	0.38	*C. jejuni*	ST45	human/poultry/ ruminant /wildfowl
15A	0	0	0	0	0	<0.11			
15B	0	0	0	0	0	<0.11			
15C	1	1	0	0	0	0.7	*C. jejuni*	(~ST2381)	water/wildfowl
16A	1	1	0	0	0	0.7	*Campylobacter spp. 1*		water/wildfowl
16B	1	1	0	0	0	0.7	*Campylobacter spp. 1*		water/wildfowl
							Campylobacter spp. 1		water/wildfowl
16C	1	2	0	0	0	4.1	*C. jejuni*	(~ST3640)*	water/wildfowl
							Campylobacter spp. 1		water/wildfowl

* Lower quality sequence result. Not included in subsequent wgMLST comparisons.
** No isolate recovered.

F.2 Campylobacter WGS

Isolates were whole genome sequenced from 32 of the river samples, with two isolates analysed from seven of the samples. High quality sequencing results were obtained from 38 of the isolates are presented. Thirty of the isolates were identified as *Campylobacter jejuni.* These were 14 different

seven gene MLST types. Three of these were new STs, although two of these new STs only differed from ST2381 and ST3640 by 1 allele each. A wgMLST comparison of these isolates (Figure 30) demonstrates that these are a genetically diverse group of isolates, with the exception of two of the isolates from river 7 which differed by only two loci but were from samples taken six weeks apart.

Two isolates were identified as *C. lari.* These were genetically distinct from each other. Of the 748 loci shared, 483 of them were different. Both *C. lari* isolates were found in rivers that also contained *C. jejuni*.

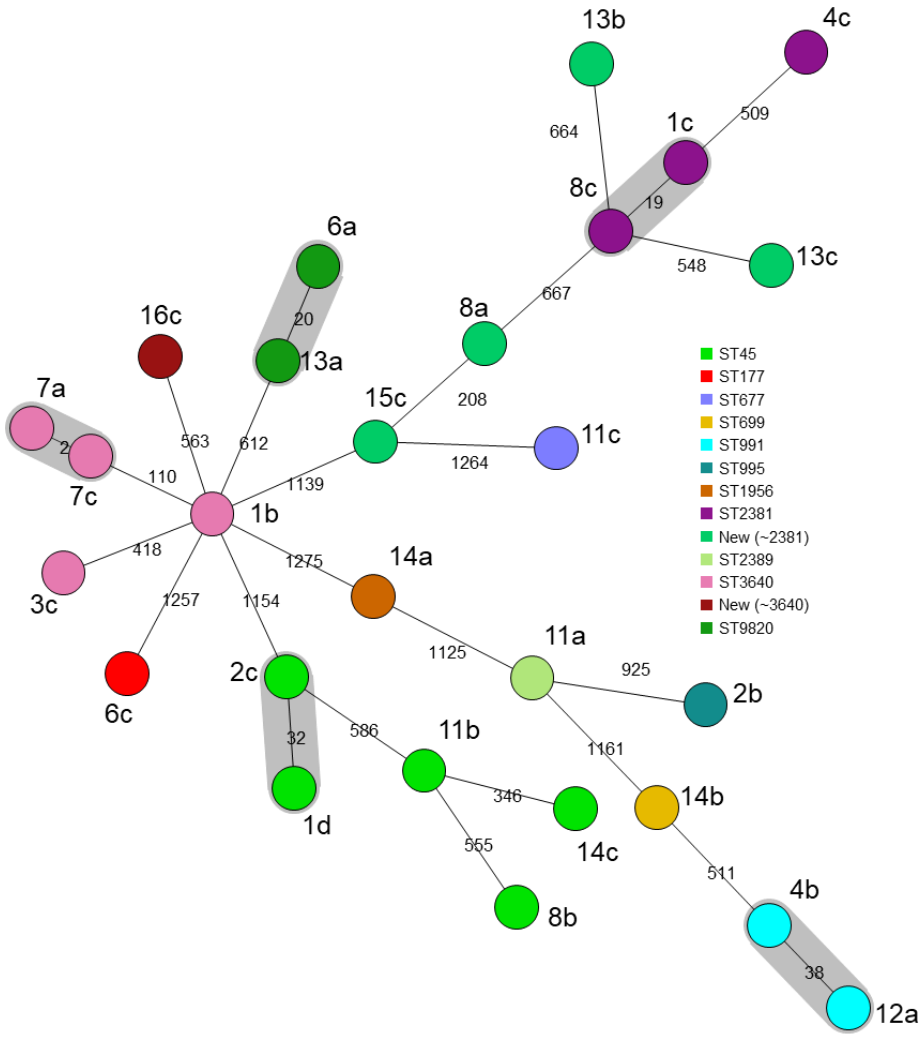

Figure 30: wgMLST comparison of *C. jejuni* isolates

Note: Labelled according to river, and smaller numbers on branches are the number of wgMLST differences between isolates. The colour relates to the MLST sequence type.

There were seven isolates sequenced which were potentially new species of *Campylobacter*. These separated into two clusters, six of which we designate *Campylobacter* spp. 1, and one

Campylobacter spp. 2. A comparison of alleles in Figure 31 suggests these are genetically diverse. *Campylobacter* spp. 1 isolates have a genome size of 2.2 million bases (*C. jejuni* are 1.6 million bases, *C. lari,* 1.45 million bases). *Campylobacter* spp. 2 isolate has a genome of 1.7 million bases. These isolates were confirmed by PCR to belong to the thermotolerant *Campylobacter,* which includes *C. jejuni* and *C. coli*.

Additional genomic analyses were performed to better understand and identify the isolates from the new *Campylobacter* spp.. The analyses used ribosomal MLST (rMLST), average nucleotide identity (ANI) and genome BLAST distance phylogeny (GBDP). Ribosomal MLST indexes the sequence variation in 53 ribosomal protein subunits and aims to provide differentiation that spans intraspecies bacterial typing and identification to the domain level (Jolley et al, 2012). ANI and GBDP evaluate sequence variation over an entire genome and have been proposed as in-silico alternatives to DNA-DNA hybridisation for determining interspecific genomic relatedness of bacteria (On et al, 2017). When the genomes from isolates of these new *Campylobacter* spp. were compared with type strains of all validly described *Campylobacter* species, all three analyses confirmed that these genomes were from the genus *Campylobacter* but were different to all validly described species in this genus.

There is no evidence that these novel *Campylobacter* spp. 1 and 2 cause disease in human, so the health risk is unknown. Four of these isolates were recovered from rivers which also contained *C. jejuni*. These new *Campylobacter* spp. were recovered from every sample taken from river 16 including two occasions where *C. jejuni* were not isolated.

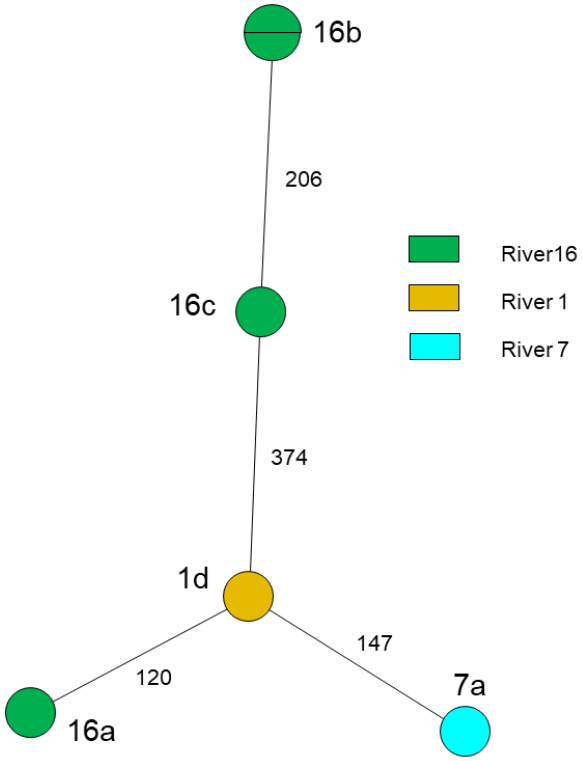

Figure 31: wgMLST comparison of the six isolates identified as new *Campylobacter* spp. 1

F.3 Isolate clusters

Comparison, of the MBiT types inferred from the WGS, with the 3693 New Zealand *Campylobacter* isolates in the MBiT library showed that most of the *C. jejuni* isolated in this pilot study clustered with isolates previously recovered from water and wildfowl, rather than ruminants, poultry or

humans (Figure 32). This suggests that many of the *Campylobacter* isolates (including *C. jejuni* isolates) found in this study have not been associated with human illness.

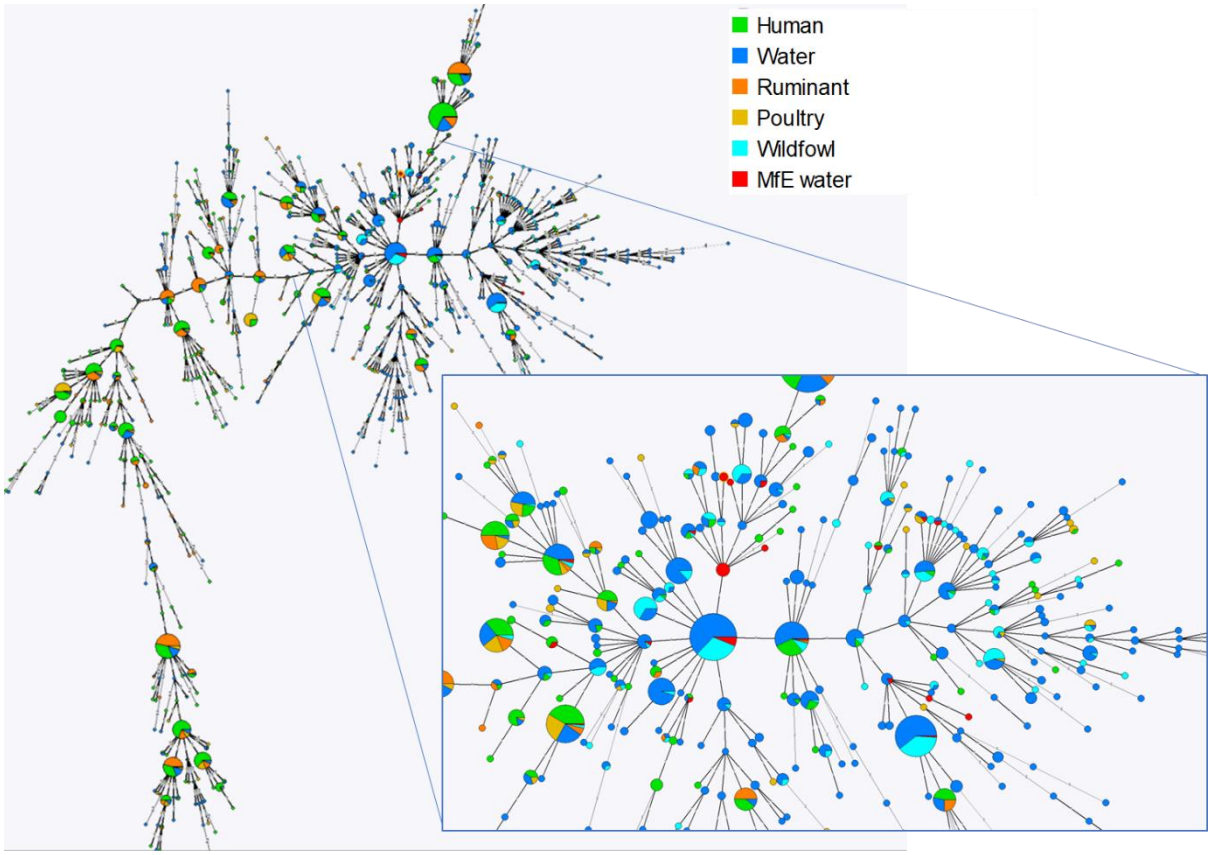

Figure 32: MBiT analysis of *C. jejuni* isolates showing the sources of the isolates

Note: MfE are the isolates from this study, shown in red

APPENDIX G: *Salmonella* and STEC WGS results

G.1 Salmonella

WGS identified isolates as *Salmonella enterica*, and serovars Enteritidis, Typhimurium, Emek or Saintpaul. The Enteritidis, Typhimurium, and Saintpaul isolates were all from rivers either dominated by human contamination, or where human contamination sometimes occurred. Where more than one isolate of *Salmonella* from a sample was sequenced, they were all the same serovar and MLST type and, when compared by wgMLST, were indistinguishable. These isolates were from the same enrichment broth, and suggest that sequencing of more than one isolate from the same enrichment broth does not offer benefits. The wgMLST analysis comparison used at least 3,000 genes from each isolate. The isolates from river 14, sampled 6 weeks apart were both *S*. Emek, and genetically very similar (Figure 33).

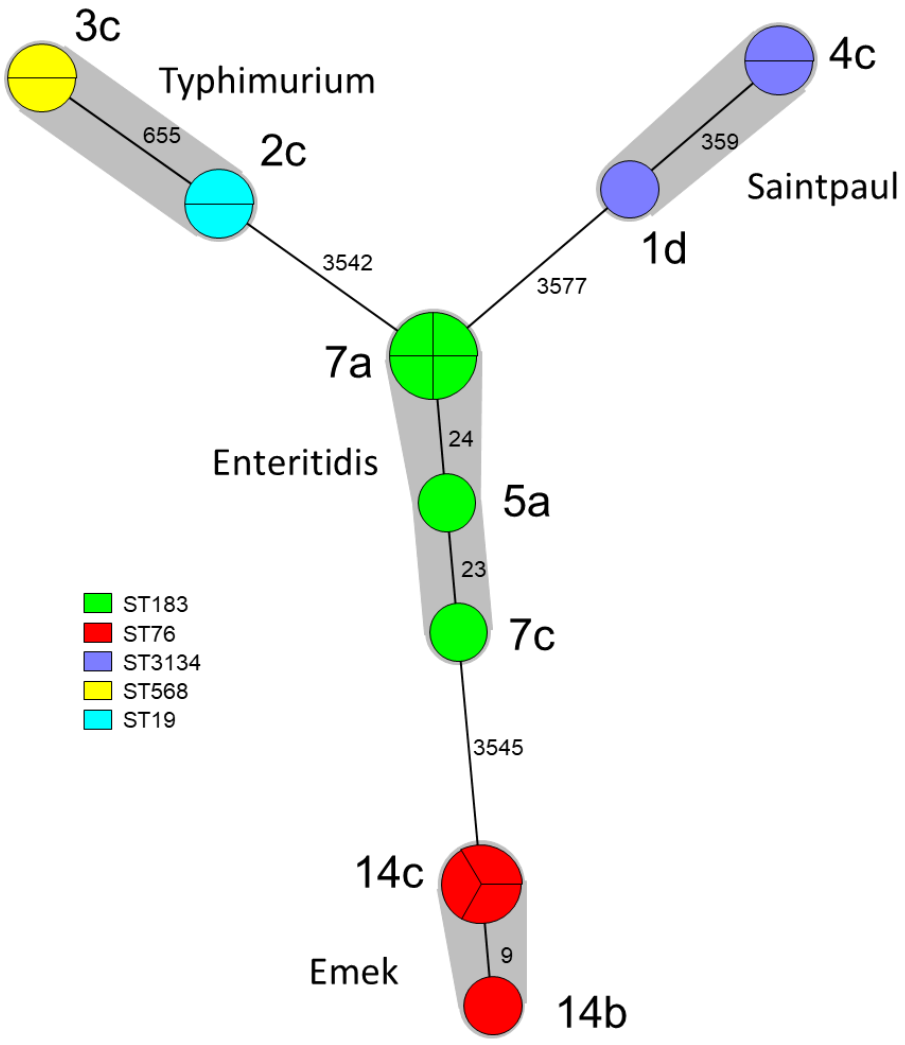

Figure 33: wgMLST comparison of *Salmonella* isolates

In 2019, there were 1153 *Salmonella* isolates confirmed among notified cases in New Zealand (ESR, 2020). The most common serotype was *S*. Typhimurium, comprising 36% of isolates. *S*. Enteritidis is the next most common with 14% of isolates. Saintpaul is a serotype generally associated with living in rural areas (King et al, 2011) and is relatively common across both human clinical samples and animals in New Zealand. There were no *S*. Emek isolates in 2019, and this is an uncommon serotype in human linked *Salmonella* isolates in New Zealand. *S*. Emek has been confirmed in bovine sources from the region where these water samples were taken (ESR, 2012). Overseas, *S*. Emek has been detected in the sewage wastewater stream and in sewage sludge in Sweden (Sahlstrom et al, 2006).

G.2 STEC

River 13b was *stx*2 positive and was confirmed as STEC, with serotype O177:H25 (Table 23). In addition two other isolates from two samples were whole genome sequenced (Table 23). The *eae* positive isolate from river 2c was confirmed as an *E. coli*, with serotype O6:H10. This isolate had none of the common virulence genes and is not predicted to be pathogenic. The second *eae* positive isolate was from river 13d and was identified as *E. albertii*. This isolate had virulence genes present, and is predicted to be pathogenic and an Enteropathogenic *E. coli* (EPEC).

E. coli O177:H25 is uncommon in New Zealand with only two clinical cases confirmed by August 2020. Overseas, *E. coli* O177 has been reported in human and one mixed animal faecal waste sample (Garcia-Aljaro et al, 2005). *E. coli* O6 has been found in humans, dogs and cats. In this study it was found in an area with human and ruminant FST after significant rainfall (38 mm within the previous 48 hours).

The other isolate with eae gene was identified by WGS as *E. albertii*, and while it does not have the shiga toxin-producing genes, based on other virulence genes is predicted to be an EPEC (Enteropathogenic *E. coli*). *E. albertii* is associated with diarrheal illness in humans and birds (Oaks et al, 2010; Gordon, 2011). A survey of birds in Australia identified that it was not widespread in poultry, but was present in 0.95% of native birds (Gordon, 2011). In Canada it was found in freshwater in 2.5% of all *Escherichia* isolates from a survey of 527 water samples: (Maheux et al, 2014). In the current study, *E. albertii* was isolated in a river which had mostly wildfowl FST. The water sample however was positive for *stx*1 and *stx*2, suggesting that STEC may have also been present in this river, although we were not able to isolate them.

Table 23: *Escherichia* isolated from MPN enrichments and characterised by WGS

River	MPN Enrichment	Identification	Serotype	Virulence genes	Pathotype
13b	*eae* + *stx2*	*E. coli*	O177:H25	17	STEC
2c	*eae*	*E. coli*	O6:H10	0	-
13d	*eae*	*E. albertii*	-	9	EPEC

APPENDIX H: River data

The water quality data for field measurements of temperature, pH and DO are shown in Figure 34.

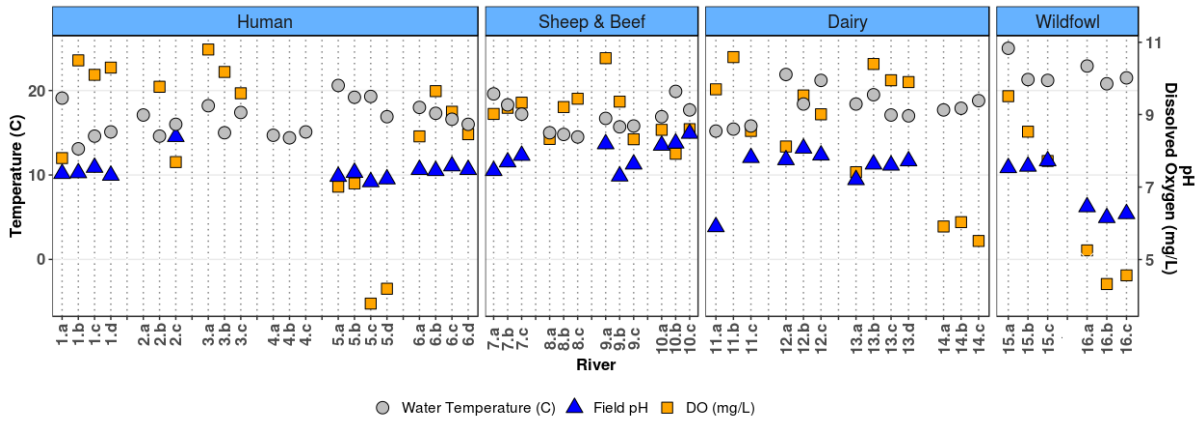

Figure 34: Temperature, pH and DO by river

Note: No data means no measurement recorded

Figure 34 shows the expected inverse association between higher temperatures and lower DO concentrations, with the actual relationship being river specific, eg rivers 6 and 14 have similar temperature (16-18°C) but different DO concentrations (8.4-9.7 and 5.5-6.0, respectively). The actual DO characteristics may be driven by local features such as natural aeration in the river which will increase DO concentrations, while the presence of organic matter can reduce DO concentrations, as oxygen is consumed as organic matter is degraded. Rivers 5, 14 and 16 had the lowest DO concentrations. Rivers 11 and 16 have pH below the usual ecological minimum of 6.5. The highest pH range was in river 11 where the first sample was impacted by a high rainfall event. A decrease in temperature from initial sampling in the first two weeks of February occurred at 11 rivers. River 16 appears the most impacted as it had higher temperatures, with low pH and DO.

Two methods of water clarity measurement were used in the field. Water clarity can be measured as the depth at which a black disk disappears and reappears from vision, or with a clarity tube where a tube is filled with water and the horizontal length at which a black target can be seen is recorded. Water clarity tubes may be used where there is extensive macrophyte growth or shallow water. The clarity tube measurements have been converted to black disk measurements using the correlations given by (Kilroy and Biggs, 2002). Generally, high water clarity is associated with low turbidity. The median clarity was 158 cm. Rivers 9 and 13 have the highest clarity (520-640 cm and 800cm respectively) (Figure 35) while rivers 5, and 10 have the lowest clarity but also have low turbidity, indicating that the characteristics of the river also influence clarity.

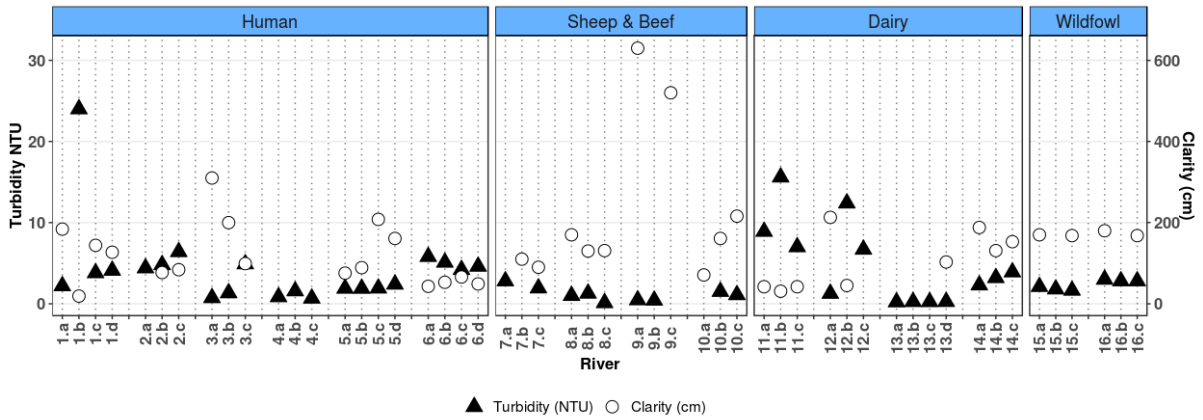

Figure 35: Clarity and turbidity by river

Note: excluding river 13 water clarity measurements at 800cm

Flow and Turbidity

Flows have been taken at nine sampling sites (rivers 1, 3, 5, 7, 8, 12, 13, 14, 16), modelled for two sites (rivers 2, 11) using data from a nearby river which had a gauge. At three sites, the flow data was 3 - 9.5 km upstream (rivers 4, 6, 9) and 3 - 6km downstream (rivers 10, 15) of the sampling site. River 1 only had hourly flow and river 7 only had 24h average flow data. River 12 was much larger in flow than the other rivers, with flows which ranged from 4.43m^3/s-22.54m^3/s, and which on two occasions had 24h average flows more than the 95th percentile. River 6, was the next largest river, with average 24h flows of 0.43 - 0.63m^3/s. River 5 had the lowest flows at 0.011m^3/s.

Turbidity is a physical property of water which is a measure of the relative clarity of water. It is affected by suspended solids, organic and inorganic matter and has been proposed as an indicator of flow. Measurements are dependent on the meter used and are often characteristic of the source. While a typical pattern is that low turbidity is associated with an increase in flow, it depends on the river. As shown in Figure 36, river 14 has much greater flow than rivers 15 and 16, but similar turbidity measurements.

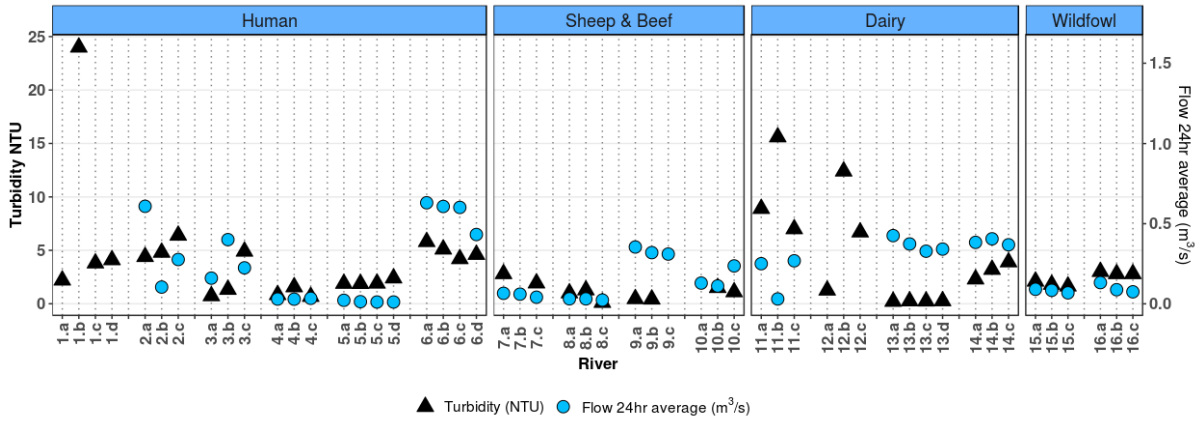

Figure 36: Turbidity and 24h average flow

Note: Excludes flow data for 3 samples above 2m^3/s. There is no flow data for river 1 and no turbidity data for 3 samples

Rainfall

Five rivers (5, 6, 13, 15, 16) had ≤1mm rain during, or up to, 72 hours before any sampling event (18 samples) and there was <2.5mm rainfall (and therefore little runoff likely) for 72 hours prior to a further 12 samples being collected. There had been rainfall within the previous 72 hours for 22 samples, of which only six were within 24 hours of sample collection. Tubidity has been proposed as a predictor of increased flow from rainfall. However, in this short study, there is no apparent association. The turbidity appears to be dependent on the characteristics of the site, as seen in river 2 where all samples have similar tubidity, despite a high recent rainfall event 2c, while river 6 has similar turbidity to river 2, but no recent rainfall (Figure 37).

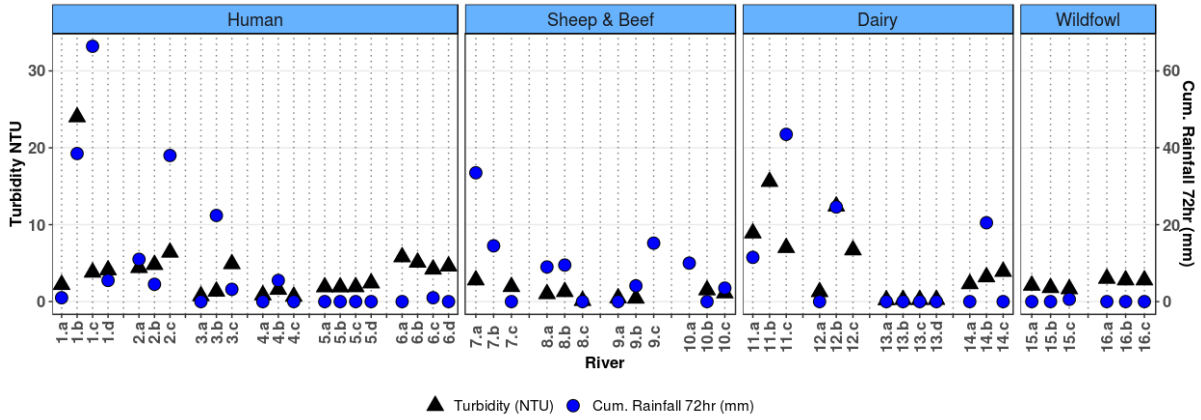

Figure 37: Turbidity and cumulative rainfall in the previous 72h by river

APPENDIX I: Faecal source information and knowledge gaps

Table 24: Information collected on pathogen prevalence and concentration in animal faeces

Knowledge gaps are highlighted in the table.

Micro-organism	Study location	Animal faecal type			
		Dairy	**Beef cattle**	**Sheep**	**Avian: (not chicken) (Swans, Canada Geese, Ducks, Geese, Gulls, pigeons)**
FIB (*E. coli* and enterococci)	New Zealand and International	Prevalence and concentration data	Prevalence and concentration data	Prevalence and concentration data	Prevalence and concentration data but no data on pigeons
Campylobacter* spp. and *C. jejuni	New Zealand and International	Prevalence and concentration data	Prevalence and concentration data	Prevalence and concentration data	Prevalence and concentration data *Campylobacter* spp.
Cryptosporidium spp.	New Zealand	Prevalence and limited concentration data	No data	Limited data on prevalence and concentration	Prevalence only
	International	Data targets *C. parvum* and *C. hominis*	Limited data on prevalence and concentration	Limited data on prevalence and concentration	Limited data on prevalence and concentration
Cryptosporidium parvum* and *C. hominis	New Zealand	Limited data on prevalence and concentration	No data	No data	No data
	International	Prevalence and concentration data	Limited data on prevalence and concentration	Limited data on prevalence and concentration	No data
Pathogenic *E. coli*	New Zealand	Prevalence only	No data	No data	No data
	International	Limited data on prevalence and concentration	Limited data on prevalence and concentration	Limited conc. data on *E. coli* O157 only	Prevalence only

Micro-organism	Study location	Animal faecal type			
		Dairy	**Beef cattle**	**Sheep**	**Avian: (not chicken) (Swans, Canada Geese, Ducks, Geese, Gulls, pigeons)**
Giardia spp.	New Zealand	Limited data on prevalence and concentration	No data	Limited data on concentration	No data
	International	Prevalence and concentration data on human infective *Giardia*	Prevalence and concentration data on human infective *Giardia*	Prevalence and concentration data on *Giardia* species. Limited concentration data on human infective *Giardia*	Limited data on prevalence and concentration
Salmonella	New Zealand	Prevalence only	No data	Prevalence only	Prevalence only
	International	Limited data on prevalence and concentration	No data	Limited data on prevalence and concentration	Limited data on prevalence and concentration

REFERENCES

Ahmed W, Payyappat S, Cassidy M, Besley C. 2019. A duplex PCR assay for the simultaneous quantification of *Bacteroides* HF183 and crAssphage CPQ_056 marker genes in untreated sewage and stormwater. *Environment International* 126**,** 252-259.

APHA. 2017a. 9223 Enzyme Substrate ColiformTest. 9223 B Enzyme Substrate Test. *Standard Methods for the Examination of Water and Wastewater.* On line 2017: American Public Health Association, American Water Works Association, Water Environment Federation

APHA. 2017b. 9230 Fecal Enterococcus/Streptococcus Groups. 9230 D Fluorogenic Substrate Enterococcus Test. *Standard Methods for the Examination of Water and Wastewater.* On line 2017: American Public Health Association, American Water Works Association, Water Environment Federation.

APHA. 2017c. 9260 Detection of Pathogenic Bacteria 9260 G *Campylobacter*. *Standard Methods for the Examination of Water and Wastewater.* On line 2017: American Public Health Association, American Water Works Association, Water Environment Federation.

APHA. 2017d. 9260 Detection of Pathogenic Bacteria. 9260 B – *Salmonella. Standard Methods for the Examination of Water and Wastewater.* On line 2017: American Public Health Association, American Water Works Association, Water Environment Federation.

Baque RH, Gilliam AO, Robles LD, Jakubowski W, Slifko TR. 2011. A real-time RT-PCR method to detect viable *Giardia lamblia* cysts in environmental waters. *Water Research* 45**,** 3175-3184.

Bell C, Kyriakides A. 2005. *A Practical Approach to the Organism and its Control in Foods*, Wiley Online Library.

Best EL, Powell EJ, Swift C, Grant KA, Frost JA. 2003. Applicability of a rapid duplex real-time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture plates. *FEMS Microbiology Letters* 229**,** 237-241.

Bradshaw JK, Snyder BJ, Oladeinde A, Spidle D, Berrang ME, Meinersmann RJ, Oakley B, Sidle RC, Sullivan K, Molina M. 2016. Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. *Water Research* 101**,** 498-509.

Cashdollar JL, Brinkman NE, Griffin SM, Mcminn BR, Rhodes ER, Varughese EA, Grimm AC, Parshionikar SU, Wymer L, Fout GS. 2013. Development and evaluation of EPA Method 1615 for detection of enterovirus and norovirus in water. *Applied and Environmental Microbiology* 79**,** 215-223.

Chern E, Siefring S, Paar J, Doolittle M, Haugland R. 2011. Comparison of quantitative PCR assays for Escherichia coli targeting ribosomal RNA and single copy genes. *Letters in Applied Microbiology* 52**,** 298-306.

Cornelius AJ, Vandenberg O, Robson B, Gilpin BJ, Brandt SM, Scholes P, Martiny D, Carter PE, Van Vught P, Schouten J. 2014. Same-day subtyping of *Campylobacter jejuni* and *C. coli* isolates by use of multiplex ligation-dependent probe amplification–Binary typing. *Journal of Clinical Microbiology* 52**,** 3345-3350.

Davies R, Wales A. 2013. *Salmonella* contamination of cereal ingredients for animal feeds. *Veterinary Microbiology* 166**,** 543-549.

Derzelle S, Grine A, Madic J, De Garam CP, Vingadassalon N, Dilasser F, Jamet E, Auvray F. 2011. A quantitative PCR assay for the detection and quantification of Shiga toxin-producing *Escherichia coli* (STEC) in minced beef and dairy products. *International Journal of Food Microbiology* 151**,** 44-51.

Devane M, Moriarty E, Weaver L, Cookson A, Gilpin B. 2020. Fecal indicator bacteria from environmental sources; Strategies for identification to improve water quality monitoring. *Water Research* 116204. DOI.org/10.1016/j.watres.2020.116204

Devane M, Nicol C, Ball A, Klena J, Scholes P, Hudson J, Baker M, Gilpin B, Garrett N, Savill M. 2005. The occurrence of *Campylobacter* subtypes in environmental reservoirs and potential transmission routes. *Journal of Applied Microbiology* 98**,** 980-990.

Devane ML, Moriarty EM, Wood D, Webster-Brown J, Gilpin BJ. 2014. The impact of major earthquakes and subsequent sewage discharges on the microbial quality of water and sediments in an urban river. *Science of the Total Environment* 485**,** 666-680.

Devane ML, Robson B, Nourozi F, Scholes P, Gilpin BJ. 2007. A PCR marker for detection in surface waters of faecal pollution derived from ducks. *Water Research,* 41**,** 3553-3560.

Donaldson K, Griffin DW, Paul J. 2002. Detection, quantitation and identification of enteroviruses from surface waters and sponge tissue from the Florida Keys using real-time RT–PCR. *Water Research* 36**,** 2505-2514.

EPA. 2013. Method 1609: Enterococci in water by taqman® quantitative polymerase chain reaction (qpcr) with internal amplification control (iac) assay [Online]. Washington, D.C.: United States Environmental Protection Agency. [Accessed].

ESR. 2012. *Human salmonella isolates, 2012, no. 1 January* [Online]. https://surv.esr.cri.nz/enteric_reference/human_salmonella.php?we_objectID=2911: ESR, Ministry of Health. [Accessed 25 August 2020].

ESR. 2020. *Human salmonella isolates, 2019* [Online]. Ministry of Health. [Accessed 21/8/2020. https://surv.esr.cri.nz/enteric_reference/human_salmonella.php?we_objectID=5083].

Eyles R, Niyogi D, Townsend C, Benwell G, Weinstein P. 2003. Spatial and temporal patterns of *Campylobacter* contamination underlying public health risk in the Taieri River, New Zealand. *Journal of Environmental Quality* 32**,** 1820-1828.

Farnleitner AH*,* Savio D*,* Sommer R*,* Reischer G*,* Kirschner A*,* Zerobin W*,* Stadler H. 2018. Integrated strategy to guide health-related microbial quality management at alpine karstic drinking water resources*. In* W. B.White*,* J. S.Herman*,* E. K.Herman*, &* M.Rutigliano *(Eds.), Karst groundwater contamination and public health: Beyond case studies (pp.* 185–192*).* Cham, the Netherlands*: Springer International.*

Faul F, Erdfelder E, Lang AG, Buchner A. 2007. G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods,* 39**,** 175-191.

Garcia-Aljaro C, Muniesa M, Blanco J, Blanco M, Blanco J, Jofre J, Blanch A. 2005. Characterization of Shiga toxin-producing *Escherichia coli* isolated from aquatic environments. *FEMS Microbiology Letters* 246**,** 55-65.

Garcia–R JC, French N, Pita A, Velathanthiri N, Shrestha R, Hayman D. 2017. Local and global genetic diversity of protozoan parasites: spatial distribution of *Cryptosporidium* and *Giardia* genotypes. *PLoS Neglected Tropical Diseases,* 11**,** e0005736. DOI: 10.1371/journal.pntd.0005736

Garrett N, Devane M, Hudson J, Nicol C, Ball A, Klena J, Scholes P, Baker M, Gilpin B, Savill M. 2007. Statistical comparison of *Campylobacter jejuni* subtypes from human cases and environmental sources. *Journal of Applied Microbiology* 103**,** 2113-2121.

Gilpin B, Hewitt J, Scholes P, Coxon S, Hayman D, Pita A, Knox M, Garcia-Ramirez JC. 2018. *Freshwater Microbiological Sciences Review (Stage One) Work Stream B: Microbial Methods. Report CSC18018*. Prepared for the Ministry for the Environment by the Institute of Environmental Science and Research Limited.

Gilpin BJ, Walshe G, Smith D, Marshall JC, French NP. 2013. Application of molecular epidemiology to understanding campylobacteriosis in the Canterbury region of New Zealand. *Epidemiology & Infection,* 141**,** 1253-1266.

Gordon DM 2011. Reservoirs of infection: The Epidemiological Characteristics of an Emerging Pathogen, *Escherichia albertii*. Canberra, Australia: The Australian National University.

Graciaa D, Cope JR, Roberts VA, Cikesh BL, Kahler AM, Vigar M, Hilborn ED, Wade TJ, Backer LC, Montgomery SP, Secor WE, Hill VR, Beach MJ, Fullerton KE, Yoder JS, Hlavsa MC. 2018. Outbreaks associated with untreated recreational water — United States, 2000–2014. *Morbidity and Mortality Weekly Report* 67**,** 701–706.

Green HC, Dick LK, Gilpin B, Samadpour M, Field KG. 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. *Applied and Environmental Microbiology* 78**,** 503-510.

Harwood VJ, Staley C, Badgley BD, Borges K, Korajkic A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Review,* 38**,** 1-40.

Haugland RA, Siefring S, Lavender J, Varma M. 2012. Influences of sample interference and interference controls on quantification of enterococci fecal indicator bacteria in surface water samples by the qPCR method. *Water Research* 46**,** 5989-6001.

Health Canada. 2012 *Guidelines for Canadian Recreational Water Quality*, 3rd Edition. Health Canada, Ottawa, Canada. Available at https://www.canada.ca/en/healthcanada/services/publications/healthy-living/guidelines-canadian-recreational-water-qualitythird-edition.html

HC. 2014. Mlfp-46 isolation of thermophilic campylobacter from food. [Online]. Health Canada. [Accessed].

Hernroth BE, Conden-Hansson A-C, Rehnstam-Holm A-S, Girones R, Allard AK. 2002. Environmental factors influencing human viral pathogens and their potential indicator organisms in the blue mussel, *Mytilus edulis*: the first Scandinavian report. *Applied and Environmental Microbiology* 68**,** 4523-4533.

Hewitt J, Bell D, Simmons GC, Rivera-Aban M, Wolf S, Greening GE. 2007. Gastroenteritis outbreak caused by waterborne norovirus at a New Zealand ski resort. *Applied and Environmental Microbiology* 73**,** 7853-7857.

Hewitt J, Leonard M, Greening GE, Lewis GD. 2011. Influence of wastewater treatment process and the population size on human virus profiles in wastewater. *Water Research* 45**,** 6267-6276.

Hill VR, Polaczyk AL, Hahn D, Narayanan J, Cromeans TL, Roberts JM, Amburgey JE. 2005. Development of a rapid method for simultaneous recovery of diverse microbes in drinking water by ultrafiltration with sodium polyphosphate and surfactants. *Applied and Environmental Microbiology* 71**,** 6878-6884.

Hoorfar J, Ahrens P, Rådström P 2000. Automated 5' nuclease PCR assay for identification of *Salmonella enterica*. *Journal of Clinical Microbiology* 38**,** 3429-3435.

Horn B, McBride G, Lake R, Muirhead R. 2018. *Freshwater Microbiological Sciences Review (Stage One) Work Stream C: Dose Response Illness and Infection Risk. Client Report FW17069*. Prepared for the Ministry for the Environment by the Institute of Environmental Science and Research Limited.

Huang W-C, Hsu B-M, Kao P-M, Tao C-W, Ho Y-N, Kuo C-W, Huang Y-L 2016. Seasonal distribution and prevalence of diarrheagenic *Escherichia coli* in different aquatic environments in Taiwan. *Ecotoxicology and Environmental Safety* 124**,** 37-41.

Ibekwe AM, Watt PM, Shouse PJ, Grieve CM. 2004. Fate of *Escherichia coli* O157:H7 in irrigation water on soils and plants as validated by culture method and real-time PCR. *Canadian Journal of Microbiology* 50**,** 1007-1014.

ISO. 2017a. ISO 10272-1:2017 Microbiology of the Food Chain - Horizontal Method for Detection and Enumeration of *Campylobacter* spp. — Part 1: Detection Method. International Organization for Standardization.

ISO. 2017b. ISO 6579-1:2017 Microbiology of the Food Chain - Horizontal Method for the Detection, Enumeration and Serotyping of *Salmonella*-Part 1: Detection of *Salmonella* spp.. International Organization for Standardization.

ISO/TS. 2012. *ISO/TS 13136:2012(e). Microbiology of Food and Animal Feed – Real-time Polymerase Chain Reaction (PCR)-based Method for the Detection of Food-borne Pathogens - Horizontal Method for the Detection of Shiga Toxin-producing Escherichia coli (STEC) and the Determination of O157, O111, O26, O103 and O143 Serogroups* [Online]. International Organization for Standardization.

Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C, Colles FM, Wimalarathna H, Harrison OB, Sheppard SK, Cody AJ. 2012. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* 158**,** 1005-1015.

Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Takeda N, Katayama K. 2003. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *Journal of Clinical Microbiology* 41**,** 1548-1557.

Kilroy C, Biggs BJ. 2002. Use of the SHMAK clarity tube for measuring water clarity: comparison with the black disk method. *New Zealand Journal of Marine and Freshwater Research* 36**,** 519-527.

King N, Lake R, Cressey P. 2011. *Risk Profile:* Salmonella *(non-typhoidal) in poultry (whole and pieces). MPI Techncial Paper No: 2015/04.* Prepared for the Ministry for Primary Industries by the Institute of Environmental Science and Research Limited.

Korajkic A, Mcminn BR, Harwood VJ. 2018. Relationships between microbial indicators and pathogens in recreational water settings. *International Journal of Environmental Research and Public Health* 15**,** 2842. DOI: 10.3390/ijerph15122842.

Kukkula M, Arstila P, Klossner M-L, Maunula L, Bonsdorff C-HV, Jaatinen P. 1997. Waterborne outbreak of viral gastroenteritis. *Scandinavian Journal of Infectious Diseases* 29**,** 415-418.

Lake R, Horn B, McBride G, Muirhead, R. 2018. *Freshwater Microbiological Sciences Review (Stage One). Work Stream C: Dose Response Recommendation for Dose Response. Client Report FW17068.* Prepared for the Ministry for the Environment by the Institute of Environmental Science and Research Limited.

Lal A, Marshall J, et al. 2018. A bayesian spatio-temporal framework to identify outbreaks and examine environmental and social risk factors for infectious diseases monitored by routine surveillance. *Spatial and Spatio-temporal Epidemiology* 25**,** 39-48.

Larned S, Whitehead A, Fraser C, Snelder T, Yang Y. 2018. *Water quality state and trends in New Zealand rivers. Analyses of national data ending in 2017.* A report prepared by National Institute of Water and Atmospheric Research for the Ministry for the Environment.

Lindsey RL, Pouseele H, Chen JC, Strockbine NA, Carleton HA. 2016. Implementation of whole genome sequencing (WGS) for identification and characterization of Shiga toxin-producing *Escherichia coli* (STEC) in the United States. *Frontiers in Microbiology* 7**,** 766. DOI: 10.3389/fmicb.2016.00766.

Maheux A, Boudreau D, Bergeron M, Rodriguez M 2014. Characterization of *Escherichia fergusonii* and *Escherichia albertii* isolated from water. *Journal of Applied Microbiology* 117**,** 597-609.

Malorny B, Paccassoni E, Fach P, Bunge C, Martin A, Helmuth R. 2004. Diagnostic real-time PCR for detection of *Salmonella* in food. *Applied and Environmental Microbiology* 70**,** 7046-7052.

Mary C, Chapey E, Dutoit E, Guyot K, Hasseine L, Jeddi F, Menotti J, Paraud C, Pomares C, Rabodonirina M. 2013. Multicentric evaluation of a new real-time PCR assay for quantification of *Cryptosporidium* spp. and identification of *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Journal of Clinical Microbiology* 51**,** 2556-2563.

McBride G, Soller J 2017. *Technical background for 2017 MfE 'Clean Water' swimmability proposals for rivers Report FWWQ1722.* Prepared for the Ministry for the Environment by the National Institute of Water and Atmospheric Research. https://niwa.co.nz/sites/niwa.co.nz/files/Swimmability%20Paper%2010%20May%202017%20FINAL.pdf

McBride G. 2003. *Preparing Exact Most Probable Number (MPN) Tables using Occupancy Theory and Accompanying Measures of Uncertainty*. Wellington: NIWA

McBride G, Till D, Ryan T, Ball A, Lewis G, Palmer S, Weinstein P. 2002. *Freshwater Microbiology Research Programme Pathogen Occurrence and Human Health Risk Assessment Analysis.* A Report prepared for the Ministry for the Environment and the Ministry of Health.

Messner MJ, Berger P, Nappier S. 2014. Fractional poisson – A simple dose-response model for human norovirus. *Risk Analysis* 34, 1820-1829.

Ministry of Health. 2019. *Guidelines for Drinking-water Quality Management for New Zealand. Updated 2019* [Online]. Wellington: Ministry of Health. [Accessed 2019 Retrieved from https://www.health.govt.nz/publication/guidelines-drinking-water-quality-management-new-zealand].

Milne J, Moriarty E, Davie T, Kitson J, Gilles T. 2018. *Freshwater Microbiological Sciences Review (Stage One). Work Stream A: Site Selection. Recommended Sampling Sites. Report CSC18004*. Prepared for the Ministry for the Environment by the Institute of Environmental Science and Research Limited.

MIMM 2008. Method 13.1 Thermotolerant *Campylobacters*: MPN Method. *Meat Industry Microbiological Methods* 5th ed. On line.

MfE. 2003. *Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas*. Wellington: Ministry for the Environment.

Monaghan RM, Smith LC, Muirhead RW. 2016. Pathways of contaminant transfers to water from an artificially-drained soil under intensive grazing by dairy cows. *Agriculture, Ecosystems & Environment* 220**,** 76-88.

Moriarty E, Milne J, Gilpin B, McBride G, Horn B, Muirhead R, Lake R, Hayman D, Coxson S, Pita A. 2018. Freshwater Microbiological Sciences Review (Stage One). Work Stream D: Survey Design and QMRA Project Plan. Report CSC18006. Prepared for the Ministry for the Environment by the Institute of Environmental Science and Research Limited.

Napier MD, Haugland R, Poole C, Dufour AP, Stewart JR, Weber DJ, Varma M, Lavender JS, Wade TJ. 2017. Exposure to human-associated fecal indicators and self-reported illness among swimmers at recreational beaches: a cohort study. *Environmental Health* 16, 103. DOI.org/10.1186/s12940-017-0308-3.

New Zealand Government 2020. *National Policy Statement for Freshwater Management*. Wellington: Ministry for the Environment.

New Zealand Government 2017. *National Policy Statement for Freshwater Management 2014 (Amended 2017).* Wellington: Ministry for the Environment.

New Zealand Government 2014. *National Policy Statement for Freshwater Management.* Wellington: Ministry for the Environment.

Oaks JL, Besser TE, Walk ST, Gordon DM, Beckmen KB, Burek KA, Haldorson GJ, Bradway DS, Ouellette L, Rurangirwa FR 2010. *Escherichia albertii* in wild and domestic birds. *Emerging Infectious Diseases* 16**,** 638-646.

Olson ME, Goh J, Phillips M, Guselle N, Mcallister TA. 1999. *Giardia* cyst and *Cryptosporidium* oocyst survival in water, soil, and cattle feces. *Journal of Environmental Quality* 28**,** 1991-1996.

On SL, Miller WG, Houf K, Fox JG, Vandamme P. 2017. Minimal standards for describing new species belonging to the families *Campylobacteraceae* and *Helicobacteraceae*: *Campylobacter*, *Arcobacter*, *Helicobacter* and *Wolinella* spp. *International Journal of Systematic and Evolutionary Microbiology* 67**,** 5296-5311.

Paton AW, Paton JC. 1998. Detection and characterization of shiga toxigenic *Escherichia coli* by using multiplex pcr assays for *stx*1, *stx*2, *eaea*, enterohemorrhagic *E. coli hly*A, rfb_{O111} and rfb_{O157}. *Journal of Clinical Microbiology* 36**,** 598-602.

Pattis I, Horn B, Armstrong B, Cressey P, Lopez L. 2020. *Annual Report Concerning Foodborne Disease in New Zealand 2019. Client Report FW20013*. Prepared for the Ministry for Primary Industries by the Institute of Environmental Science and Research Limited.

PHE 2015. Identification of *Salmonella* species. ID 24 Issue 3. *UK Standards for Microbiology Investigations.* Public Health England.

Phiri BJ, Pita AB, Hayman DT, Biggs PJ, Davis MT, Fayaz A, Canning AD, French NP, Death RG. 2020. Does land use affect pathogen presence in New Zealand drinking water supplies? *Water Research* 116229. DOI.org/10.1016/j.watres.2020.116229.

Puig M, Jofre J, Lucena F, Allard A, Wadell G, Girones R. 1994. Detection of adenoviruses and enteroviruses in polluted waters by nested PCR amplification. *Applied and Environmental Microbiology* 60**,** 2963-2970.

Rainforth H, Harmsworth G. 2019. *Kaupapa Māori Freshwater Assessments. A summary of iwi and hapū-based tools, frameworks and methods for assessing freshwater environments*, Perception Planning Ltd.

Reischer GH, Kasper DC, Steinborn R, Mach RL, Farnleitner AH. 2006. Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. *Applied and Environmental Microbiology* 72**,** 5610-5614.

Sahlstrom L, De Jong B, Aspan A. 2006. *Salmonella* isolated in sewage sludge traced back to human cases of salmonellosis. *Letters in Applied Microbiology* 43**,** 46-52.

Savio D, Stadler P, Reischer GH, Kirschner AK, Demeter K, Linke R, Blaschke AP, Sommer R, Szewzyk U, Wilhartitz IC. 2018. Opening the black box of spring water microbiology from alpine karst aquifers to support proactive drinking water resource management. *Wiley Interdisciplinary Reviews: Water,* 5**,** e1282. DOI.org/10.1002/wat2.1282

Seemann T, Goncalves Da Silva A, Bulach DM, Schultz MB, Kwong JC, Howden BP. accessed 2020. *Nullarbor Github* [Online]. Available: https://github.com/tseemann/nullarbor [Accessed].

Shanks OC, Kelty CA, Sivaganesan M, Varma M, Haugland RA. 2009. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology* 75**,** 5507-5513.

Shanks OC, White K, Kelty CA, Hayes S, Sivaganesan M, Jenkins M, Varma M, Haugland RA. 2010. Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology* 76**,** 1359-1366.

Soller JA, Bartrand T, Ashbolt NJ, Ravenscroft J, Wade TJ. 2010a. Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination. *Water Research* 44**,** 4736-4747.

Soller JA, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ. 2010b. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Research* 44**,** 4674-4691.

Tani N, Dohi Y, Kurumatani N, Yonemasu K. 1995. Seasonal distribution of adenoviruses, enteroviruses and reoviruses in urban river water. *Microbiology and Immunology* 39**,** 577-580.

USEPA. 2012. *2012 Recreational water quality criteria*. Report 820-F-12-058. Office of Water. Environmental Protection Agency.

USEPA. 2018. *2017 Five-Year Review of the 2012 Recreational Water Quality Criteria*. Report 823-R-18-001. Washington D.C.: Office of Water Office of Science and Technology, United States Environmental Protection Agency.

Vereen Jr E, Lowrance RR, Jenkins MB, Adams P, Rajeev S, Lipp EK. 2013. Landscape and seasonal factors influence *Salmonella* and *Campylobacter* prevalence in a rural mixed use watershed. *Water Research* 47**,** 6075-6085.

WHO 1999. *Health-Based Monitoring of Recreational Waters: The feasibility of a new approach (The 'Annapolis Protocol')*. World Health Organization, Geneva.

Wolf S, Hewitt J, Greening GE. 2010. Viral multiplex quantitative PCR assays for tracking sources of fecal contamination. *Applied and Environmental Microbiology* 76**,** 1388-1394.

Wong T, Devane ML, Hudson JA, Scholes P, Grace Savill M, Klena JD. 2004. Validation of a PCR method for *Campylobacter* detection on poultry packs. *British Food Journal* 106**,** 642-650.

INSTITUTE OF ENVIRONMENTAL SCIENCE AND RESEARCH LIMITED

Kenepuru Science Centre
34 Kenepuru Drive, Kenepuru, Porirua 5022
PO Box 50348, Porirua 5240
New Zealand
T: +64 4 914 0700 **F:** +64 4 914 0770

Mt Albert Science Centre
120 Mt Albert Road, Sandringham, Auckland 1025
Private Bag 92021, Auckland 1142
New Zealand
T: +64 9 815 3670 **F:** +64 9 849 6046

NCBID – Wallaceville
66 Ward Street, Wallaceville, Upper Hutt 5018
PO Box 40158, Upper Hutt 5140
New Zealand
T: +64 4 529 0600 **F:** +64 4 529 0601

Christchurch Science Centre
27 Creyke Road, Ilam, Christchurch 8041
PO Box 29181, Christchurch 8540
New Zealand
T: +64 3 351 6019 **F:** +64 3 351 0010

www.esr.cri.nz