# AN ASSESSMENT OF THE HUMAN HEALTH RISKS POSED BY PATHOGENS THROUGH VARIOUS WATER TYPES

MARONEL STEYN

# AN ASSESSMENT OF THE HUMAN HEALTH RISKS POSED BY PATHOGENS THROUGH VARIOUS WATER TYPES

Dissertation submitted by

**Maronel Steyn**

in fulfilment of the requirements for the Degree:

**MAGISTER TECHNOLOGIAE: ENVIRONMENTAL HEALTH**

in the

School of Environmental Development and Agriculture

within the

Faculty of Health and Environmental Sciences

of the

Technikon Free State

Study Leader: Prof. P. Jagals (D.Tech)
Co-study Leader: Mrs. B. Genthe (M.Sc.)

BLOEMFONTEIN
January 2003

Central University of Technology, Free State

# DECLARATION OF INDEPENDENT WORK

I, MARONEL STEYN, Identity Number [redacted] and Student Number [redacted] do hereby declare that this research project, submitted to the Technikon Free State for the degree MAGISTER TECHNOLOGIAE: ENVIRONMENTAL HEALTH, is my own independent work.

This work has not been submitted before to any institution by myself or, to the best of my knowledge, any other person in fulfilment of requirements for the attainment of any qualification.

.....
**SIGNATURE OF CANDIDATE**

2003-03-06
**DATE**

# ACKNOWLEDGEMENTS


- To my Father in Heaven - for carrying me through this. I could not walk this path alone.
- To my parents – Mom, Dad – thanks for all the prayers, your love, and teaching me life-lessons in preparation for this.
- To Prof. P. Jagals. Paul thanks for being my study leader, my biggest encouragement, but most of all my friend.
- Special thanks to Mrs. B. Genthe. Bettina – not only for your subject knowledge and guidance, but most of all for being there beyond the 95$^{th}$ percentile.
- To the NRF and Technikon Free State for your financial support.
- To Johann - thank you for your love and endless patience - it kept me strong.
- To Janine, Christo and Johan - thanks for your support.
- Corinne, Irms, and Liezl – thanks for being true friends in every sense of the word.
- Lea, Jemina and Catherine - thanks for the long hours spent in the lab.
- To the rest of my family and friends (especially the $H_2O$–babies) – thanks for understanding "neglect" – I'll make it up to you.

# OPSOMMING

Die doel van die studie was om te bepaal of mense wat by verskeie waterverbruikspunte binne die Mangaung Munisipale gebied aan water, wat waarskynlik 'n infeksie deur die bakteriële patogeen *Salmonella* tot gevolg kan hê, blootgestel word, soos vorige mikrobiologiese moniteringsprogramme in die gebied (gebasseer op die indikator organisme *E. coli*) inderdaad gesuggereer het. Die bepaling het geskied deur die toepassing van 'n Water-verwante Kwantitatiewe Mikrobiologiese Risiko Bepalingsproses (WVKMRB) in gekose gebiede waar gebruik van water vir huishousdelike en ontspanningsverwante aktiwiteite kan lei tot die potensiële inname van besoedelde water.

Die WVKMRB-proses het bestaan uit die Waarneembare Nadeel- en - Effekvlak (WNEV), asook die Kwantitatiewe Mikrobiologiese Risikobepaling (KMRB)-benaderings. Die WNEV-benadering (WNEVB) is gebaseer op die teenwoordigheid van *E. coli* in die toetswaters om die *moontlike* infeksie-risiko te bepaal, terwyl die KMRB die *waarskynlike* infeksie-risiko (deur die werklike getalle van die patogeen *Salmonella*) in die verskeie watertipes voorspel. Hierdie benaderings is toegepas in beide onbehandelde oppervlak-water, wat hoofsaaklik gebruik word vir ontspanningsdoeleindes, en in gestoorde water (onbehandelde fonteinwater en behandelde munisipale water) wat vir huishoudelike doeleindes gebruik is.

The (WNEVB) het die *E. coli* getalle teen verskeie watergebruiksgrense (vervat in watergehalte riglyne) gemeet, terwyl die KMRB die waarskynlike infeksie-risiko ($W_i$) bereken het deur middel van die vier risikobepalingstappe van gevaarbepaling, blootstellingbepaling, dosis-en-effek bepaling en risikokarakterisering. Die resultate van die WNEVB en die KMRB is vergelyk om te bepaal of die gebruik van WNEVB alleenlik (soos gebruik deur Omgewingsgesondheidspraktisyns), op 'n betroubare en aaneenlopende basis, infeksie-risiko kan voorspel. Waterinname was die enigste blootstellingsroete wat ondersoek is. Inname-volumes vir die verskeie watergebruike was as volg: Vir die onbehandelde oppervlakwater wat gebruik is vir ontspanningsverwante doeleindes in die Renosterspuit

Kwartinêre Opvangsgebied (RK van 100mℓ, 50mℓ en 10mℓ gebruik, afhangende van die intensiteit van kontak met die water. Aangepasde inname-volumes is na verbruikers in verskillende ouderdomsgroepe, vanaf jong kinders (1,318 mℓ) tot bejaardes (865 mℓ), geëkstrapoleer.

Vir die onbehandelde oppervlakwater in die RKO is *E. coli* en *Salmonella*-resultate vir die drie monsternemingspunte (RS1, RS2 en BS) gekombineer, en die gemiddeld gebruik om risiko te bereken. 'n Enkele, asook seisoenale (242 dae) risiko is bereken op blootstelling aan die gemiddeld sowel as die 95$^{ste}$ persentiel. Onbehandelde oppervlakwater in die RKO het deur middel van beide die WNEVB en die KMRB, 'n infeksie-risiko aangedui vir ontspanningsgebruikers, selfs vir 'n enkele blootstelling. Die WNEVB het deurentyd infeksie-risiko vir 'n enkele blootstelling óf oor- óf onderskat, en kon dus die risiko nie op 'n aaneenlopende basis in die water voorspel nie. Die toepassing van die volledige WVKMRB-proses word daarom aanbeveel vir toekomstige gebruik.

'n Enkele sowel as jaarlikse (365 dae) infeksie-risiko blootsteling is bereken vir die houer-gestoorde (onbehandelde fontein en gesuiwerde munisipale) water (gebruik vir huishoudelike doeleindes), gebaseer op die gemiddeld en redelike maksimum-verwagte dosis (95$^{ste}$ persentiel). Beide die WNEV- en die KMRB-benaderings het 'n infeksie-risiko na slegs 'n enkele blootstelling aangedui. Die gebruik van die WNEVB alleen was nie betroubaar nie, aangesien die risiko van 'n enkele insident óf oor- óf onderskat is. Die toepassing van die volledige WVKMRB-proses word ook in hierdie geval vir toekomstige gebruik aanbeveel, in ag genome die verskeie onsekerhede wat deur die loop van die studie ontwikkel het. Dit is onseker tot watter mate *E. coli* die infeksie-risko deur *Salmonella* kan bepaal, aangesien dit die moontlike voorkoms van verskeie patogene aandui. 'n Verdere tekortkoming is die gebrek aan geskikte risiko-limiete vir die studiegebied. Die finale aanbeveling is dus dat die volledige WVKMRB-proses wat tydens hierdie studie ontwikkel is, op 'n gereelde grondslag toegepas moet word om die risiko wat met die inname van water gepaardgaan, te bepaal, met dien verstande dat dit gebaseer is op 'n groter verskeidenheid van patogene en geassosieerde indikator mikro-organismes.

# SUMMARY

The aim of this study was to determine whether people exposed to the waters at various points of use in selected areas of the Mangaung Municipality, were indeed subjected to a probability of infection by the bacterial pathogen *Salmonella* as previous microbiological monitoring programmes in the area (based on the indicator organism *E. coli*) had suggested. Applying a Water-related Quantitative Microbial Risk Assessment (WRQMRA) process determined this where domestic and recreational water-use activities lead to the potential ingestion of polluted water.

The WRQMRA consisted of the Observed-adverse-effect-level (OAEL) and the Quantitative Microbial Risk Assessment (QMRA) approaches. The OAEL approach (OAELA) was based on the occurrence of *E. coli* to determine the *possible* risk of infection, while the QMRA predicted the *probable* risk of infection by *Salmonellae* numbers. Both these approaches were applied to untreated surface waters, used mainly for recreational purposes, as well as to container-stored water (untreated spring water and treated municipal supply water) used for domestic purposes.

The OAELA measured the *E. coli* numbers against various water quality guideline limits for the various water uses, while QMRA calculated the probable infection risk ($P_i$) by applying the four risk assessment steps of hazard assessment, exposure assessment, dose-response assessment and risk characterisation. The results of the OAELA and QMRA were compared to determine whether the use of the OAELA alone (applied by Environmental Health Practitioners) could, on a continual basis, reliably predict the risk of infection.

Ingestion was the only exposure route investigated, based on selected volumes for the various water-uses. For recreational use of the untreated surface waters in the Renoster Spruit Quarternary Catchment (RSQC), 100mℓ, 50mℓ and 10mℓ were used depending on the level of contact with the water. Modified ingestion volumes were extrapolated for various consumer age groups ranging from infants (1,318 mℓ) to the elderly (865 mℓ).

For the waters in the RSQC, the *Salmonellae* occurrences at three sampling sites (RS1, RS2 and BS) were combined and used to calculate the risk. A single-exposure, as well as seasonal (242-day) exposure to the mean, as well as 95$^{th}$ percentile risk was calculated. For untreated surface waters in the RSQC, both the OAEL and the QMRA approaches indicated a risk of infection to recreational users even for a single exposure event. However, the OAELA either over- or underestimated the risk of infection for singular exposure events and therefore could not predict a continual risk of infection. It is recommended that the full WRQMRA process be used in future.

A single-exposure, as well as annual (365-days) risk of infection was calculated for the container-stored (untreated spring and treated municipal supply) water applied for domestic purposes based on the mean and reasonable maximum (95$^{th}$ percentile) expected dose. Both the OAELA and the QMRA indicated a risk of infection after even a single exposure. However, the OAELA inconsistently over- or underestimated the risk on single sampling events, therefore not reliably indicating the risk of infection on its own. The full WRQMRA process is again recommended, considering several uncertainties that developed throughout the study. It was uncertain to what extent *E. coli* could indicate the risk of infection by *Salmonellae*, since it is an indicator of the potential presence of many other pathogens as well. Suitable risk limits lacked for the study area. It is recommended that the entire WRQMRA process developed for this study be applied more often in assessing risk posed by ingestion of water, but with provision for a wider range of pathogens and associated indicator micro-organisms.

# CONTENTS LIST


| | | |
|---|---|---|
| **List of Tables** | | **i** |
| **List of Figures** | | **iii** |
| **Chapter 1: INTRODUCTION** | | **1-13** |
| 1 | STUDY RATIONALE | 1 |
| 2 | THE WATER-RELATED QUANTITATIVE MICROBIAL RISK ASSESSMENT (WRQMRA) PROCESS | 4 |
| 2.1 | Hazard assessment – the first step in this WRQMRA process | 6 |
| 2.2 | The Observed-Adverse-Effect-Level (OAEL) Approach (OAELA) | 7 |
| 2.3 | Quantitative microbial risk assessment (QMRA) | 8 |
| 2.4 | Uncertainty analysis | 9 |
| 3 | THE SCOPE OF THIS STUDY | 10 |
| 4 | RESEARCH AIM | 12 |
| 4.1 | Objectives | 12 |
| **Chapter 2: APPLICATION OF THE WRQMRA PROCESS** | | **14-41** |
| 1 | THE STUDY AREA | 14 |
| 2 | HAZARD ASSESSMENT | 16 |
| 2.1 | The microbiological hazard and its indicator | 16 |
| 2.1.1 | *Salmonellae* as microbial hazard | 17 |
| 2.1.2 | *E. coli* as microbiological indicator organism group | 18 |
| 2.2 | Uncertainties associated with the Hazard Assessment step | 19 |
| 3 | EXPOSURE ASSESSMENT | 20 |
| 3.1 | Water-use areas and activities related to possible hazards | 21 |
| 3.1.1 | The Renoster Spruit Quarternary Catchment | 22 |
| 3.1.2 | Water stored in containers in households | 22 |
| 3.2 | The occurrence of *Salmonellae* and *E. coli* | 23 |
| 3.2.1 | Observed-adverse-effect-levels (OAEL's) based on *E. coli* numbers | 23 |
| 3.2.2 | The levels of *Salmonellae* in waters of the study area | 25 |
| 3.3 | Ingested water volumes | 25 |
| 3.3.1 | Intentional ingestion (based on daily intake per person) | 26 |
| 3.3.1.1 | Locally-sourced water | 27 |
| 3.3.1.2 | Daily water ingestion volumes | 27 |
| 3.3.1.2.1 | *Environmental factors* | 28 |
| 3.3.1.2.2 | *Social factors* | 28 |
| 3.3.1.2.3 | *Modified South African ingestion volumes for this study* | 30 |
| 3.3.2 | Involuntary ingestion | 30 |
| 3.4 | Dose | 32 |
| 3.5 | Uncertainties associated with the Exposure Assessment step | 33 |

| | | |
|---|---|---|
| 4 | DOSE-RESPONSE | 34 |
| 4.1 | The *β–Poisson* (distributed) dose-response model | 35 |
| 4.2 | Dose-response parameters | 36 |
| 4.3 | Uncertainties associated with the Dose-response Assessment step | 36 |
| 5 | RISK CHARACTERISATION | 37 |
| 5.1 | High end risk descriptors | 39 |
| 5.2 | Uncertainties associated with the Risk Characterisation step | 39 |
| 6 | *E. COLI* AS REASONABLE INDICATORS OF INFECTION RISK | 40 |
| 6.1 | Uncertainties associated with *E. coli* as indicator of infection risk | 40 |

**Chapter 3: THE RENOSTER SPRUIT QUARTERNARY CATCHMENT 42-76**

| | | |
|---|---|---|
| 1 | OCCURRENCE OF *E. COLI* AND *SALMONELLAE* IN WATERS OF THE RSQC | 43 |
| 1.1 | RS2 | 43 |
| 1.1.1 | *E. coli* and *Salmonellae* occurrence at RS2 in the Renoster Spruit | 44 |
| 1.2 | BS | 44 |
| 1.2.1 | *E. coli* and *Salmonellae* occurrence at BS in the Bloem Spruit | 46 |
| 1.3 | RS1 | 46 |
| 1.3.1 | *E. coli* and *Salmonellae* occurrence at RS1 in the Renoster Spruit | 47 |
| 1.4 | General *E. coli* and *Salmonellae* occurrence in the RSQC | 47 |
| 1.5 | Mean occurrence of *E. coli* and *Salmonellae* in the RSQC | 50 |
| 1.5.1 | Confidence intervals of the mean occurrence in the RSQC | 52 |
| 1.6 | Uncertainty analyses | 53 |
| 2 | THE OAEL APPROACH (OAELA) FOR *E. COLI* IN THE RSQC | 54 |
| 2.1 | Numbers of *E. coli* in the RSQC | 54 |
| 2.2 | Risk based on OAELA in the RSQC | 55 |
| 2.2.1 | Irrigation and intermediate body contact in the RSQC | 55 |
| 2.2.2 | Full-body contact with water in the RSQC | 56 |
| 2.2.3 | Drinking untreated water from the RSQC | 57 |
| 2.2.4 | Potential raw water extraction from the RSQC | 58 |
| 2.3 | Uncertainty analyses | 59 |
| 3 | THE QMRA APPROACH FOR *SALMONELLAE* IN THE RSQC | 60 |
| 3.1 | Numbers of *Salmonellae* in the RSQC | 60 |
| 3.2 | Ingestion volumes associated with activities observed in the RSQC | 61 |
| 3.3 | Dose | 62 |
| 3.4 | Probable risk of *Salmonellae* infection in the RSQC based on dose - response models and parameters | 63 |
| 3.5 | Characterising $P_i$ | 64 |
| 3.6 | Practicable application of the QMRA | 66 |
| 3.7 | Uncertainty analyses | 68 |
| 4 | PROBABILITY COMPARED TO POSSIBLE RISK OF INFECTION | 70 |
| 4.1 | Uncertainty analyses | 75 |

**Chapter 4: CONTAINER-ST** 77-105

| | | |
|---|---|---|
| 1 | OCCURRENCE OF *E. COLI* AND *SALMONELLAE* IN WATER STORED IN CONTAINERS | 78 |
| 1.1 | F1 | 79 |
| 1.1.1 | *E. coli* and *Salmonellae* occurrence at F1 | 79 |
| 1.2 | C1 | 80 |
| 1.2.1 | *E. coli* and *Salmonellae* occurrence at C1 | 80 |
| 1.3 | General *E. coli* and *Salmonellae* occurrence | 81 |
| 1.4 | Uncertainty analyses | 84 |
| 2 | THE OAEL APPROACH (OAELA) FOR *E. COLI* | 85 |
| 2.1 | Numbers of *E. coli* in untreated spring water stored in containers | 86 |
| 2.2 | Numbers of *E. coli* in treated municipal supply water stored in containers | 86 |
| 2.3 | Risk based on OAELA in the container-stored waters | 87 |
| 2.3.1 | Drinking untreated spring water from F1 | 88 |
| 2.3.2 | Drinking container-stored (municipal supply) water at C1 | 88 |
| 2.4 | Uncertainty analyses | 89 |
| 3 | THE QMRA APPROACH FOR *SALMONELLAE* | 90 |
| 3.1 | Numbers of *Salmonellae* in untreated spring water | 90 |
| 3.2 | Numbers of *Salmonellae* in treated municipal supply stored in containers | 91 |
| 3.3 | Intentional daily ingestion volumes from locally-sourced water | 92 |
| 3.4 | Dose | 92 |
| 3.5 | Probable risk of *Salmonellae* infection for container-stored water based on dose-response models and parameters | 94 |
| 3.6 | Characterising $P_i$ | 95 |
| 3.7 | Uncertainty analyses | 98 |
| 4 | PROBABILITY COMPARED TO POSSIBLE RISK OF INFECTION | 100 |
| 4.1 | Uncertainty analyses | 103 |

**Chapter 5: SUMARY, CONCLUSIONS AND RECOMMENDATIONS** 106-108

| | | |
|---|---|---|
| 1 | SUMMARY | 106 |
| 2 | CONCLUSIONS | 107 |
| 3 | RECOMMENDATIONS FOR FUTURE WORK | 107 |

**Chapter 6: REFERENCE LIST** 109-117

**APPENDICES:**
Appendix A: Bacterial Pathogen Analyses – *Salmonella spp.*
Appendix B: Bacterial Pathogen Analyses – *Escherichia coli*
Appendix C: Statistical analyses
Appendix D: Modified Daily Water Ingestion Volumes
Appendix E: RSQC Data
Appendix F: Container-stored Drinking Water Data

# LIST OF TABLES

## Chapter 2: THE APPLICATION OF THE WRQMRA PROCESS 14-41


| | | |
|---|---|---|
| Table 2.1 | A review of the Water-related Quantitative Microbial Risk Assessment process | 14 |
| Table 2.2 | The selected microbial hazards and summary of considerations for inclusion | 16 |
| Table 2.3 | Water-use areas, sampling points, activity-associated ingestion and potential users | 21 |
| Table 2.4 | Observed-adverse-effect-levels (OAEL's) for *E. coli* in water | 25 |
| Table 2.5 | Intentional daily ingestion volumes from locally sourced water for South Africans | 30 |
| Table 2.6 | Involuntary ingestion volumes based on intensity of water contact per event | 32 |
| Table 2.7 | The β-*Poisson* (distributed) dose-response model for calculating probability of infection | 35 |
| Table 2.8 | Dose-response parameters to calculate probability of infection for this study | 36 |
| Table 2.9 | Probability of infection ($P_i$) expressed in terms of US-EPA maximum annual drinking water ingestion risk limit for 10,000 of population | 38 |


## Chapter 3: THE RENOSTER SPRUIT QUARTERNARY CATCHMENT 42-76


| | | |
|---|---|---|
| Table 3.1 | *E. coli* and *Salmonellae* occurrence (per 100mℓ) at RS2 in the Renoster Spruit | 44 |
| Table 3.2 | *E. coli* and *Salmonellae* occurrence at BS in the Bloem Spruit /100mℓ | 46 |
| Table 3.3 | *E. coli* and *Salmonellae* occurrence (per 100mℓ) at RS1 in the Renoster Spruit | 47 |
| Table 3.4 | Specific occurrence of *E. coli* and *Salmonellae* in the RSQC | 48 |
| Table 3.5 | Mean *E. coli* and *Salmonellae* occurrence (per 100mℓ) in the RSQC | 51 |
| Table 3.6 | Observed water use activities and activity-related ingestion volumes in the RSQC | 61 |
| Table 3.7 | Expected dose of *Salmonellae* associated with exposure to waters of the RSQC at a range of ingested volumes | 62 |
| Table 3.8 | Probability of *Salmonellae* infection ($P_i$) based on a single exposure to the mean, maximum and minimum dose | 63 |
| Table 3.9 | Single and seasonal risk based on $P_i$ | 64 |
| Table 3.10 | Varying levels of $P_i$ based on intensity and frequency of water-use | 66 |

| | | |
|---|---|---|
| Table 4.1 | *E. coli* and *Salmonellae* occurrence (per 100mℓ) at F1 | 79 |
| Table 4.2 | *E. coli* and *Salmonellae* occurrence (per 100mℓ) at C1 | 80 |
| Table 4.3 | General occurrence of *E. coli* and *Salmonellae* at F1 and C1 /100mℓ | 81 |
| Table 4.4 | Expected dose of *Salmonellae* in treated and untreated waters stored in containers | 92 |
| Table 4.5 | Mean probability of *Salmonellae* infection ($P_i$) based on a single exposure event | 94 |
| Table 4.6 | Single and seasonal risk based on $P_i$ for untreated water stored in containers (F1) | 96 |
| Table 4.7 | Single and seasonal risk based on $P_i$ for treated water stored in containers (C1) | 97 |

# LIST OF FIGURES


**Chapter 1: INTRODUCTION** 1-13

Figure 1.1 Conceptual framework of the Water-related Quantitative Microbial Risk Assessment process 5

**Chapter 2: THE APPLICATION OF THE WRQMRA PROCESS** 14-41

Figure 2.1 Sampling sites within Middle-Modder River tertiary catchment 15

**Chapter 3: THE RENOSTER SPRUIT QUARTERNARY CATCHMENT** 42-76

Figure 3.1 Sampling sites within the Renoster Spruit Quarternary Catchment 43

Figure 3.2 Log *Salmonellae* occurrence in relation to log *E. coli* at specific sampling points in the Renoster Spruit Quarternary Catchment 49

Figure 3.3 Log occurrence of *E. coli* and *Salmonellae* at the three sampling sites of the Renoster Spruit Quarternary Catchment during the summer period 50

Figure 3.4 Log *E. coli* and *Salmonellae* occurrence in the Renoster Spruit Quarternary Catchment 51

Figure 3.5 Log mean *E. coli* occurrence (with upper and lower confidence interval [95% Ci] limits) in the Renoster Spruit Quarternary Catchment for the summer period 55

Figure 3.6 Comparing the occurrence of *E. coli* to OAEL's for *irrigation* and *intermediate body contact* in the Renoster Spruit Quarternary Catchment 56

Figure 3.7 Comparing the occurrence of *E. coli* to a NOAEL and a higher risk limit for *full-body immersion* in the Renoster Spruit Quarternary Catchment 57

Figure 3.8 Comparing the occurrence of *E. coli* to a NOAEL and a higher risk limit for drinking untreated water in the Renoster Spruit Quarternary Catchment 58

Figure 3.9 Comparing the occurrence of *E. coli* to the LOAEL and a higher risk limit for raw-water extraction for treatment from the Renoster Spruit Quarternary Catchment 59

Figure 3.10 Log mean Salmonellae occurrence (with upper and lower confidence interval [95% Ci] limits) in the Renoster Spruit Quarternary Catchment for the summer period 61

Figure 3.11 Expected dose and probability of Salmonellae infection ($P_i$) related to different water intake volumes 64

Figure 3.12 Single and seasonal exposure per 10,000 of the population related to different water intake volumes 65

Figure 3.13 *E. coli*, *Salmonellae* and associated risk (mean, as well as 95th percentile) measured in the Renoster Spruit Quarternary Catchment 72

**Chapter 4: CONTAINER-STORED WATER 77-105**

Figure 4.1 Sampling sites for untreated spring water (F1) and container-stored municipal supply water (C1) 78

Figure 4.2 *Salmonellae* occurrence in relation to *E. coli* at specific sampling sites 82

Figure 4.3 *E. coli* and *Salmonellae* occurrence in container-stored water at sampling site F1 83

Figure 4.4 *E. coli* and *Salmonellae* occurrence in container-stored water at sampling site C1 84

Figure 4.5 Log *E. coli* occurrence (with upper and lower [95% Ci] limits) at sampling site F1 for the sampling period 86

Figure 4.6 Log *E. coli* occurrence (with upper and lower [95% Ci] limits) at sampling site C1 for the sampling period 87

Figure 4.7 Comparing the occurrence of *E. coli* to a NOAEL and a LOAEL for drinking untreated water at sampling site F1 88

Figure 4.8 Comparing the occurrence of *E. coli* to a NOAEL and a LOAEL for drinking untreated water at sampling site C1 89

Figure 4.9 Log Salmonellae occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site F1 for the 2001/02 sampling period 90

Figure 4.10 Log Salmonellae occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site C1 for the 2001/02 sampling period 91

Figure 4.11 Expected dose and probability of infection ($P_i$) related to daily intentional water intake volumes 95

Figure 4.12 Single and annual $P_i$ per 10,000 of the population related to daily intentional water intake volumes 98

Figure 4.13 *E. coli*, *Salmonellae* and associated risk (with mean and 95th Percentile) measured at sampling site F1 101

Figure 4.14 *E. coli*, *Salmonellae* and associated risk (with mean and 95th Percentile) measured at sampling site C1 103

# INTRODUCTION

More than a billion (World Summit on Sustainable Development (WSSD, 2002)) of the world's population use water from any available source without (or with limited) treatment, for domestic- (drinking and body-washing), recreation- (cooling down in summer), as well as livelihood-purposes such as fish harvesting (Haas and Eisenberg, 2001; Jagals, 2000; Montaigne and Essick, 2002).

Outbreaks of waterborne disease such as diarrhoea occur when people somehow ingest polluted water from natural resources and other supplies. Infectious waterborne diseases occur continuously all over the world, and while it is not possible to identify infectious agents in all cases, microbiological agents are believed to dominate in such disease outbreaks (Haas and Eisenberg, 2001). It has become imperative to develop and apply methods to assess the risk posed by water that people use. In other words, it has become necessary to predict the risk of water (containing pathogenic microorganisms) causing diarrhoeal infections in humans when ingested (Haas et al., 1999).

This study focussed on the applicability of a water-related quantitative microbial risk assessment process for use by environmental and public health practitioners to predict the risk of infection posed by microorganisms present in environmentally polluted water in an urban and peri-urban environment.

## 1 STUDY RATIONALE

Ingestion of faecally polluted water is long recognised as a cause of diarrhoea (Department of Water Affairs and Forestry (DWAF), 2002; Genthe and Franck, 1999; Haas and Eisenberg, 2001; Jagals, 1997; Pretorius and De Villiers, 2002; Water Research Commission (WRC), 1998). Infectious diseases related to ingestion of faecally polluted water are caused by bacterial, viral and parasitic pathogens. These diseases include

gastroenteritis, hepatitis A, salmonellosis and cholera (Craun, 1986; Genthe and Rodda, 1999), to mention a few.

Analyses of the health-related microbiological quality of various water types in the Mangaung local municipal area (within the Middle-Modder River tertiary catchment - South-eastern Free State, South Africa), showed that resource waters were often heavily faecally polluted by urban discharges (Griesel and Jagals, 2002; Jagals, 1997; Jagals et al., 1995), posing a possible risk of infection to users. Municipal supply water stored in containers by households also posed risk of infection of intestinal disease (Bokako, 2000, Nala, 2002).

However, assumptions of these risks were based on the occurrence of microbiological indicator organisms instead of actual pathogens. Such an "indicator" approach does not provide a quantitative value for the microbiological water-borne health hazards that threaten water users (Du Preez et al., 2001), since it can merely indicate, by indicator organisms being present in certain numbers in water, the risk of infection by diarrhoea-causing pathogens potentially occurring in the same water if ingested by a person.

According to Genthe and Rodda (1999), indicator organism counts generally tend to underestimate water-related health risks. This may lead to underestimation of the *probability* that users, through ingestion of faecally polluted water, may be infected by diarrhoea-causing pathogens. By implication, this means that e.g., *Escherichia coli* (*E. coli*) can only indicate the *possible* risk of infection by pathogens.

A reasonable approach towards estimating (predicting) a more *probable* risk of infection in people using microbiologically contaminated waters, needs to be followed. According to Anderson (2001), Genthe and Rodda (1999), as well as Haas et al. (1999), a quantitative microbial risk assessment, based on actual pathogen numbers present (occurring) in water, is such an approach.

The two approaches of *possible* and *probable* risk of infection used in this study are based on the "Weight-of-Evidence Class" classification used by Risk*Assistant™ (1995) for cancer research, which classifies carcinogens according to their potential of causing cancer in

humans. A carcinogen with a hi causing cancer, is referred to as a *probable* carcinogen, while a *possible* carcinogen are less likely to cause cancer.

The same approach was followed for the pathogen and indictors in this study. This study combined these two approaches in a water-related quantitative microbial risk assessment (WRQMRA) process based on the occurrence of an actual pathogen group (*probable risk*) (and its association with an indicator group (*possible risk*)) in the waters of the Middle-Modder River tertiary catchment.

The rationale for this was to understand more comprehensively the extent of association between the two approaches. Environmental health practitioners (EHP's) generally apply data on the measured occurrences of, for instance, *E. coli* in water (by comparing this to various levels of documented *risks of infection*) to do some form of hazard assessment or even assessment and expression of health risk (Gray, 2001; Koren, 1991).

While the key focus of this study was to assess *probable* risk of infection by means of a risk assessment process, the study was also done to form some impression of association, if any, between occurrence of indicator organisms and the occurrence of at least one pathogen in the waters of the study area. Another consideration was that the *probable* infection risk assessment of the WRQMRA process may be complex and expensive. It was therefore decided to compare the results of the assessment to the *possible* infection risk indicated by *E. coli* occurrences in the same samples (a less complex and cheaper process). The rationale for this was that this part of the WRQMRA process could be most useful to communicate the *possible* infection risk to water users on regular bases.

Comparing the application value of the *possible* versus the *probable* risk of infection in the study area waters also provided information environmental health practitioners could use. It could assist them in deciding what approach would be best practice to follow, to form a reasonable impression of the water quality and the danger it posed, thereby enabling them to optimally direct resources to manage water pollution in the high-risk areas described in earlier and ongoing studies in the Mangaung local municipal area.

## 2 THE WATER-RELATED MICROBIAL RISK ASSESSMENT (WRQMRA) PROCESS

The WRQMRA process for this study was developed because the need was identified to assess both the applicability of the possible (indicator), as well as the probable risk (pathogen) of infection posed by ingesting waters in the study area.

The process had four main components (Figure 1.1):

1 Hazard assessment whereby the water uses in the study area, as well as pollution sources and levels of occurrence of risk agents, and / or their indicators, were assessed.

2 Assessment of the indicator organism occurrence that can at best, indicate the possible risk. This component is the Observed-Adverse-Effect-Level (OAEL) approach (OAELA) based on indicator occurrence numbers and expressing possible infection risk from related water quality guidelines.

3 The third component dealt with the probable risk and is referred to as the Quantitative Microbial Risk Assessment (QMRA) approach, based on doses of an actual pathogen ingested, assessing responses to such doses, and characterising probable risk of infection.

4 The final step in the process was to test for associations between the two risk components and to validate the one process as a supportive- or replacement process for the other based on the *uncertainties* expected to arise from the application of the whole process.

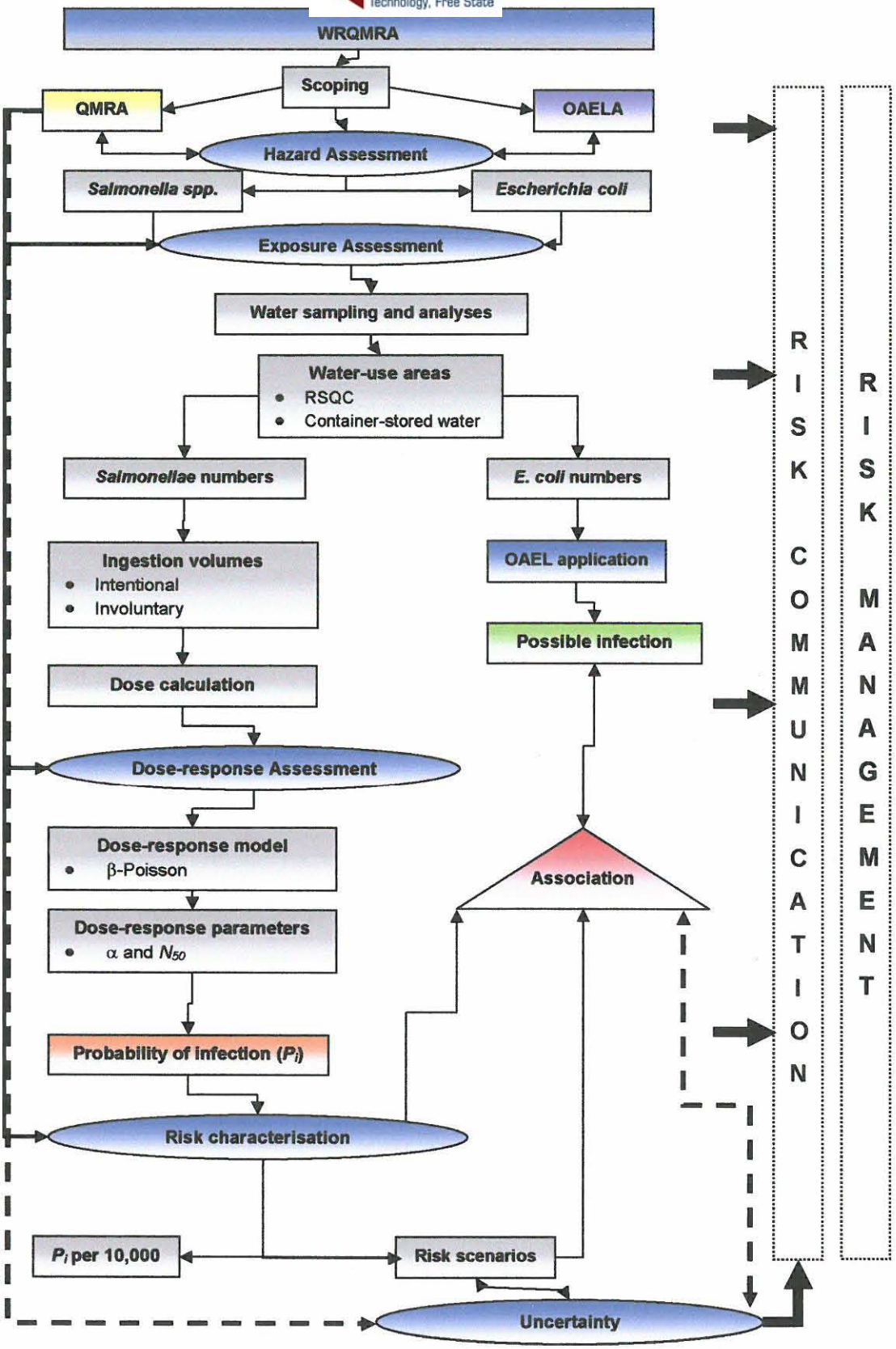



**Figure 1.1:** Conceptual framework of the Water-related Quantitative Microbial Risk Assessment process

## 2.1 Hazard assessment – this WRQMRA process

This first step in this study was a lesser process than the typical comprehensive hazard assessment step in health risk assessment, which usually consists of identifying (unknown) disease-causing (hazardous) agents, usually in a particular geographical area, and determining particular adverse health effects (harm) associated with the specific agent in exposed individuals (Covello and Merkhofer, 1993; Haas et al., 1999).

The WHO (1998) defines a health hazard as a set of circumstances that could cause harm. These harmful effects may range from asymptomatic infections, to morbidity (disease or illness) and even to mortality (death) (Haas et al., 1999). This would imply that water in a particular assessed area would have to be screened for (all) possible microbiological pathogens, whether known at that stage or not, and their health effects in the exposed community investigated (International Life Science Institute (ILSI)-Risk Science Institute (RSI) (ILSI-RSI), 1996).

Because of limited resources, this is not possible in hazard studies. For instance, microbiological indicators - instead of actual pathogens - often form the bases for environmental health-related hazard assessments. An example is the Hazard Analysis and Critical Control Point (HACCP) system in the food industry where assessment of possible microbiological food hazards is based on the presence of *E. coli* (Food Safety and Inspection Service (FSIS), 1996). While this approach would not provide clear answers on the infectivity of contaminated foods, the FSIS (1996) deems that where processes are under control for *E. coli,* the potential presence of enteric pathogens will be minimised. This implies that the presence of *E. coli* would at least indicate critical points in the preparation process where hazards are likely to occur.

In the water industry, the application of HACCP in treatment processes is still in its infancy. Early application of the HACCP system still tends to use microbial indicators, or at the very least, use indicators in conjunction with pathogen or toxicity analyses (Haas and Eisenberg, 2001). Determining indicator occurrence in treatment systems provides more affordable

and quicker tools for water quali ntify areas of potential hazard. Well-established indicator-guidelines exist that can indicate at what level problems (hazards) could be expected to occur at critical points in the treatment system.

The basic principle of the abovementioned is that high levels of *E. coli* in carrier media such as food and water indicate possible health risks to process managers, even though health effects are not confirmed. Chapter 2 discusses Hazard Assessment (HA) in more detail.

## 2.2 The Observed-Adverse-Effect-Level (OAEL) Approach (OAELA)

Health-related water quality guidelines such as those proposed by the Department of Water Affairs and Forestry (DWAF, 1996a and b), the National Microbial Monitoring Programme (NMMP) (DWAF, 2002), and the World Health Organisation (WHO, 1996) use risk-like expressions (e.g. "chance for infection" or "risk of microbial infection") to describe various levels of risk of gastrointestinal infections associated with pathogens in microbiologically polluted water. These risk-like statements would normally be based on the occurrence of microbiological indicators such as *Escherichia coli* in waters potentially to be ingested. Such ingestion (exposure) is usually not measured but is associated with water–uses such as drinking, recreation and other domestic purposes such as food preparation and body washing (Jagals, 2000; Theron, 2000).

This approach assumes that if *E. coli* occurs in water at or below certain guideline risk-levels, associated pathogenic microorganism levels would be such that no or gross-parameter levels such as low, medium and / or high adverse health effects (gastrointestinal disease e.g. diarrhoea) would be observed in groups that might ingest the water (OAEL approach). For example: Volume (1) on Domestic Water Use of the South African Water Quality Guidelines (SAWQG) (DWAF, 1996a) assumes short-term continuous exposure to water containing ≥ 10 faecal coliforms per 100 mℓ would pose a slight risk of microbial infection. Griesel and Jagals (2002), Kindzierzki and Jackson (1998), Kolluru et al. (1996), as well as the WHO (1998) all infer that this approach is the same as applying an OAELA.

The use of indicator organisms is justified because of historic difficulties (e.g., high cost,

danger to assessment personnel hniques) to detect pathogens in water (Genthe and Rodda, 1999; Haas et al., 1999). Indicator microorganisms are generally not pathogenic but merely *indicate* the potential for pathogenic microorganisms to be present in water. This implies that the absence of indicators in resource water, e.g., a river or dam, well or a borehole, cannot guarantee the actual absence of pathogens in these same sources. Nor can the *presence* of indicator organisms (traditionally absent from apparently "cleaner" water, e.g., well, borehole or piped water supply) be an indication that pathogens are present in such water (Genthe and Kfir, 1995; Haas et al., 1999; Payment and Franco, 1993; Payment et al., 2000).

Indicator levels tested in this study will be measured against the OAEL's obtained from various guideline documents, of which the results could at best, be described as *possible* risks. The expression "*possible risk of infection*" will therefore be used throughout this study when referring to the risk assessment value of indicator microorganisms.

## 2.3 Quantitative microbial risk assessment (QMRA)

Conversely, the Quantitative Microbial Risk Assessment (QMRA) part of the WRQMRA process, based on actual pathogen presence (*Salmonellae*), is a particularly useful tool to predict the *probable risk of infection* (Genthe and Rodda, 1999; Haas and Eisenberg, 2001; ILSI-RSI, 1996).

Risk assessment may be defined as the qualitative or quantitative characterisation (description) and estimation (evaluation) of potentially adverse health effects (morbidity, mortality), associated with exposures to hazardous substances (risk agents), processes, action or events (Covello and Merkhofer; 1993; Rose, 1997). The risk agents exposed to may be biological, chemical, or physical and may be released from point or non-point risk sources (e.g., broken sewer pipes, drain blockages) to the air, soil, water, and food (environmental carrier media) (Kolluru et al., 1996).

However, a risk assessment (RA) process may be very comprehensive (and often very expensive) and may vary substantially according to the design of a particular study. A

typical comprehensive RA mode veral steps such as identifying the hazard (e.g. toxic substance), assessing the magnitude of exposure of the affected individual or population to such substance, the response of the persons to the dose of the substance and the characterisation of the risk (e.g. health outcomes such as morbidity or mortality) (Covello and Merkhofer, 1993; Haas and Eisenberg, 2001; Haas et al, 1999; Kolluru et al., 1996; Skivington, 1997).

QMRA is the application of RA principles to estimate the consequences from an actual exposure to infectious *microorganisms*. The use of QMRA in the water environment requires direct measurement of pathogens to determine whether water that human populations are exposed to, through various water-use activities and ingestion related to such activities, may be the source of exposure to microbial infections in such populations (Haas et al., 1999).

Application of a comprehensive QMRA process would have been ideal to provide environmental health practitioners with essential information regarding the actual health risk in the exposed target population (Haas and Eisenberg, 2001). For a water-related QMRA process, combinations of steps such as hazard assessment, exposure assessment (ending in a dose), dose-response assessment, and risk characterisation have been applied in several studies (Du Preez et al., 2001; Genthe and Rodda, 1999; Haas and Eisenberg, 2001; Haas et al. 1999; ILSI-RSI, 1996).

However, because of a lack of resources, this study focussed on, and applied a limited number of the elements usually contained in a QMRA process. These are discussed in Chapter 2.

## 2.4 Uncertainty analysis

Since information such as dose-response relationships, exposure magnitudes, etc. is almost invariably incomplete, it is also necessary to ascertain the potential error (uncertainty) involved in risk assessment, especially when associations between the results of a QMRA and an OAELA are investigated.

Uncertainty analysis is a major c ical risk assessment process. It is an evaluation of all assumptions, and all components of the assessment process used, that may create uncertainties in the expression of risk experienced throughout the course of a study such as this one (Covello and Merkhofer, 1993; Genthe and Rodda, 1999). Haas et al. (1999) defines uncertainty as the "factors of imprecision and inaccuracy that limit the ability to exactly quantify the risk". For this study, according to the recommendations of Craun (1993), the effect of uncertainty on the confidence interval associated with the characterised risk was evaluated, and expressed qualitatively, as well as, wherever possible, quantitatively (e.g., upper and lower limits of the 95% confidence interval). Detailed discussions of uncertainty analyses will follow in each subsequent chapter in this dissertation.

## 3 THE SCOPE OF THIS STUDY

Shortages of resources (e.g., finances, time, data, and fieldworkers) impeded implementing comprehensive water-related risk models for this study. It was therefore important that a scoping task be undertaken prior to performing the WRQMRA, since scoping would determine the extent to which the process would be implemented. Scoping for this study was done according to the recommendations by Haas et al. (1999), and directed the following principal issues:

- Risk agent selection: The study was not to include a comprehensive hazard assessment. *Salmonellae* were selected as the pathogen and *E. coli* as the indicator microorganism based on knowledge about existing hazards in water.
- The analysis methodology to determine risk agent occurrence in various water types was investigated and selected according to availability of technology, as well as the supporting institutional infrastructure.
- Ingestion (of water) was the only exposure route considered.
- Actual exposure was not investigated, but was rather based on approximated

intentional and involuntary  related to water-use activities.

- Water ingestion volumes applied were derived from literature on previous studies done elsewhere, based on circumstances similar to the study area.
- The actual number of people exposed through the various water uses was not addressed. Hypothetical risk scenarios were used instead.
- The mean, maximum (95$^{th}$ percentile) and minimum expected dose, as well as single exposure doses were to be applied to determine the probability of infection.
- Mean and 95$^{th}$ percentile risk, as well as single probability of single exposure events were to be extrapolated to risks over longer exposure periods (eight-month 2001 / 02 summer season (242-days) and annual (365-days) risk).
- Probability of infection ($P_i$) was to be estimated and measured against the United States-Environmental Protection Agency (US-EPA) (1994) guideline for an acceptable risk limit for consumption of drinking water of 1 in 10,000 of the population per annum (Regli et al., 1991). Scenario populations were also to be suggested to illustrate application of the WRQMRA process.
- Ingestion related to recreational, as well as domestic water-use activities of mostly developing country scenarios, were to dominate the investigation.
- Indicator microorganism occurrence was to be evaluated against observed-adverse-effect-levels (OAEL's) from guideline documents, and the *possible* risk of infection determined.
- The possible risk of infection based on indicator microorganisms was to be compared to the predicted *probability* of infection based on the occurrence of actual pathogens.
- The sole use of indicator microorganisms, to predict the possible human health risks associated with the various water uses, was then discussed based on the uncertainties involved in both instances.

# 4 RESEARCH AIM

The aim of this study was to determine whether people exposed to the waters at various points of use in the study area, were indeed subjected to a probability of infection (by *Salmonellae*) as previous microbiological indicator-based monitoring programmes had suggested. This was done by applying the WRQMRA process in selected areas where domestic and recreational water-use activities could lead to the potential ingestion of polluted water. The study also assessed whether *E. coli* alone, could be used to predict the risk of infection reliably.

## 4.1 Objectives

These objectives summarises the scope of the study (from which these were derived):

- To select a pathogen group associated with adverse human health (hazards identified).
- To select an indicator microorganism group capable of indicating the potential presence of the selected pathogen in water.
- To investigate the association between the pathogen and indicator occurrence in water.
- To compare the indicator occurrence to guideline levels (OAEL's) and determine possible infection risk.
- To assess potential exposures of users by calculating probable dose based on ingestion of the microbial hazard (related to water-use activities e.g. direct ingestion, vegetable gardening, water contact through fishing and recreation), based on the following:
  - the numbers of pathogenic microorganisms in the various water sources.
  - approximated daily ingestion volumes of the contaminated water.

- To determine whether a pa activity (e.g., domestic, recreational etc.) posed a probable risk of infection based on dose-response information.
- To compare the probable infection risk assessment (pathogen-based) (QMRA approach) with the possible infection risk based on microbiological indicator occurrence (OAELA).
- To describe the uncertainties surrounding each step of the assessment process.

# APPLICATION OF THE WRQMRA PROCESS

Application of this process resulted in an assessment of the probability of infection in the population. Table 2.1 reviews the WRQMRA process. Sections 1 – 5 of this chapter discusses the WRQMRA process, with each of its assessment steps, the expected outcomes for each step, as well as uncertainties surrounding each step of the process. Section 6 discusses the use of *E. coli* as an indicator of the possible risk of infection, as well as the associated uncertainties.

**Table 2.1:** A review of the Water-related Quantitative Microbial Risk Assessment process

| WRQMRA steps | Outcomes expected at each step | |
|---|---|---|
| **Hazard assessment** | • Bases for selecting *Salmonellae* (pathogen) and *E. coli* (microbiological indicator)<br>• Uncertainties | |
| **QMRA Step 1: Exposure assessment** (based on ingestion) | • Assessing the occurrence of *Salmonellae* (providing data for QMRA)<br>• Pathogen counts / 100mℓ<br>• Hypothetical ingestion volumes<br>  ○ Intentional ingestion (modified)<br>  ○ Unintentional ingestion<br>• Dose (n x mℓ)<br>  ○ Uncertainties surrounding this step | • Assessing the occurrence of *E. coli*<br>• *E. coli* counts (per 100mℓ) compared to OAEL's for various water uses<br>• Uncertainties |
| **QMRA Step 2: Dose-response assessment** | • Modelling of probability of infection<br>• Hypothetical parameters (from literature on studies elsewhere)<br>• Uncertainties surrounding this step | |
| **QMRA Step 3: Characterising probability of infection** | • Infection risk estimate<br>  ○ Expressing $P_i$ per 10 000 of the population<br>  ○ Expressing $P_i$ per single and seasonal exposure<br>• Risk scenarios<br>• Possible vs. probable risk estimate<br>  ○ Assessing and discussing association<br>• Uncertainties surrounding this step | |

## 1 THE STUDY AREA

The study was conducted in the Mangaung Local Municipal area in the South-eastern Free State, which includes the urban and peri-urban developments of Bloemfontein, Botshabelo, and Thaba N'chu (Figure 2.1).

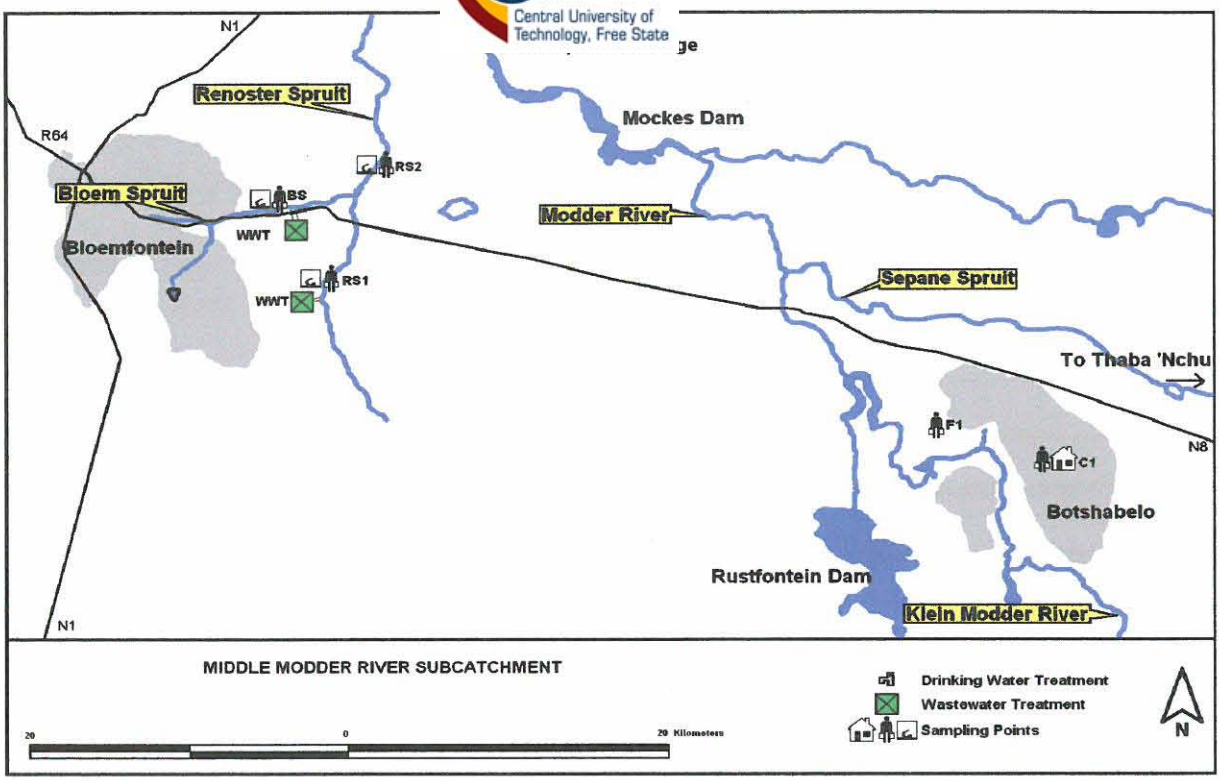



**Figure 2.1:** Sampling sites within Middle-Modder River tertiary catchment

In catchment management terms, this sub-catchment forms part of the Middle-Modder River tertiary catchment. Substantial temperature variations and erratic flash rainfall patterns, predominantly in late summer, characterise the climate (Jagals, 2000; Pretorius and De Villiers, 1999).

A variety of point or non-point faecal pollution sources contributes to the microbiological contamination of surface water within the study area (Griesel and Jagals, 2002; Jagals, 1997). Such contamination may cause the health–related microbiological quality of the water in the area to deteriorate, which may lead to a risk of infection by pathogenic microorganisms and possible gastrointestinal diseases. As part of the hazard assessment step, the following faecal pollution sources in the area are listed:

- Treated effluent from two wastewater treatment facilities (one in the Bloem Spruit, and one in the Renoster Spruit).
- Diffuse urban surface run-off from surrounding urban areas (Bloemfontein city, Botshabelo and Thaba N'chu).

- Livestock farming activities ıam from selected sampling sites.

The WRQMRA process was applied in two selected water-use areas (details in Figure 2.1 as well as Sections 3.1), which determined the water-use activities and associated ingestion volumes.

# 2 HAZARD ASSESSMENT

For reasons discussed in Section 2.1 of Chapter 1, a comprehensive hazard assessment as described by Covello and Merkhofer (1993) and by Haas et al. (1999) was not considered. Hazard assessment (HA) in this study was based on the following function:

- HA discussed the bases of selecting, as microbiological hazards, the known pathogen group (*Salmonellae*), and its associated indicator, *E. coli* for this study.

## 2.1 The microbiological hazard agent and its indicator

Table 2.2 summarises the considerations for selecting the *Salmonellae* group as the hazard agent and *E. coli* as the associated indicator.

**Table 2.2:** The selected microbial hazards and summary of considerations for inclusion

| Microorganisms | References | Considerations |
|---|---|---|
| **Indicator:** *E. coli* | • DWAF, 1996a, 1996b<br>• Griesel, 2001<br>• Haas et al., 1999<br>• Jagals, 2000 | • *E. coli* indicates faecal pollution of water and therefore indicates the potential presence of bacterial pathogens such as *Salmonellae* |
| **Pathogen:** *Salmonella spp.* | • Liversage, 2001<br>• Potgieter, 2002<br>• CDC 2000<br>• Du Preez, et al., 2001<br>• Genthe and Rodda, 1999<br>• Giannella, 2001<br>• Haas and Eisenberg, 2001<br>• Haas et al., 1999<br>• KidsHealth, 2000<br>• Oxoid Corporation, 1990<br>• Oxoid Corporation, 1997<br>• Polo et al., 1998<br>• Standard Methods, 1998<br>• WHO, 1996<br>• Yates, 1999 | • Personal communication<br>• Literature review<br>• Availability of dose-response models and parameters<br>• Availability of equipped laboratories for analyses<br>• Safety and simplicity of analyses methods<br>• Analysis costs |

The bacterium *Salmonella spp.* form part of the family *Enterobacteriaciae*. Of the 2,300 serotypes that occur, roughly 200 cause human disease (Centres for Disease Control and Prevention (CDC), 2000; Standard Methods, 1998). *Salmonellae* cause infection of the gastrointestinal system in humans, a self-limiting illness called salmonellosis (Benenson, 1995). Clinically the disease is limited to the stomach and intestines (diarrhoea, abdominal cramps and fever), but the organisms can spread to other parts of the body, such as blood and bone and cause enteric fevers, including typhoid fever (Giannella, 2001; KidsHealth, 2000).

Although 90% of all recorded salmonellosis cases have been food-borne, water-borne outbreaks have been documented (Haas et al., 1999). The pathogen is usually found in wastewater (in numbers of 23 – 80,000 per litre) and in agricultural run-off (Polo et al., 1998; Yates, 1999). Presence of *Salmonellae* is more common in summer than winter (CDC, 2000). Salmonellosis is common in developing countries with poor sanitation facilities and personal hygiene practices. Over the past 10 years travellers from the United States of America (USA) to developing countries (e.g. Asia, Africa and Latin America) have been especially at risk (CDC, 2000).

Typhoid fever, on the other hand, is a life-threatening febrile disease, caused by the species *Salmonella typhi,* with about 400 cases occurring in the US each year. This disease is common in developing countries such as South Africa, affecting approximately 12.5 million people yearly (CDC, 2000; Giannella, 2001; KidsHealth, 2000).

Personal communications with the management of the Botshabelo Hospital (Liversage, 2001), and personnel working in the laboratory of the Local Municipality of Mangaung (Potgieter, 2002), indicated that infections by *Salmonellae* bacteria were prominent amongst people in the area hospitalised with gastro-intestinal disease. Genthe and Rodda (1999) also identified the need to apply a risk assessment model to pathogenic bacteria such as *Salmonellae* in South Africa.

For this step of the study, literature d that the microbiological pathogen group *Salmonellae* is hazardous to human health since these could cause the onset of a diarrhoeal syndrome in people. *Salmonellae* have been implicated in waterborne disease outbreaks worldwide (Haas et al., 1999; Standard Methods, 1998; WHO, 1996). It has the potential to be hazardous, even in waters containing low numbers of these microorganisms (Du Preez, et al., 2001; Giannella, 2001; Polo et al., 1998).

Another critical consideration for selecting *Salmonellae* was availability of dose-response models and parameters. Haas and Eisenberg (2001) and Haas et al. (1999) suggested parameters to model the probability of infection (Section 4.2).

Further considerations were costs of analyses and suitably equipped laboratories to detect these pathogens in water samples accurately and safely. Methodologies to detect its occurrence in water are well documented (Du Preez, et al., 2001; Oxoid Corporation, 1990, 1997; Standard Methods, 1998). The detection methods applied were reported to be relatively safe, simple, and affordable.

### 2.1.2 *E. coli* as the microbiological indicator organism group

*Escherichia coli* is a common inhabitant of all warm-blooded animals, but only certain strains are capable of producing illness in humans (Benenson, 1995). At least four types of *E. coli* are pathogenic to humans, which includes enterotoxigenic, enterohemorrhagic (*E. coli* 0157:H7), enteropathogenic and enteroinvasive. *E. coli* is historically used to indicate faecal pollution of water and therefore indicates the potential for pathogenic microorganisms to occur in water (Griesel, 2001; Haas et al., 1999).

The health-related microbiological indicator organisms selected for this study had to indicate the potential presence of the selected pathogens. This study therefore used *E. coli* to indicate the potential presence of pathogenic bacteria such as *Salmonellae* in water (DWAF, 1996a and 1996b; Jagals, 2000).

To add value to the interpretations of the large databases built by the previous studies done in the area (Bokako, 2000; Griesel, 2001; Griesel and Jagals, 2002; Jagals, 1994, Jagals et

al. 1997; Nala, 2001; Pretorius, 19 De Villiers, 2002), this study also investigated associations between *E. coli* and *Salmonellae* in the waters of the study area.

*E. coli* formed the bases of previous water quality studies in the area (Bokako, 2000; Griesel, 2001; Nala, 2002; Pretorius and De Villiers, 1999). These studies included several "risk of infection" statements, based on water quality guidelines such as the South African Water Quality Guidelines (SAWQG) (DWAF, 1996a & 1996b), the Quality of Domestic Water Supplies: Assessment Guide (WRC, 1998), and the World Health Organisation (WHO) Guidelines for drinking water quality (1996). These risk values, however, might not effectively indicate actual risk (Genthe and Rodda, 1999; Haas et al., 1999).

## 2.2 Uncertainties associated with the Hazard Assessment step

- The bacterial hazard agents (*Salmonellae*) were selected on the strength of personal communication with epidemiologists and health services managers in the area (Liversage, 2001, Potgieter, 2002), instead of being confirmed by means of analytical or epidemiological processes. This could be deemed a rather subjective selection process.
- It is not known whether the species (e.g., *Salmonella spp.*) chosen for the specific study area were the most suitable. While the selected microorganisms may be the primary cause of waterborne disease outbreaks elsewhere, other disease-causing microorganisms may be more prominent in the study area.
- Costs and capacity limited the inclusion of more organism types e.g. viruses and protozoan parasites. The selected hazards might therefore not be sufficiently representative of the microbial risk in the study area.
- The pathogen investigation limited to *Salmonellae* caused uncertainty as to whether the possible risk of infection suggested by previous studies were realistic as the waters could not be examined for the many pathogens potentially present in the water.
- Previous studies reporting high indicator microorganism concentrations based on

guideline values, indicated a fection by pathogens (Bokako, 2000, Nala, 2002). It is uncertain to what extent indicator microorganisms (such as *E. coli*) can predict or indicate risk of infection.

? It is uncertain to what extent *E. coli* occurrence is associated with (co-occurs) with *Salmonellae* occurrence and therefore, to what extent *E. coli* would indicate the actual presence of this pathogen group in water.

## 3 EXPOSURE ASSESSMENT

This step within the risk assessment process typically includes the following components (Covello and Merkhofer, 1993; Du Preez et al., 2001; Genthe and Rodda, 1999; Haas and Eisenberg, 2001; Haas et al., 1999; Kolluru et al., 1996; Medema, 2002):

- descriptions of the intensity, frequency and duration of exposure through various environmental media e.g., air, food, soil, and water;
- routes of exposure;
- the numbers and characteristics of the exposed population;
- consideration of conditions that might affect the consequences (e.g., age, immune status).

A shortage of resources in general did not allow for comprehensive application of all of these for this study. For example, determining the immune status of a person involves an epidemiological process that may be very expensive and time consuming.

This study applied a simplified three-step exposure assessment:

- Determination of the *Salmonellae* numbers (quantified from the organism *occurrence* detailed in Section 3.2) in water resources in the selected water use areas (analytical study phase).
- Approximations of *ingestion volumes* of water, associated with the water-use activities in the selected water-use areas, were derived and modified from Bourne et al. (1987; 1992), DWAF (1996a and b), Genthe and Rodda (1999), Haas et al. (1999), Medema

et al. (2001) and Roseberry 992) Section 3.3). Ingestion of water (intentional or not) was considered as the only route of exposure for this study.

- Calculation of the *dose* (the number of pathogens per volume unit of water ingested) (Section 3.4).

## 3.1 Water-use areas and activities related to exposure to possible hazards

This study focussed on the risk posed by water (containing *Salmonellae*) ingested by people while participating in certain activities in the water-use areas (Table 2.3).

**Table 2.3:** Water-use areas, sampling points, activity-associated ingestion and potential users

| Water-use area | Sampling point (Figure 2.1) | Ingestion during: | Potential user(s) |
|---|---|---|---|
| **Ingestion of untreated surface water receiving faecally polluted surface water (Chapter 3)** | **Renoster Spruit Quarternary Catchment:**<br>• **RS2:** Downstream final sampling site<br>• **RS1:** Draining agricultural areas, informal settlements and treated effluent to RS2<br>• **BS:** Draining formal residential and industrial areas as well as treated effluent to RS2 | **Domestic use:**<br>• Drinking and cooking<br>• Body-washing<br>• Vegetable gardening<br>**Recreational use** (unintentional ingestion):<br>• Full–body immersion<br>• Intermediate contact through i.e. fish harvesting<br>• Golf course irrigation | • People from informal peri-urban settlements<br>• Destitute people<br>• Golfers |
| **Ingestion of treated supply water and untreated spring water stored in containers at home (Chapter 4)** | **Informal residential area:**<br>• **C1:** Storing and using treated municipal supply at home<br>• **F1:** Storing and using untreated water from spring at home | **Domestic use:**<br>• Drinking and cooking<br>• Body-washing<br>• Vegetable gardening | **C1:** Established community in dense low socio-economic urban development<br>**F1:** Peri-urban community with low socio-economic status |

The total study area supports a diversity of human settlements based on economy and population (rural, peri-urban, developed urban areas, etc.), as well as extensive agricultural activities (Jagals, 2000). It therefore made sense to select two areas where the manner in which the populations used water (water-use activity), formed the basis of the water-related exposure (activity-related ingestion) in each particular population.

Chapters 3–4 provide detailed discussions of these areas, the respective sampling points, and the risk posed by the water used, to communities in the proximity of each point.

### 3.1.1 The Renoster Spruit Qu ıent (Chapter 3)

People are often compelled, through socio-economic circumstances such as poverty, to use untreated water (Sobsey et al., 2002) from sources proven (Griesel, 2001) to be faecally polluted. The Renoster Spruit Quarternary Catchment (RSQC) is such an area.

The RSQC forms part of the larger Middle-Modder River tertiary catchment. Within the RSQC, three sampling sites (RS1, RS2 and BS) were selected to represent the health-related microbiological quality of water within this area.

Water sampled at the point RS2 (Figure 2.1) represented the worst possible microbiological quality of water draining from Bloemfontein city, as reflected by points BS and RS1 (upstream). These discharges included treated effluent from two wastewater treatment facilities (WWTF's).

The respective *E. coli* and *Salmonellae* data, measured at each separate site, were combined and their means used in the WRQMRA process as representative of the risk posed by the health-related microbiological quality of the surface water of the whole of the RSQC.

### 3.1.2 Water stored in containers in households (Chapter 4)

Large groups of people in settlements within the area, whether supplied with treated water sources (e.g. communal taps <200m from home and constantly available) or not, collected, stored and used water from other available sources, with or without limited treatment, for domestic purposes. Most of these families lived in small permanent houses constructed of bricks, while other impoverished families lived in simple temporary shacks. Previous studies on the microbiological quality of domestically-stored water in the area (Bokako, 2000; Jagals et al., 1997; Nala, 2002) indicated that the quality of such water deteriorated during storage and handling and concluded that health risks can be expected.

Sampling site F1 (Figure 2.1) was an *unprotected spring* (accessible from all surface ends by humans and animals without any form of pollution- or other control) from which people collected untreated water for domestic purposes, and stored it in containers at home. Community-owned livestock also drank water directly at this site. Sampling site C1 (Figure

2.1), was at a household where tr ed from a communal tap, was stored in containers at home.

## 3.2 The occurrence of *Salmonellae* and *E. coli*

To determine the occurrence of *Salmonellae* and *E. coli,* the quality of water was monitored at various locations (details in Section 3.1 and Chapters 3-4) throughout the sub-catchment during the summer months from September 2001 to April 2002.

A summer-season sampling period made sense as the majority of water-use activities investigated for this study took place in warm (summer) weather conditions. The summer rainfall contributed to sporadic elevated levels of faecal pollution of the waters, which played a role in widening the upper limit for the 95% confidence interval of the mean risk in surface waters by widening the occurrence range for *Salmonellae* and *E. coli.*

Water samples were taken at least monthly and the numbers of microorganisms occurring in the various waters determined. *Salmonellae* were detected by means of a three-tube most-probable-number (MPN) technique (Appendix A), and *E. coli* by a selective medium and membrane filtration technique (Appendix B) (Oxoid Corporation, 1990; Standard Methods, 1998).

### 3.2.1 Observed adverse-effect-levels (OAEL's) based on *E. coli* numbers

From a health risk perspective, if indicators occur in water at or below / above levels in guidelines e.g. such as those stipulated in the South African Water Quality Guidelines (DWAF, 1996a and b), associated pathogenic microorganism levels should be such that no- or particular adverse effects (such as e.g. low, medium or high) are observed in the exposed group.

Griesel and Jagals (2002), Kindzierzki and Jackson (1998), Kolluru et al. (1996), as well as the WHO (1998), infer that these indicator occurrence levels are observed-adverse-effect-levels (OAEL's).

Guidelines apply combinations of (-adverse-effect-levels) e.g. NOAEL's (no-observed-adverse-effect-levels) and LOAEL's (lowest-observed-adverse-effect-levels) etc. (Blumenthal et al., 1999; Haas et al., 1999; WHO, 1998) as minimum risk levels for some water uses (e.g., the National Microbial Monitoring Programme (NMMP) (DWAF, 2002).

An example of a NOAEL is the WHO (1996) *E. coli* guideline for domestic use of water, stating that *E. coli* should not be detectable in any 100mℓ (and therefore no health effects are expected). A LOAEL on the other hand is the "low potential health risk" posed by faecal coliforms (≤ 130 *E. coli* or faecal coliforms per 100mℓ) as proposed by the South African Water Quality Guidelines for Recreational water use (DWAF, 1996b).

Another complicating factor with regards to the use of *E. coli* is that their guideline values have generally been relegated in favour of faecal coliforms because of the historically complex multi-step methodologies required to detect *E. coli* (Griesel and Jagals, 2002), as well as the time-consuming analyses methodologies.

*E. coli* is a more reliable indicator of faecal pollution than faecal coliforms and therefore of possible risk. For this study *E. coli* was detected with a simple single step procedure (Appendix B) described by Jagals et al. (2001) as a reliable method.

It is not uncommon to find *E. coli* being used in the place of faecal coliforms since several current guidelines often use risk-like expressions based on the occurrence of both in water. This comparative approach is, for instance, followed by the guidelines in the NMMP (DWAF, 2002) which use "faecal coliforms or *E. coli*" as a single phrase.

The numbers of *E. coli* detected in the waters of the Middle-Modder River sub-catchment during this study were compared to guideline OAEL's for the various water-uses to indicate the possible microbiological infection risks to consumers. For this study, Table 2.4 was compiled for typical OAEL's for *E. coli* (from DWAF, 2002, Jagals, 2000, and WHO, 1998).

**Table 2.4:** Observed-adverse-effe r *E. coli* in water

| Water use category | | *E. coli* (faecal coliforms ) per 100 mℓ |
|---|---|---|
| **Domestic** (Intentional as well as unintentional ingestion risk) | Drinking | ≤1 [1] Insignificant chance for infection (LOAEL)<br>0 [6] Not detectable (NOAEL) |
| | Food Preparation | ≤1 [1] Insignificant chance for infection (LOAEL) |
| | Bathing (contact risk) | ≤10[1] Insignificant effects (LOAEL) |
| | Laundry (contact risk) | ≤10[1] Insignificant effects (LOAEL) |
| **Agri- and horticulture** (Unintentional ingestion risk) | Vegetable and salad crops eaten uncooked, sports fields, public parks | ≤1,000[3] No potential risk (NOAEL)<br>≤1,000[4] Low potential health risk (LOAEL) |
| **Livelihood Fishing** (Unintentional ingestion risk) | Harvesting impoundments with baited fishing lines | ≤1,000[3] (LOAEL)<br>Increased contact with water during activities |
| **Recreation** (Unintentional ingestion risk) | Full-body immersion (unlimited ingestion inferred) | ≤130[2][5] (LOAEL)<br>Risk of gastrointestinal effects expected |
| | Intermediate contact (limited ingestion inferred) | ≤1,000[3] [5] (LOAEL)<br>Gastrointestinal effects are indicated |
| **Raw water extraction** | Intended for treatment for ultimate potable purposes | ≤2,000[4] (LOAEL)<br>Low potential health risk indicated |

1 Quality of Domestic Water Supplies - Volume 1: Assessment Guide (WRC, 1998)

2 South African Water Quality Guidelines - Vol. 2: Recreational Use (DWAF; 1996b)

3 Guidelines for wastewater reuse in agriculture and aquaculture: recommended revisions based on new research evidence (Blumenthal et al., 1999)

4 National Microbial Monitoring Programme (DWAF, 2002)

5 Draft Guidelines for Safe Recreational-water Environments: Coastal and Fresh-waters (WHO, 1998)

6 Drinking Water Quality Guidelines (WHO, 1996)

### 3.2.2 The levels of *Salmonellae* in waters of the study area (QMRA approach)

Pathogenic microorganisms such as the *Salmonellae* should not be detectable in a litre of water ingested by people (Venter et al, 1996; WHO, 1996). Should these occur, a hazard is certainly constituted, although the risk posed by this hazard might not yet be characterised, since it is still to be established whether people actually ingested the water (exposure).

The results in Chapters 3 to 4 showed that *Salmonellae* occurred in considerable numbers in some of the water areas. This is a hazard, and by implication a health risk.

## 3.3 Ingested water volumes

In order for pathogenic microorganisms in water to cause diarrhoeal disease in humans (to be hazardous), a dose of bacteria, or their toxins, should first be ingested with the water (Ward and Akin, 1984). The major exposure mechanism (route) chosen for this study was

water ingestion, either intentionally ır for survival) or involuntary (during recreation and other domestic-related purposes such as body-washing), since this is the route evaluated by available risk models (Genthe and Rodda, 1999).

Infection is a pre-requisite for disease and can occur in any population, depending on the dose (Rose, 1997). To determine dose, the number of pathogenic microorganisms per water volume unit (e.g., 20 *Salmonellae* /100mℓ), as well as the volume of water ingested (e.g., 100mℓ, 50mℓ or even 1,000mℓ), must be established.

Section 3.2 had already dealt with the occurrence (numbers) of *Salmonellae*. This section deals with *ingested water volumes*. An area-specific survey on the volume of water that people ingest per day fell outside the scope of this study. Actual exposure of people (the actual volumes of water ingested) in the study area was therefore not measured. In other words, actual water use was not formally investigated e.g., duration and frequency of a particular (domestic or recreational) water-use event or activity. Instead, documented ingestion volumes from local and international studies were used.

Ingestion was divided into two categories i.e., intentional and involuntary, since assumptions about, and observations of water-use activities in the study area, formed the bases for describing the water ingestion that was applied during the exposure assessments in this study (Discussed in Chapters 3 – 4).

### 3.3.1 Intentional ingestion (based on daily intake per person)

Intentional ingestion implies daily water intake by people to sustain life. Water-related health-endpoints, such as diarrhoea, are the common result of ingesting microbiologically contaminated water. The associated risk is therefore largely dependant on the daily intake of water. It was therefore critical, for the purposes of this study, to clearly define what daily intake of water meant.

Ershow and Cantor (1989), as well as Roseberry and Burmaster (1992) (corroborated by several other authors) defined the intake of water (liquid) as:

- tap water (taken and consumed from a source or supply near the household);

- beverages made of water (wa[illegible]pplies e.g., commercial beverages such as tea, coffee, etc.) and
- water used to prepare food (e.g., stew).

Of these abovementioned categories, "tap water" is the category that, in this study, is the closest match for locally sourced water ingested through drinking, or otherwise ingested with locally prepared foods and beverages.

#### 3.3.1.1 Locally-sourced water

Not all the waters sourced are from taps. For the purposes of this study, the phrase "locally-sourced water" was used to indicate the daily intake of water from local sources. This intake category included:

- direct water ingestion;
- water ingested through beverages such as coffee, tea, and soup and
- water used for preparation of foods such as stews, porridge, etc. that are prepared from substantial water volumes.

Water bound in food, as well as in commercial beverages such as carbonated cool drinks, is assumed generally free from harmful microbiological agents, and was therefore not considered in the risk model.

For this study, total daily intake volumes of locally-sourced water were considered.

#### 3.3.1.2 Daily water ingestion volumes

Various authors use an adult water ingestion rate of 2,000 millilitres per head per day (mℓ/hd.d), based on human feeding studies (Genthe and Rodda, 1999, Haas and Eisenberg, 2001; Haas et al., 1999). This parameter appears popular since it includes drinking and food-related consumption and therefore does not underestimate the daily risk of infection. Several countries and a number of international organisations have adopted this value (Regli et al., 1991; Rose, 1997; Rose and Gerba, 1991).

For this study, however, it was uncertain whether 2,000 mℓ/hd.d would be applicable since

the target households consisted n hed black families living in sub-standard housing with varying levels of access to water supply. To determine dose based on daily water intake, locally-sourced water ingestion volumes that reflect South African conditions needed to be compiled for this study.

Unfortunately, South African studies on the daily volume of water people would actually ingest, appear limited. Bourne et al. conducted two studies (1987; 1992) that surveyed the daily consumption of water by individuals in black, coloured, and white households in the Greater Cape Town area. Theron (2000) previously reported on total daily water volumes collected and stored per black household in the Botshabelo area.

It made sense to use data from these studies to compile ingestion volumes for this study, should certain generic social and environmental factors appear to be reasonably similar e.g. age, gender, housing type, and access to local water sources.

#### 3.3.1.2.1 *Environmental factors*

Bourne et al. (1992) (for the Cape Town area) as well as Theron (2000) (previously reporting on total daily water volumes stored per household for this Mangaung study area), reported that housing types (e.g., house, shack, hostel) had no significant influence on the daily water collection and ingestion volumes of the black people in these particular areas. Neither did the type of access to water supply (inside tap, outside tap and communal tap), not even when it involved walking considerable distances with heavy filled containers (Bokako, 2000; Bourne et al., 1992; Theron, 2000). The effect of housing- and water access types appeared quite similar for the Cape Town and Mangaung areas.

#### 3.3.1.2.2 *Social factors*

Bourne et al. (1992) stated that the ingestion-data for black people in the Cape Town study were questionable for various reasons (e.g., underreporting because consumption of tap water appeared to be below the level of consciousness). They also concluded that an adult daily intake of 2,000mℓ/hd.d, as traditionally used by the WHO and US-EPA, is a useful approximation for local (South African) use in the light of the Cape Town studies.

Since it was decided to use modifi gestion volumes, this study did not follow this approach, but instead applied the ingestion volumes from the Cape Town studies after certain adaptations (Appendix D).

Although Bourne et al. (1992) questioned the validity of the ingestion-data for black people in the Cape Town study, these were nevertheless applied in combination with the data for white, and coloured people since the daily intake volumes did not differ significantly between the groups. The data was tested for significant differences with the Kruskal-Wallis Analysis of Variance on Ranks test (Appendix C) using SigmaStat Version 2.03 (1997). The differences in the mean values among the groups were not great enough to exclude the possibility that the difference is due to random sampling variability. There were no statistically significant differences ($P = 0.496$).

However, these means for the groups could not simply be used, since the Bourne et al. (1992) study did not report data for daily ingestion by black infants. Conversely, this particular study included intake data for elderly black persons, which the Bourne et al. (1987) study on whites and coloureds did not.

Since diarrhoea notoriously affects infants, children, and the elderly (sensitive sub-populations) more severely than adults (Bourne and Coetzee, 1996; Haas et al., 1999), it made sense, from a risk assessment perspective, to categorise the exposed populations in the study area into age groups. It was therefore necessary to extrapolate values for the infants and the elderly by using all three race-based data sets.

Another complicating factor was that the Bourne et al. studies (1987; 1992) used quite a number of age groups. It was reasoned that fewer, rather than more, permutations of age- versus risk, would simplify the reporting in this study, while still adding value to the risk assessment process. An international age-grouping (Roseberry and Burmaster, 1992), that contained fewer categories than the South African studies, was therefore applied. Fitting regression curves to the data enabled extrapolation to derive hypothetical daily intakes for the age groupings infants up to the elderly (Appendix D).

Table 2.5 shows the daily water consumption per age group (e.g. children between ages 1 and 11, etc.). Mean ingestion data from both the Bourne et al. (1987; 1992) studies were grouped into five age categories per race group, per gender, and a polynomial cubic regression curve fitted (R = 0.99) to the means (Appendix D). The dependent variable y (daily intakes of locally-sourced water) was calculated from $y = y0 + a*x + b*x^2 + c*x^3$. For the last data point beyond the 65 ≤ age parameter, a generic life expectancy endpoint 75 years was entered to extrapolate ingestion volumes for the elderly. For the first data point, infants (birth to 1 years of age), the modified ingestion volume was not acceptable, since it appeared to be too low.

Instead, for the infants (birth to 1 year of age) the mean ingestion volume for infants (boys and girls) (150 mℓ/kg/day), was calculated from the (yellow) Weight / Age (kg) chart (Robinson et al., 1982; WHO, 1994), applied by the Department of Health at municipal clinics all over South Africa. In order to determine the mean ingestion volume, the mean bodyweight (8.9 kg) used to represent the weight for infants (boys and girls) was read from the Department of Health chart and multiplied by 150 mℓ/kg/day. For this study, it was assumed that babies younger than six months were still breastfed and water consumption was not considered.

**Table 2.5:** Intentional daily ingestion volumes from locally sourced water for South Africans

| Locally-sourced water intake | Infants 0 ≤ age < 1# | Children 1 ≤ age < 11 | Adolescents 11 ≤ age < 20 | Adults 20 ≤ age < 65 | Elderly 65 ≤ age |
|---|---|---|---|---|---|
| Volumes = mℓ/hd.d | 1,318 | 630 | 773 | 952 | 865 |

# Mean ingestion compiled from 50th percentile weight/age (kg) clinic chart multiplied with mean ingestion of 150 mℓ/kg/day

### 3.3.2 Involuntary ingestion

People generally do not enter resource water with the intention to drink from it. Nevertheless, certain water-use activities (e.g., recreational activities such as swimming) lead to unintentional or accidental (involuntary) ingestion, with persons often not even aware

of the risks involved in areas wheinated. Involuntary water intake rates depend primarily on the intensity of contact with the water.

The South African Water Quality Guidelines (DWAF, 1996b), the WHO (1998), and several other authors identified types of contact and associated intake volumes, through use of water environments:

- Extensive direct (full body) contact involves full body immersion and a significant risk of swallowing water e.g., swimming and bathing.
- Meaningful (intermediate) direct contact appears to involve a lower risk of swallowing water e.g., skiing, wind surfing, body-washing.
- Limited contact – e.g., wading, boating, rowing, fishing.
- No contact – where enjoyment is of aesthetic beauty of the water environment.

Specific surveys on the volume of water that people ingest per event, during actual environmental water-use activities, fell outside the scope of this study. Ingestion volumes from local and international studies were used instead (Genthe and Rodda, 1999; Medema et al., 2001; WHO, 1998).

The ingestion volumes used to represent the use of water for doing laundry and fishing was not found in literature, but instead based on observations and associated level of contact with the water.

Table 2.6 summarises involuntary ingestion volumes based on the level of water contact used for this study. Involuntary ingestion of water during events was established at 100 mℓ, 50 mℓ, and 10 mℓ depending on the intensity of contact during specific activities.

Chapters 3-4 discuss various hypothetical events (scenarios), which will illustrate the application of the values in Table 2.6.

**Table 2.6:** Involuntary ingestion volu ntensity of water contact per event

| Contact intensity | Full-body immersion | Intermediate | Other |
|---|---|---|---|
| **Intake volumes** | **≤ 100 mℓ swallowed per event**<br>• DWAF, 1996b<br>• WHO, 1998<br>• Genthe and Rodda, 1999<br>• Haas et al., 1999 | **50 mℓ swallowed per event**<br>• Medema et al., 2001 | **10 mℓ accidental gulping**<br>• Genthe and Rodda, 1999<br>• Medema et al., 2001 |
| **Events** | • Social swimming activities<br>• Sporting swimming e.g. triathlon<br>• Children playing in water<br>• Body-washing in resource water | • Repeated immersion during skiing, wind-surfing, canoeing | • Laundry<br>• Fishing<br>• Ingestion related to irrigation in agri- and horticulture (e.g. golf courses) |

## 3.4 Dose

This process is the end-product of the exposure assessment step (Covello and Merkhofer, 1993; Genthe and Rodda, 1999; Medema, 2002).

Dose depends on the volume of water ingested, as well as the concentration of microbial agents per volume unit. The greater the dose exposed to, the greater the chance of adverse health consequences. Dose therefore describes the intensity of exposure (Teunis et al., 1997).

To determine dose during a comprehensive risk assessment study would normally comprise two parts, i.e., an experimental study and a social (epidemiological) study (Covello and Merkhofer, 1993). The experimental work would be to assess the pathogen numbers per volume unit of e.g., the selected water type. The social part would establish the daily volume of water ingested per person in the exposed user group (Payment, 1997; Payment et al., 1997; Teunis et al., 1997).

The social study component (e.g., usually an observational study comprising of observation sheets and questionnaires), to determine the daily volumes of (how much) water ingested by the various users was instead concluded from human consumption surveys, as well as guidelines on recreational water safety (Section 3.3).

For this study, dose was calculate *ıellae* numbers detected in the respective water samples potentially ingested in the hypothetical volumes of water shown in Tables 2.5 and 2.6.

## 3.5 Uncertainties associated with the Exposure Assessment step

- During the analytical phase, factors such as applying the optimum analytical method, correct application of technique, human error during counting or calculation of dose etc., although considered throughout, could have played a role in unrealistically estimating (over- or underestimating) the risk.
- The culturability of microorganisms differs. It might be that the microorganisms chosen for this study were present in the water, but very difficult to culture. This would have given biased results.
- Variation in consumption volumes. The volumes depicted by the studies done on the daily consumption of water by different age and race groups applied in this study, may not be representative of the study population. The daily ingestion volume of an individual depends on various factors such as availability of water for consumption, individual preference, type of exposure, and source application.
- Conditions that may affect the consequences such as age, gender, etc. (Bourne et al., 1987; Bourne et al., 1992) were not investigated but only extrapolated from associated data. This limited risk statements on the numbers of infants and the elderly (sensitive sub-populations), actually exposed to the risk of contracting diarrhoeal disease.
- Cultural and racial differences may also influence not only the volume of water consumed daily, but also their activities associated with exposure e.g., recreational water contact.
- Uncertainty exists around seasonal variations. Summer and winter data was not compared to determine the quantitative effect this had on the probability of infection.

As one of the water uses in[ ] l activities such as swimming, and swimming in rivers and streams mostly occur within summer months, this could have had a substantial impact on the end result because winter-risk would inevitably be over-estimated.

- Rainfall, although noted, was not considered in terms of the influence it might have had on the occurrence and survival of the selected microorganisms.
- Die-off and sedimentation of microorganisms was not investigated. These could have had an impact on the concentrations, since certain microorganisms die-off or sediment much quicker than others, preventing their inclusion during sampling. This would prevent a true indication of the actual occurrence of microorganisms in the waters.
- The numbers and characteristics (immunity) of people exposed were not investigated. This would prevent an explicit risk statement towards how many people are actually exposed to a health risk.
- It is uncertain to what extent use of only the 50$^{th}$ percentile mean body weight for infants (for all race groups) over- or underestimated the risk of infection.
- Hypothetical ingestion volumes from the Cape Town area could be inapplicable in the Mangaung study area.
- Modified ingestion volumes may not be entirely applicable because of the limited reference base for studies of this nature in South Africa.

## 4 DOSE-RESPONSE

This component of the QMRA process predicted the possible response of the exposed individuals to the calculated *Salmonellae* dose and therefore required dose-response information. This study applied already-existing dose-response models reported in international literature (Covello and Merkhofer, 1993; Genthe and Rodda, 1999; Haas and Eisenberg, 2001; Haas et al., 1999; Rose and Gerba, 1991; Rose et al., 1991). These

models can calculate daily, monthl  infection based on parameters available for particular disease-causing organisms.

## 4.1 The *β-Poisson* (distributed) dose-response model

The objective of dose-response assessments is to develop a relationship $P_I$ (probability of infection) between the level of microbial exposure (through dose), and being infected, to such an extent that the occurrence of an adverse consequence (diarrhoea) is likely. To characterise the relationship of $P_I$, dose-response assessments particularly consider infection (Haas et al., 1999). Haas (1983), Rose and Gerba (1991), as well as Rose et al. (1991), evaluated several dose-response models. Haas and Eisenberg (2001) recommended the *β-Poisson* (distributed) model, which best described the probability of infection $P_I$ for *Salmonella spp.* (Table 2.7). This model was applied for the study.

**Table 2.7:** The *β-Poisson* (distributed) dose-response model for calculating probability of infection

| **Daily risk of infection (Formula 1)** (Haas and Eisenberg, 2001) $P_I = 1 - \left[1 + \frac{d}{N_{50}}(2^{\frac{1}{\alpha}} - 1)\right]^{-\alpha}$ | | **Average risk over longer exposure periods (Formula 2)** (Genthe and Rodda, 1999) $P_x = 1 - \left[1 - P_I(d)\right]^x$ | |
|---|---|---|---|
| $P_I$ | = probability (risk) of infection | $P_x$ | = probability (risk) of one or more infections over period x |
| $d$ | = dose or exposure (number of organisms) | $x$ | = number of days of exposure |
| $\alpha$ | = parameter characterised by dose-response relationship | $P_I(d)$ | = daily risk, using geometric mean for $d$ |
| $N_{50}$ | = median infectious dose | $d$ | = geometric mean organism numbers ingested daily over period x |

Du Preez et al. (2001) and Genthe and Rodda (1999), have used the *β-Poisson* (distributed) dose-response model to calculate the probability of bacterial infection that users of South African waters might get after a single exposure (Formula 1) as well as repeated exposures (Formula 2). This model best accounts for variation in the pathogenicity of organisms such as enteric bacteria (*Salmonellae*), as well as the sensitivity in hosts (Haas et al., 1999). This model applies a distribution of values ($\alpha$ and $N_{50}$ parameters), which are available for *Salmonellae* (Section 4.2, Table 2.8 below).

## 4.2 Dose-response paramet

The dose, as well as information on health end-points, is required to derive dose-response parameters. Information on health endpoints is obtained through epidemiological studies, from which parameters are derived and applied in models such as the *β-Poisson* (distributed) model to predict the probable risk of *Salmonellae* infections in other areas at different stages based on dose. Several authors (Haas and Eisenberg, 2001; Haas et al., 1999; Rose and Gerba, 1991) reported dose-response parameters for *Salmonellae* infections, of which the latest parameters, suggested by Haas et al. (1999), were used.

**Table 2.8:** Dose-response parameters to calculate probability of infection for this study

| Parameters | *Salmonellae* |
|---|---|
| $\alpha$ | 0.3126 |
| $N_{50}$ | 23,600 |

## 4.3 Uncertainties associated with Dose-response assessment step

- Actual response after exposure (ingestion) was not measured in the study. The characterisation of the relationship of $P_I$ was based on modelling – the actual infection rate was not measured by e.g., an increase in serum antibody increases or other inflammatory reactivity in human subjects.
- Reported dose-response information is based on human experiments with healthy adult volunteers. From a public health perspective, this age group is not the most important group. The risk of infection to newborns, children, elderly persons and other risk groups such as immuno-compromised persons, may be underestimated (Rose, 1997; Teunis et al., 1997).
- Subgroups in the human population with different susceptibilities to disease were not investigated. Not all individuals who become infected (i.e., replication or growth of virus in host) will develop clinical illness. The very young and old, and the immuno-compromised, (AIDS patients, transplant recipients, and those on chemotherapy)

(Rose, 1997) are at the greate itcome of most illnesses (Gerba and Rose, 1993).

- Dose-response information used is based on well-characterised laboratory strains of pathogens. The intrinsic infectivity (and potential to cause morbidity) may differ between lab maintained cultures and indigenous pathogens (Haas et al., 1999).
- The intrinsic uncertainty of the dose-response models applied. Dose-response information is based on static models that calculated the probability of individual infection or disease as a result of a single exposure event. These models does not address properties which are unique to infectious disease transmission, such as secondary (person-person) disease transmission.
- A great deal of uncertainty is introduced through the transmission pathway of the selected pathogens. The role that water plays in the transmission of the selected enteric pathogens was not investigated.
- The dose-response information (models and parameters) applied in this study are from international literature and may not reflect the true probability of infection in South Africa, or more specifically in the study area.

## 5 RISK CHARACTERISATION

The whole idea of health risk characterisation is to predict and assess future impacts of a risk agent on an exposed population. Risk characterisation is a description of the probability of infection (or mortality) under the conditions of exposure described, based on the dose-response (model) parameters used. However, for the purposes of this study, only the relationship of $P_i$ (probable risk of infection) was characterised. Haas and Eisenberg (2001) justify focussing only on infection as an initial step in a complete risk assessment process, by arguing that limiting infections may introduce elements of conservatism from a public health perspective. This would provide applicable information to public health practitioners to

understand the risks involved and ate and manage the possible risk of the exposed community.

To characterise $P_I$ as a product of the QMRA process, this step integrated the information from the exposure (to the dose) and dose-response assessment steps into a risk statement. $P_I$ was quantitatively estimated by means of models, calculated from international dose-response parameters, hypothetical ingestion volumes, and locally measured occurrence (numbers) of the risk agents.

A risk statement, based on recommendations of the US-EPA (1994) for consumption of drinking-water, was then applied. The latter specifies that human populations should not be subjected to the risk of infection by enteric disease greater than $10^{-4}$ or 1 case / 10,000 persons / year (Regli et al., 1991). This annual risk limit is equal to a $P_I$ of 0.0001 or a probability of infection of 0.01% (Table 2.9).

The risk expression involved the risk of infection after a single exposure to doses calculated from the mean *Salmonellae* occurrences, as well as probabilities of infection after exposure throughout the 2001/02 summer season. Use of exposure scenarios involved the risks to populations smaller than 10,000 for which risk statements were also made.

**Table 2.9:** Probability of infection ($P_I$) expressed in terms of US-EPA maximum annual drinking water ingestion risk limit for 10,000 of population

| *PI* | % Probability | Risk of infection per 10,000 population |
|---|---|---|
| 1 | 100% | $\frac{10,000}{10,000}$ |
| 0.1 | 10% | $\frac{1,000}{10,000}$ |
| 0.01 | 1% | $\frac{100}{10,000}$ |
| 0.001 | 0.1% | $\frac{10}{10,000}$ |
| 0.0001 | 0.01% | $\frac{1}{10,000}$ |
| 0.00001 | 0.001% | $\frac{1}{100,000}$ |

The mean, as well as 95$^{th}$ percentile (Section 5.1 below) *Salmonellae* occurrence was used to characterise $P_I$. Medema et al. (2001) expresses the probability of infection as percentage

risk, while Haas et al. (1999) uses population of 10,000 expression. This study characterised risk in terms of $P_i$, and expressed it as a fraction of 10,000 of the population, as well as percentage (%) $P_i$.

## 5.1 High-end risk descriptors

The use of the 95$^{th}$ percentile for this study is based on the US-EPA Guidance for Risk Characterisation (1995), which applies high-end risk descriptors as estimates of the individual risk for those persons at the upper end of the risk distribution. Risk*Assistant™ (1995) refers to this high-end risk descriptors as the "Reasonable Maximum Exposure" (RME). The aim of RME is to communicate estimates of the individual in the population with the highest exposure, but not beyond the true distribution. For example, an elderly person with the highest daily water intake would not be reasonable, since this is not likely to occur.

When large populations are assessed, a large number of individuals may be included within this "high end" (e.g., above the 90$^{th}$ or 95$^{th}$ percentile). The use of the 95$^{th}$ percentile therefore makes provision for sensitivity within the population (or subgroups) as certain people within populations may be more highly exposed than others (e.g., the immuno-compromised, elderly, infants, etc. depending on the situation). However, sensitive subgroups within the population was not investigated for this study. Use of the 95$^{th}$ percentile for this study aimed to indicate what level of risk of infection are posed by the water types investigated (e.g., exceeding the 95$^{th}$ percentile or not).

## 5.2 Uncertainties associated with the Risk Characterisation step

A number of sources of uncertainty in this estimation of risk of *Salmonellae* associated with the three water-use areas have already been discussed in Sections 2.2, 3.5, and 4.3. The following additional uncertainties were identified:

? Not all people infected will become clinically ill. Risk estimates based on $P_I$ as an endpoint may overestimate the number of clinical illness cases, and thus constitutes a worst case scenario. For example, it does not mean that if 10,000 people became

infected all 10,000 will becor ice have diarrhoea, thus overestimating the disease risk.

- Substitution of less-than-values or non-detect results with values such as the detection limit or half the detection limit will have resulted in overestimation of the risk estimates.
- A large source of uncertainty is the method used to calculate seasonal and yearly risk. Seasonal and yearly risk was calculated from the product of daily risks and an overall geometric mean was used instead of a distribution of monthly geometric means, which according to Haas (personal communication) cited in Genthe and Rodda (1999), results in risk estimates up to 2.5 orders of magnitude greater than those obtained from a distribution of monthly geometric means.
- The existence of other epidemiological states of the disease process may also affect risk estimates. For example, the post-infection status accounts for individuals previously exposed to the pathogen. This conferred immunity may take on different forms from long-term and complete protection to short-term and partial protection.

## 6 *E. COLI* AS REASONABLE INDICATORS OF INFECTION RISK

The possible risk of infection indicated by the *E. coli* OAEL's (guideline limits per 100 mℓ) was correlated with the probable risk of infection (*Pi*) indicated by the *Salmonellae* numbers detected per 100mℓ. Conclusions were then made about the extent to which the *E. coli* occurrences in the study-waters would indicate comparable risk of infection from *Salmonellae* to $P_I$.

### 6.1 Uncertainties associated with *E. coli* as indicator of infection risk

- *E. coli* may die-off or settle at different rates than *Salmonellae*, which may influence the level of detection.
- *E. coli* may over– or under estimate the risk of *Salmonellae* infection as it is not known to what extent *E. coli* occurrence in water is associated with *Salmonellae* occurrence.

- Certain pollution events suc ay have not have a similar effect on the occurrence of *E. coli* as on that of the *Salmonellae*, thereby influencing their respective mean risk descriptions.

- While this study only focused on the co-occurrence of *Salmonellae* and *E. coli*, many other pathogen groups this study did not test for, may be indicated by *E. coli* in the same waters. The associated estimation of *Salmonellae* by the occurrence of *E. coli* may therefore be an underestimation of the total pathogenicity of the test waters since all other waterborne pathogens (e.g., viruses, protozoan parasites) and their associated indicators, were not investigated.

# INFECTION RISK ASSOCIATED WITH INGESTION OF FAECALLY POLLUTED SURFACE WATER

This particular study area fell within the Renoster Spruit quarternary catchment (RSQC), which is part of the Middle-Modder River tertiary catchment (Figure 2.1, Chapter 2). The urban and peri-urban areas of the Mangaung local authority lies within the RSQC (Figure 3.1).

Water-uses for this area varied considerably, ranging from remote possibilities of water ingestion, such as golfers playing on courses irrigated with faecally-polluted waters, to destitute people using the untreated water for domestic purposes such as laundry, body washing and, on occasion, even for drinking and food preparation (Griesel and Jagals, 2002). Faecal pollution of the surface waters of the RSQC generally constituted poorly treated effluents from wastewater treatment facilities and other faecally polluted discharges from the urban areas generated by blocked sewers and polluted surface run-off, especially after rainfall events (Griesel and Jagals, 2002; Jagals, 1997; Pretorius, 2002).

This chapter is divided into four sections. Section 1 discusses the characteristics of three sampling points RS1, RS2 and BS. These sites did not represent specific direct stream-water user-activities at each point, but were rather selected for their proximities to the various activities associated with the various water uses. These points therefore represented the surface water quality of the whole of the RSQC. This section also discusses the occurrence of *E. coli* and *Salmonellae* in the surface waters of the RSQC jointly (mean general occurrence), as well as separately (mean occurrence per sample point). The latter establishes which area in the RSQC contributes the highest microorganism numbers and related risk to people most likely to apply water from that point for particular use.

Section 2 discusses the evaluation of *E. coli* against OAEL's (OAEL approach) to determine

the possible risk of infection, whi s the probable risk of infection based on the numbers of *Salmonellae* found in the RSQC (QMRA approach). Section 4 discusses the possibility that *E. coli,* as indicator organism, could accurately predict the probable risk of infection posed by the pathogen, *Salmonellae,* to human users of waters in the RSQC. Appendix C describes the use of the means, confidence intervals as well as the 95th percentile as high-end risk descriptor used in the various tables and figures of this chapter.

# 1 OCCURRENCE OF *E. COLI* AND *SALMONELLAE* IN WATERS OF THE RSQC

The Renoster Spruit is the main drain of the RSQC. Of its several minor tributaries, the Bloem Spruit is the most prominent (Figure 3.1). The sampling regime comprised the three sampling sites RS1, RS2 and BS, representing various water uses in their proximity.

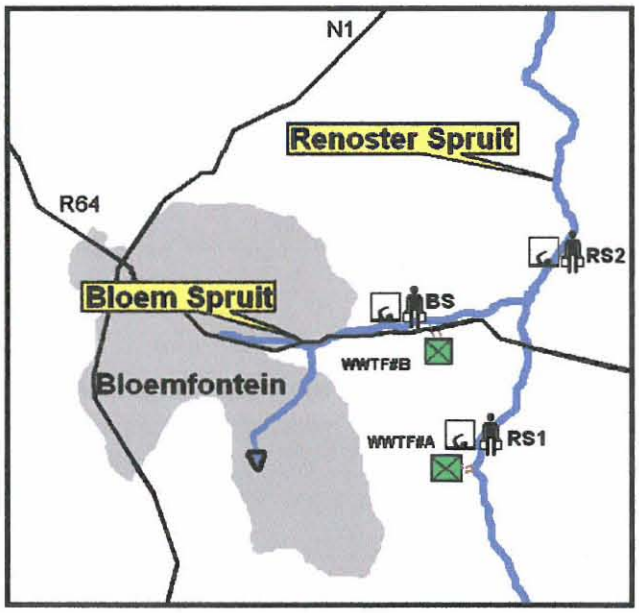



**Figure 3.1:** Sampling sites within the Renoster Spruit Quarternary Catchment

## 1.1 RS2

**Site description:** This site comprehensively represented the health-related microbiological quality of water in the RSQC immediately downstream from Bloemfontein city. Faecally-polluted urban discharges draining to this point included:

- Final effluent from two [boilerplate] nent facilities – one in the Bloem Spruit and one in the upper reaches of the Renoster Spruit.
- Polluted urban run-off from poorly sanitised settlements on the peri-urban fringes within the upper reaches of the Renoster Spruit (represented by RS1).
- Polluted urban run-off drained by the Bloem Spruit from urban areas generally serviced by full waterborne sewage systems (represented by BS).

- **Water uses observed:**
  - Young males (schoolchildren) swimming and playing in the stream.
  - People fishing in the Spruit, especially in downstream impoundments.
  - Sprinkler-irrigation of health-sensitive crops (vegetable gardening).

### 1.1.1 *E. coli* and *Salmonellae* occurrence at RS2 in the Renoster Spruit

Table 3.1 shows the *E. coli* and *Salmonellae* counts /100mℓ tested in the water samples taken at RS2 throughout the 2001/02 summer period. The data were log-transformed to reduce the variance (Appendix C) and, since the data were then lognormal, tested for significant relationships using the Pearson Product Moment Correlation (Appendix C) coefficient (SigmaStat, 1997).

**Table 3.1:** *E. coli* and *Salmonellae* occurrence (per 100mℓ) at RS2 in the Renoster Spruit

| Occurrence | *E. coli* | *Salmonella spp.* | Pearson Product Moment Correlation |
|---|---|---|---|
| Geometric mean | 11,995 | 117 | Correlation Coefficient: r = 0.424 |
| Minimum | 1,000 | 36 | P Value: P = 0.102 |
| Maximum | 290,900 | 1,001 | |
| Number of Samples | 16 | | There were no significant relationships between any pairs of variables (EC vs. Sal) (P > 0.050) |
| | Log-transformed data | | |
| Mean | 4.08 | 2.07 | |
| Confidence interval (95% Ci) | 0.35 | 0.23 | |
| Ci Upper Limit | 4.43 | 2.30 | |
| Ci Lower Limit | 3.73 | 1.83 | |
| 95th Percentile | 5.14 | 2.91 | |

## 1.2 BS

- **Site description:** This site was in the Bloem Spruit approximately 4 km upstream from RS2. BS received:

- Urban surface run-off in city.
- Final effluent from the Bloem Spruit wastewater treatment facility (WWTF#B, Figure 3.1). The health-related microbiological quality of the effluent varied and often tested above legal requirements (Griesel, 2001).
- Faecally polluted (untreated sewage, raw abattoir effluent) overflows from storm water draining points (often observed at this point during the study period). These were from major defects in sewer systems, which included sewer blockages and pipe breaks (Pretorius, 2002).

- **Water uses observed:** predominantly recreational activities but also some activities related to domestic uses:
  - The risk of infection related to ingestion of water used for irrigation (from the Bloem Spruit) was represented by this site. The health-related microbiological water quality and associated risk of infection of Bloem Spruit water at this site was concluded by Griesel and Jagals (2002), as well as Jagals et al. (1997), to be of similar quality to that of the WWTF#B effluent. Since this site is in close proximity of two golf courses being irrigated with final effluent from the WWTF#B, from a risk perspective, water from this site was therefore considered comparable to water irrigated on these golf courses from the effluent maturation ponds.
  - School children, especially young males from nearby informal settlements, were often observed participating in recreation-related activities that included full body immersion (swimming, splashing and playing). These activities took place in the stream itself, or in tertiary treatment ponds of the WWTF adjacent to the streams.
  - Destitute homeless people living under bridges and in enclosed stormwater channels in the Bloem Spruit system used the water for domestic purposes.

### 1.2.1 *E. coli* and *Salmonella* BS in the Bloem Spruit

Table 3.2 shows the un-transformed and log-transformed results at BS. The log-transformed data were normally distributed. *E. coli* counts /100mℓ at this site were higher than at the other two sites.

**Table 3.2:** *E. coli* and *Salmonellae* occurrence at BS in the Bloem Spruit /100mℓ

| Occurrence | *E. coli* | *Salmonella spp.* | Pearson Product Moment Correlation |
|---|---|---|---|
| Geometric mean | 64,656 | 138 | Correlation Coefficient: r = 0.664 |
| Minimum | 2,670 | 36 | P Value: P = 0.00502 |
| Maximum | 1,200,000 | 4,383 | |
| Number of Samples | 16 | | The *E. coli*: *Salmonellae* pairs of variables were significantly related (P < 0.050). The r and P values were positive, which indicated that the variables tended to increase together |
| | Log-transformed data | | |
| Mean | 4.81 | 2.14 | |
| Confidence interval (95% Ci) | 0.42 | 0.28 | |
| Ci Upper Limit | 5.23 | 2.42 | |
| Ci Lower Limit | 4.39 | 1.86 | |
| 95th Percentile | 5.99 | 3.07 | |

## 1.3 RS1

**Site description:** This site was in the Renoster Spruit approximately 8 km upstream from point RS2 (Figure 3.1).

- The principal contributor to perennial stream-flow at this point was final effluent from WWTF#A. The health-related microbiological quality of the effluent varied and often tested outside legal requirements (Griesel, 2001).
- Rapidly-expanding informal settlements were encroaching the Renoster Spruit. These areas were largely unsewered, with extensive ventilated improved pit latrine (VIP) installation programmes being implemented.
- Livestock farming activities upstream from the sampling site contributed to the faecal pollution and potential presence of *Salmonellae* (Griesel, 2001; Jagals, 1997).

**Water uses observed:**

- Again, predominantly, schoolchildren (young males) from nearby black informal settlements used the water in the Renoster Spruit for recreational purposes that lead to full-body immersion.

- People from nearby i ts were observed fetching water in containers from the Spruit. Water fetched in containers was stored at home and used with limited (or no) treatment for potable purposes and food preparation.
- Women were periodically observed doing laundry in the Spruit and pre-school children often used the opportunity to play in the water.
- People from time to time bathed themselves in the stream after doing laundry or fishing etc.

### 1.3.1 *E. coli* and *Salmonellae* occurrence at RS1 in the Renoster Spruit

Table 3.3 shows the un-transformed and log-transformed results at RS1. The log-transformed data were normally distributed. *Salmonellae* occurred in greater numbers at this site than at the other two.

**Table 3.3:** *E. coli* and *Salmonellae* occurrence (per 100ml) at RS1 in the Renoster Spruit

| Occurrence | *E. coli* | *Salmonella spp.* | Pearson Product Moment Correlation |
|---|---|---|---|
| Geometric mean | 29,716 | 289 | Correlation Coefficient: r = 0.463 |
| Minimum | 4,040 | 73 | P Value: P = 0.0712 |
| Maximum | 227,900 | 949 | |
| Number of Samples | 16 | | There were no significant relationships between *E. coli* and *Salmonellae* pairs (P > 0.050) |
| | Log-transformed data | | |
| Mean | 4.47 | 2.46 | |
| Confidence interval (95% Ci) | 0.25 | 0.16 | |
| Ci Upper Limit | 4.73 | 2.62 | |
| Ci Lower Limit | 4.22 | 2.30 | |
| 95th Percentile | 5.33 | 2.90 | |

## 1.4 General *E. coli* and *Salmonellae* occurrence in the RSQC

The *E. coli* results were similar to those reported over several years by Jagals (1997) as well as Griesel and Jagals (2002), who found *E. coli* counts up to $10^6$ in prior studies done on faecal pollution of the waters in the RSQC. Human activities and associated environmental impacts within the study area (e.g. discharges from wastewater treatment facilities and blocked sewers) contributed significantly to the high levels (Pretorius, 2002).

This study was the first to test the surface waters of the RSQC for the occurrence of *Salmonellae* in order to assess whether these pathogenic microorganisms, often associated

with the presence of *E. coli* (DWA lunter, 2002), would also occur in the surface waters of the RSQC. Polo et al. (1998) indicated that although *Salmonellae* are often detected in the absence of faecal indicators in water, it is generally accepted that *Salmonellae* are present when indicator organisms are present in high densities (Hunter, 2002). The *E. coli* reported in the prior RSQC studies could therefore indicate the potential presence of *Salmonellae* although it could not be certain to what extent *Salmonellae* would co-occur with *E. coli*.

Table 3.4, as well as Figure 3.2 summarises the occurrence of *E. coli* and *Salmonellae* in the waters of the RSQC (Appendix E).

**Table 3.4:** Specific occurrence of *E. coli* and *Salmonellae* in the RSQC

| Occurrence n = 16 samples each | | RS2 | BS | RS1 | Kruskal-Wallis ANOVA on Ranks |
|---|---|---|---|---|---|
| *E. coli* | Geomean | 11,995 | 64,565 | 29,716 | Significant Difference (P = 0.031) RS2 significantly lower than BS and RS1 (Dunn's MCT) |
| | Minimum | 1,000 | 2,670 | 4,040 | |
| | Maximum | 290,900 | 1,200,000 | 227,900 | |
| | Log-transformed data | | | | |
| | Mean | 4.08 | 4.81 | 4.47 | |
| | 95% Ci | 0.35 | 0.42 | 0.25 | |
| | Ci UL | 4.43 | 5.23 | 4.72 | |
| | Ci LL | 3.73 | 4.39 | 4.22 | |
| | 95th Percentile | 5.14 | 5.99 | 5.33 | |
| *Salmonella spp.* | Geomean | 117 | 138 | 289 | Significant Difference (P = 0.020) RS1 numbers significantly higher than RS2 but not BS (Dunn's MCT) |
| | Minimum | 36 | 36 | 73 | |
| | Maximum | 949 | 4,383 | 949 | |
| | Log-transformed data | | | | |
| | Mean | 2.07 | 2.14 | 2.46 | |
| | 95% Ci | 0.23 | 0.28 | 0.16 | |
| | Ci UL | 2.30 | 2.42 | 2.62 | |
| | Ci LL | 1.84 | 1.86 | 2.30 | |
| | 95th Percentile | 2.91 | 3.07 | 2.90 | |

Geomean = Geometric Mean; Ci = Confidence interval; UL = Upper limit; LL = Lower limit

Table 3.4 shows that *Salmonellae* as well as *E. coli* counts were significantly lower at RS2 than at the other two sampling sites. *Salmonellae* were higher for RS1 than for BS but not significantly so. Conversely, *E. coli* were higher at BS than RS1 but also not significantly.

The significant differences for *Salmonellae* shown in Table 3.4 must be seen in perspective. Multiple comparison tests (MCT's) indicated that *Salmonellae* were significantly more at RS1 than at RS2 but while higher than at BS, not significantly so.

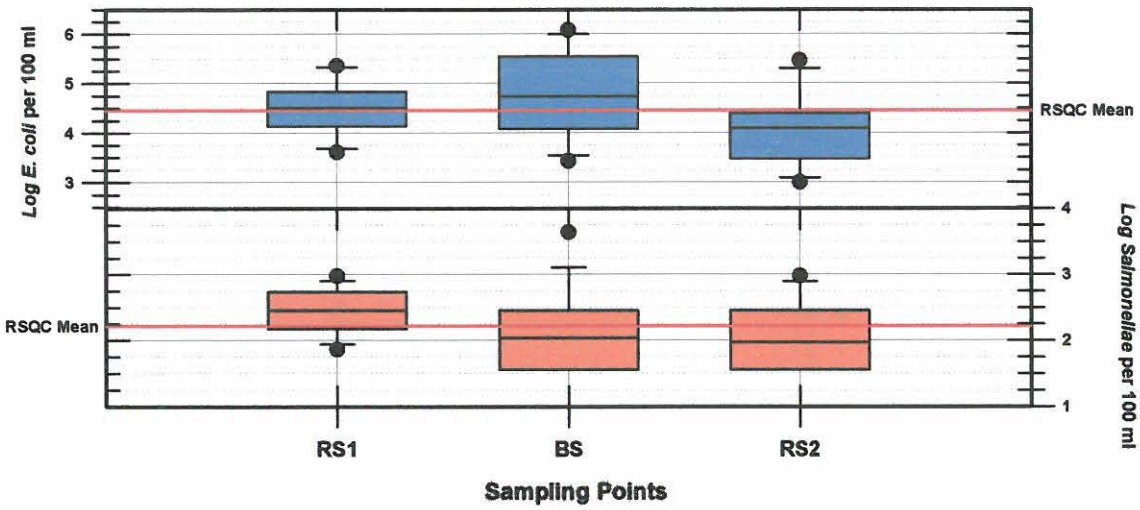



**Figure 3.2:** Log *Salmonellae* occurrence in relation to log *E. coli* at specific sampling points in the Renoster Spruit Quarternary Catchment

This implies that the major contributor of *Salmonellae* in the Renoster Spruit is the upper reaches of the Spruit and not so much the Bloem Spruit. This may be the result of agricultural run-off from concentrated livestock farming on small-holdings in the area as well as poorly treated effluent from waste water treatment facility WWTF#A. Conversely, the Bloem Spruit was the major contributor of *E. coli*.

What is noticeable though is that one would expect the highest *E. coli* occurrence to be from the same source as the highest *Salmonellae* occurrence, since *E. coli* indicates the potential presence of this pathogen. This suggests that *E. coli* is perhaps not the best indicator of the presence of *Salmonellae* in water as they do not co-occur. The red lines, in Figure 3.2, showing the log mean organism occurrence in the whole of the RSQC, confirmed a tendency that not one of the three sites was constantly more polluted than the other.

While Figure 3.2 shows that of the three sites, RS2 had the lowest numbers of *E. coli* and *Salmonellae*, the highest number of organisms isolated on any given sampling date could have come from any one of the sites as the data shown by Figure 3.3 suggests. The *E. coli* results from previous studies (Jagals, 1997, Griesel, 2001) confirmed this. However, these prior studies focussed on the occurrence of *E. coli* as indicator microorganisms, and not that of actual pathogens such as *Salmonellae*.

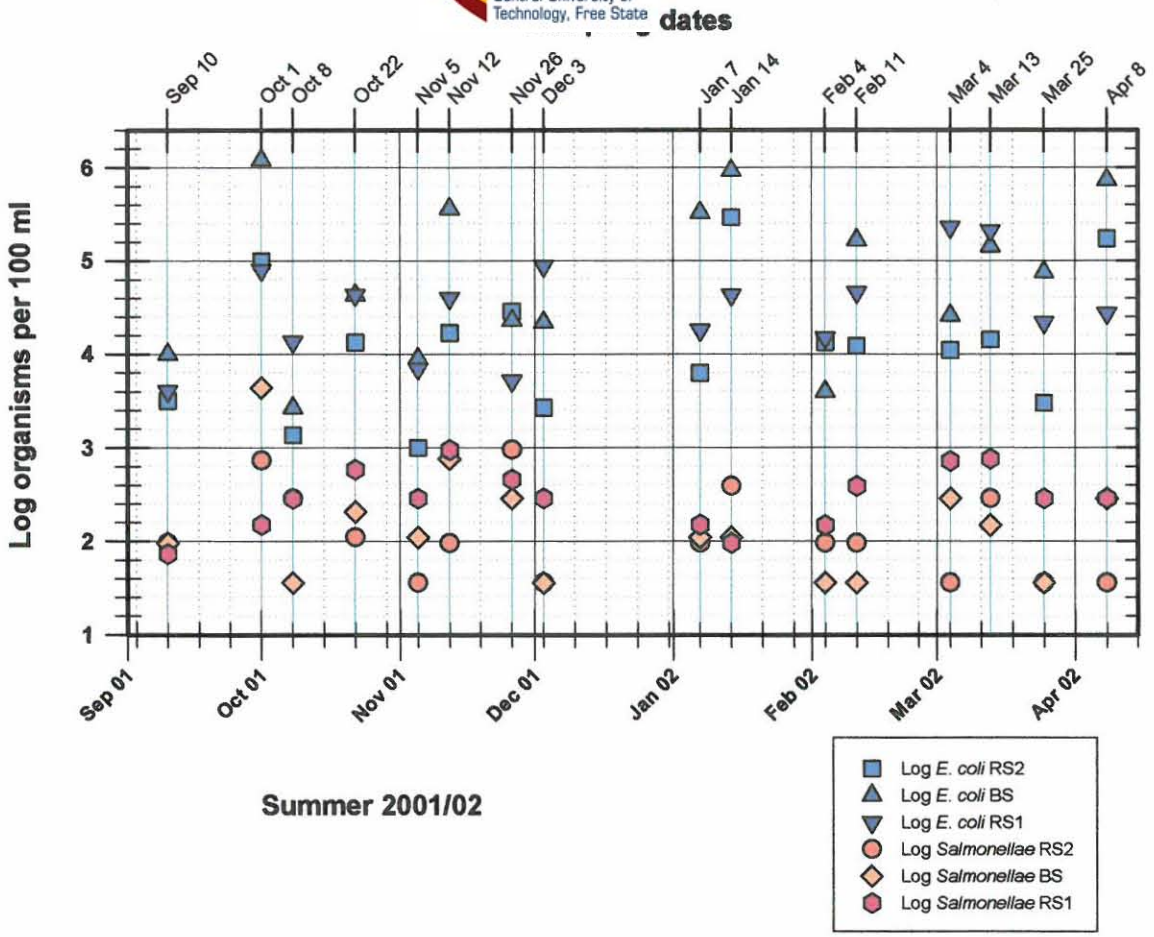



**Figure 3.3:** Log occurrence of *E. coli* and *Salmonellae* at the three sampling sites of the Renoster Spruit Quarternary Catchment during the summer period

From Tables 3.1 to 3.3 as well as Figures 3.2 and 3.3, it is indeed evident that *Salmonellae* occurred whenever *E. coli* occurred in samples taken from the surface waters of the RSQC, but it was apparent that *Salmonellae* and *E. coli* did not generally co-occur. In other words, the organisms did not consistently occur at levels in relation to another. This is probably the reason why data at two of the three sites tested no association between *Salmonellae* and *E. coli.*

## 1.5 Mean occurrence of *E. coli* and *Salmonellae* in the RSQC

Table 3.5 summarises the mean occurrence data for *E. coli* and *Salmonellae* in the surface waters at the three sampling sites in the RSQC (Appendix E). The log-transformed data from each sampling point were combined per sampling date (event) and the mean, 95% confidence interval of the mean, as well as the 95$^{th}$ percentile (US-EPA, 1995) calculated

(Chapter 2, Section 5.1). The re ance (through log transformation) had resulted in the data being smoothed, which resulted in a more visible and calculable co-variance.

**Table 3.5:** Mean *E. coli* and *Salmonellae* occurrence in the RSQC

| Occurrence | | *E. coli* | *Salmonella spp.* | Pearson Product Moment Correlation |
|---|---|---|---|---|
| RSQC Number of samples 48 | Geomean | 28,444 | 167 | Correlation Coefficient: r = 0.5 |
| | Minimum | 1,000 | 36 | P Value = 0.000297 |
| | Maximum | 1,200,000 | 4,383 | There was a significant relationship between the *E. coli*: *Salmonellae* pairs of variables, characterised by the tendency to increase together (P < 0.050) |
| | | Log-transformed data | | |
| | Mean | 4.45 | 2.22 | |
| | 95% Ci | 0.21 | 0.14 | |
| | Ci UL | 4.66 | 2.36 | |
| | Ci LL | 4.24 | 2.08 | |
| | 95th Percentile | 5.33 | 2.75 | |

Geomean = Geometric Mean; Ci = Confidence interval; UL = Upper limit; LL = Lower limit

Figure 3.4 shows the average (the symbols) for the log organism counts /100mℓ of all three sampling sites for a specific sampling date, while the purple solid line represents the *combined* mean occurrence for the whole of the RSQC.

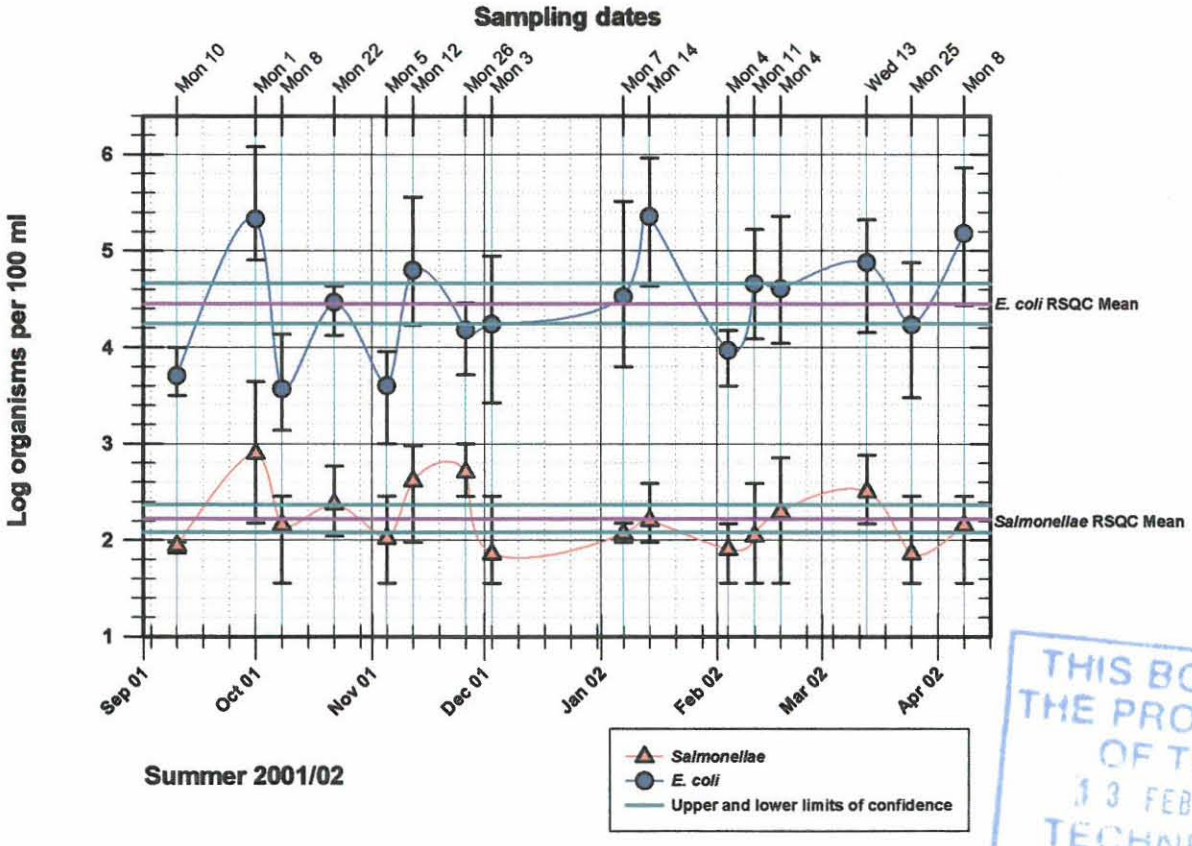



**Figure 3.4:** Log *E. coli* and *Salmonellae* occurrence in the Renoster Spruit Quarternary Catchment

Figure 3.4 shows that although t[Central University of Technology, Free State] of an association between the occurrence of *E. coli* and *Salmonellae* at the different sampling sites, the combined *E. coli* and *Salmonellae* occurrence for the RSQC followed a qualitatively similar trend. The error bars show the respective maximum / minimum (max / min) measurement for each sampling date. The upper bar represented the highest level of occurrence measured and conversely, the lower bar the minimum occurrence for the particular sampling date.

### 1.5.1 Confidence intervals of the mean organism occurrences in the RSQC

Figure 3.4 shows 95% confidence intervals (between the dark-cyan upper and lower limit lines) of the mean occurrences of both organism groups. The considerable variation was recorded, since all outliers (Appendix C) were used to calculate the confidence intervals. The assumption for this study was that the occurrence and consequent risks were most likely to occur within these confidence intervals. According to the US-EPA (1995), however (Chapter 2, Section 5.1), the high end of the risk distribution in the risk characterisation should also be included. Since the upper limit of the 95% confidence interval of the mean would not necessarily include a high-end such as the 95$^{th}$ percentile, the latter was also included for this study.

It was within the respective 95% confidence intervals of the mean, as well as 95$^{th}$ percentile occurrence that:

- The ***possible*** risk of infection posed by the occurrence of *E. coli* could be evaluated against Observed-Adverse-Effect-Levels (OAEL's) found in various water quality guidelines based on the OAEL approach (OAELA) (Chapter 2, Section 3.2.1).
- The ***probable*** risk of infection posed by the occurrence of *Salmonellae* could be assessed based on the QMRA approach (Chapter 2, Section 3.2.2).

The outlier data points outside the confidence intervals were only considered for extraordinary single exposure risk characterisations discussed in Section 4 of this chapter.

## 1.6 Uncertainty analyse

A number of sources of uncertainty in applying the WRQMRA process have already been discussed in Chapter 2 (Sections 2.2, 3.5, 4.3, 5.2 and 6.1). The following additional sources of uncertainty were identified:

- Shortage of resources (cost of analyses, time, etc.) limited the number of pathogen and indicator species chosen to test the health-related microbiological quality of untreated surface water in the RSQC. Limiting analyses to *Salmonellae* (pathogen) and *E. coli* (indicator) occurrence caused uncertainty in expressing the actual risk of infection for the area.
- It is uncertain to what extent the *E. coli* group indicates the potential presence of *Salmonellae* in water. *E. coli* is an indicator microorganism group that indicates the potential presence and subsequent risk of infection by a whole range of pathogens. *Salmonella spp.* is only one pathogen group whose occurrence is potentially indicated. Actual occurrence of *Salmonellae* should therefore be investigated.
- Three sampling sites RS1, RS2 and BS were chosen to give an indication of the health-related microbiological quality of untreated surface water in the RSQC. It is uncertain to what extent only three selected sampling sites represented the whole of the RSQC.
- The water in the vicinity of the three sampling sites chosen to represent the RSQC was known (from previous studies in the area) to be faecally polluted, which could have caused biased results.
- While RS1 was the major contributor of *Salmonellae* in the RSQC, sampling site BS contributed more *E. coli*. This caused further uncertainty in the use of *E. coli* as indicator of *Salmonellae*, since there was not much of an association (low correlation coefficient) between *E. coli* and *Salmonellae* occurrence in the area.

? According to Standard Met rithmetic mean gives the best indication of central tendency of data for risk assessment, since it tends to overestimate the risk and therefore ensure safety. Use of the geometric mean reflects a more conservative mean probability of infection while use of the 95$^{th}$ percentile, from a health risk prevention perspective, reflects a more protective probability, since it includes people at the high-end of the risk distribution.

## 2 THE OAEL APPROACH (OAELA) FOR *E. COLI* IN THE RSQC

The OAELA establishes a possible risk of infection based on the occurrence of *E. coli* compared to various OAEL's and higher guideline risk limits. Section 2.1 discusses the general occurrence of *E. coli* while Section 2.2 evaluates this occurrence against OAEL's and higher guideline limits, for the various observed water uses, in order to conclude the infection potential of the water.

### 2.1 Numbers of *E. coli* in the RSQC

Figure 3.5 summarises the mean log *E. coli* occurrence component in the RSQC previously illustrated by Figure 3.4 above.

It shows more clearly the 95% upper and lower confidence intervals as well as the minimum and maximum occurrences (ranges) for each sampling date. The confidence intervals in Figure 3.5 form the focus of the OAELA discussions in the following sections.

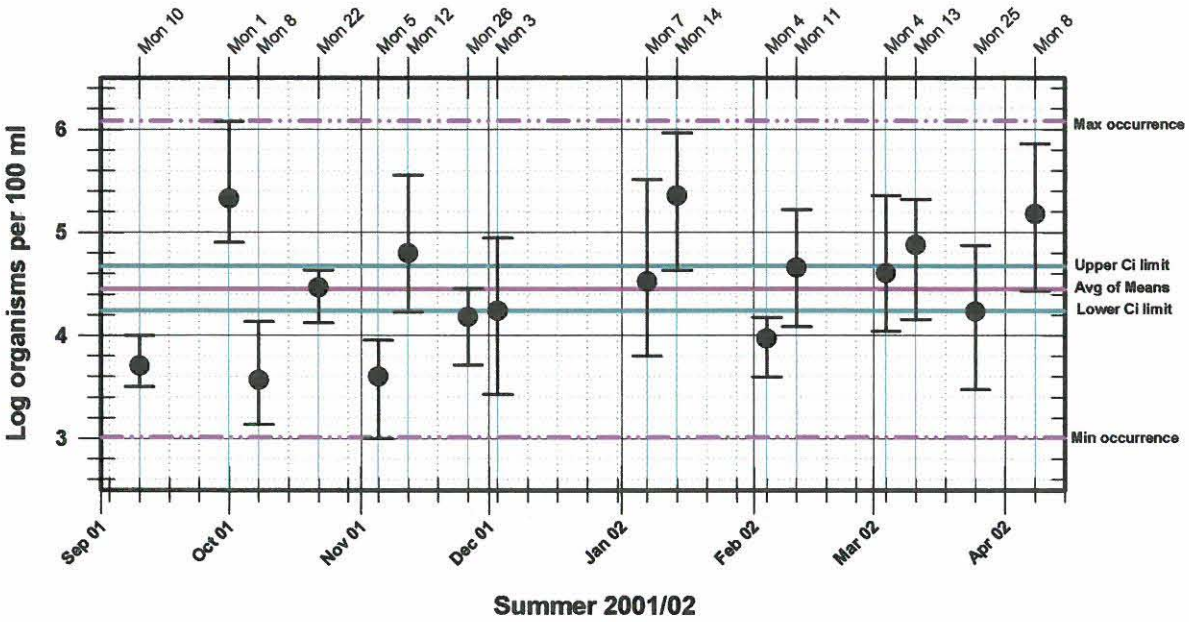



**Figure 3.5:** Log mean *E.coli* occurrence (with upper and lower confidence interval [95% Ci] limits) in the Renoster Spruit Quarternary Catchment for the summer period

## 2.2 Risk based on OAELA in the RSQC

Figures 3.6, 3.7, 3.8 and 3.9 shows the log *E. coli* occurrence in the RSQC compared to guideline (Chapter 2; Table 2.4) OAEL's (in thick solid lines), for various water uses observed in the area. *E. coli* occurrences were evaluated against No-OAEL's (NOAEL's) or Lowest-OAEL's (LOAEL's).

In addition to the OAEL's, higher risk limits (indicated by dash-dotted lines) were also applied for each of the water uses to give a further impression of the possible level of risk of gastrointestinal infection indicated by the *E. coli* occurrence (DWAF, 1996a and b; 2002; WHO, 1996). The OAEL lines represents the maximum (upper) limit below which the water is of acceptable (no-to-low risk) health-related microbiological quality for each of the uses. The higher risk lines, on the other hand, correspond to the minimum / lowest level / limit above which high possibility of health risks for the various uses could be expected.

### 2.2.1 Irrigation and intermediate body contact in the RSQC (Figure 3.6)

Figure 3.6 shows that during the summer of 2001/02, surface waters in the major streams of the RSQC were, from an indicator-based health-related microbiological perspective, not

suitable for irrigation of crops tha aw. The waters were also not suited for activities involving intermediate body contact with the water.

The log mean *E. coli* numbers exceeded not only the NOAEL, but also the high potential risk limit, throughout the season. This indicated a high possibility of health risks by pathogenic microorganisms to people applying the waters for fishing, canoeing and similar activities.

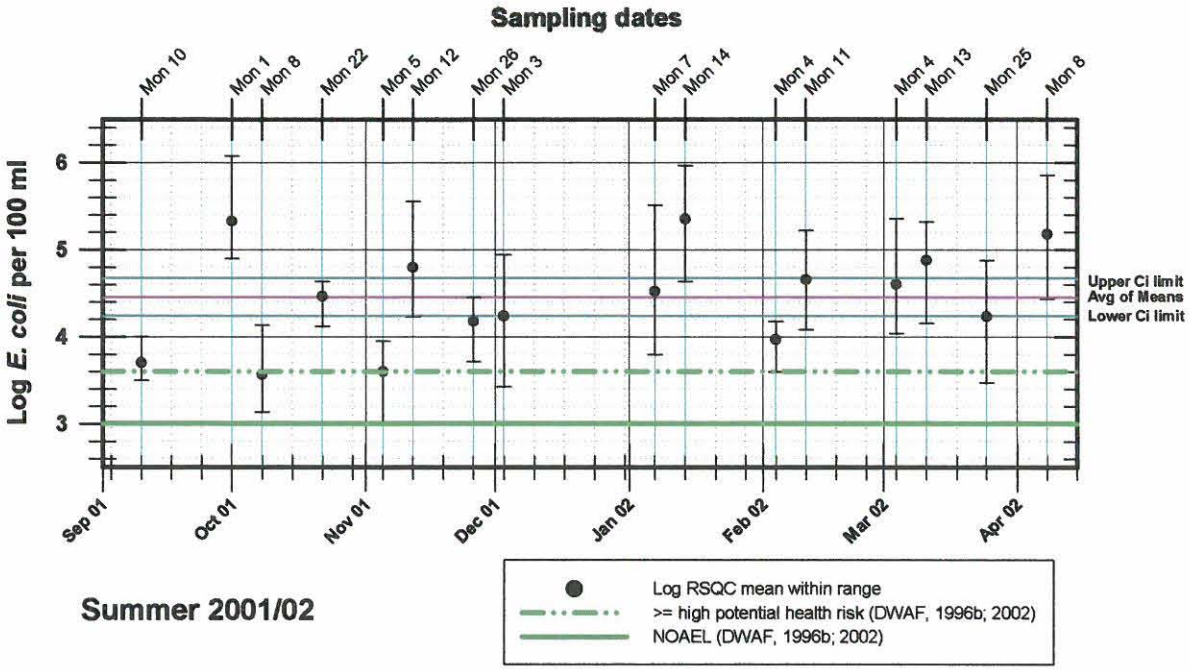



**Figure 3.6:** Comparing the occurrence of *E.coli* to OAEL's for *irrigation* and *intermediate body contact* in the Renoster Spruit Quarternary Catchment

### 2.2.2 Full-body contact with water in the RSQC (Figure 3.7)

Figure 3.7 shows that the waters of the RSQC were even more unsuitable for recreational water-use activities that involved full-body immersion e.g., swimming. This continued throughout the season.

Risks of health effects associated with contact recreational water use increase as *E. coli* levels increase (DWAF, 1996b).

According to the US-EPA, in areas where water, containing more than 1,000 *E. coli* per 100 mℓ is used for full-body immersion during recreation, gastrointestinal illness can be expected to increase approximately in accordance with the following relationship (DWAF, 1996b):

$y =$ [illegible] $\log x$)

where $y$ = illness rate /100 000 persons and

$x$ = number of *E. coli* /100 mℓ ($x \geq 3$).

This formula is based on epidemiological studies investigating exposure and calculating the health effects associated with varying water quality. This formula is applied in Section 4 of this Chapter.

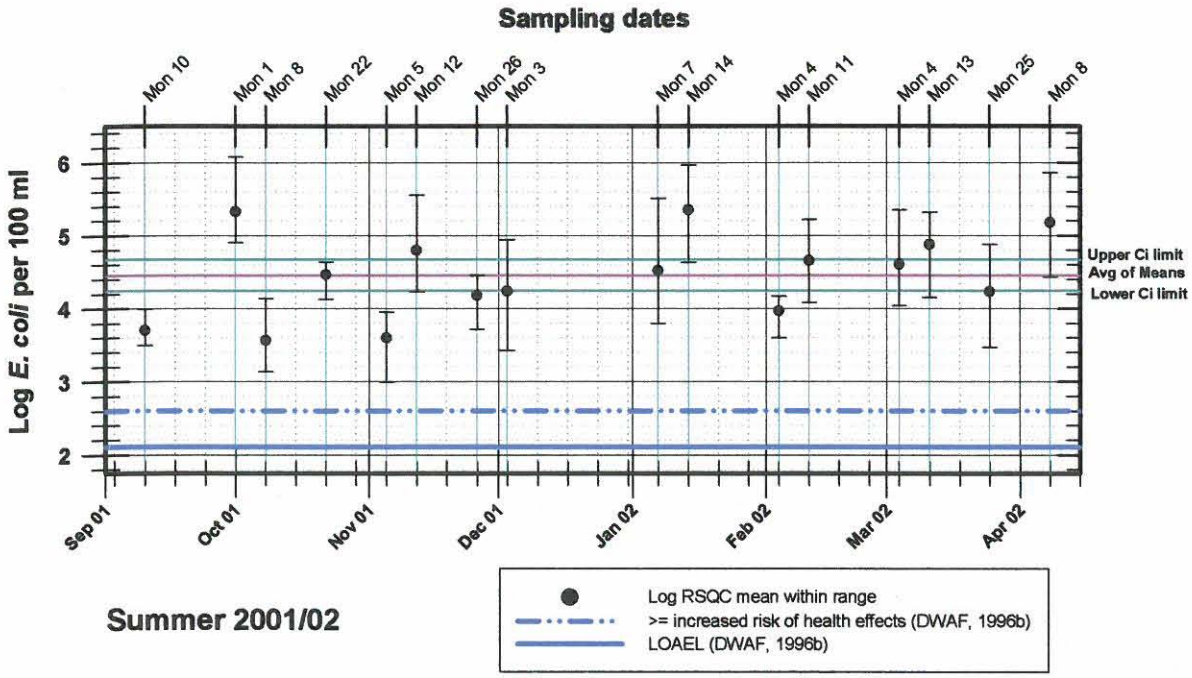



**Figure 3.7:** Comparing the occurrence of *E.coli* to the LOAEL and a higher risk limit for *full-body immersion* in the Renoster Spruit Quarternary Catchment

### 2.2.3 Drinking untreated water from the RSQC (Figure 3.8)

Figure 3.8 shows that water in the RSQC is totally unfit for human consumption from an *E. coli* indicator perspective.

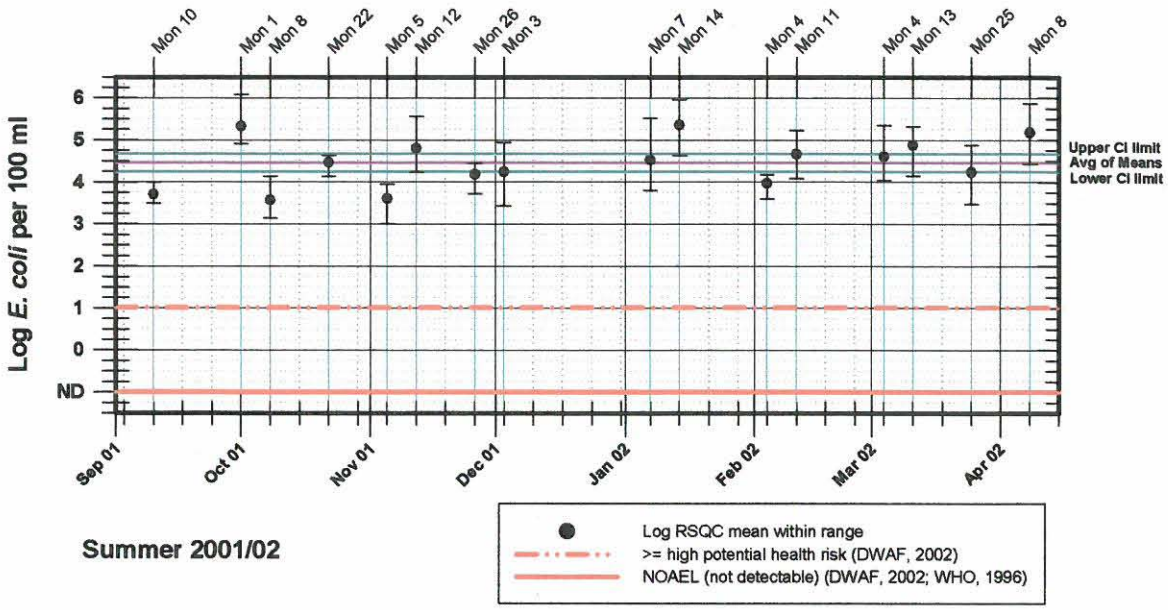



**Figure 3.8:** Comparing the occurrence of *E.coli* to a NOAEL and a higher risk limit for drinking untreated water in the Renoster Spruit Quarternary Catchment

### 2.2.4 Potential raw water extraction from the RSQC (Figure 3.9)

The Implementation Manual for the National Microbial Monitoring Programme for surface waters in South Africa (DWAF, 2002), states that water with < 2000 counts/100 mℓ is associated with a low potential health risk and could be used for drinking after limited treatment (home treatment). Should authorities however decide to extract water from the RSQC for full (conventional) treatment (flocculation, sedimentation, filtration and disinfection), Figure 3.9 shows that the water is of such inferior quality that, according to the guidelines of the DWAF (2002), it is barely treatable. This implies that environmental health practitioners (EHP's), as well as community health-workers that might want to implement awareness creation programmes at households wanting to apply home treatment, would find it very difficult to ensure a safe product.

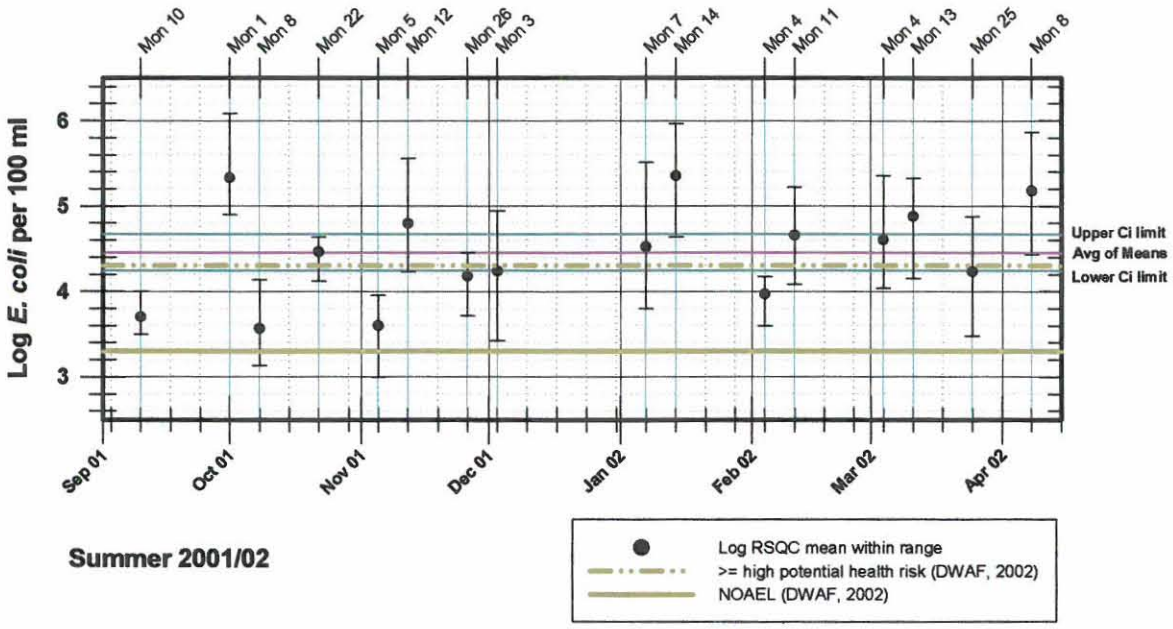



**Figure 3.9:** Comparing the occurrence of *E.coli* to a NOAEL and a higher risk limit for raw-water extraction for treatment from the Renoster Spruit Quarternary Catchment

## 2.3 Uncertainty analyses

- Combining the results for each separate sampling site (uncertainty partly discussed in Section 1.6 above) may not be realistic in terms of the risk approach, since it may lead to over- or under estimation of the occurrence and subsequent risk at each separate site. It might be that for a specific sampling event, sampling site RS1 for example had a much greater *Salmonellae* occurrence and resultant risk than sampling site BS. More importantly, indicator counts were shown not to correlate with *Salmonellae* counts at two out of the three sampling sites investigated. Combining the results and estimating the risk would underestimate the risk for RS1 and overestimate the risk for BS for that specific sampling event.

- It is also uncertain to what extent this over- or under estimation would cancel each other out in terms of the outlying values considered.

- Since *E. coli* indicates the presence of faecal pollution of water and therefore shows the potential presence of various pathogens in water, it is uncertain to what extent it

could indicate the potential *onellae* and the probable risk of infection posed by *Salmonellae*. *E. coli* tends to overestimate the risk of infection by *Salmonellae* since it indicates the presence of a whole range of pathogenic microorganisms.

? It is therefore uncertain to what extent the possible risk of infection (based on the OAELA) could indicate the probable risk of infection by *Salmonellae* (QMRA approach).

## 3 THE QMRA APPROACH FOR *SALMONELLAE* IN THE RSQC

This section discusses the application of the four QMRA steps (described in Chapter 2) on the surface waters of the RSQC. This process was applied to determine whether the numbers of *Salmonellae* (selected hazard) in these waters would lead to a probability of infection, if people were to ingest these waters.

As part of the exposure assessment step for this section, Section 3.1 (to follow) summarises the occurrence of *Salmonellae* in the waters of the RSQC (discussed in Section 1 above). Section 3.2 discusses the various ingestion volumes associated with the water-use activities. Section 3.3 deals with expected doses and Section 3.4 with the infection probability based on dose-response models and parameters (discussed in Chapter 2, Section 4). These are applied in risk scenarios in Section 3.5 to illustrate the use of QMRA to determine probability of *Salmonellae* infection for various water uses over different exposure periods. Risk statements (characterising the probability of infection) follow in Section 3.6.

### 3.1 Numbers of *Salmonellae* in the RSQC

Figure 3.10 shows the combined log mean *Salmonellae* concentrations in the RSQC per 100mℓ (solid purple line), as well as the log mean (●) concentrations for each sampling date.

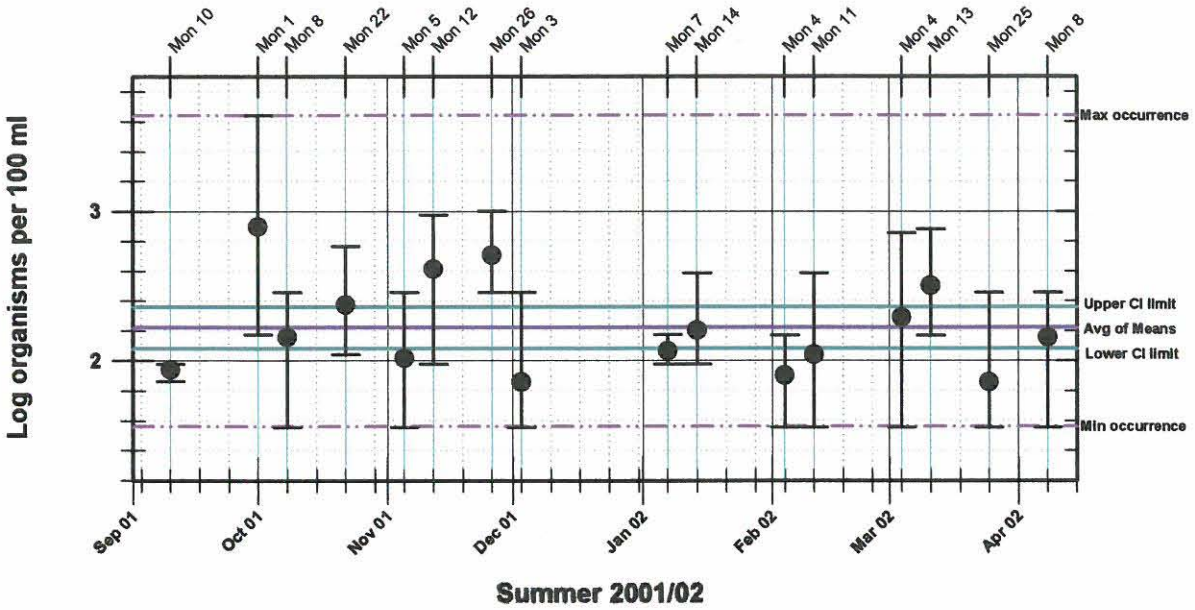



**Figure 3.10:** Log mean *Salmonellae* occurrence (with upper and lower confidence interval [95% Ci] limits) in the Renoster Spruit Quarternary Catchment for the summer period

## 3.2 Ingestion volumes associated with activities observed in the RSQC

The water-use activities previously observed in the area (in the vicinity of the sampling sites), were associated with possible volumes of water thought to be ingested during contact with the water (Table 3.6). For example, a person swimming in the stream was assumed to have more contact with the water than a person exposed to the water droplets thought to be ingested while doing laundry in the stream. The volume of 100mℓ ingestion therefore corresponds with swimming, and 10mℓ with that of doing laundry.

**Table 3.6:** Observed water use activities and activity-related ingestion volumes in the RSQC

| Water-use activity | RS1 | BS | RS2 | Water ingestion volume (mℓ) |
|---|---|---|---|---|
| Fishing | | | ✓ | 10mℓ |
| Laundry | ✓ | | | 10mℓ# |
| Drinking | ✓ | ✓ | | 10mℓ; 50mℓ; 100mℓ# |
| Body-washing | ✓ | | | 100mℓ |
| Swimming (full-body immersion) | ✓ | ✓ | ✓ | 100mℓ# |

# See discussion in Section 3.2

Important to notice is that, although destitute and homeless people could potentially make use of the RSQC waters for domestic purposes such as intentional drinking, it is not done (observed) by rule, when other waters are available or accessible. Instead of using the

intentional (voluntary) ingestion v oup as discussed in Chapter 2 (Table 2.5, Section 3.3.1.2.3) for this chapter, the involuntary ingestion volumes of 10mℓ, 50mℓ and 100mℓ was therefore related to these water-uses.

## 3.3 Dose

Dose, for this study, is calculated by the following equation:

**X /100mℓ x fraction /100mℓ ingested**

where **X** is the counts/100mℓ

and the **fraction /100mℓ** is the volume of water (per 100mℓ) actually ingested.

Table 3.7 shows the geometric mean (Appendix C) dose, as well as the upper (maximum expected) and lower (minimum expected) dose for ingestion based on the 95$^{th}$ percentile and the minimum *Salmonellae* occurrence in the waters of the RSQC respectively.

Ingestion volumes of 100mℓ, 50mℓ and 10mℓ are assumed (Chapter 2, Section 3.3.2), reflecting potential water ingestion through water-use activities in the RSQC (discussed in Section 3.2 above).

**Table 3.7:** Expected dose of *Salmonellae* associated with exposure to waters of the RSQC at a range of ingested volumes

| *Salmonellae* occurrence (a) per 100mℓ (c) | | Ingestion volume (b) | Expected dose (a*b)/100 (c) | | |
|---|---|---|---|---|---|
| | | | Mean dose | Maximum dose (upper limit) | Minimum dose (lower limit) |
| Geometric mean | 167 | 100mℓ | 167 | 883 | 36 |
| 95$^{th}$ Percentile (maximum) | 883 | 50mℓ | 84 | 442 | 18 |
| Minimum occurrence level | 36 | 10mℓ | 17 | 88 | 4 |

Table 3.7 shows that users of untreated water from the RSQC would have ingested a mean number of 167 *Salmonellae* organisms had they ingested 100 mℓ of the water. At an ingestion volume of 100 mℓ, the potential user could expect to ingest not more than 883, and not less than 36 *Salmonellae* for the 2001/02 summer season on any random day at any of the three sampled sites in the RSQC.

## 3.4 Probable risk of *Sal*[illegible]*on* in the RSQC based on dose-response models and parameters

The probability of infection ($P_i$) was calculated according to the *β-Poisson* distribution model, using dose-response parameters from Haas and Eisenburg (2001), based on exposure to the doses summarised in Table 3.7.

$P_i$ in Table 3.8, as well as Figure 3.11, principally reflects the possible risk posed by a single exposure to the mean dose associated with volume of water ingested (100mℓ, 50mℓ and 10mℓ).

**Table 3.8:** Probability of *Salmonellae* infection ($P_i$) based on a single exposure to the mean, maximum and minimum dose

| Ingestion volume (mℓ) | Expected dose | | Probability of infection ($P_i$) (Single exposure) |
|---|---|---|---|
| 100 | Mean | 167 | 0.0174 |
| | Maximum | 883 | 0.0801 |
| | Minimum | 36 | 0.0039 |
| 50 | Mean | 84 | 0.0089 |
| | Maximum | 442 | 0.0435 |
| | Minimum | 18 | 0.0019 |
| 10 | Mean | 17 | 0.0018 |
| | Maximum | 88 | 0.0094 |
| | Minimum | 4 | 0.0004 |

Also shown are the infection risk levels based on ingestion of water varying in quality from less contaminated (minimum expected dose based on true minimum *Salmonellae* occurrence) to more polluted water (maximum expected dose based on 95$^{th}$ percentile).

If a person, for instance, involuntarily ingested 10mℓ of water per single exposure event, the risk of infection would range between $P_i$ of 0.0004 to 0.0094 with a mean probable risk of 0.0018. In other words, depending on the water use (with which a volume is associated, hence a dose), for a person that ingests 10 mℓ of reasonably uncontaminated water, the minimum expected $P_i$ would be 0.0004. Conversely, a $P_i$ of 0.0801 could be expected at the maximum expected dose level should 100mℓ of the most polluted water be ingested.

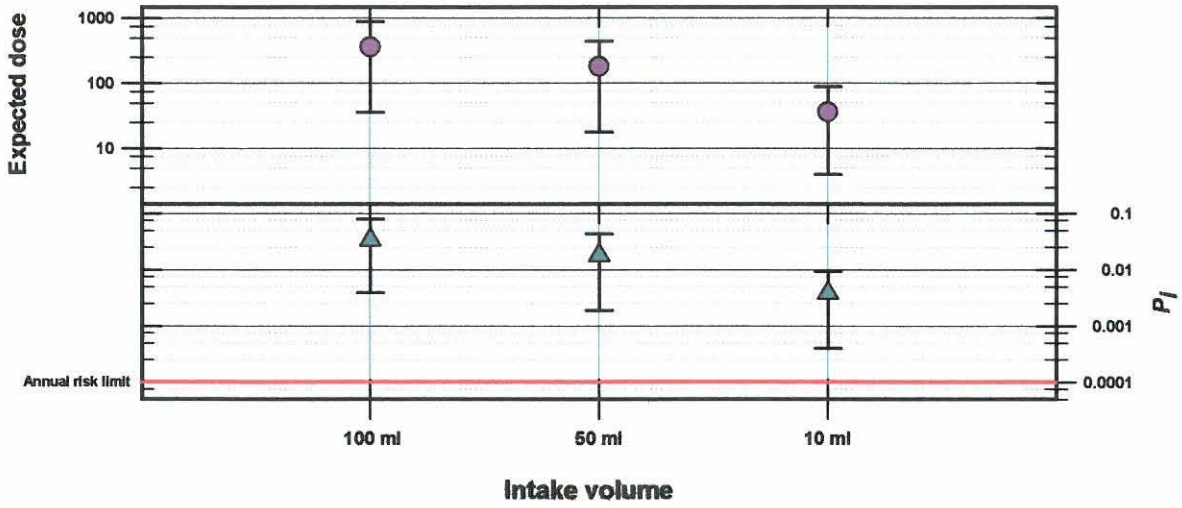



**Figure 3.11:** Expected dose and probability of *Salmonellae* infection ($P_i$) related to different water intake volumes

## 3.5 Characterising $P_I$

The single- as well as seasonal-exposure risk of infection was calculated with the *β-Poisson* distribution model (Formulas 1 and 2) discussed in Chapter 2, Section 4.1. $P_I$ for the summer season was based on 242 days (the total number of days for the whole eight month summer period). The reasons for a summer season sampling period were explained in Chapter 2, Section 3.2. Sections 3.3 and 3.4 (above) summarised the calculation of the expected dose and subsequent probable risk of infection based on the mean, maximum (95$^{th}$ percentile) and minimum (true minimum detected) *Salmonellae* occurrence and water-use related ingestion volumes for the RSQC. In Table 3.9, as well as Figure 3.12, $P_I$ is characterised into a risk statement.

**Table 3.9:** Single and seasonal risk based on $P_I$

| Ingestion volume (mℓ) | Expected level | $P_I$ Single exposure | % $P_I$: Single exposure | Infections per 10,000 after single exposure | % $P_I$: Summer season | Infections per 10,000 after seasonal exposure |
|---|---|---|---|---|---|---|
| 100 | Mean | 0.0174 | 1.74% | 174 | 98.65% | 9,860 |
| | Maximum | 0.0801 | 8.01% | 801 | 100% | 10,000 |
| | Minimum | 0.0039 | 0.39% | 39 | 60.88% | 6,088 |
| 50 | Mean | 0.0089 | 0.89% | 89 | 88.46% | 8,846 |
| | Maximum | 0.0435 | 4.35% | 435 | 100% | 10,000 |
| | Minimum | 0.0019 | 0.19% | 19 | 37.54% | 3,754 |
| 10 | Mean | 0.0018 | 0.18% | 18 | 35.39% | 3,539 |
| | Maximum | 0.0094 | 0.94% | 94 | 89.79% | 8,979 |
| | Minimum | 0.0004 | 0.04% | 4 | 9.01% | 901 |

For a single exposure, as well as ı 2001/02 summer season (242 days), $P_i$ is expressed in percentages, as well as for a fraction of a population of 10,000. This corresponds to the format used by the US-EPA (1994) to express maximum acceptable annual risk limit (0.01% or 1 infection per 10,000 of the population) for consumption of drinking-water (Regli et al., 1991), to which $P_i$, for this study, is compared.

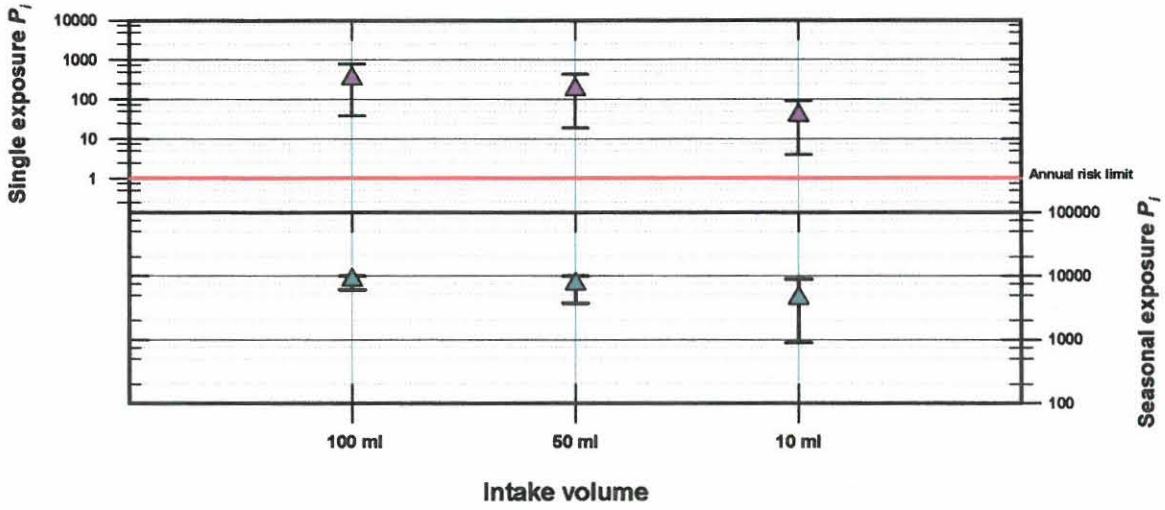



**Figure 3.12:** Single and seasonal exposure $P_i$ per 10,000 of he population related to different water intake volumes

A $P_i$ of 1 equals a 100% probability of infection (Chapter 2, Section 5, Table 2.9). For example, Table 3.9 shows that the percentage probability of infection (0.04%) from only a single exposure to the minimum expected dose at an ingestion volume of 10mℓ, already exceeded the maximum acceptable annual risk limit (0.01% or 1 per 10,000) suggested by the US-EPA (1994). As exposure to more contaminated water at higher ingestion volumes occurs, the closer $P_i$ gets to 1. This implies increasing risk of infection. For example, at any time during the summer at a volume of 100mℓ ingestion, the maximum percentage $P_i$ for a single exposure is 8.01%.

For the whole 242-day summer period, the maximum risk at e.g. 100mℓ is 100%, which implies that a whole population of 10,000 would have been infected. The 242-day infection risk therefore portrays the worst-case-scenario (ultimate maximum risk limit), as it is unlikely that people will be exposed to the same numbers of *Salmonellae* and the same volumes of water for each of the 242 days. It is also unlikely that people would, except perhaps for

daily activities such as body washi ıumption, participate in these activities and therefore be exposed to such a risk on a daily basis, throughout the season.

## 3.6 Practicable application of the QMRA

To illustrate the probability of infection in a more practicable fashion, three risk scenarios were created, based on water-use activities related to ingestion of specific volumes of water. These scenarios apply the mean, maximum, and minimum expected dose of *Salmonellae* for single, as well as multiple (e.g. seasonal) exposures to indicate the probable risk of infection with a 95% certainty. The infection risks were calculated based on the dose-response ratios established for *Salmonellae* with healthy volunteers (Haas and Eisenberg, 2001). The following scenarios were used to determine the probable risk of infection, based on the mean *Salmonellae* occurrence in the surface waters of the RSQC and the results are shown in Table 3.10:

1. A person takes a sip of water from a stream in the RSQC. This person does it twice a week throughout the summer season (35 weeks). He intentionally ingests 50 mℓ of the stream water on each occasion.

2. A child plays in one of the tributaries within the RSQC. The child accesses the stream 20 times (distributed evenly) during the summer season (242 days). The child ingests 100 mℓ of water on each occasion.

3. A woman washes her laundry in the Renoster Spruit. She washes her clothes at least 3 times a month throughout the season (eight months). On each occasion, while she washes her clothes, she involuntary ingests 10 mℓ of the stream water.

**Table 3.10:** Varying levels of $P_i$ based on intensity and frequency of water-use

| Scenario | 1. Sip of water | 2. Playing child | 3. Laundry |
|---|---|---|---|
| No. of times exposed | 70 | 20 | 24 |
| Ingestion volume per occasion | 50 | 100 | 10 |
| Mean $P_i$ per occasion | 0.0089 | 0.0174 | 0.0018 |
| Cumulative infection risk for the number of occasions | 0.4645 | 0.2967 | 0.0424 |
| 95$^{th}$ Percentile $P_i$ per occasion | 0.0434 | 0.0801 | 0.0094 |
| Cumulative infection risk for the number of occasions | 0.9557 | 0.8117 | 0.2025 |

Analyses of the infection risk bas sure scenarios showed that the mean risk was relatively low ($P_i$ of 0.0018 – 0.0089) in the case of a single exposure at the *Salmonellae* levels measured. However, it was evident that $P_i$ levels, reflected by the various activities in surface waters of the RSQC, were well above the acceptable maximum annual risk, as used for the consumption of drinking-water (0.01% or 1 infection per 10,000 persons per year).

In the case of repeated exposure, the mean risk was clearly higher ($P_i$ of 0.04-0.47). At the high end of the risk spectrum (95$^{th}$ percentile), people are exposed to a considerable risk of *Salmonellae* infection, even after only a single exposure. Persons, who use the surface waters of the RSQC regularly within the scope of the uses identified for this study, therefore run a considerable risk of contracting *Salmonellae* infection.

These risk expressions appear considerable if based on a population size of 10,000 people, which is somewhat unrealistic for the situation. The following refers to individual risks and not general population risks:

- The number of people intentionally drinking (taking a few sips) water from the untreated surface waters of the RSQC (approximately 20), is a small sub-population of a bigger general urban population of approximately 500,000 within the RSQC. From Table 3.10 a cumulative $P_i$ of 0.465 implies that 9.3 persons out of 20 drinking water from a stream in the RSQC, would be infected based on the mean *Salmonellae* occurrence. Based on the 95$^{th}$ percentile cumulative risk (Table 3.10), 19.1 out of the 20 people would have been infected with *Salmonellae*.
- This scenario represents a child from an informally settled community near the particular RSQC tributary. The hypothetical community is approximately 10,000 strong of which 3,000 children (exposed individuals) could potentially access the stream to play. Staff from the local primary health-care clinic would like to know what the probabilities would be for the children to become infected with *Salmonellae* when they play in the stream. The "stream-playing" portion of the population is therefore

30%, which implies that 89 on of 10,000 stand a chance of being infected, based on the mean cumulative infection risk. The 95th percentile cumulative infection risk calculated to a probability of 2,435 children getting infected.

- 2,000 households are assumed in the community of 10,000 (family size of 5). This implies that at least 2,000 persons (exposed individuals) use the river water to do laundry. This calculates to a $P_i$ of 0.0084 (based on mean cumulative infection risk), which, although lower than the children's risk, could still be unacceptable in terms of the local health burden, as well as primary health care delivery. For local municipal management the implication of this would be to increase delivery of available and accessible treated water to households, as well as provide safer (in terms of health-related microbiological water quality) recreational facilities. Based on the 95th percentile cumulative infection risk, 405 people are likely to be infected with *Salmonellae*.

The probability of infection predicted by these scenarios were all well above the annual acceptable risk limit, again indicating that the surface waters of the RSQC constituted a considerable risk of *Salmonellae* infection.

## 3.7 Uncertainty analyses

- *Salmonellae* are just one of the pathogen groups potentially present in the surface waters of the RSQC. The probable risk of infection indicated by the numbers of *Salmonellae* present in the RSQC may therefore be a gross underestimation of the actual risk in the area.

- The water-use activities (and frequency etc. thereof) were largely assumed and the ingestion volumes associated with the activities assumed based on the level of contact supposed. It is uncertain which activities really took place in the area, and if they took place, whether the assumed volumes associated with the activities (and the level of contact) were realistic.

- The number of people per ... actually exposed to the waters in the RSQC, was not investigated. It is therefore uncertain what population size was actually at risk of infection.
- It is uncertain to what extent the water-use activities (and frequency thereof) chosen for this study is realistic as well as representative for the study area, since, although certain activities have been previously observed to take place, water-use was not investigated for this study.
- The ingestion volumes associated with the water-use activities was not investigated but based on literature, observations and the assumed level of contact with the water.
- The mean as well as 95$^{th}$ percentile doses, over the summer period, were used. These doses were then applied to describe the probable risk for a single exposure as well as for exposure throughout the summer period of 242 days. The use of these doses could cause over-or under estimation of the actual infection risk for the area.
- Calculating the probable risk of infection for the 242-day summer period causes major uncertainty in that it is unlikely that people would be exposed to the same dose (same number of *Salmonellae* and the same volume of water ingested) for 242 days.
- Use of the US-EPA (1994) maximum acceptable annual risk limit caused further uncertainty. This limit is based on consumption of drinking-water and most likely based on a developed country situation. The limit is based on a population of 10,000, which is unrealistic for the situation in the RSQC.
- Certain risk scenarios were created and applied to practicably explain the risk of infection for the RSQC. These scenarios are based on assumption of water-use and the ingestion volumes associated with the activities. It is uncertain whether the volumes and activities applied in the scenarios are realistic and representative of the situation in the RSQC.
- The seasonal risk of infection is based on the risk of infection calculated for a single

exposure and extrapolated isure period. This causes uncertainty to whether this is a total over- or under estimation of the risk of infection, since it is unlikely that people will be exposed to the same dose (numbers of *Salmonellae* per volume of water ingested) on a daily basis throughout the season of 242 days.

? It is uncertain to what extent the QMRA approach used for this study could predict the actual risk of infection to users of untreated surface water in the RSQC.

## 4 PROBABILITY COMPARED TO POSSIBLE RISK OF INFECTION

Section 1 of this chapter shows no clear (statistically significant) associations in the *occurrences* of *E. coli* and *Salmonellae* in the waters of the RSQC. However, this was based on the series of *single* occurrences at the different sampling sites per date for each organism group. This section investigates whether the *mean* as well as *95th percentile* occurrences of *E. coli* could have reliably indicated (using the OAELA) the mean risk posed by *Salmonellae* (using the QMRA approach).

The purpose of this was to form an impression of whether the typical data (for the summer season), such as collected for this study, would have provided environmental health practitioners (EHP's) with a reliable estimation of the actual risk posed by a pathogen group such as *Salmonella spp*.

An EHP would tend to follow the OAEL approach based on indicator occurrence. A health worker would typically look at the trend of the *E. coli* occurrence in the RSQC (taking into account outlying - low and high - values), in order to make a decision on the possible risk of infection for the area over the summer 2001/02 period.

Figure 3.13 illustrates the seasonal, as well as the mean single event log *E. coli* and *Salmonellae* occurrence in the waters of the RSQC for the 2001/02 summer. The figure includes the infection probability ($P_i$) for *Salmonellae* based on an ingestion of 100mℓ (highest ingestion volumes constitute highest risk). The mean, as well as 95th percentile is used to compare risk of infection for *E. coli* and *Salmonellae*.

It is evident that *E. coli* occurred initially above all the OAEL's, for the mean as well as the 95$^{th}$ percentile high risk limit. The 95$^{th}$ percentile risk, the mean risk, as well as the single event risks posed by *Salmonellae* is clearly above the acceptable risk for consumption of drinking water suggested by the US-EPA (1994) (0.01% or 1 infection in 10,000 of the population). This section is therefore not about whether a risk has occurred or not, but simply to illustrate whether the QMRA approach could add value to the typical OAELA that environmental health practitioners would typically follow.

To determine whether *E. coli* (OAEL approach) would have predicted the occurrence of diarrhoeal disease in people, would require a full epidemiological investigation based on deductions made from an occurrence-and-effect type of study. The EPA had provided such an *E. coli* model (the formula referred to in Section 2.2.2) to quantify the risk of infection for people exposed to water by full-body immersion.

The risk of gastro-intestinal infection indicated by the log *E. coli* counts /100mℓ in the RSQC waters was calculated as follows:

$$y = -150.5 + 423.5 (\log x)$$

where $y$ = illness rate/100 000 persons, and

$x$ = *E. coli* counts /100 mℓ.

This study, however, investigated the ingestion of water through any means (from washing of clothes to direct (intentional) ingestion). Although the above formula is based on ingestion through full-body contact recreation (DWAF, 1996b), epidemiological studies are, fortunately, based on similar principles for various uses and inclined the use of this formula for the other water-uses investigated.

For this study therefore, the results concluded from application of the OAELA for full-body contact (whether found to be a reliable method or not), is therefore assumed applicable for the other water-uses investigated for this study.

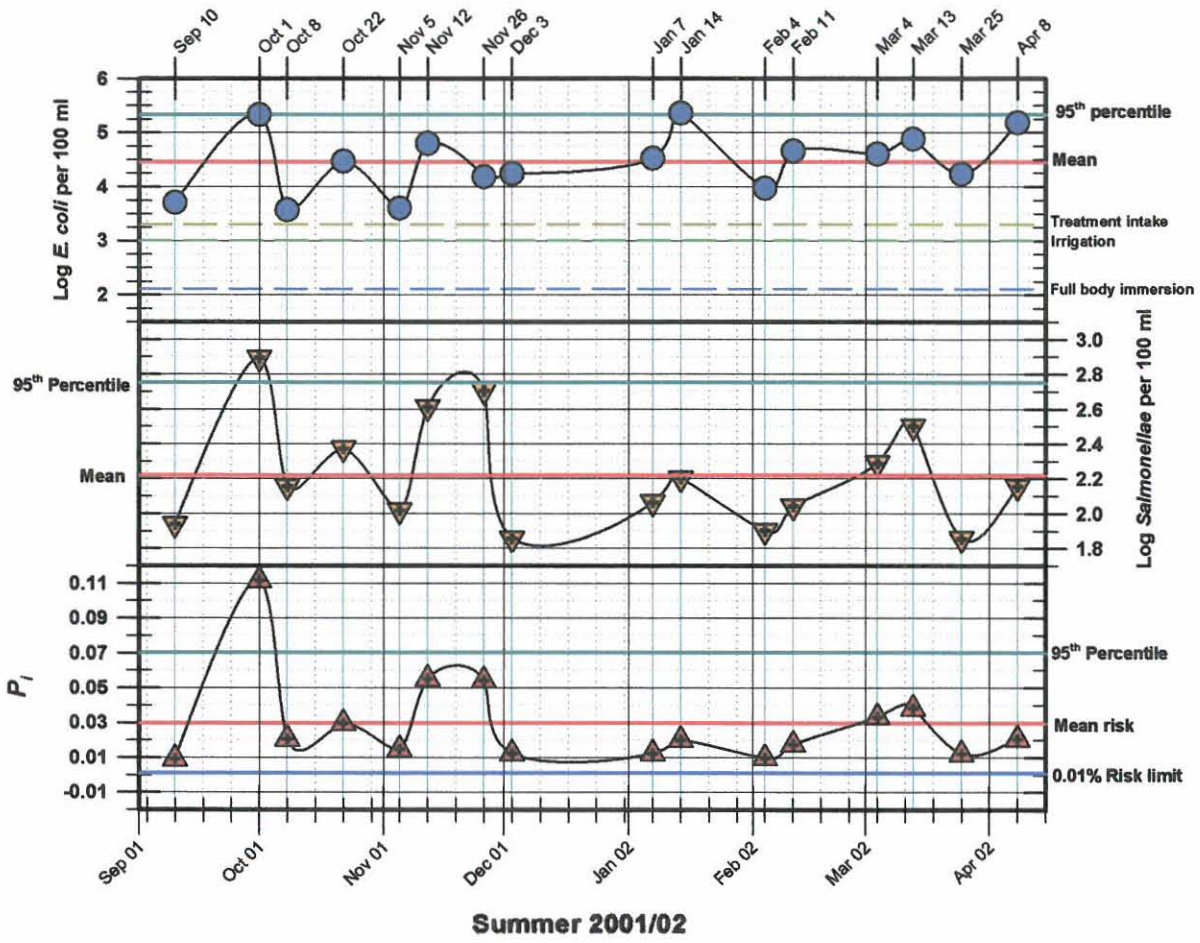



**Figure 3.13:** *E coli*, *Salmonellae* and associated risk (mean, as well as 95th percentile) measured in the Renoster Spruit Quarternary Catchment

Based on this formula and with the mean occurrence of *E. coli* (*x*) in the RSQC at 28,444 (log 28,444 = 4.45), gastrointestinal illness can be expected at 1,734.08 out of 100,000 of the population, which equals 173 out of 10,000 of the population.

This risk of infection corresponds to the probable risk of *Salmonellae* infection for the RSQC (174 out of 10,000) based on a single exposure to the mean *Salmonellae* occurrence per 100mℓ assessed by this study (Table 3.9).

This is however assumed a coincidence, since one would assume that *E. coli* should tend to overestimate the risk of infection (compared to the probable risk of infection for *Salmonellae*), as *E. coli* indicates the potential presence of various pathogens (and not only *Salmonellae*) and subsequent risks of infection, in water. What is noteworthy though is that

*E. coli* indicated a possible risk of er 10,000, while the QMRA approach concluded a much higher probable risk of 801 infections per 10,000 when the 95$^{th}$ percentile *E. coli* occurrence was compared to the 95$^{th}$ percentile probability of *Salmonellae* infection. Based on this, *E. coli* and the OAELA underestimated the risk of *Salmonellae* infection at the 95$^{th}$ percentile level.

For single event risk prediction, from a health perspective, one typically looks at the worst-scene scenario based on the high outliers. This approach portrays the worst-case scenario expected, in order to over-, rather than underestimate the risk of infection for the whole season.

On visual appraisal of Figure 3.13 alone, it was not clear whether *E. coli* could reliably indicate the risk of *Salmonellae* infection. High outlying values (above or below the 95% confidence interval of the mean, but within the 95$^{th}$ percentile) observed on various sampling dates, were therefore used to calculate the single event risks. High outlying log mean *E. coli* counts were observed on various sampling dates. The probability of infection for a single sampling event (1 October 2001) was above that of the 95$^{th}$ percentile, and therefore did not correspond to a "reasonable" infection risk. For the purposes of this discussion this outlier was not used, since it is deemed an unlikely risk. Instead, the 12 November 2001 and 13 March 2002 were chosen to represent two high outlying occurrence values (in the same log phase). The mean *E. coli* occurrences in the RSQC for the two dates (based on an ingestion of 100mℓ) did not differ by much – 6.26 x $10^4$ (log 62,600 = 4.80) and 7.54 x $10^4$ (log 75,400 = 4.88) respectively. When applied in the US-EPA formula discussed above, these occurrences calculated to a risk of gastrointestinal infections of 188 and 191 per 10,000 of the population respectively.

However, when compared to the differences in the mean risk of *Salmonellae* infection based on an ingestion of 100mℓ on each of the specific dates, it is evident that the OAELA could not successfully indicate the risk of infection. The mean *Salmonellae* risk on 12 November 2001 calculated to 552 infections per 10,000, compared to the much lower mean risk of 384 infections per 10,000 on 13 March 2002. Although both the OAELA and the

QMRA approach indicated a risk *coli* numbers on 12 November 2001 indicated a risk of infection comparable to that of the 13 March 2002. This was however not indicated by the QMRA approach. In this instance, for example, EHP's using only *E. coli* data, would have been unable to reliably predict the expected number of cases (underestimating the risk of infection) and warn clinic personnel.

The same trend occurred for two low outlying values on 8 October 2001 and 5 November 2001. Again, there was not much difference in the log mean *E. coli* numbers ($5.88 \times 10^3$ and $5.67 \times 10^3$ respectively) and risk of gastro-intestinal infection (based on the formula) on the respective dates (144.6 and 143.9 per 10,000 respectively), but the QMRA approach based on actual *Salmonellae* occurrences per 100mℓ ingested, indicated a difference in the probable infection risk. The risk of *Salmonellae* infection for 8 October equalled 207, while the risk was 149 per 10,000 of the population for 5 November 2001, a difference of 62.4 infections. Although not a log-phase difference, a difference of 62.4 infections in a population could have a substantial impact. In other words, at a lower level of organism occurrence, in terms of human health protection (aim of EHP's), depending on the population at risk (numbers, age, gender, etc.) the OAELA would not indicate a reliable risk of infection, when compared to the QMRA approach.

The OAELA based on the occurrence of *E. coli*, was expected to overestimate the risk of infection for *Salmonellae* (since it indicates presence of *various* pathogens present in water), but did not give a true indication of the probability of *Salmonellae* infection, by underestimating the infection risk.

For this study, the OAELA could not satisfactorily indicate the risk of infection by *Salmonellae* by the waters of the RSQC. However, application of the full WRQMRA approach could be recommended bearing in mind the uncertainties involved in application of this process.

Sections 1.6, 2.3 and 3.7 of this chapter already discussed a number of sources of uncertainty in applying the WRQMRA process in the RSQC. These all influenced the uncertainty of using *E. coli* as an indicator of the risk of infection in the area.

- The association between *E. coli* and *Salmonellae* or rather the lack thereof caused uncertainty in whether *E. coli* could indicate the presence of *Salmonellae* and the associated risk of infection.
- Both the OAEL and QMRA approaches indicated a risk of infection to users of untreated waters of the RSQC. Comparing the possible risk of infection by *E. coli* with the probable risk of infection by *Salmonellae* per sampling event, caused uncertainty to whether the costs of analyses, etc by the QMRA approach was justified.
- Comparison of the OAELA and the QMRA approach based only on the infection risk associated with ingestion of 100mℓ, caused uncertainty as to whether the OAELA would give a more-, or less reliable prediction of the infection risk based on smaller ingestion volumes.
- It is uncertain to what extent the results obtained from application of the US-EPA (DWAF, 1996b) formula based on epidemiological studies for full-body contact recreation could be applied to compare the OAEL and the QMRA approaches for the other water-uses discussed in this chapter.
- Acceptable risk limits in terms of the risk of infection should be established for South Africa, and more specifically the study area. It is uncertain to what extent, the fact that an acceptable risk limit has not yet been established for the area, affected the study outcomes.
- However, if singe sample event risks (with high and low outlying values) are compared it seems as if *E. coli* would have given either a over- or under estimation of the risk on certain sampling events.

- It is uncertain whether the EPA based on full-body immersion could reliably demonstrate the risk of gastro-intestinal infection indicated by *E. coli* alone.
- It is uncertain which of the two approaches (OAEL and QMRA) could best indicate the risk of infection in the RSQC as various uncertainties have been identified which developed around application of these approaches.
- It is uncertain whether the resources (time, money, etc.) involved in applying the full WRQMRA process was justified in terms of the study-outcomes.

# RISK OF INFECTION ASSOCIATED WITH INGESTION OF UNTREATED AND TREATED WATER STORED IN CONTAINERS AT HOME

In various parts of South Africa, people do not readily have access to in-house municipal supply water, and are compelled to fetch treated water from yard- or communal taps daily and store it in containers within their homes (Bokako, 2000; Genthe and Seager, 1996; Nala, 2002; Theron, 2000). In other instances where treated supply is not available or accessible, people have little other choice but to fetch water from any available / accessible source, whether treated or not, and store it in containers within their homes (Sobsey et al, 2002).

Previous studies done on municipal (treated) supply water stored in containers (in this study area) indicated a risk of infection to consumers, based on indicator organism numbers present in these waters. Bokako (2000), as well as Nala (2002) found that the quality of the treated supply was of good microbiological quality, but once tapped, the handling and storage practises caused deterioration of the microbiological quality to such an extent that human health effects were possible in consumers.

This chapter focuses on an area where people fetched treated as well as untreated water and stored it in containers at home for daily domestic use such as drinking and cooking (intentional ingestion), body-washing and laundry. As with Chapter 3, this chapter particularly focused on the probable risk of infection posed by *Salmonellae* potentially present in these waters. The chapter is divided into 4 sections. Section 1 discusses the characteristics of the two sampling sites C1 (representing treated supply) and F1 (untreated spring water), as well as the general occurrence of *E. coli* and *Salmonellae* at these sites. Section 2 discusses the evaluation of *E. coli* against OAEL's (OAEL approach) to determine the *possible* risk of infection, while Section 3 covers the *probable* risk of infection based on

the numbers of *Salmonellae* found in untreated waters stored in containers (QMRA approach). Section 4 discusses whether *E. coli,* as indicator organism, could accurately predict the probable risk of infection posed to human users by the *Salmonellae* pathogens detected in these container-stored waters. Uncertainties involved throughout application of the Water-related Quantitative Microbial Risk Assessment (WRQMRA) process, are discussed after each section. Appendix C describes the use of the means, confidence intervals, as well as the 95$^{th}$ percentile as high-end risk descriptor used in the various tables and figures of this chapter, while Appendix D discusses the extrapolation (by means of curve-fitting) of modified daily ingestion volumes of water per age group based on previous studies found in literature.

## 1 OCCURRENCE OF *E. COLI* AND *SALMONELLAE* IN WATER STORED IN CONTAINERS

Since the recreational water-use activities described in Chapter 3 required a summer sampling period, the samples for this chapter were taken on the same days throughout 2001/02. The two sampling sites C1 and F1 are shown in Figure 4.1.

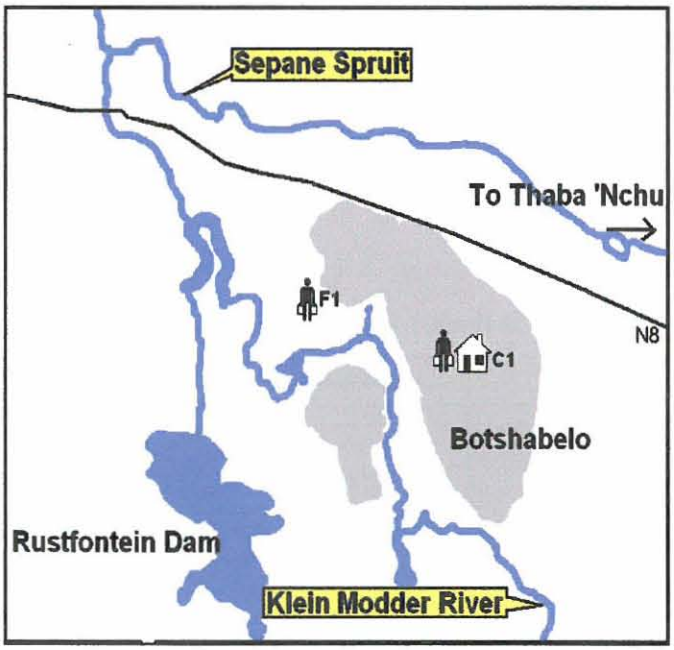



**Figure 4.1:** Sampling sites for untreated spring water (F1) and container-stored municipal supply water (C1)

## 1.1 F1 (Figure 4.1)

**Site description:** This site is an unprotected spring in the hills adjacent to the far outskirts of Botshabelo. This site is situated approximately 400 metres from the perimeter of an informal peri-urban settlement that have limited treated water supply and inadequate sanitation. For most of the residents in this community, this was the closest supply of water for domestic purposes.

**Water uses observed:**

- People from the nearby rural community fetched water from this site and used it for various domestic purposes such as drinking, food preparation, body washing, laundry, and vegetable gardening.
- During school holidays, children were observed playing with the water.
- Community-owned livestock was also observed drinking water directly at this site.

### 1.1.1 *E. coli* and *Salmonellae* occurrence at F1

Table 4.1 shows the numbers of *E. coli* and *Salmonellae* tested in the water samples taken at F1 throughout the 2001/02 sampling period (Appendix F). The data were log-transformed to reduce the variance (Appendix C), and the non-parametric Spearman Rank Order Correlation (SigmaStat, 1997) coefficient (Appendix C) used to test for significant relationships.

**Table 4.1:** *E. coli* and *Salmonellae* occurrence (per 100mℓ) at F1

| Occurrence | *E. coli* | *Salmonella spp.* | Spearman Rank Order Correlation |
|---|---|---|---|
| Geometric mean | 104 | 0.28 | Correlation Coefficient: r = 0.424 |
| Minimum | 7 | 0 | P Value: P = 0.102 |
| Maximum | 25,300 | 584 | |
| Number of Samples | 16 | | There were no significant relationships between any pairs of variables (EC vs. Sal) (P > 0.050) |
| | Log-transformed data | | |
| Mean | 2.02 | <1 | |
| Confidence interval (95% Ci) | 0.47 | 0.60 | |
| Ci Upper Limit | 2.49 | 0.05 | |
| Ci Lower Limit | 1.54 | 0 | |
| 95th Percentile | 3.71 | 2.50 | |

## 1.2 C1 (Figure 4.1)

**Site description:** In most areas of Botshabelo, people do not readily have access to water in their houses or yards and need to store water in containers in their houses after fetching it from a treated municipal supply anything up to 300 m away. C1 represented the health-related microbiological quality of container-stored municipal supply water at such a household.

- This particular sampling site was chosen because in previous studies, analyses of the microbiological quality of the stored water (Bokako, 2000; Nala, 2002), as well as general household-member attitudes towards container hygiene and domestic water handling practices (Theron, 2000) constantly produced worse results than all samples of other households.

**Water uses observed:**

- The container-stored water was primarily used for drinking (intentional ingestion), and other domestic purposes (e.g., cooking, vegetable gardening and laundry).

### 1.2.1 *E. coli* and *Salmonellae* occurrence at C1

Table 4.2 shows the un-transformed and log-transformed results at C1 (Appendix F). Since the water at this site represents treated supply water stored in containers, microorganism groups were expected to occur in lower numbers at this site. This was indeed the case.

**Table 4.2:** *E. coli* and *Salmonellae* occurrence (per 100mℓ) at C1

| Occurrence | *E. coli* | *Salmonella spp.* | Spearman Rank Order Correlation |
|---|---|---|---|
| Geometric mean | 1.16 | 0.26 | Correlation Coefficient: r = 0.601<br>P Value: P = 0.0137<br><br>The *E. coli*: *Salmonellae* pairs of variables were significantly related (P < 0.050).<br>The r and P values were positive, which indicated that the variables tended to increase together |
| Minimum | 0.1 | 0.1 | |
| Maximum | 13,300 | 1,898 | |
| Number of Samples | 16 | | |
| | Log-transformed data | | |
| Mean | 0.06 | <1 | |
| Confidence interval (95% Ci) | 0.66 | 0.58 | |
| Ci Upper Limit | 0.73 | 0.01 | |
| Ci Lower Limit | 0 | 0 | |
| 95th Percentile | 2.14 | 1.93 | |

Table 4.3 and Figure 4.2 show the general occurrence of the microorganism groups tested in the water samples taken at F1 and C1 throughout the sampling period. The non-parametric Mann-Whitney Rank Sum test (SigmaStat Version 2.03, 1997) was used to test for significant differences in the log-transformed occurrence data, regardless of whether normality was passed or failed.

The table shows *E. coli* results similar to those reported over several years by Bokako (2000) as well as Nala (2002), who found *E. coli* counts up to $10^1$/100mℓ in prior studies done on treated municipal supply water stored in containers at home. However, the previous studies did not include the untreated spring water stored in containers.

This study was the first to test untreated and treated water stored in containers at home (within the study area) for the occurrence of *Salmonellae* in order to assess whether these pathogenic microorganisms, often associated with the presence of *E. coli* (Hunter, 2002), would also occur in the container-stored water.

**Table 4.3:** General occurrence of *E. coli* and *Salmonellae* at F1 and C1 /100mℓ

| Occurrence n = 16 samples each | | F1 | C1 | Mann-Whitney Rank Sum Test |
|---|---|---|---|---|
| *E. coli* | Geomean | 104 | 1.16 | **Significant Difference** (P = <0.001) |
| | Minimum | 7 | 0.1 | |
| | Maximum | 25,300 | 13,300 | C1 significantly lower than F1 |
| | | **Log-transformed data** | | |
| | Mean | 2.02 | 0.06 | |
| | 95% Ci | 0.47 | 0.66 | |
| | Ci UL | 2.49 | 0.73 | |
| | Ci LL | 1.54 | <1 | |
| | 95th Percentile | 3.71 | 2.14 | |
| *Salmonella spp.* | Geomean | 0.28 | 0.26 | **No Significant Difference** (P = 0.985) |
| | Minimum | 0.1 | 0.1 | |
| | Maximum | 584 | 1,898 | There are no statistically significant difference in numbers between F1 and C1 |
| | | **Log-transformed data** | | |
| | Mean | <1 | <1 | |
| | 95% Ci | 0.60 | 0.58 | |
| | Ci UL | 0.05 | 0.01 | |
| | Ci LL | 0 | 0 | |
| | 95th Percentile | 2.50 | 1.93 | |

Geomean = Geometric Mean; Ci = Confidence interval; UL = Upper limit; LL = Lower limit

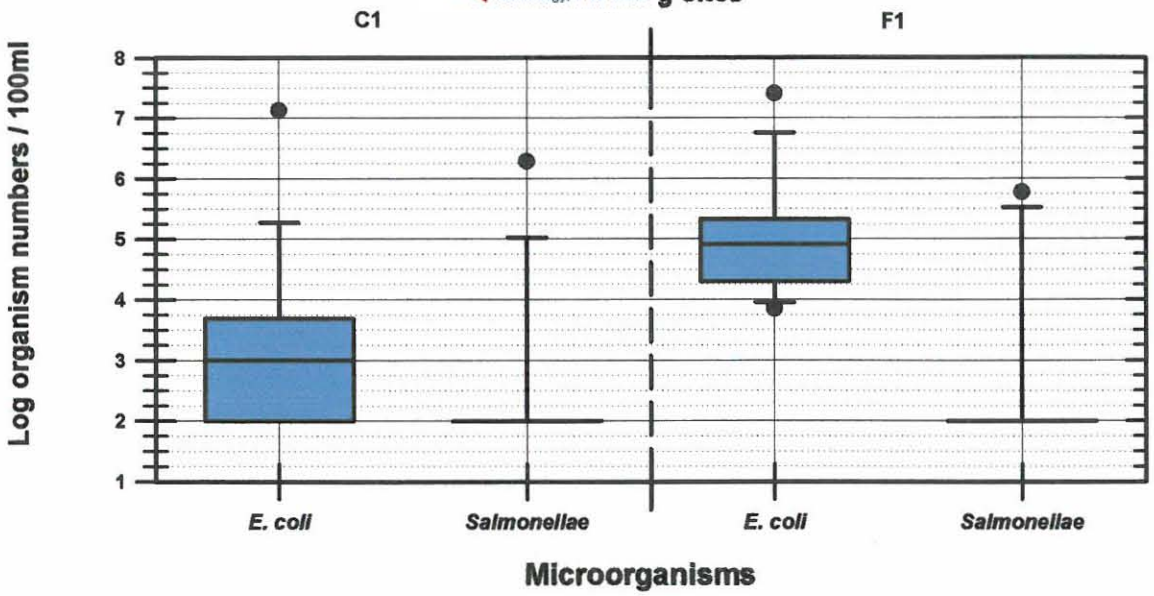



**Figure 4.2:** *Salmonellae* occurrence in relation to *E. coli* at specific sampling sites

Polo et al. (1998) indicated that although *Salmonellae* are often detected in the absence of faecal indicators in water, it is generally accepted that *Salmonellae* are present when indicator organisms are present in high densities (Hunter, 2002). The *E. coli* reported in the prior studies done on container-stored water could therefore indicate the possible presence of *Salmonellae* although it could not be certain to what extent *Salmonellae* would co-occur with *E. coli.*

Table 4.3 shows that while *E. coli* were significantly higher at F1 than C1, *Salmonellae* were also higher for F1 than for C1 but not significantly so. From Tables 4.1 to 4.3 as well as Figure 4.2, it was apparent that *Salmonellae* and *E. coli* did not generally co-occur in samples taken from the waters stored in containers. In other words, the organism groups did not consistently occur at levels in relation to one another, since *Salmonellae* were not detected every time *E. coli* occurred. This caused uncertainty whether the level and extent of *E. coli* occurrences in the container-stored water, in general, could predict the occurrence of *Salmonellae* in order to predict a *possible risk* of infection comparable to a *probable risk* (of infection).

Figure 4.2, however, does not give a clear indication of the co-occurrence of *E. coli* and *Salmonellae* on each sampling date. Figures 4.3 and 4.4 more clearly depicts the co-

occurrences, together with the c of the occurrences per sampling date measured at F1 and C1 respectively.

Figure 4.3 shows the organism numbers measured at F1 and Figure 4.4 the same for sampling site C1 on each sampling date (symbols). The 95% confidence intervals are between the dark-cyan upper and lower limit lines of the mean occurrences (purple lines).

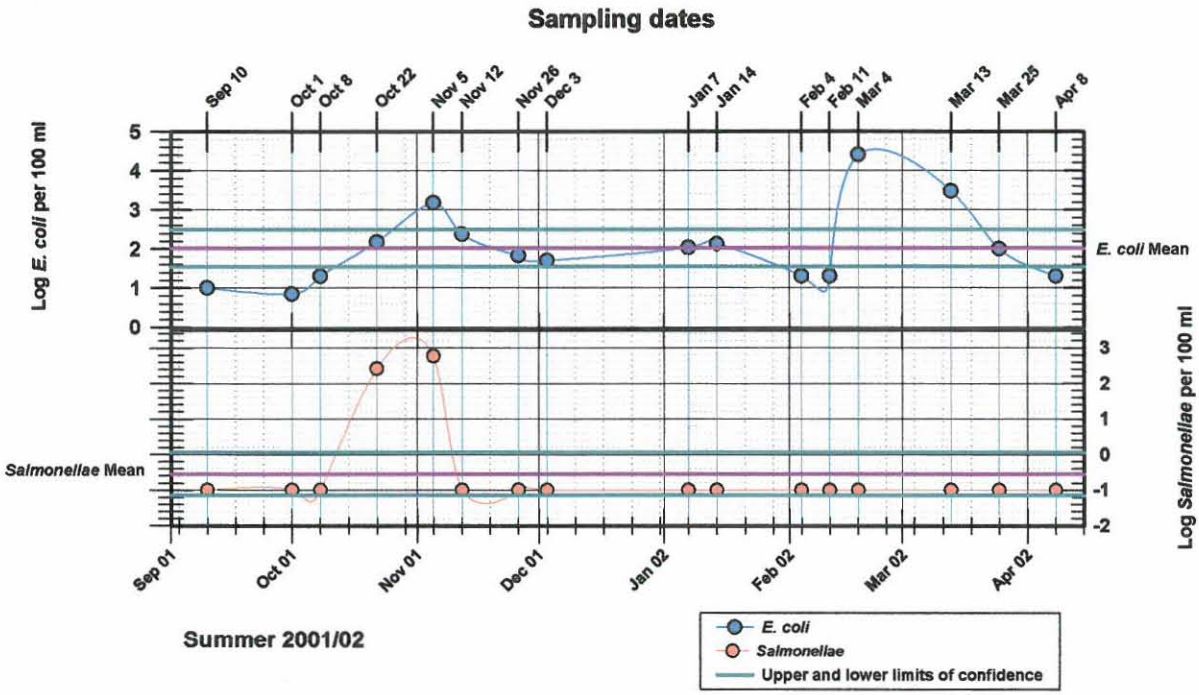



**Figure 4.3:** *E. coli* and *Salmonellae* occurrence in container-stored water at sampling site F1

In both instances (F1 and C1) the varying upper and lower measurements (outliers) were NOT discarded (Appendix C).

As in Chapter 3, these outliers contributed to calculating the confidence intervals even though they could be deemed as chance occurrences and it was assumed that the occurrence and consequent risks were most likely to occur within the confidence intervals. As with the previous chapter, it was within the respective 95% confidence intervals of the mean and 95$^{th}$ percentile occurrence, that the ***possible*** risk of infection posed by the occurrence of *E. coli* could be evaluated against Observed-Adverse-Effect-Levels (OAEL's) found in domestic water quality guidelines based on the OAELA (Chapter 2, Section 3.2.1).

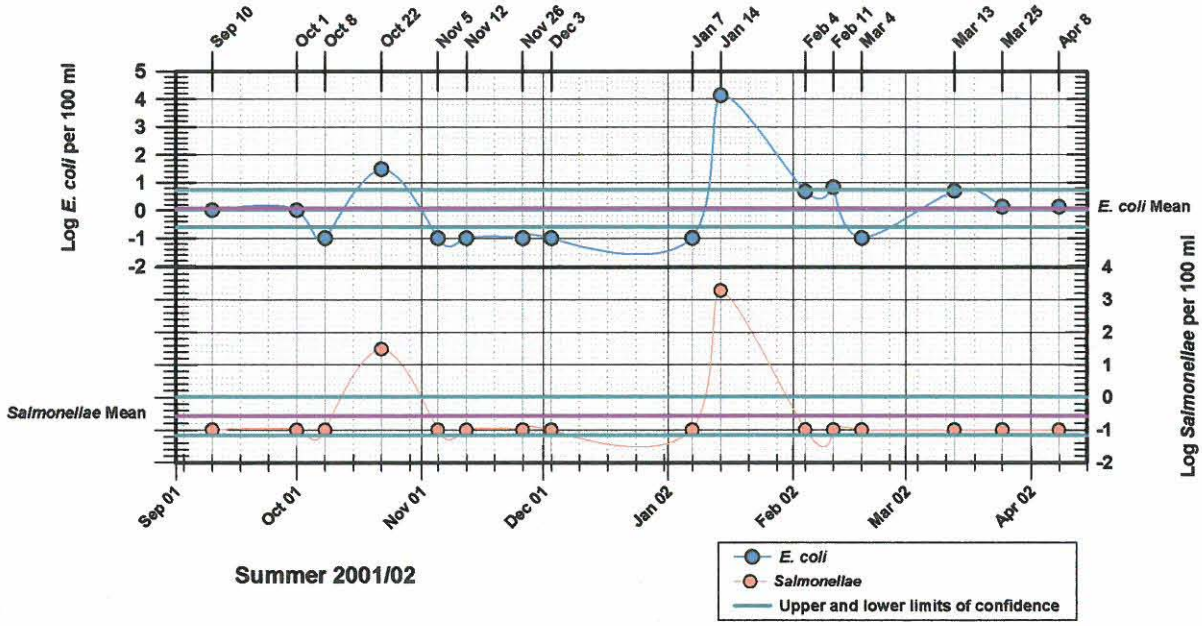



**Figure 4.4:** *E. coli* and *Salmonellae* occurrence in container-stored water at sampling site C1

The QMRA approach (Chapter 2, Section 3.2.2) assessed the ***probable*** risk of infection posed by the *Salmonellae* occurrences on average (mean) as well as at the 95$^{th}$ percentile level.

## 1.4 Uncertainty analyses

A number of sources of uncertainty in applying the WRQMRA process have already been discussed in Chapter 2 (Sections 2.2, 3.6, 4.3, 5.2, and 6.1). The following additional sources of uncertainty were identified:

- Shortage of resources (cost of analyses, time, etc.) limited the number of pathogen and indicator species chosen to test the health-related microbiological quality of untreated and treated water stored in containers at home. Limiting analyses to *Salmonellae* (pathogen) and *E. coli* (indicator) occurrence caused uncertainty in expressing the actual risk of infection for container-stored water.
- It is uncertain to what extent the *E. coli* group indicates the potential presence of *Salmonellae* in container-stored water. *E. coli* is an indicator microorganism group that indicates the potential presence and subsequent risk of infection by a whole

range of pathogens. *Salmo* ne pathogen group whose occurrence is potentially indicated. Actual occurrence of *Salmonellae* should therefore be investigated.

- A single sampling site was chosen to represent the health-related microbiological quality of untreated, as well as treated water stored in containers and used for domestic purposes. It is uncertain to what extent only one sampling site (in each instance) could represent the quality of container-stored water in general.

- While F1 was the major contributor of both *E. coli* and *Salmonellae*, C1 and F1 did not differ significantly with regards to *Salmonellae* numbers. This caused further uncertainty in the use of *E. coli* as indicator of *Salmonellae,* since there was not much of an association (low correlation coefficient) between *E. coli* and *Salmonellae* occurrence in the container-stored water.

- Log-transformed occurrence data was used instead of un-transformed occurrence data. Log-transformation caused the data to be normally distributed and removed a great deal of variance. It is uncertain to what extent the use of this smoothed log-transformed data would cause over- or under estimation of the risk of infection.

- According to Standard Methods (1998), the arithmetic mean gives the best indication of central tendency of data for risk assessment, since it tends to overestimate the risk and therefore ensure safety.

## 2 THE OAEL APPROACH (OAELA) FOR *E. COLI*

The OAELA compares the *E. coli* occurrences to various OAEL's and higher guideline risk limits in order to establish a ***possible*** risk of infection. Section 2.1 discusses the general occurrence of *E. coli*. Section 2.2 evaluates the indicator occurrence against OAEL's and higher guideline limits, for domestic purposes, in order to conclude the infection possibility of the container-stored water.

## 2.1 Numbers of *E. coli* in ng water stored in containers

Figure 4.5 summarises the log *E. coli* component in the untreated spring water illustrated by Figure 4.3 above.

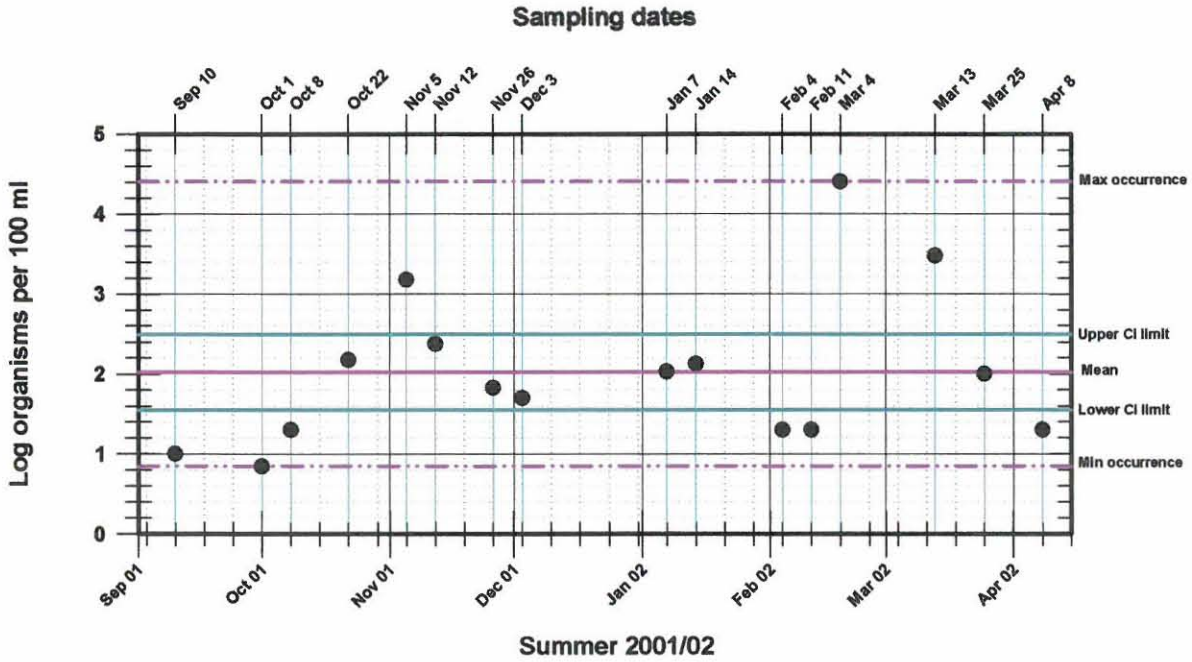



**Figure 4.5:** Log *E.coli* occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site F1 for the sampling period

It portrays more clearly the 95% upper and lower confidence intervals, which form the focus of the OAELA discussions in the following sections. It also shows the minimum and maximum occurrences (ranges) for each sampling date over the 2001/02 sampling period.

## 2.2 Numbers of *E. coli* in treated municipal supply water stored in containers

Figure 4.6 summarises the log *E. coli* occurrence component (previously in Figure 4.4). The 95% confidence intervals as well as the minimum and maximum occurrences (ranges) are shown clearly. The confidence intervals in this figure form the focus of the OAELA discussions in the following sections.

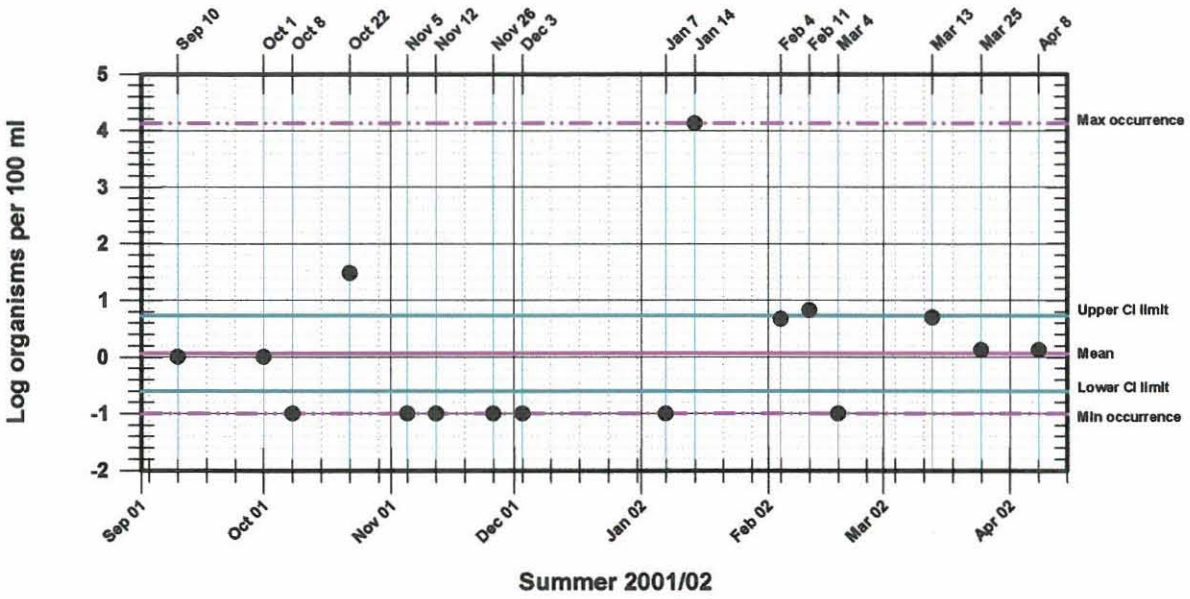



**Figure 4.6:** Log *E.coli* occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site C1 for the sampling period

## 2.3 Risk based on OAELA in the container-stored waters

This chapter focussed only on ingestion (consumption) of the water, whether treated or not, related to domestic water use (intentional ingestion). Risks of health effects associated with ingestion of water are said to increase as *E. coli* levels (per 100 mℓ) increase (DWAF, 1996a and b). As the *E. coli* density increases, the volume of water needed to be ingested in order to cause ill effects, decreases. However, OAELA's generally apply the indicator microorganism (*E. coli* for this study) counts per 100 mℓ.

Figures 4.7 and 4.8 compare the *E. coli* counts /100mℓ in water from C1 and F1 to guideline OAEL's (thick solid lines) for domestic use of the water (Chapter 2; Table 2.4).

In addition to the OAEL's, higher risk limits (indicated by dash-dotted lines) were also applied. This was done to give a further impression of the possible level of risk of gastrointestinal infection indicated by the *E. coli* occurrence in waters used for domestic purposes.

The OAEL lines represents the maximum (upper) limit below which the water is of acceptable (no-to-low risk) health-related microbiological quality for domestic purposes.

The higher risk lines, on the other hand, refer to the minimum / lowest level / limit above which high potential health risks for the various domestic purposes could be expected.

### 2.3.1 Drinking untreated spring water from F1 (Figure 4.7)

Figure 4.7 shows that the untreated water fetched at F1, was totally unfit for human consumption from an *E. coli* indicator perspective.

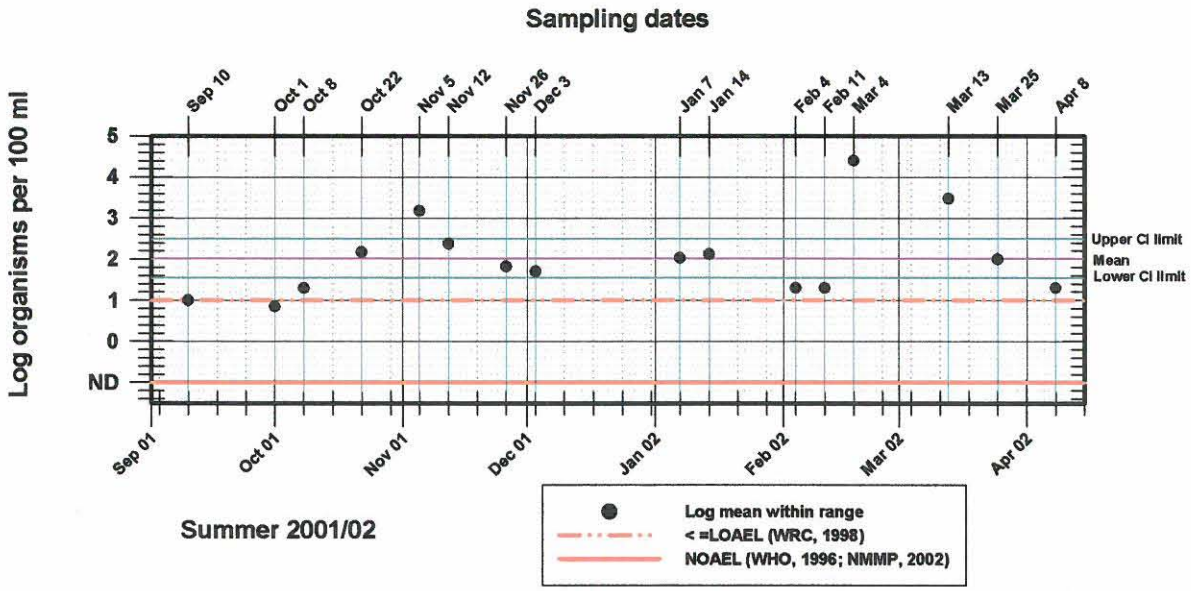



**Figure 4.7:** Comparing the occurrence of *E.coli* to a NOAEL and a LOAEL for drinking untreated water at sampling site F1

### 2.3.2 Drinking container-stored (municipal supply) water at C1 (Figure 4.8)

From a health-related microbiological indicator perspective (based on *E. coli* for this study), the container-stored water was often unfit for human consumption but not frequently above the higher risk limit. Figure 4.8 shows that people ingesting this water would not be exposed to a high potential health risk, since the upper and lower confidence limits of the mean are still within the higher risk limit.

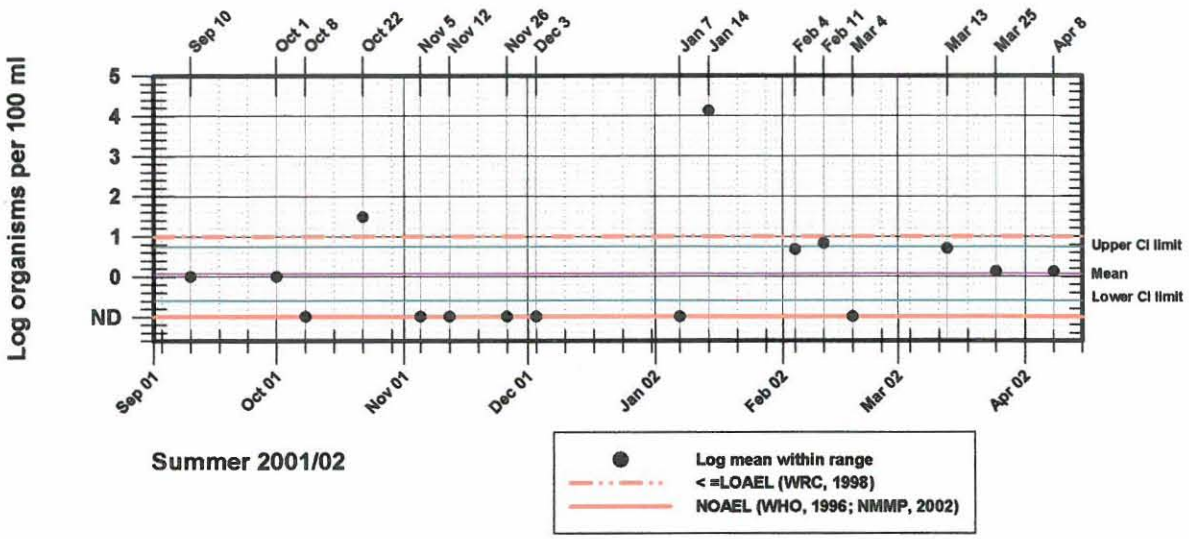



**Figure 4.8:** Comparing the occurrence of *E.coli* to a NOAEL and a LOAEL for drinking untreated water at sampling site C1

## 2.4 Uncertainty analyses

? Application of log-transformed data caused the occurrence band to become narrower. Since the narrower occurrence band will cause the expected risk band to be narrower, it is uncertain what influence this will have on the expression of the risk of infection.

? Since *E. coli* indicates the presence of various pathogenic microorganism groups, it would tend to overestimate the possible risk of infection by *Salmonellae* alone, when compared to guideline values, since these levels are based on a whole range of pathogens.

? It is uncertain to what extent a possible risk of infection indicated by *E. coli* numbers in the water, indicate a risk of infection by *Salmonellae*.

? It is therefore uncertain to what extent the possible risk of infection (based on the OAELA) could indicate the probable risk of infection by *Salmonellae.*

? From a health perspective, it is uncertain to what extent the NOAEL's and LOAEL's could be used to compare the *E. coli* occurrences in container-stored (municipal treated) water and especially untreated water from a spring used for consumption.

The QMRA approach determines whether the numbers of *Salmonellae* (selected hazard) in these waters would lead to a probability of infection, if people were to consume these waters. This section discusses the application of the four QMRA steps (described in Chapter 2) on the container-stored treated and untreated waters.

Section 3.1 of this chapter sums up the occurrence of *Salmonellae* in the container-stored waters (discussed in Section 1 above) as part of the exposure assessment step of the QMRA approach. Section 3.2 discusses the various ingestion volumes associated with intentional ingestion (consumption) of water depending on the age group (discussed in Chapter 2, Table 2.5). Section 3.3 deals with expected doses and Section 3.4 with the probable risk of infection based on dose-response models and parameters (discussed in Chapter 2, Section 4). Risk statements, characterising the probabilities of infection, follow in Section 3.5.

## 3.1 Numbers of *Salmonellae* in untreated spring water

Figure 4.9 shows the occurrence of *Salmonella spp.* in the untreated water at F1 for the 2001/02 sampling period.

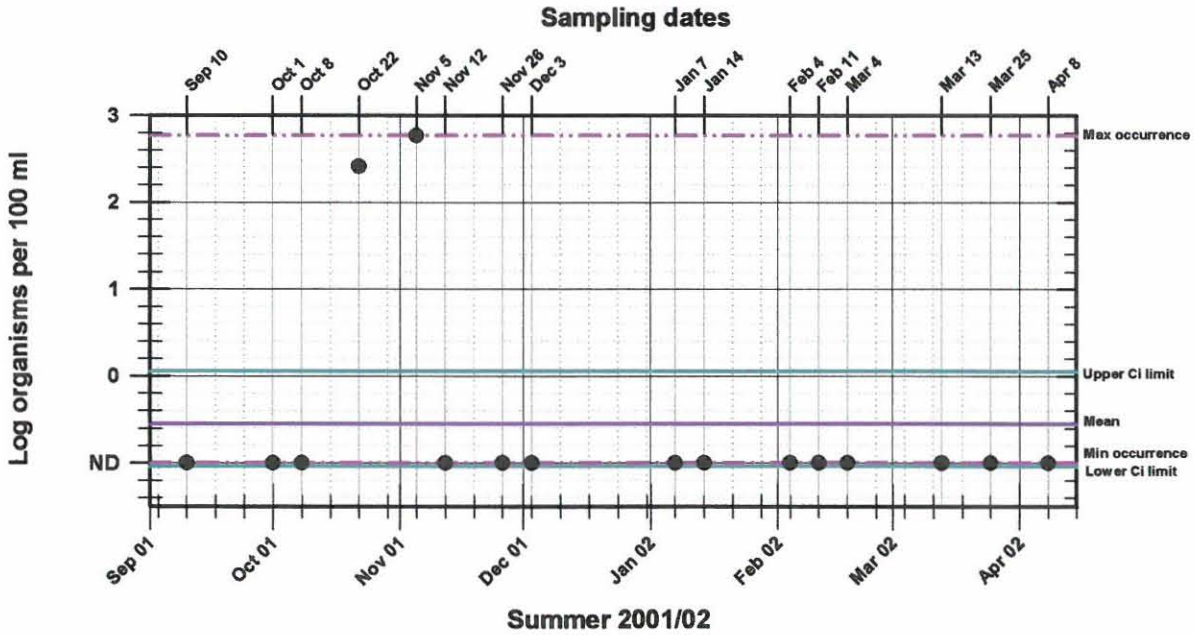



**Figure 4.9:** Log *Salmonellae* occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site F1 for the 2001/02 sampling period

The upper and lower limits of the intervals of the mean indicate the expected limits within which *Salmonellae* (and the associated risk) is most likely to occur. The figure also shows two high peak events (outliers) outside the expected occurrence band within the sampling period.

## 3.2 Numbers of *Salmonellae* in treated municipal supply stored in containers

Figure 4.10 shows the occurrence of *Salmonella spp.* in the treated municipal supply water stored in containers sampled for the 2001/02 summer period. The expected limits within which *Salmonellae* (and the associated risk) was most likely to occur is indicated by the upper and lower limits of the 95% confidence intervals of the mean. As for the untreated water (Figure 4.9), this figure also shows two high peak events (outliers) outside the expected occurrence band within the sampling period.

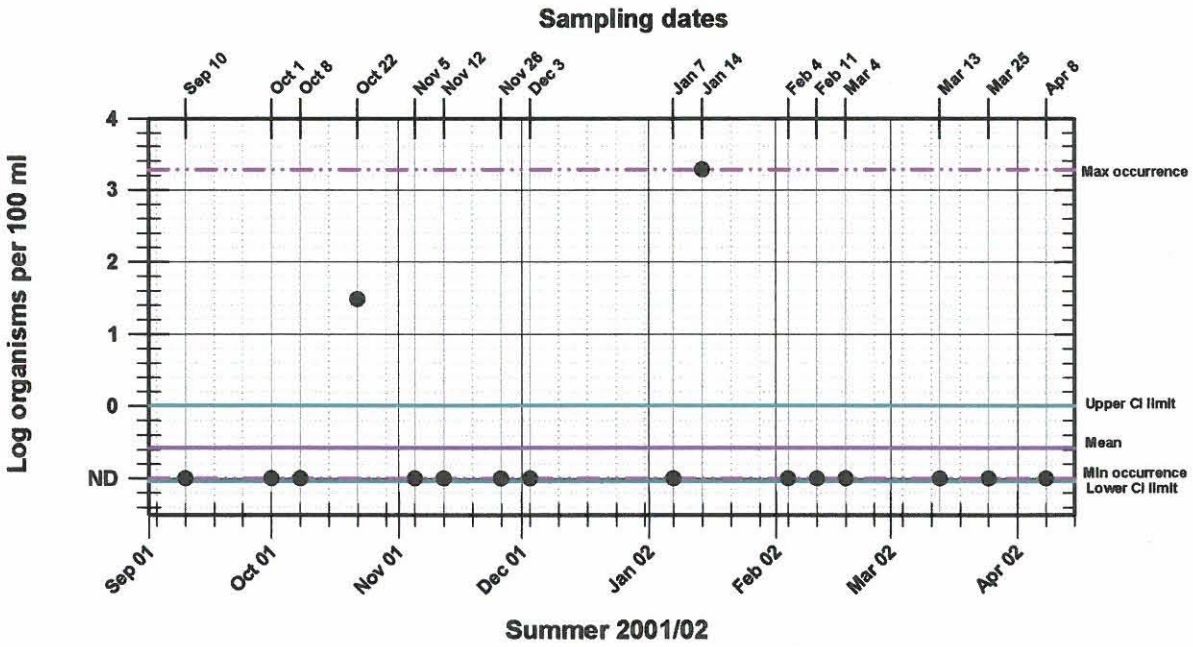



**Figure 4.10:** Log *Salmonellae* occurrence (with upper and lower confidence interval [95% Ci] limits) at sampling site C1 for the 2001/02 sampling period

## 3.3 Intentional daily ingestion volumes from locally-sourced water

The intentional daily ingestion volumes (discussed in Chapter 2, Table 2.5) from locally sourced water were related to specific consumer age-groups. The daily ingestion volumes

(Bourne et al., 1987; 1992) were  sumer age-groups (Roseberry and Burmaster, 1992). The infants had the highest daily intake (1,318 mℓ) volume, with the children between 1 and 11 years of age with the lowest intake volume (630 mℓ) (Appendix D1 and D2).

## 3.4 Dose

The expected dose on exposure is directly related to the numbers of *Salmonellae* numbers ingested per volume unit of water. This implies that the more *Salmonellae* occurred in the container-stored waters, the higher the number a potential user would be likely to ingest. From a *dose* perspective, this also implies that the greater the volume of water ingested, or the greater the number of *Salmonellae* per volume unit ingested, the greater the dose. Dose is calculated with the following equation:

**X /100mℓ x fraction /100mℓ ingested**

where **X** is the counts/100mℓ and

the **fraction /100mℓ** is the volume of water (per 100mℓ) actually ingested.

Table 4.4 shows the geometric mean (Appendix C) dose, as well as the upper (maximum expected) and lower (minimum expected) doses for ingestion based on the 95$^{th}$ percentile and true minimum *Salmonellae* occurrences, respectively, in the container-stored waters.

**Table 4.4:** Expected dose of *Salmonellae* in treated and untreated waters stored in containers

| Sample site | *Salmonellae* occurrence (a) per 100mℓ (c) | | Age groups | Age in years | Ingestion volume mℓ/hd.d (b) | Expected dose (a*b)/100 (c) | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Mean dose | Maximum dose (UL) | Minimum dose (LL) |
| F1 | Geometric mean | 0.2810 | Infants | 0 ≤ age < 1 | 1,318 | 3.70 | 4,481 | 1.32 |
| | | | Children | 1 ≤ age < 11 | 630 | 1.77 | 2,142 | 0.63 |
| | 95$^{th}$ Percentile | 340 | Adolescents | 11 ≤ age < 20 | 773 | 2.17 | 2,628 | 0.77 |
| | Minimum | 0.1 | Adults | 20 ≤ age < 65 | 952 | 2.68 | 3,237 | 0.95 |
| | | | Elderly | 65 ≤ age | 865 | 2.43 | 2,941 | 0.87 |
| C1 | Geometric mean | 0.2646 | Infants | 0 ≤ age < 1 | 1,318 | 3.49 | 6,550 | 1.32 |
| | | | Children | 1 ≤ age < 11 | 630 | 1.67 | 3,130 | 0.63 |
| | 95$^{th}$ Percentile | 497 | Adolescents | 11 ≤ age < 20 | 773 | 2.05 | 3,842 | 0.77 |
| | Minimum | 0.1 | Adults | 20 ≤ age < 65 | 952 | 2.52 | 4,731 | 0.95 |
| | | | Elderly | 65 ≤ age | 865 | 2.29 | 4,299 | 0.87 |

The ingestion volumes discussed in Section 3.4.1, were applied, reflecting potential daily water ingestion (intentional) per consumer age-group (Section 3.3 above).

Table 4.4 shows that users of untreated spring water at F1 would have ingested a mean number of $2.810 \times 10^{-1}$ *Salmonellae* organisms had they ingested 100 mℓ of the water. Had an infant ingested this water at a volume of 1,318 mℓ, he would be exposed to a mean number of 3.7 *Salmonellae*. The same infant could expect to ingest not more than 4,481 and not less than 1.32 *Salmonellae* for the 2001/02 sampling period on any given day at F1.

Had the same infant consumed 1,318 mℓ from the treated source stored in containers (C1), the infant would have ingested a mean number of 3.49 *Salmonellae*. On any given day during the sampling period, this infant could expect to ingest not more than 6,550 and not less than 1.32 *Salmonellae*.

## 3.5 Probable risk of *Salmonellae* infection for container-stored water based on dose-response models and parameters

The *β-Poisson* distribution model, together with dose-response parameters from Haas and Eisenburg (2001) was used to calculate the probability of infection ($P_i$), based on the exposure to the doses summarised in Table 4.4.

$P_i$ in Table 4.5 and Figure 4.11 principally reflect the possible risk posed by a single exposure to the mean dose associated with the consumer age-group (1,318mℓ for infants, 630mℓ for children, etc.). Also shown are the infection risk levels based on ingestion of water varying in quality from less contaminated (minimum expected dose based on true minimum *Salmonellae* detected) to more polluted water (maximum expected dose based on the 95$^{th}$ percentile).

If an elderly person, for instance, ingested 865mℓ (Appendix D) of water at C1 once, the risk of infection would range between $P_i$ of $1.0 \times 10^{-4}$ to $2.48 \times 10^{-1}$ with a mean probable risk of $2.0 \times 10^{-4}$.

**Table 4.5:** Mean probability of *Sal* $P_i$) based on a single exposure event

| Ingestion volume (mℓ/d) | Dose level | F1 | | C1 | |
|---|---|---|---|---|---|
| | | Expected dose | Probability of infection ($P_i$) (Single exposure) | Expected dose | Probability of infection ($P_i$) (Single exposure |
| Infants 1,318 | Mean | 3.70 | 0.0004 | 3.49 | 0.0004 |
| | Maximum | 4,481 | 0.2540 | 6,550 | 0.3096 |
| | Minimum | 1.32 | 0.0001 | 1.32 | 0.0001 |
| Children 630 | Mean | 1.77 | 0.0002 | 1.67 | 0.0002 |
| | Maximum | 2,142 | 0.5194 | 3,131 | 0.2053 |
| | Minimum | 0.63 | 0.0001 | 0.63 | 0.0001 |
| Adolescents 773 | Mean | 2.17 | 0.0002 | 2.05 | 0.0002 |
| | Maximum | 2,628 | 0.1833 | 3,842 | 0.2326 |
| | Minimum | 0.77 | 0.0001 | 0.77 | 0.0001 |
| Adults 952 | Mean | 2.68 | 0.0003 | 2.52 | 0.0003 |
| | Maximum | 3,237 | 0.2096 | 4,731 | 0.2618 |
| | Minimum | 0.95 | 0.0001 | 0.95 | 0.0001 |
| Elderly 865 | Mean | 2.43 | 0.0003 | 2.29 | 0.0002 |
| | Maximum | 2,941 | 0.1973 | 4,299 | 0.2482 |
| | Minimum | 0.87 | 0.0001 | 0.87 | 0.0001 |

In other words, depending on the consumer age-group and associated ingestion volume (which determines dose), for a person that ingests 630mℓ (lowest ingestion volume for all age groups) of reasonably uncontaminated water (C1), the minimum expected $P_i$ would be $1.0 \times 10^{-4}$. Conversely, a $P_i$ of $3.09 \times 10^{-1}$ could be expected at the maximum expected dose level should 1,318mℓ (highest ingestion volume for all age groups) of the most polluted water be ingested.

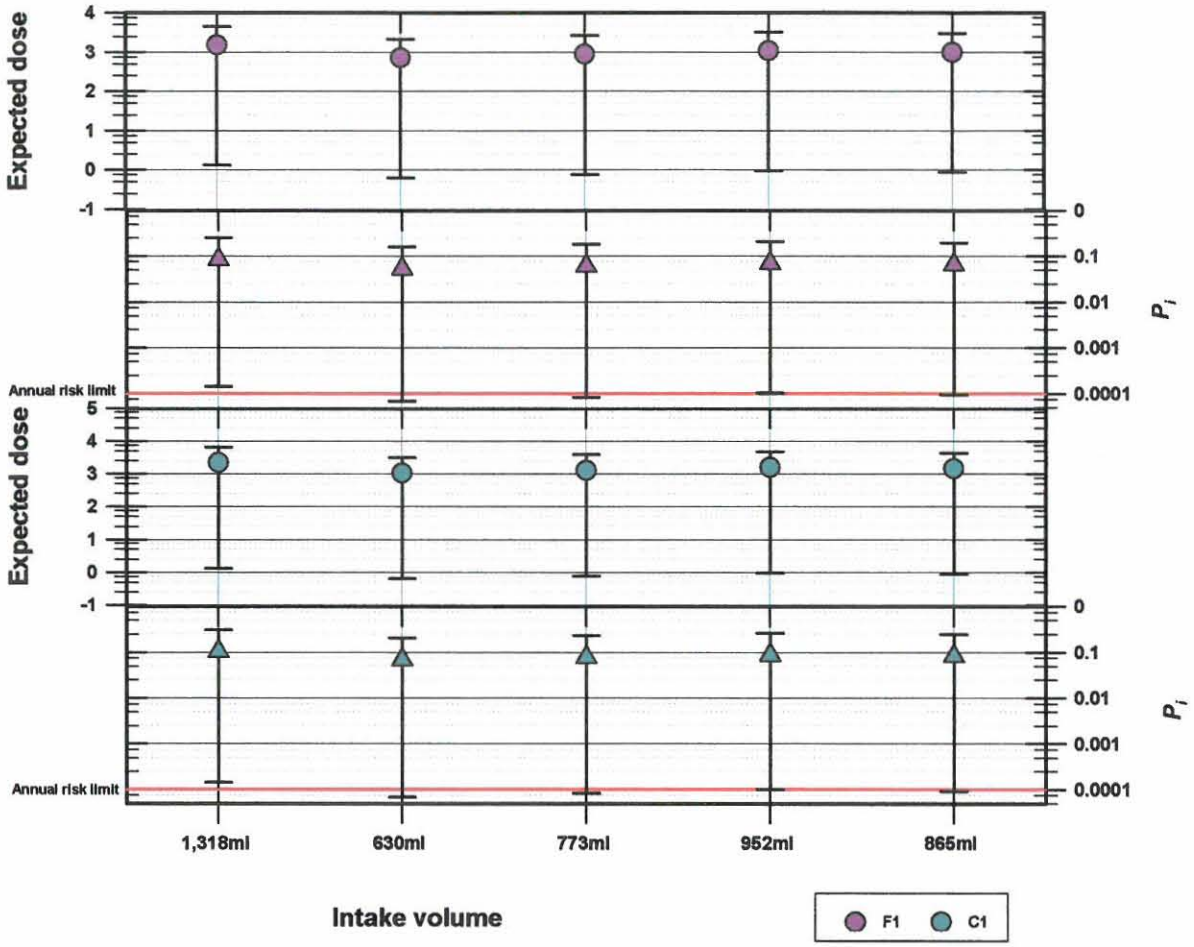



**Figure 4.11:** Expected dose and probability of infection ($P_i$) related to daily intentional water intake volumes

## 3.6 Characterising $P_i$

The single- as well as annual (365 days) exposure risk of infection (Chapter 2, Section 3.2) was calculated with the β-Poisson distribution model (Formulas 1 and 2) discussed in Chapter 2, Section 4.1. $P_i$ was measured for the 2001/02 summer season, but because seasonality, etc. was not tested but assumed not to have an influence on the container-stored water, $P_i$ for a single exposure was extended to a yearly risk of infection. $P_i$ for the annual exposure was therefore based on 365 days (the total number of days in a year).

For a single exposure, as well as for the annual (365 day) exposure, $P_i$ is expressed in percentages, as well as for a fraction of a population of 10,000. This corresponds to the US-EPA (1994) maximum acceptable annual risk limit (0.01% or 1 infection per 10,000 of the population) for consumption of drinking-water (Regli et al., 1991), to which $P_i$, for this

study, is compared. Sections 3. summarised the calculation of the expected dose and subsequent probable risk of infection based on the mean, 95th percentile (maximum) and minimum *Salmonellae* occurrence, as well as the consumer age-group ingestion volumes for F1 and C1. Tables 4.6 and 4.7, as well as Figure 4.12 characterise $P_i$ into a risk statement.

**Table 4.6:** Single and seasonal risk based on $P_i$ for untreated water stored in containers (F1)

| Ingestion volume (mℓ) | Expected level | $P_i$ Single exposure | % $P_i$: Single exposure | Infections per 10,000 after single exposure | % $P_i$: Annual risk (365 days) | Infections per 10,000 after annual exposure |
|---|---|---|---|---|---|---|
| Infants 1,318 | Mean | 0.0004 | 0.04% | 4 | 13.62% | 1,362 |
| | Maximum | 0.2540 | 25.40% | 2,540 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 5.08% | 508 |
| Children 630 | Mean | 0.0002 | 0.02% | 2 | 6.76% | 676 |
| | Maximum | 0.5194 | 51.94% | 5,194 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 2.46% | 246 |
| Adolescents 773 | Mean | 0.0002 | 0.02% | 2 | 8.23% | 823 |
| | Maximum | 0.1833 | 18.33% | 1,833 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.01% | 301 |
| Adults 952 | Mean | 0.0003 | 0.03% | 3 | 10.04% | 1,004 |
| | Maximum | 0.2096 | 20.96% | 2,096 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.70% | 370 |
| Elderly 865 | Mean | 0.0003 | 0.03% | 3 | 9.17% | 917 |
| | Maximum | 0.1973 | 19.73% | 1,973 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.36% | 336 |

A $P_i$ of 1 equals a 100% probability of infection (Chapter 2, Section 5, Table 2.9). For example, Table 4.6 shows that the percentage probability of infection (0.01%) from only a single exposure to the minimum expected dose at an ingestion volume of 1,318 mℓ, already equalled the maximum acceptable annual risk limit (0.01%) suggested by the US-EPA (1994).

As exposure to more contaminated water, at higher ingestion volumes occurs, the closer $P_i$ gets to 1. This implies increasing risk of infection. For example, at any time during the year at a volume of 1,318mℓ ingestion, the maximum percentage $P_i$ for a single exposure is 25.40% (2,540 infections per 10,000).

**Table 4.7:** Single and seasonal ri[illegible]eated water stored in containers (C1)

| Ingestion volume (mℓ) | Expected level | $P_i$ Single exposure | % $P_i$: Single exposure | Infections per 10,000 after single exposure | % $P_i$: Annual risk (365 days) | Infections per 10,000 after annual exposure |
|---|---|---|---|---|---|---|
| Infants 1,318 | Mean | 0.0004 | 0.04% | 4 | 12.88% | 1,288 |
| | Maximum | 0.3096 | 30.96% | 3,096 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 5.08% | 508 |
| Children 630 | Mean | 0.0002 | 0.02% | 2 | 6.38% | 638 |
| | Maximum | 0.2053 | 20.53% | 2,053 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 2.46% | 246 |
| Adolescents 773 | Mean | 0.0002 | 0.02% | 2 | 7.77% | 777 |
| | Maximum | 0.2326 | 23.26% | 2,326 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.01% | 301 |
| Adults 952 | Mean | 0.0003 | 0.03% | 3 | 9.48% | 948 |
| | Maximum | 0.2618 | 26.18% | 2,618 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.70% | 370 |
| Elderly 865 | Mean | 0.0002 | 0.02% | 2 | 8.66% | 866 |
| | Maximum | 0.2482 | 24.82% | 2,482 | 100% | 10,000 |
| | Minimum | 0.0001 | 0.01% | 1 | 3.36% | 336 |

The same example can be used to describe the risk at sampling site C1 (Table 4.7). For instance, the percentage probability of infection (0.01%) from only a single exposure to the minimum expected dose at an ingestion volume of 1,318mℓ, equalled the maximum acceptable annual risk limit (0.01%) suggested by the US-EPA (1994). As exposure to more contaminated water, at higher ingestion volumes occurs, the closer $P_i$ gets to 1, which implies increasing risk of infection. At any time during the year at a volume of 1,318mℓ ingestion, the maximum percentage $P_i$ for a single exposure at C1 is 30.96%.

This indicates that the mean risk of infection is lower for C1 than for F1, but that the infection risk based on the 95$^{th}$ percentile is higher for C1 than for F1. This phenomenon may be result of the high outlying *Salmonellae* occurrences at C1, included in calculating the 95$^{th}$ percentile.

The yearly infection risk calculated for this study, portrays the worst-scene-scenario (ultimate maximum risk limit), as it is unlikely that people will be exposed to the same numbers of *Salmonellae* and the same volumes of water for each of the 365 days.

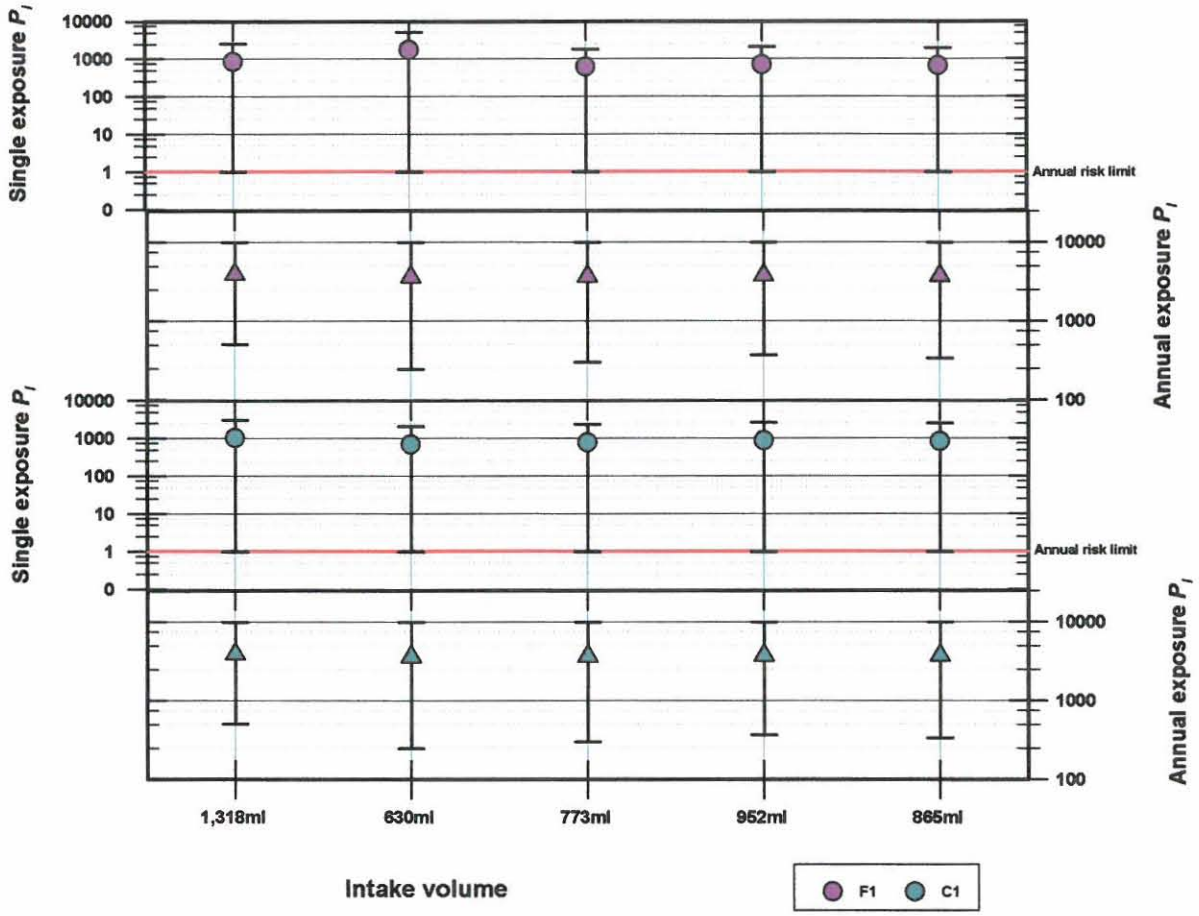



**Figure 4.12:** Single and annual $P_I$ per 10,000 of the population related to daily intentional water intake volumes

## 3.7 Uncertainty analyses

? *Salmonellae* are just one of the pathogen groups potentially present in the waters stored in the containers. The probable risk of infection indicated by the numbers of *Salmonellae* present in the container-stored water may therefore be a gross underestimation of the actual risk in the area.

? The modified ingestion volumes, Chapter 2 (Table 2.5) were extrapolated from ingestion volumes found in literature. These volumes were also associated with a certain consumer age-grouping (also based on literature). It is uncertain whether the modified ingestion volumes, as well as the age-groupings were realistic for the area.

- The number of people per a[illegible]xposed to the container-stored waters was not investigated. It is therefore uncertain what population size was actually at risk of infection.

- The mean as well as 95$^{th}$ percentile dose (over the year) was used. These doses were then applied to describe the probable risk for a single exposure as well as for exposure throughout the year (365 days). Application of these doses could either over- or under estimate the risk of infection.

- Calculating the probable risk of infection for the 365-days causes major uncertainty in that it is unlikely that people would be exposed to the same dose (same number of *Salmonellae* and the same volume of water ingested) for 365 days.

- Use of the US-EPA (1994) maximum acceptable annual risk limit caused further uncertainty. Although this limit is based on consumption of drinking-water, it is most likely based on a developed country situation and not that of a developing country such as South Africa.

- The annual risk of infection is based on the risk of infection calculated for a single exposure and extended over a longer exposure period. This causes uncertainty to whether this is a total over- or under estimation of the risk of infection, since it is unlikely that people will be exposed to the same dose (numbers of *Salmonellae* per volume of water ingested) on a daily basis throughout the year.

- It is uncertain to what extent the QMRA approach used for this study could predict the actual risk of infection to users of treated and untreated water stored in containers at home.

- Since C1 represented treated municipal supply water stored in containers in households, it was assumed that the health-related microbiological quality at C1 would be superior to that of the untreated spring water. However, $P_i$ based on the 95$^{th}$ percentile occurrence and associated dose, was higher for C1 than for F1. It is

uncertain whether this is a ...e health-related microbiological quality of treated and untreated container-stored water and whether this indicated that Salmonellae died-off or was not picked up due to sedimentation.

? It is uncertain to what extent the 95th percentile is influenced by high outlying occurrences, resulting in higher 95th percentiles and therefore in higher doses and subsequent risks of infection.

## 4 PROBABILITY COMPARED TO POSSIBLE RISK OF INFECTION

Section 1 of this chapter discusses how, based on the series of *single* occurrences per date for each organism group, no clear associations existed for the *occurrences* of *E. coli* and *Salmonellae* in the container-stored water. However, based on visual appraisal of Figure 4.13 and 4.14, the *E. coli* and *Salmonellae* occurrences at sampling sites F1 and C1 followed a similar trend. This section investigates whether the *mean* as well as *95th percentile* occurrences of *E. coli* could have reliably (overestimation rather than underestimation) indicated (using the OAELA) the *mean* and *95th percentile* risk posed by *Salmonellae* (using QMRA).

As with Chapter 3, this section is therefore not about whether a risk has occurred or not, but to illustrate whether a QMRA could add value to the typical OAELA that environmental health practitioners would typically follow.

An EHP would tend to follow the OAEL approach based on indicator occurrence. A health worker would typically look at the trend of the *E. coli* occurrence in the spring water (taking into account outlying low and high values), to make a decision on the possible risk of infection for the sampling site over the 2001/02 sampling period.

Figure 4.13 illustrates the yearly, as well as the mean single event *E. coli* and *Salmonellae* occurrence in the untreated spring water stored in containers for the 2001/02 sampling period. The occurrences for both microorganisms at the 95th percentile are also shown. The graph includes the mean, as well as 95th percentile infection probability ($P_i$) for

*Salmonellae*. However, *Pi* was ( one of the consumer age-groups, the infants, since they have the highest daily intake volume of all the consumer age-groups investigated. $P_i$ for the infants therefore constituted the worst-scene scenario risk. The same is shown for the treated municipal supply water stored in containers by Figure 4.14.

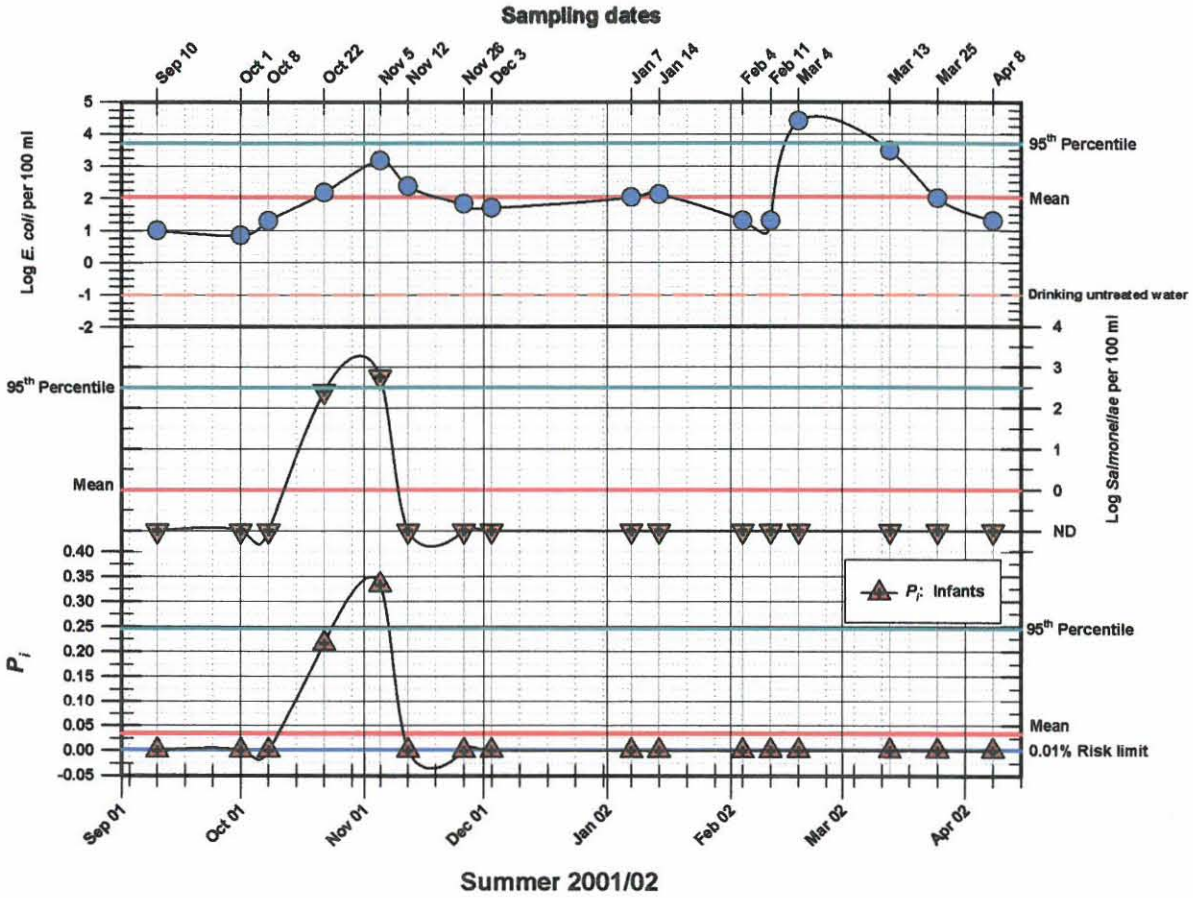



**Figure 4.13:** *E coli*, *Salmonellae* and associated risk (with mean and 95th Percentile) measured at sampling site F1

From Figure 4.13 it is evident that *E. coli* occurred in numbers substantially above all the OAEL's. The mean and 95th percentile risks, as well as certain single event risks posed by *Salmonellae* were also clearly above the acceptable risk for consumption of drinking water suggested by the US-EPA (1994) (0.01% or 1 infection in 10,000 of the population).

High *E. coli* outliers were observed on the 5 November 2001 and 4 March 2002. However, when compared to the probability of infection directly related to the *Salmonellae* occurrence in the spring water, a risk is only indicated for 5 November 2001. What is noticeable though is the fact that the outlying *E. coli* occurrences on 4 March 2002 were much higher

(approximately one order of magn 5 November 2001, giving the impression of a higher risk. In this instance, for example, EHP's would, by looking at the *E. coli* occurrences alone, have overestimated the risk of infection in March. In other words, EHP's would have been unable to reliably (consistently) predict the risk had they only followed the OAELA. However, the probability of infection on the 5 November 2001 was above that of the 95$^{th}$ percentile, which is used to indicate the reasonable maximum exposure. This outlying $P_i$ was therefore not included in further comparing the possible-, with the probable risk of infection. The *E. coli* occurrence on 4 March 2002 exceeded the 95$^{th}$ percentile. This data point was therefore also discarded as an outlier since people were not likely to be exposed to such a high risk. If the outlier on 13 March 2002 is used, however, EHP's following the OAELA, would tend to overestimate the risk of infection by *Salmonellae*.

Figure 4.14 shows a similar tendency for the treated municipal supply water. Again high outlying *E. coli* occurrences are shown for two sampling events (22 October 2001 and 14 January 2002). When compared to the probability of infection directly related to the *Salmonellae* occurrence in the municipal supply water, a risk is only indicated for 14 January 2002. For this day, the highest outlying *E. coli* occurrence (also above the 95$^{th}$ percentile) corresponded with the probability of infection indicated for the same sampling event. Again, EHP's would not have been able to reliably (consistent) indicate the risk of infection based on the *E. coli* occurrence (OAELA) alone. In other words, EHP's would either over- or underestimate the risk and would therefore not be able to give a constant overestimation (for protection of health) of the probable *Salmonellae* infection.

It is evident that the OAELA alone could not indicate the risk of infection in the container-stored municipal supply or spring water for all the sampling dates. However, application of the full WRQMRA approach could be recommended bearing in mind the uncertainties involved in application of this process.

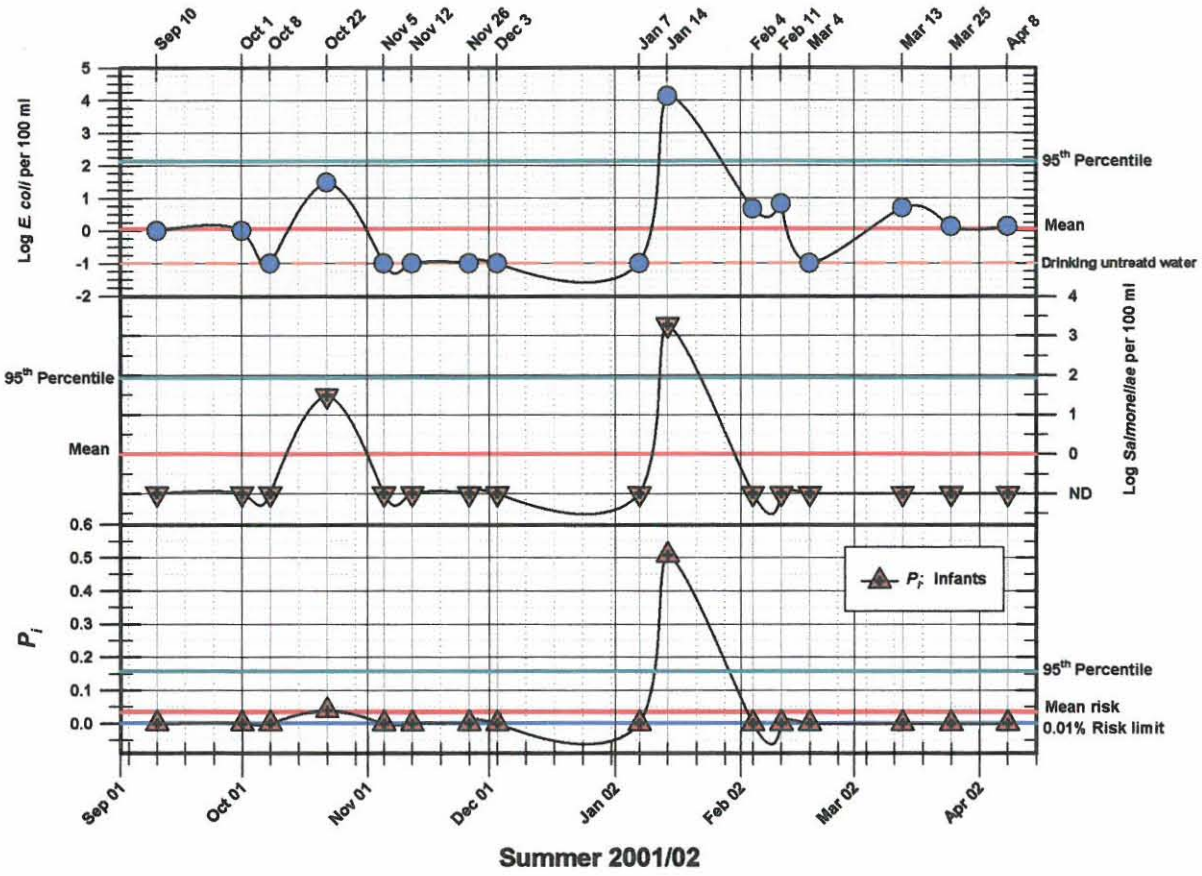



**Figure 4.14:** *E coli*, *Salmonellae* and associated risk (with mean and 95th Percentile) measured at sampling site C1

## 4.1 Uncertainty analyses

? Sections 1.4, 2.4 and 3.6 of this chapter already discussed a number of sources of uncertainty in applying the WRQMRA process in the container-stored waters. These all influenced the uncertainty of using *E. coli* as an indicator of the risk of infection in the area.

? The association between *E. coli* and *Salmonellae* or rather the lack thereof caused uncertainty in whether *E. coli* could indicate the presence of *Salmonellae* and the associated risk of infection.

? Both the OAEL and QMRA approaches indicated a risk of infection to users of untreated waters of the RSQC. Comparing the possible risk of infection by *E. coli* with

the probable risk of infectio ber sampling event, caused uncertainty to whether the costs of analyses, etc by the QMRA approach was justified.

- Acceptable risk limits in terms of the risk of infection should be established for South Africa, and more specifically the study area. It is uncertain to what extent, the fact that an acceptable risk limit has not yet been established for the area, affected the study outcomes.

- However, if single sample event risks (with high and low outlying values) are compared it seems as if *E. coli* would have given an over estimation of the risk on certain sampling events.

- It is uncertain to what extent this over estimation is due to the fact that *E. coli* indicates the potential presence (and therefore the possible infection risk) of ***various*** pathogens and to what extent this overestimation can be justified in terms of risk management in an area.

- It is uncertain which of the two approaches (OAEL and QMRA) could best indicate the risk of infection in the untreated and treated waters, as various uncertainties have been identified which developed around application of these approaches.

- As *Salmonellae* and the probable risk of infection are directly related, it might be that the methodology failed on the sampling events where outlying *Salmonellae* occurrences are seen. This might be a total over estimation of the risk of gastrointestinal infection. This causes further uncertainty, since this could mean that *E. coli* could in fact have reliably indicated the risk of infection.

- It is uncertain whether the waters (treated municipal supply and untreated spring water) would pose a risk of infection to users if applied for other domestic purposes such as body-washing, laundry, etc.

- Although different volumes [illegible]-group was used for the risk prediction, it is uncertain to what extent this risk prediction is reliable in terms of sensitive subpopulations such as infants and the elderly since this was not investigated.
- It is uncertain whether the resources (time, money, etc.) involved in applying the full WRQMRA process was justified in terms of the study-outcomes.

# SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

## 1 SUMMARY

The Water-related Quantitative Microbial Risk Assessment (WRQMRA) process comprised the Observed-adverse-effect-level- (OAEL) and Quantitative Microbial Risk Assessment (QMRA) approaches. These approaches were applied to predict the risk of infection based on ingestion of ① the untreated surface waters of the Renoster Spruit Quarternary Catchment (RSQC) (South-eastern Free State) (Chapter 3) and ② water stored in containers (untreated spring and treated municipal supply) at households near Botshabelo, used for drinking (Chapter 4).

The OAEL approach was based on the occurrence of *E. coli*, of which the numbers were compared to the values in various guidelines (No-observed-adverse-effect-levels (NOAEL's) and Lowest-observed-adverse-effect-levels (LOAEL's)) for the different water uses (e.g., ingestion, full-body immersion, intermediate body contact, irrigation). *E. coli* occurred in considerable numbers in the RSQC and exceeded recommended safety limits for all the intended purposes. These indicated possible risks of infection to people using the surface waters of the RSQC without some form of treatment. Considerable *E. coli* counts were also found in the container-stored water, which indicated a possible risk of infection to people using these waters for drinking and food preparation.

The QMRA approach was based on the occurrence of organisms from the pathogen group *Salmonellae*. The mean occurrence, as well as the 95$^{th}$ percentile was used to calculate the expected dose, based on various water-use activities and associated ingestion volumes. The expected doses (mean, minimum and maximum) were applied in a mathematical dose-response model to determine the probable risk of infection ($P_i$). $P_i$ was expressed as a fraction of the population, as well as a percentage risk and compared to the maximum acceptable annual risk limit for drinking water consumption suggested by the US-EPA

(0.01% or 1 infection per 10,000 Risk scenarios were applied in Chapter 3 to more practicably characterise the probable risk to individuals.

## 2 CONCLUSIONS

A considerable risk of *Salmonellae* infection existed for users of the RSQC waters, users of the untreated spring water, as well as for those drinking water from the treated municipal supply stored in containers at home. Both the OAEL and QMRA approaches indicated potential risk of infection to users of the surface waters of the RSQC and container-stored drinking water throughout the 2001/02 season. It seemed as if the OAELA, based on *E. coli* tend to give a reliable indication of the risk, and since this is a much simpler and more cost-effective method, its use could very well be continued in future to indicate risk of infection. However, the OAELA alone could not give a reliable indication of the risk of *Salmonellae* infection, since it consistently either over- or underestimated the risk of *Salmonellae* infection.

Depending on the desired level of prevention of ill health (e.g., diarrhoeal disease in a community), environmental health practitioners and other health workers would need a quantitative estimate of the people at risk of infection. Such a quantitative prediction is an advantage of the QMRA if added to the OAEL approach – in other words applying a full WRQMRA as suggested by this study.

## 3 RECOMMENDATIONS FOR FUTURE WORK

A considerable number of uncertainties were involved in applying both these approaches. A few recommendations for future application of the WRQMRA process is listed based on these uncertainties:

- A wider range of pathogenic microorganisms (e.g., viruses, protozoan parasites), and for that matter associated indicator microorganisms where possible, should be included, since this could eliminate much of the uncertainty for future application of the WRQMRA process.

- An acceptable risk limit sh[illegible] for South Africa, and more specifically the study area, in order to more practicably characterise the risk of *Salmonellae* (and other pathogenic microorganism) infections.
- Actual daily ingestion volumes should be investigated throughout South Africa and established for all race and age groups, since the modified ingestion volumes extrapolated for this study were based on only a limited number of studies done previously.
- More research should be done on the actual volumes of water ingested during various water-use activities, such as those described in the RSQC (e.g., laundry, fishing, body-washing) (Chapter 3).
- Research should be conducted and concluded on the quantitative association between *Salmonellae* and *E. coli* occurrences in water, especially in the study area, since little information is available, in the study area specifically, to verify occurrence for these microorganisms.
- The application of the WRQMRA process (based on the above-mentioned recommendations) should be practiced more often and in more areas throughout South Africa, keeping in mind the uncertainties involved throughout the process application.
- The importance of Risk Assessment and application of the QMRA approach in addition to the application of the OAEL approach (typically applied by environmental health practitioners) should be work-shopped and communicated to all health workers on all levels (e.g., community workers, managers).
- Environmental health managers should communicate the risk of infection predicted in this study, to community health workers in the study area, in order to manage prevention of the probable *Salmonellae* risk.

# REFERENCE LIST

**Anderson, JM.** (2001). Prospects for international guidelines for water recycling. Water 21, August. pp. 16-21.

**Beliaeff, B. and Mary, JY.** (1993). The "Most Probable Number" estimate and its confidence limits. Wat. Res. Vol. 27 (5). pp. 799-805.

**Benenson, AS.** (1995). Control of Communicable Diseases Manual. (Ed.). (16th Edn.). An official report of the American Public Health Association, Washington DC. 576 pp.

**Bergey's Manual®.** (1994). Bergey's Manual® of Determinative Bacteriology. (9th Edn.). Holt, JG., Krieg, NR., Sneath, PHA., Staley, JT. and Williams, ST. (Eds.). International Edition. Williams and Wilkins Publishers, Baltimore. 787 pp.

**Biolab Catalogue.** (1997). Version 2. Merck.

**bioMérieux.** (2000). API 20 E – Identification system for *Enterobacteriaceae* and other Gram-negative rods. Catalogue No. 07584 B – 04/98.

**Blumenthal UJ., Peasey A., Ruiz-Palacois, G. and Mara, DD.** (1999). Guidelines for wastewater reuse in agriculture and aquaculture: Recommended revisions based on new research evidence. Water and Environmental Health at London and Loughborough (WELL) study. June. Task No. 68 (Part 1).

**Bokako, TC.** (2000). Health risk related to water supply and consumption in a marginalised urban area. M. Tech: Environmental Health Dissertation. Technikon Free State, Bloemfontein.

**Bourne, DE. and Coetzee, N.** (1996). An atlas of potentially water-related diseases in South Africa. Volume 1. Mortality 1990. WRC Report No. 584/1/96. Water Research Commission, Pretoria. 49 pp.

**Bourne, LT., Bourne, DE., Watermeyer, GS. and Klopper, JML.** (1987). A liquid consumption survey of individuals in greater Cape Town. WRC Report No. 74/2/87. Water Research Commission, Pretoria. 162 pp.

**Bourne, LT., Bourne, DE., Steyn, [illegible] I, C.** (1992). Liquid Consumption Patterns Among the Black Population of Cape Town. WRC Report No. 334/1/92. Water Research Commission, Pretoria. 51 pp.

**Centres for Disease Control and Prevention (CDC).** (2000). Division of Bacterial and Mycotic Diseases. Disease information: Salmonellosis. Internet: http://www.cdc.gov/ncidod/dbmd/diseaseinfo/salmonellosis_t.htm.

**Covello, VT. and Merkhofer, MW.** (1993). Risk Assessment Methods. Approaches for Assessing Health and Environmental Risks. Plenum Press, New York. 309 pp.

**Craun, GF.** (1986). Waterborne disease in the United States. CRC Press, Inc. Boca Raton, Florida, USA.

**Craun, G.F.** (1993). Safety of Water Disinfection: Balancing Chemical and Microbial Risks. (Ed.). International Life Sciences Institute (ILSI) Press, Washington, DC. 690 pp.

**Department of Water Affairs and Forestry (DWAF).** (1996a). South African Water Quality Guidelines. Vol. 1: Domestic Water Use. (2nd Edn.). Department of Water Affairs and Forestry, Pretoria.

**Department of Water Affairs and Forestry (DWAF).** (1996b). South African Water Quality Guidelines. Vol. 2: Recreational Water Use. (2nd Edn.). Department of Water Affairs and Forestry, Pretoria.

**Department of Water Affairs and Forestry (DWAF).** (2002). National Microbial Monitoring Programme for Surface Water. Implementation Manual. Department of Water Affairs and Forestry, Pretoria, South Africa.

**Du Preez, M., Venter, SN., Theron, J., Matlala, M., Gericke, M. and Singmin, Y.** (2001). Enteropathogens in water; rapid detection techniques, occurrence in South African waters and evaluation of epidemic risks (health-related). WRC Report No. 741/1/01. Water Research Commission, Pretoria. 41 pp.

**Ershow, AG. and Cantor, KP.** (1989). Total water and Tap water Intake in the United States: Population-based estimates of Quantities and Sources. Life Sciences Research office, Bethesda.

**Food Safety and Inspection Service (FSIS).** (1996). Food Safety and Inspection Service. Part 2. Pathogen Reduction; Hazard Analysis and Critical Control Point (HACCP) Systems; Final Rule. Federal Register. Vol. 61. (144). 51 pp.

**Genthe, B. and Franck, M.** (199 , ssing microbial water quality in small community water supplies: an $H_2S$ strip test. WRC Report No. 961/1/99. Water Research Commission, Pretoria. 33 pp.

**Genthe, B. and Kfir, R.** (1995). Studies on microbiological drinking water quality guidelines. WRC Report No. 469/1/95. Water Research Commission, Pretoria. 38 pp.

**Genthe, B. and Rodda, N.** (1999). Application of Health Risk Assessment Techniques to Microbial Monitoring Data. WRC Report No. 470/1/99. Water Research Commission, Pretoria. 81 pp.

**Genthe, B. and Seager, J.** (1996). The effect of water supply, handling, and usage on water quality in relation to health indices in developing communities. WRC Report No. 562/1/96. Water Research Commission, Pretoria. 63 pp.

**Gerba, CP. and Rose, JB.** (1993). Comparative Environmental Risk Assessment. Chapter 9. Estimating Viral Disease Risk from Drinking Water. pp. 117-135.

**Giannella, RA.** (2001). Book of Medical Microbiology. Chapter 21 – *Salmonella*. Internet: http://gsbs.utmb.edu/microbook/ch021.htm.

**Glantz, SA.** (1997). Primer of Biostatistics. (4th Edn.). McGraw-Hill, New York.

**Gray, A.** (2001). World Health and Disease. (Ed.). Health and Disease Series. (3rd Edn.). Alden Group, Oxford. 352 pp.

**Griesel, M.** (2001). The effect of various urban discharges on the microbiological water quality in catchment systems: an environmental health-related impact study. M. Tech: Environmental Health Dissertation. Technikon Free State, Bloemfontein.

**Griesel, M. and Jagals, P.** (2002). Faecal indicators in the Renoster Spruit system of the Modder-Riet River catchment and implications for human users of the water. Water SA. Vol. 28 (2). pp. 227-234.

**Haas, CN.** (1983). Estimation of risk due to low doses of microorganisms: a comparison of alternative methodologies. Am. J. Epidemiol. Vol. 13. pp. 545-552.

**Haas, C. and Eisenberg, JNS.** (2001). Chapter 8: Risk Assessment. *In:* Fewtrell, L. and Bartram, J. (Eds.). Water Quality: Guidelines, Standards and Health. Assessment of risk and risk management for water-related infectious disease. World Health Organisation (WHO) Water Series. IWA Publishing, London. 424 pp.

**Haas, CN., Rose, JB. and Gerba, [illegible]** antitative Microbial Risk Assessment. John Wiley & Sons, Inc., New York. 449 pp.

**Helsel, DR. and Hirsch, RM.** (1995). Statistical Methods in Water Resources. Elsevier Science BV. Amsterdam, The Netherlands.

**Hunter, PR.** (2002). Drinking water and diarrhoeal disease due to *Escherichia coli*. Journal of Water and Health. In Press.

**International Life Science Institute (ILSI)-Risk Science Institute (RSI) (ILSI-RSI).** (1996). A Conceptual Framework to Assess the Risks of Human Disease Following Exposure to Pathogens. Risk Analysis. Vol. 16 (6). pp. 841-848.

**Jagals, P.** (1994). The effect of diffuse effluents on the quality of water in the Modder River. MDip Tech Dissertation. Technikon Free State, Bloemfontein.

**Jagals, P.** (1997). Stormwater runoff from typical developed and developing South African urban developments: definitely not for swimming. Wat. Sci. Tech. Vol. 35. (11-12).

**Jagals, P.** (2000). The impacts of polluted urban run-off on the Modder River Catchment: A microbiological perspective. D. Tech: Environmental Health Thesis. Technikon Free State, Bloemfontein.

**Jagals P., Grabow, WOK. and De Villiers, JC.** (1995). Evaluation of indicators for assessment of human and animal faecal pollution of surface run-off. Wat. Sci. Tech. Vol. 24 (5 – 6). pp. 235 – 241.

**Jagals, P., Grabow, WOK. and Williams, E.** (1997). The effects of supplied water quality on human health in an urban development with limited basic subsistence facilities. Water SA. Vol. 22 (3). pp. 235-238.

**Jagals P., Grabow, WOK., Griesel, M. and Jagals, C.** (2000). Evaluation of Selected Membrane Filtration and Most Probable Number Methods for the Enumeration of Faecal coliforms, *Escherichia coli* and Enterococci in Environmental Waters. Kluwer Academic Publishers, Boston, USA. Quantitative Microbiology. Vol. 2 (2). pp. 129-140 (12).

**Katzenellenbogen, JM., Joubert, G. and Karim, SSA.** (1997). Epidemiology. A manual for South Africa. (Eds.). Oxford University Press. Cape Town, South Africa. 295 pp.

**KidsHealth**. (2000). The Nemours Foundation. Salmonellosis. Internet: http://www.kidshealth.org/parent/infections/bacterial_viral/salmonellosis_prt.htm.

**Kindzierzki, WB. and Jackson,** uation of Canadian drinking water guidelines using probability modelling of population exposure. Wat. Sci. Tech. Vol. 38 (6). pp. 229-236.

**Kolluru, RV., Bartell, SM., Pitblado, RM. and Stricoff, RS.** (1996). Risk Assessment and Management Handbook. For Environmental, Health, and Safety Professionals. (Eds.). McGraw-Hill, Inc, New York.

**Koren, H.** (1991). Handbook of Environmental Health and Safety: Principles and Practices. (2nd Edn.). Lewis Publishers, Chelsea.

**Liversage, G.** (2001). Correspondence. Botshabelo. Free State Province, South Africa.

**Medema, GJ.** (2002). Correspondence. KIWA Water Research, The Netherlands.

**Medema, GJ., Ketelaars, HAM. and Hoogenboezem, W.** (2001). *Cryptospordium* and *Giardia*: occurrence in sewage, manure and surface water. (Eds.). Association of River Waterworks – RIWA, Amsterdam. 171 pp.

**Merck Corporation**. (1996). Coliforms show their true colours. Catalogue. No. 1.10426. Merck KgaA, 64271 Darmstadt, Germany.

**Microsoft® Excel.** (2002). Microsoft® Excel Help Files: Percentiles.

**Millipore Corporation**. (1992). Water Microbiology. Laboratory and Field Procedures. Catalogue No. AD323 Rev.3/92 92-217. Bedford, Massachusetts.

**Montaigne, F. and Essick, P.** (2002). Water Pressure. Journal of the National Geographic Society. September. pp. 2-32.

**Nala, NP.** (2002). The impact of an Educational Intervention on the Microbiological Infection Risk posed by water stored in households. M. Tech: Environmental Health Dissertation. Technikon Free State, Bloemfontein.

**OXOID Corporation.** (1990). The OXOID Manual. 6th Edition. Bridson, EY (Ed.). Alphaprint, Alton.

**OXOID Corporation.** (1997). Food-borne Pathogens. Monograph Number 1. *Salmonella*. Post, DE (Ed.).

**Payment, P.** (1997). Epidemiolc ...strointestinal and respiratory diseases: incidence, fraction attributable to tap water and costs to society. Wat. Sci. Tech. Vol. 35 (11). pp. 7-10 (4).

**Payment, P. and Franco, E.** (1993). *Clostridium perfringens* and Somatic Coliphages as Indicators of the Efficiency of Drinking Water Treatment for Viruses and Protozoan Cysts. Appl. Environ. Microbiol. Vol. 59 (8). pp. 2418-2424.

**Payment, P. Siemiatycki, J., Richardson, L., Renaud, G., Franco, E. and Prevost, M.** (1997). A prospective epidemiological study of gastrointestinal health effects due to the consumption of drinking water. International Journal of Environmental Health Research. Vol. 7 (1). pp. 5-31 (27).

**Payment, P., Berte, A., Prevost, M., Menard, B. and Barbeau, B.** (2000). Occurrence of pathogenic microorganisms in the Saint Lawrence River (Canada) and comparison of health risks for populations using it as their source of drinking water. Can. J. Microbiol. Vol. 46. pp. 565-576.

**Pearson, JCG. and Turton, A.** (1993). Statistical Methods in Environmental Health. Chapman and Hall, London. 184 pp.

**Polo, FL., Figueras, MJ., Inza, I., Sala, J., Fleisher, JM. and Guarro, J.** (1998). Relationship between presence of *Salmonella* and indicators of faecal pollution in aquatic habitats. FEMS Microbiol. Lett. Vol. 160. pp. 253-256.

**Potgieter, E.** (2002). Personal interview. Bloemfontein.

**Pretorius, E.** (1996). An investigation into the effects of various levels of sanitation on surface water quality in a typical developing community. M. Tech Dissertation. Technikon Free State, Bloemfontein.

**Pretorius, E.** (2002). The impact of socio-economic and human behavioural factors on the water of the Fontein Spruit Catchment – A water Management Model Study in a developing community. D. Tech Civil-Engineering Thesis. Technikon Free State, Bloemfontein.

**Pretorius, E. and De Villiers, GDT.** (1999). An integrated approach to the management of water quality in a developing South Africa. *In:* Ellis, JB. (Ed.). Impacts of Urban Growth on Surface Water and Groundwater Quality. Proceedings of international symposium held by the International Union of Geodesy and Geophysics (IUGG). 18-30 July. Birmingham, UK. International Association of Hydrological Sciences (IAHS) Publication no. 259.

**Pretorius, E. and De Villiers, GL ..** (2002). ..npact of Informal Living Conditions on water Quality in the Bloemfontein Municipality. South African Geographical Journal. Vol. 84 (2). pp. 199-207.

**Regli, S., Rose, JB., Haas, CN. and Gerba, CP.** (1991). Modelling the risk from *Giardia* and viruses in drinking water. Journal of the American Water Works Association. Vol. 83. pp. 76-84.

**Risk*Assistant™.** (1995). Hampshire Research Institute. Thistle Publishing.

**Robinson, CH., Lawler, MR., Chenoweth, WL. And Garwick, AE.** (1982). Normal and Therapeutic Nutrition. (17th Edn.). Macmillan Publishing Company, New York. pp. 673-677.

**Rose, JB.** (1997). Environmental Ecology of *Cryptosporidium* and public health implications. Annu. Rev. Public Health. Vol. 18. pp. 135-161.

**Rose, JB. and Gerba, CP.** (1991). Use of Risk Assessment for Development of Microbial Standards. Wat. Sci. Tech. Vol. 24. (2). pp. 29-34.

**Rose, JB., Haas, CN. and Regli, S.** (1991). Risk Assessment and the control of waterborne giardiasis. Am. J. Public Health. Vol. 81. pp. 709-713.

**Roseberry, AM. and Burmaster, DE.** (1992). Log-normal distributions for water intake by children and adults. Risk Analysis. Vol. 12 (1). pp. 9-104.

**South African Bureau of Standards (SABS).** (1984). South African Standard Specification for Water for Domestic Supplies. SABS Method 241-1984. South African Bureau of Standards, Pretoria. 15 pp.

**South African Bureau of Standards (SABS).** (1987). Bacteriological quality of water by the membrane filter method. SABS Method 221. South African Bureau of Standards, Pretoria. 6 pp.

**SigmaPlot® 8.** (2002). Exact graphs for exact sciences: Users Manual. SPSS Science Inc. Chicago.

**SigmaStat® Version 2.03.** (1997). Statistical software: Users Manual. SPSS Science Inc. Chicago.

**Skivington, P.** (1997). Risk As: r Quality Management. WRC Project No. TT 90/97. Water Research Commission, Pretoria. 27 pp.

**Sobsey, MD., Handzel, TR. and Venczel, LV.** (2002). Chemical Disinfection and Safe Storage of Household Drinking Water in Developing Countries to Reduce Waterborne Disease. International Water Association (IWA) Conference. Melbourne, Australia.

**Standard Methods.** (1998). Standard Methods for the Examination of Water and Wastewater. (20th Edn.). Clesceri, LS. Eaton, AD. and Greenberg, AE. (Eds.). American Public Health Association, Washington DC.

**Teunis, PFM., Medema, GJ., Kruidenier, L. and Havelaar, AH.** (1997). Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. Wat. Res. Vol. 31 (6). pp. 1333-1346.

**Theron, J.** (2001). Laboratory Manual in Microbiology and Plant Pathology. (Ed.). Department of Microbiology and Plant Pathology. University of Pretoria, South Africa. 252 pp.

**Theron, L.** (2000). Application of a water-related environmental health epidemiological process: A guide for environmental health practitioners. M.Tech. Environmental Health Dissertation. Technikon Free State, Bloemfontein.

**Tillet, HE.** (1993). Potential inaccuracy of microbiological counts from routine water samples. Wat. Sci. Tech. Vol. 27. pp. 15-18.

**United States Environmental Protection Agency (US-EPA).** (1994). National Primary drinking water regulations: Enhanced surface water treatment requirements; proposed rule. Fed. Reg. 59. pp. 38832-38858.

**United States-Environmental Protection Agency (US-EPA).** (1995). Guidance for Assessing Chemical Contaminant Data for use in Fish Advisories. Vol. 2: Risk Assessment and Fish Consumption Limits. (3rd Edn.). Appendix D: Guidance on Risk Characterization. Internet: http://www.epa.gov/waterscience/fishadvice/Volume2

**Venter, SN., Steynberg, MC., Du Plessis, G., De Wet, CME., Hohls, D., Rodda, N. and Kfir, R.** (1996). Tools for microbial water quality assessment of South African rivers. WRC Report No. 380/1/96. Water Research Commission, Pretoria. 780 pp.

**Wadsworth, HM.** (1989). Handbook of Statistical Methods for Engineers and Scientists. McGraw-Hill, New York.

**Ward, RL. and Akin, EW.** (1984, ...ious Dose of Animal Viruses. CRC Press. Rev. Environ. Contr. (14). pp. 297-310.

**Water Research Commission (WRC).** (1998). Quality of Domestic Water Supplies. Vol. 1: Assessment Guide. (2nd Edn.). Department of Water Affairs and Forestry & Department of Health, Pretoria.

**World Health Organisation (WHO).** (1994). An Evaluation of Infant Growth. WHO Working group on Infant Growth. Nutrition Unit. World Health Organisation, Geneva.

**World Health Organisation (WHO).** (1996). Guidelines for drinking-water quality. (2nd Edn.). Volume 2. Health criteria and other supporting information. World Health Organisation, Geneva. 973 pp.

**World Health Organisation (WHO).** (1998). Guidelines for Safe Recreational-water Environments: Coastal and Fresh-waters (Draft). For Consultation, Geneva. EOS/DRAFT/98.14. 251 pp.

**World Summit on Sustainable Development (WSSD).** (2002). World Summit on Sustainable Development. Outreach – 2002. Stakeholder Forum. Issue III. Internet: www.earthsummit2002.org

**Yates, MV.** (1999). Pathogens in reclaimed water. Food, Agricultural and Biological Engineering. University of California, Riverside. Internet: http://www.geoflow.com.wastewater.pathogens.htm.

Appendix A

# BACTERIAL PATHOGEN ANALYSES – *Salmonella spp.*

## MOST PROBABLE NUMBER (MPN) TECHNIQUE

The methods described by Oxoid Corporation (1990, 1997), Standard Methods (1998) as well Haas et al. (1999), were used as basis for the equipment and procedures for the MPN technique used for *Salmonellae* analyses.

## 1 EQUIPMENT

### 1.1 Glassware

- 250-mℓ screw-cap Schott® bottles.
- 20-mℓ test tubes fitted with screw-caps.
- 25-mℓ and 250-mℓ sterile graduated cylinders.

### 1.2 Pipettes

- 0.1-mℓ and 1-mℓ Finnpipette® adjustable pipettes (calibration errors were checked not to exceed 2.5%).
- Sterile disposable pipette tips (for dispensing 0.1-mℓ and 1-mℓ).
- Standard graduated glass pipettes for larger volumes.

### 1.3 Incubators

- Labocon and Scientific incubators fitted with fans for air circulation (temperatures varied within 0.5°C accuracy, especially within test tube racks and stacks of incubated plates).
- Water baths (25-ℓ) equipped with gabled covers (to aid temperature maintenance within 0.2°C of setting) and uniformly distributed heating elements in steel inner jacket (to ensure constant temperature distributions) were used.

### 1.4 Other

- Inoculation needles
- Gas flame
- Ethanol

## 2. PREPARATION / PROCEDURES

### 2.1 Sterilisation

Steam sterilisation of equipment was done in an autoclave at 121°C for 15 minutes. Glassware was steam-sterilised at 121°C for 30 minutes as described by Standard Methods (1998) and Millipore Corporation (1992).

### 2.2 Resuscitation or pre-enrichment media - Buffered Peptone Water (BPW)

- 20 g of the powder was added to 1 litre of distilled water with a low mineral content / conductivity.
- The suspension was mixed well and distributed into the final containers (sterile 250-mℓ Schott® bottles and screw-capped test tubes).
- The media was steam-sterilised by autoclaving at 121°C for 15 minutes.

### 2.3 Enrichment media - Rappaport-Vassiliadis (RV) Enrichment Broth

- 30 g of the powder was added to 1 litre of distilled water with a low mineral content / conductivity.
- The suspension was gently heated and frequently stirred until the powder was totally dissolved.
- 10-mℓ volumes were dispensed, distributed into sterilised screw-capped test tubes and steam-sterilised by autoclaving at 115°C for 15 minutes.

## 2.4 Plating media - XLD Me

- 53 g of the powder was added to 1 litre of distilled water with a low mineral content / conductivity.
- The suspension was gently heated and frequently stirred until the medium boiled.
- On boiling, it was immediately removed from the heat and transferred to a water bath at 50°C.
- The medium was poured into 60 mm petri-dishes as soon as the medium cooled.

## 2.5 Pre-enrichment dilutions

The sterile 250-mℓ screw-capped Schott® bottles were prepared according to the number of selected dilutions:

- 225-mℓ Buffered Peptone Water (BPW) has already been added to each bottle, and sterilised (Section 2.2).
- Water samples were vigorously shaken to homogeneously mix the contents.
- 25-mℓ of water sample was immediately transferred from the Whirlpack® sample to the 1st bottle of 225-mℓ BPW ($10^{-1}$ dilution).
- Using a sterile pipette, 1-mℓ of the $10^{-1}$ dilution was aseptically transferred to the 1st tube containing a 9-mℓ volume of BPW to provide a $10^{-2}$ dilution.
- Using a fresh sterile pipette, 1-mℓ of the $10^{-2}$ dilution was aseptically transferred to the 2nd tube containing a 9-mℓ volume of BPW to provide a $10^{-3}$ dilution.
- Subsequent dilutions were made up in a similar manner. The number of dilutions to inoculate varied according to the presumed level of pollution of the water to be tested. Since a three-tube MPN technique was applied, each dilution was done in triplicate.
- Pre-enrichment dilutions were incubated at 35°C for 16 – 20 hours.

## 2.6 Enrichment dilutions

The sterile test tubes were prepared in a rack according to the number of selected dilutions. The number of dilutions to inoculate varied according to the presumed level of pollution of the water to be tested. Since a three-tube MPN was applied, each dilution was done in triplicate.

- A volume of 10-mℓ sterile Rappaport-Vassiliadis (RV) broth has already been added to each tube and sterilised.
- The 250-mℓ Schott® bottles and screw-cap test tubes were vigorously shaken to homogeneously mix the contents of the pre-enrichment media.
- 0.1-mℓ of the pre-enriched BPW media ($10^{-1}$ dilution) was aseptically (Theron, 2001) transferred from the 1st 250-mℓ Schott bottle to the 1st tube of 10-mℓ broth ($10^{-1}$ dilution). This was done for all three tubes in the range.
- Using a fresh sterile pipette, 0.1-mℓ of the pre-enriched BPW media ($10^{-2}$ dilution) was aseptically (Theron, 2001) transferred to the 2nd tube containing a 10-mℓ volume of RV Broth to provide a $10^{-2}$ dilution. This was done for all three tubes in the range.
- Subsequent dilutions were made up in a similar manner.
- Test tubes were incubated at 42°C ± 1°C for 24-48 hours.

## 2.7 Inoculation of petridishes (streak plate method) (Theron, 2001)

- The work surface was sterilised with ethanol.
- The inoculation needle was sterilised by holding it in the gas flame (until the whole needle was red-hot).
- The needle point was cooled off by pressing it lightly on a clean agar plate.
- The needle was inserted into one of the test tubes (containing RV broth) with the highest dilutions. Aseptic techniques was used (flaming of test tube lip before and after inserting inoculation needle into test tube) throughout.
- Streaking was done unto a clean XLD agar plate, by streaking the needle back and forth along one side of the plate.

- The needle was flamed an ht of the previous streak. The needle was then drawn (at an angle of 120°) through the previous streak. Without flaming the needle, 2 to 3 more streaks were drawn parallel to the first, without touching the first streak.
- The previous step was repeated 2 to 3 more times on the same plate.
- The needle was flamed again and the whole process repeated (for each of the three tubes per dilution), continuing with the next lower dilution.
- The plates were inverted and incubated at 35°C for 18-24 hours.

# 3 TECHNIQUES FOR ORGANISM RECOVERY AND ENUMERATION

## 3.1 Sampling

Samples were taken in 800-mℓ sterile Whirlpacks® from the various water environments and placed in cooler bags (7°C–10°C), for transportation to the laboratory. The samples were analysed within 6 hours of collection.

## 3.2 Dilutions for Most Probable Number

Before calculation of the Most Probable Number (MPN) can be done, a characteristic number of positive dilutions must be selected from the readings of the test tubes (corresponding to positive plates) and the rest discarded. Red colonies with black centres were counted as positive for *Salmonellae* (Oxoid Corporation, 1990).

Where three or more dilutions have been inoculated, a characteristic number of the lowest three figures (ending with 0 wherever possible) must be retained (Beliaeff and Mary, 1993; Standard Methods, 1998). The plates with positive single colonies (together with their corresponding positive RV broth containing tubes (Section 2.6)) were retained. Where less than three dilutions were used, note and retain the number of positive tubes (in this case corresponding to positive plates) from all dilutions (Standard Methods, 1998; Jagals, 2000).

Example 1: Polluted surface wa

| | |
|---|---|
| $10^{-1}$ | 3 + out of 3 |
| $10^{-2}$ | 3 + out of 3 |
| $10^{-3}$ | 2 + out of 3 |
| $10^{-3}$ | 1 + out of 3 |
| | Retain 3; 2; 1. |

Example 2: Drinking water

| | |
|---|---|
| $10^{-1}$ | 1 + out of 3 |
| $10^{-1}$ | 1 + out of 3 |
| $10^{-2}$ | 0 + out of 3 |
| | Retain 1; 1; 0. |

## 3.3 Calculation of Most Probable Number

The MPN is a statistical estimation of the density of microorganisms, assumed to correspond to a Poisson distribution in the volumes inoculated (Beliaeff and Mary, 1993; Standard Methods, 1998). This statistical density estimates can either be read from a density table, such as the one suggested by Standard Methods (1998) or Beliaeff and Mary (1993), or the numbers can be calculated based on a simple formula. For this study, the calculations were done on computer based on Thomas's formula from Standard Methods (1998). In order to apply the formula, the positive plates were assumed representative of the positive tubes containing RV broth (Section 2.6), and recorded as such. The formula is:

$$\text{MPN / 100 m}\ell = \frac{\text{no. of positive tubes x 100}}{\sqrt{(\text{m}\ell \text{ sample in negative tubes x m}\ell \text{ sample in all tubes}}}$$

The formula was programmed in a MS Excel® (2002) spreadsheet. The analyst entered

- the number of positive test tubes per dilution;
- the number of tubes per dilution (e.g., 3; 5), as well as
- dilutions (expressed in decimals e.g. $10^{-1} = 0.1$; $10^{-2} = 0.01$).

Counts were expressed as number of organisms per 100 mℓ.

# 4 ANALYTICAL QUALITY

According to Standard Methods (1998), it is especially important that laboratories performing only a limited number of microbiological testing, exercise strict quality control procedures. A quality assurance programme was established to ensure accuracy of results obtained during this study. The guidelines proposed for minimal quality control programmes, recommended by Standard Methods (1998), were followed during the study.

- The sterility of the media, dilution and rinse water, glassware and equipment was checked with sterile water as a sample during each sample series analyses.
- The media used in the study was checked by testing for known positive and negative control cultures of *Salmonellae*.

## 4.1 Control cultures for the microbiological testing of *Salmonella spp.*

Stock cultures of *Salmonella typhimurium* (positive control (ATCC 14028) – culture acquired from Oxoid) and *Escherichia coli* (negative control (ATCC 25922) – culture acquired from Oxoid) were made up according to the prescriptions provided (Oxoid Corporation, 1990; Standard Methods, 1998).

## 4.2 Procedures for medium check

Sterile distilled water were spiked with pure cultures and applied through the whole series of pre-enrichment (BPW), selective enrichment (RV broth) and inoculation of the culturing media (XLD medium).

# 5 COLONY VERIFICATION

The growth medium, XLD has been found to be not very selective (Oxoid Corporation, 1997). One of the reasons is the vast array of species and sub-species often to be found in a single pathogenic organism group or species as well as in the multitude of non-pathogenic organism groups. Amongst these variants, one will inevitably find non-pathogenic organisms

that find the selectivity of the spec imodating and may even manifest in the colours prescribed to the analyst for identification.

To establish the reliability of detected pathogen numbers, as well as the selectivity of the media for detecting the selected pathogen group, a verification programme was followed according to Standard Methods (1998).

Standard Methods (1998) recommends that at least 10 colonies per month be picked randomly from known positive samples and verified. Only colonies counted as *Salmonellae* on the selective growth media were selected.

Verification was done be means of API®, a multi-test identification system gallery by bioMérieux (2000). This identification system consists of strips with a characteristic number of micro-tubes containing dehydrated substrates. These substrates support specific enzymatic activity or fermentation of sugars. Each micro-tube is inoculated with a dense bacterial suspension made up of the original selected colony, which at the same time reconstitutes the substrates. Metabolic end products are produced during incubation, which produces spontaneous colour changes or revealed colours afterwards by the addition of reagents. The various reactions are then coded and read into a Reading Table. The identification is obtained from an Identification Table or a computerised Analytical Profile Index.

# 6 CONFIRMATION PROCEDURE FOR *SALMONELLAE*

## 6.1 XLD Medium

Red colonies with black centres (*Salmonellae* on XLD Medium) that the analyst counted as the positive colonies for *Salmonellae* on the specific media, were selected. The colony morphology was carefully noted and included colour, size, shape, composition, and edge appearance. A note was also made of the number of colonies counted from every particular plate, as well as the number taken for verification on the API® 20E identification system

The single colonies identified as p *ellae* were picked up and grown on MacConkey® Agar (Oxoid Corporation, 1990; Standard Methods, 1998) to strip the colonies of their colour. This part of the process was the last step in which the colonies were touched with the metal-eye of an inoculum needle. Further removal of the isolated colony from the MacConkey® Agar to be used for identification on the API® 20E strip, was done with sterile swabs to exclude possible interferences from the metal eye of an inoculum needle with the oxidase test. Before inoculating the API® 20E strip an oxidase test was done, since *Salmonellae spp.* are predominantly oxidase negative (Bergey's Manual®, 1994).

## 6.2 API® 20E Multi-test galleries (bioMérieux®)

API® 20E are standardised identification systems for *Enterobacteriaceae* and other non-fastidious gram-negative rods. The systems use 20 miniaturised biochemical tests (respectively) in the strips with a related database. These systems can be used to identify a substantial number of species that included the *Salmonellae spp.* used in this study.

## 6.3 Preparation of the inoculum

Homogenous bacterial suspensions, of the selected (and purified) colonies were made according to the prescriptions contained in the manual provided with the commercial identification kit (bioMérieux, 2000).

## 6.4 Inoculation of the strips

The micro-tubes on the prepared strips were filled according to prescription and incubated for 18-24 hours at 35-37°C.

## 6.5 Reading the strips

After the incubation time, the spontaneous colour reactions from each strip were recorded. Reagents were added to the prescribed tubes and the colour reaction recorded. All these recording were done on result sheets provided with the kit.

## 6.6 Identification

The pattern of each of the reactions obtained was hand-coded, on the result sheets, into a numerical profile. These numerical profiles are read into the ANALYTICAL PROFILE INDEX as a number. The Index then provides the name of the species that matches the code. It is important to note that if *Salmonellae* is positively identified with this identification system, serological identification should be performed to confirm the bacterial identification. Serological confirmation however, did not fall within the scope of this study.

Appendix B

# BACTERIAL INDICATOR ANALYSES – *Escherichia coli*

## MEMBRANE FILTRATION TECHNIQUE

For *E. coli* analyses, the membrane filtration method described by South African Bureau of Standards (SABS) (1984 and 1987), Millipore Corporation (1992) and Standard Methods (1998) was used.

## 1 EQUIPMENT

### 1.1 Filter & vacuum assembly

The membrane filtration system consisted of the following:

- 3 x Millipore® 3-place PVC manifolds.
- 9 x 47-mm diameter Millipore® glass filter holder sub-assembly, comprising
  - glass funnels (± 250-mℓ capacity),
  - fritted glass base support for filter membrane, and
  - clamp to secure funnel on base after loading filter membrane.
- 2 x EDWARDS® 1.5 two-stage 220/240 V 50/60 Hz vacuum/pressure pump.
- Two sets of 1-litre vacuum filter glass flasks for moisture traps before the vacuum pumps.
- The assembly is connected by means of silicone rubber tubing.

### 1.2 Pipettes

- 0.1-mℓ and 1-mℓ Finnpipette® adjustable pipettes (calibration errors were checked not to exceed 2.5%).
- Sterile disposable pipette tips (for dispensing 0.1 mℓ and 1 mℓ).
- Standard graduated glass pipettes for larger volumes.

### 1.3 Membrane filters

- Sterile Millipore® HA-type 0.45 µm pore size membranes, 47 mm in diameter, white, and grid-marked.

### 1.4 Incubators

- Labocon and Scientific incubators fitted with fans for air circulation (temperatures varied within 0.5°C accuracy, especially within stacks of incubated plates).
- Water baths (25 ℓ) equipped with gabled covers, to aid temperature maintenance within 0.2°C of setting, and uniformly distributed heating elements in steel inner jacket, to ensure constant temperature distributions were used.

### 1.5 Counting

- A ZEISS® stereo microscope was used to count the colonies on membrane filters.

## 2 PREPARATION / PROCEDURES

### 2.1 Sterilising

Equipment was steam-sterilised in an autoclave at 121°C for 20 minutes after each completed filtration session of all samples. Dry sterilisation of equipment was done between each sample filtration session in an oven at 180°C for 10 minutes. Between indicator-group filtration (within each sample filtration session), the sub-assemblies were immersed in boiling water for 10 min to decontaminate. Forceps were immersed in alcohol and flamed before filter-handling between batches.

### 2.2 Phosphate buffer

Stock phosphate buffer magnesium chloride solutions were prepared according to Standard Methods (1998). Working solutions of buffer were made up by adding 1.25 mℓ of phosphate (34 g $KH_2PO_4$ / ℓ distilled water) buffer and 5 mℓ of magnesium chloride solution (81.1 g $MgCl_2 \cdot 6H_20$ / ℓ distilled water) to 1 litre of reagent grade water and autoclaved to sterilise.

## 2.3 Dilutions

A dilution procedure was followed to achieve the ideal colony range of between 20 and 60.

- A volume of 90-mℓ sterile phosphate buffer was prepared per sample.
- Samples were vigorously shaken to homogeneously mix the contents.
- 10 x 1-mℓ extractions, from various areas and depths in the sample, were aseptically transferred from the sample to the prepared volume of phosphate buffer, to prepare a 100-mℓ of $10^{-1}$ diluted sample.
- 1 mℓ of $10^{-1}$ dilution was aseptically transferred to a 9-mℓ volume of sterile phosphate buffer to provide a $10^{-2}$ dilution.
- Subsequent dilutions were made up in a similar manner.

## 2.4 Chromocult® Coliformen Agar (Merck Corporation, 1996)

- 26.5 g of the powder were suspended in 1-litre of distilled water.
- The mixture was gently boiled in a flowing water bath while gently being stirred until the powder was totally dissolved (the media does not require autoclaving).
- The media was cooled to 40-50°C and the Cefsulodin solution (10 mg in 2 mℓ of distilled water) was added to the 1-litre of medium by gently shaking to homogenise. (Cefsulodin solution was added to eliminate the expected accompanying flora, especially *Pseudomonas* spp. and *Aeromonas* spp.)
- The liquid was poured into 90 mm petri dishes.

# 3 TECHNIQUES FOR ORGANISMS RECOVERY AND ENUMERATION

## 3.1 Sampling

Samples were taken in 800-mℓ sterile Whirlpacks® from the various water environments and placed in cooler bags (7°C–10°C), for transportation to the laboratory. The samples were analysed within 6 hours of collection.

Three sets of Millipore® 3-place vacuum manifolds, complete with filter holder sub-assemblies were used. The electric vacuum pumps evacuated through a dual moisture trap system comprising 1-litre vacuum flasks. Before each session of filter plating, each glass assembly was separately wrapped in tin foil and steam sterilised at 121°C for 15 minutes. Constant sterilisation and decontamination of the glass sub-assemblies was done during filtration sessions between samples to avoid cross contamination. Filter plating of the same sample was done in decreasing dilution order to avoid contamination.

A sterile phosphate buffer was used for diluting samples and rinsing funnels after filtration (Millipore Corporation, 1992). Pre-sterilised membrane filters were used. Membranes were loaded with a sterile forceps, grid side up, onto the fritted glass support base of the funnel holder, and the funnel clamped onto the filter base.

The sample was then re-mixed by vigorously shaking the bottle for several seconds. 20-30 mℓ of sterile buffer was poured into the funnel and a volume of sample was pipetted into the buffer.

For clear water, volumes of between 10-mℓ and 100-mℓ sample were pipetted. For turbid water, 1 mℓ of undiluted sample or sample dilute was pipetted onto the filter. For turbid samples dilutions of up to $10^{-4}$ were prepared and filtered within 20 minutes. For first run-off, especially after long dry periods, as well as for faecal- and wastewater-related samples, dilutions of up to $10^{-7}$ were prepared. All sample portions, suspended in dilution, were filtered within 30 minutes to avoid inactivation or multiplication of organisms in the dilution.

Vacuum was applied while slowly swirling the manifold unit to ensure uniform suspension of the sample in the volume of buffer during filtering. The funnel walls were rinsed repeatedly (3 times) with approximately 30 mℓ of sterile buffer. Buffer was drawn into a syringe and ejected through a sterile Sterivex® (Millipore®) filter to avoid contamination.

Vacuum was broken and the membrane lifted with a sterile forceps, and put grid side up,

onto a selective medium in petri  o trapped air under the membrane. The dishes were marked, inverted and incubated at 37°C for 24 hours, ± 2 hours (Millipore Corporation, 1992; SABS, 1984 & 1987; Standard Methods, 1998).

## 3.2 Dilutions

All samples were filtered in triplicate (3 filters) per dilution. Dilutions were made up to ideally achieve counts of between 20 to 60 colonies per plate (Standard Methods; 1998). Tillet (1993) described various factors that could lead to inaccuracies or unacceptable variation in counts of the same sample at the point of sampling and in the laboratory. Even vigorous mixing of a sample in the laboratory, before extraction, could not prevent variation in counts due to natural random distribution of organisms in such a sample.

Dilution procedures in the laboratory should ideally be adapted to minimise variations, while diluting from the sample (Section 2.3). Undiluted sample applications varied between 1 mℓ and 100 mℓ. These applications were single extractions by pipette or decanted into sterile 100 mℓ measuring cylinders from the raw sample, after the sample had been vigorously shaken. Organism distribution in the water body at the sampling point is considered achieved by turbulence in the stream.

## 3.3 Counting the microorganisms

After incubation for appropriate periods of time, colonies in various shades of dark-blue-to-violet colonies (Merck Corporation, 1996) were counted. To achieve reliable statistical quantification the final count per 100 mℓ per sample was calculated as follows (Standard Methods, 1998):

$$\frac{\frac{[(\text{Filter 1} + \text{filter 2} + \text{filter 3}) / 3] \times 100}{\textit{Sample size}}}{\text{Sample dilute}}$$

A formula was programmed in a MS Excel® (2002) spreadsheet. The analyst enters ① the counts from each of the 3 filters (membranes), ② sample size (maximum 1 mℓ for diluted

samples) as well as ③ the dilutic ).1; 0.01; etc. (minimum 1 mℓ for undiluted samples). Counts are expressed as colony forming units (cfu) / number of organisms per 100 mℓ.

# 4 ANALYTICAL QUALITY CONTROL

As with Appendix A, a quality assurance programme was established, according to the recommendations of Standard Methods (1998), and followed for this study:

- The sterility of the media, filters, dilution and rinse water, glassware and equipment was checked with sterile water as a sample during each sample series analyses.
- The media used in the study was checked by testing for known positive and negative control cultures of *E. coli*.

## 4.1 Control cultures for microbiological testing of *E. coli*

Stock cultures of *E. coli* (positive control – culture acquired from SABS), *Enterobacter aerogenes* and *Citrobacter freundii* (negative control – culture acquired from SABS) were made up (bioMérieux, 2000; Merck Corporation, 1996; Standard Methods, 1998).

## 4.2 Procedures for medium check

Volume units of 1 mℓ of the positive and negative stock culture solutions were filtered through membranes. The membranes were placed on petri dishes containing the selective growth media. Stock culture analyses were done at least once a month for the duration of the project to check a specific medium.

The specific colony colour identification and distinction was standardised by the analyst group (making sure everyone see and understand the same colour – including the various nuances / shades) and used to identify the indicator organisms tested for on the media.

The actual selectivity / specificity of the selective growth media, Chromocult® Coliformen Agar has been found to be inconsistent. One of the most common reasons is the vast array of species and sub-species that often find the selectivity of a specific medium accommodating and which may even manifest in the colours prescribed to the analyst for identification.

Verification was again done by means of the API® identification system by bioMérieux (2000) described for *Salmonellae spp.* in Appendix A. Before verification began, the coloured selected colonies were first stripped of the colouration caused by the selective substrates of the growth medium to eliminate all possible interference with the functions of the Identification System Galleries.

# 6 CONFIRMATION PROCEDURE OF *E. COLI*

## 6.1 Chromocult® Coliformen Agar

Deep blue-to-violet colonies (*E. coli* on Chromocult® Coliformen Agar) that the analyst would count as the coloured *E. coli* colonies on a given specific media were selected. Nearly the same procedure, as described for *Salmonellae spp.* verification in Appendix A, was applied for *E. coli* verification, with the following differences:

- Membrane-grown colonies to detect *E. coli* on the Chromocult® Coliformen medium were only partially picked up. The remaining colony was used for intermediate *E. coli* verification with KOVACS' indole reagent according to the user manual (Merck Corporation, 1996). This was necessary because it was feared that the indole reaction might influence further refinement of the selected colony.
- To obtain pure and strong single *E. coli* colonies, the colonies were picked up from the membranes with inoculum needles, streaked out on the same selective medium, and incubated at the prescribed temperature.

- Single colonies on the se[Central University of Technology, Free State]e then streaked out and grown on Plate Count Agar (Biolab Catalogue, 1997; Standard Methods, 1998) to strip the colonies of their colour (*Salmonellae* grown on MacConkey® Agar).

## 6.2 API® 20E Multi-test galleries (bioMérieux®)

*E. coli* verification was done with the same API® 20E multi-gallery strips used for *Salmonellae*. The procedures used for preparation of the inoculum, preparation and reading of the strips, as well as the identification of the microorganisms are described in Appendix A.

Appendix C

# STATISTICAL ANALYSES

Statistical processing of the data for this study comprised mainly of two approaches. Bacterial enumeration that generated the one group of data (1st approach) is dealt with in each of the appendices (Appendix A and B) that discuss the materials and methods for each analysis. This Appendix mainly deals with the data (2nd approach) related to the occurrence of *E. coli* and *Salmonellae* in the selected waters.

## 1 CHARACTERISTICS OF MICROBIOLOGICAL WATER QUALITY DATA

Environmental health data varies considerably, which necessitates the use of statistics (Pearson and Turton, 1993). According to Helsel and Hirsch (1995) and Standard Methods, (1998), reasons for variation in microbiological water quality data are often the following:

- Outliers (observations considerably higher or lower than the most of the data) occur often -although infrequently - with outliers for microbiological data of water resources more common on the high side.
- Data are not distributed evenly (normally) around the mean. Many statistical tests assume that data follow a normal distribution while water data often do not.
- A lower bound of zero – no negative values are possible.
- Data reported only below or above some threshold (censored data).
- Seasonal patterns.
- Autocorrelation or co-occurrence. Consecutive observations under similar circumstances tend to strongly correlate with each other. The most common kind of autocorrelation in water resources is that high values will tend to follow high values in circumstances such as intermittent high volumes of intensive rainfall.
- Dependence on other uncontrolled variables. Values strongly co-vary with discharges, rainfall or some other variable.

For this study, these characteristi[c] following approaches to analyse and interpret the data.

## 2 DATA DISTRIBUTION

The *occurrence* data sets of both *E. coli* and *Salmonellae* in this study varied widely in their distributions around their respective means. Such wide distributions are often not normally distributed around the means (Helsel and Hirsch, 1995). Application of most statistical techniques in the field of water resource management generally assumes that data sets have symmetrical distributions such as the normal (gaussian) curve (Pearson and Turton, 1993). In chemical water-quality analyses, the distribution of most analytical results follows the Gaussian (normal) curve, which has symmetrical distribution of values about the mean (parametrical data). Such (parametrical) data are suited for *parametrical statistical* tests.

However, microbiological water-quality data distributions are often not symmetrical. Bacterial counts often have a skewed distribution (non-parametrical data) because of the occurrence of more low counts than high counts in a given monitoring set (Standard Methods, 1998). For this study, an added dimension was heavily polluted waters that often yielded more high counts than low counts. Data sets can show positive skewness (e.g. polluted urban run-off) as well as negative skewness (e.g., unpolluted river water).

Problems can occur with the reliability of data interpretation where statistical procedures such as parametric tests, which assume normality, are directly employed (Helsel and Hirsch, 1995) on data that do not follow a symmetrical distribution about the mean (the Gaussian curve). These particular data sets therefore required *non-parametrical testing*.

Non-parametrical kinds are more robust than parametrical tests, although the latter are more popular in statistical testing since these are deemed more refined (Glantz, 1997, Helsel and Hirsch, 1995).

Both parametrical as well as non-parametrical tests were used for this study and the instances clearly communicated to the reader.

## 2.1 Normality and data trar

Because data for this study were assumed generally of a non-symmetrical data distribution around the mean, it was decided to transform the *E. coli* and *Salmonellae* occurrence data to their respective logs. This would ensure distributions that would have close to normal (log-normal) distribution characteristics and be more consistent in variance (Helsel and Hirsch, 1995; Standard Methods; 1998), therefore providing more options for parametrical as well as non-parametrical testing.

While many of the data sets showed normal distributions after log-transformation, several data sets within the same frame remained skewed albeit much closer to normal than before transformation. Failure of the normality test indicated the presence of outlying points, which caused the data to vary widely and to be inconsistently distributed (non-parametric). According to Helsel and Hirsch (1995), parametric test methods lose considerable power to detect differences in non-normal data, while non-parametric testing display considerable power of accuracy in non-normal as well as normal data testing and display.

For this study, therefore, regardless of whether the data sets failed or passed normality (using the Kolmogorov-Smirnov normality test – SigmaStat Version 2.03 (1997)), non-parametric statistical test procedures were applied throughout the study, with the only exception discussed in Section 7.2.of this appendix.

Data reports in this document are generally done in tables in the text. Some elements of the data, such as reporting on the probability of *Salmonellae* infection ($P_i$), were reported in their normal values for general clarity for the reader. The rest were generally reported in their $\log_{10}$ values.

## 2.2 Outliers (extreme values)

Outliers are observations (values) which are quite different (odd / unusual) from others in the data set (Glantz, 1997; Pearson and Turton, 1993). While it is often found that analysts would discard extreme values (Helsel and Hirsch, 1995), this procedure was not followed in

This study deals with risk assessn verestimation tendency of the mean is more suited, as it would generally not be conducive to underestimation and therefore protects population safety (Standard Methods, 1998). The mean was therefore preferred over the median in this study. Two applications of the mean were applied during analysing the microbiological water quality data for this study.

### 4.1 Mean (or average)

The arithmetic mean is the sum of all the data in a set, divided by the sample size. If data is normally distributed, the mean is at the centre of the distribution. The mean was therefore used in this study to determine the central value for all log-transformed data.

### 4.2 Geometric mean

In data that vary considerably, the mean most suited for a realistic central value is the geometric mean (Standard Methods, 1998). The geometric mean is calculated by log-transforming the data. This removes much of the variance, creating more "normally" distributed data, although the data may not necessarily be symmetrical yet. The *arithmetic* mean of these logarithms of data in a set is calculated and the value then transformed back (antilog) to its original unit. The resultant mean is the *geometric mean*. In this study, the geometric mean was used in results and discussion areas where reports were based on untransformed data sets.

## 5 MEASURES OF DATA SPREAD

For the Quantitative Microbial Risk Assessment (QMRA) approach of this study, the expected mean, minimum and maximum dose was calculated. The mean dose was based on the mean (discussed in Section 4 above) *Salmonellae* occurrence, while the lowest (minimum discussed in Section 5.1 below) *Salmonellae* occurrence was applied to calculate the minimum expected dose. The *Salmonellae* occurrence at the 95$^{th}$ percentile (Section 5.3 below) determined the maximum expected dose. The 95% confidence intervals of the mean (Section 5.2) are also discussed.

## 5.1 Range

The range of data is displayed by the subtracting the *minimum* from the *maximum*. The minimum is the smallest value in a population (e.g. lowest number of *Salmonellae* detected), and the maximum the largest value (highest number of *Salmonellae*). The minimum was applied in the QMRA approach to calculate the minimum expected dose (Katzenellenbogen et al., 1997).

## 5.2 Confidence intervals (Ci)

It was not possible to sample all the water in the study area all the time. In order to establish whether the limited number of samples taken over the eight months of the study could be deemed as reasonably representative of the health-related microbiological water quality in the study areas, and more importantly, whether a risk could be predicted within reasonable margins, confidence intervals around the mean were established for all the data sets.

*Confidence intervals* are used to make certain predictions about the data. A 95% confidence interval defines a range between a statistical boundary (limit) and the mean of a data set. Confidence intervals are generally applied in pairs – an upper and lower interval around a central value (the mean) - within which we can be 95% certain to find the data spread, as well as the mean of the population (Helsel and Hirsch, 1995; Pearson and Turton, 1993; Glantz, 1997).

These intervals also give some idea of the accuracy of the value estimates. Wide intervals suggests a poor estimate of the mean in the population and vice versa. The former means that the data varied considerably at that sampling point, which implies that a wide occurrence was recorded throughout the summer season, which would then reflect a wide range of risk.

Mean *E. coli* and *Salmonellae* occurrences, as well as their respective 95% confidence intervals (Ci's), were calculated for each water use area. The various risks were calculated within these intervals. Outlying values were deemed as chance occurrences but not discarded. These odd data points outside these intervals (outliers) were used in calculating

the confidence intervals. Some o[Central University of Technology, Free State]ed for extraordinary single exposure risk characterisations during illustrative risk scenarios created for this study.

## 5.3 The 95th Percentile

The 95th percentile is that value in a data set, below which 95% of the data lies. It is defined as a function that could establish a threshold of acceptance (Microsoft® Excel, 2002).

Percentiles are used to monitor compliance with water quality standards (Helsel and Hirsch, 1995). For instance, the mean and some percentile of data in a set should not exceed a standard imposed on a particular contaminant where exceedance would imply an acute risk of the population taking ill when ingesting the medium (Glantz, 1997).

The South African Water Quality Guidelines (DWAF, 1996) suggests that the 95th percentile would be a prudent level against which to test exceedance of the results of water quality monitoring for any given contaminant. In other words, for this study, the 95th percentile gave an indication of not only the spread of 95% of the data, but also the spread of the risk of infection. In other words in this study the 95th percentile established the threshold of acceptance of *Salmonellae* infection (Microsoft® Excel, 2002). Risk*Assistant™, (1995) defines this threshold as the "reasonable maximum exposure" (RME). The maximum expected dose and risk of infection was calculated from the *Salmonellae* occurrence at the 95th percentile.

# 6 ANALYSES OF VARIANCE (ANOVA)

ANOVA tests whether data sets differ significantly (Helsel and Hirsch, 1995; Wadsworth, 1989). Various ANOVA tests are available to compare two or more data sets.

According to Wadsworth (1989) and Helsel and Hirsch (1995), testing with the parametric, traditional Student-T test lose power to detect differences in non-normal data. However, non-parametric testing displays considerable power in non-normal, as well as normal data testing and display. Therefore, non-parametric ANOVA testing were generally used for this study,

with the test type clearly indicated t discussions in the text. The following tests were used (Glantz, 1997; Helsel and Hirsch, 1995; Wadsworth, 1989):



## 6.1 Mann-Whitney Rank-sum test

The Mann-Whitney rank-sum test is a non-parametric procedure. This test procedure compares two groups, even if their respective data sizes may be unequal, to establish whether data in one group tends to differ from data in another group by being larger, smaller, or larger and /or smaller.

## 6.2 Kruskal-Wallis ANOVA on Ranks (based on rank transformation)

This non-parametric test compares data for several (three or more ) different experimental groups (e.g. the *E. coli* data at three sampling sites within the RSQC) that may be affected by a single factor (the occurrence levels of microorganisms).

Whenever more than two groups are compared, the interest is not only whether the groups differed, but also which differed from the others. Multiple comparison tests (MCT's) are applied where significant differences were encountered. The following MCT's were used:

- The Tukey multiple comparison test because it is the more conservative of several available MCT's (SigmaStat, 1997) and therefore less likely to determine that a given difference is significantly different. It is the recommended test for all pairwise comparisons.
- Dunn's multiple comparison test because it is the more conservative of several available MCT's (SigmaStat, 1997) and therefore less likely to determine that a given difference is significantly different.

# 7 CORRELATION AND REGRESSION

Correlation and regression measure the strength of association between two continuous variables (Glantz, 1997; Wadsworth, 1989). For this study the two continuous variables were *E. coli* and *Salmonellae.*

Scientists and engineers often use ialise the shape and describe the behaviour of data (Glantz, 1997; Helsel and Hirsch, 1995; Wadsworth, 1989), for instance whether *E. coli* and *Salmonellae* would co-occur in the same water sample. From this, scientists would often strive to establish whether one variable (e.g. the indicator organism group *E coli* as the independent variable, could predict the occurrence of the other e.g. the pathogen group *Salmonellae* as the variable dependent on the occurrence of *E. coli*) in the same water sample.

## 7.1 Correlation

Of interest for this study was to see whether one variable (e.g. *E coli*) increased as the other variable (the pathogen group *Salmonellae*) increased (correlation), or whether their patterns of variation were totally unrelated (weak or no correlation). Measures of correlation generally lie between $-1 \leq 0 \leq 1$. When there is no correlation between two variables, the correlation coefficient $r = 0$. The stronger the association between two variables (increase = increase 1; increase = decrease –1), the closer $r = 1$ (Glantz, 1997; Helsel and Hirsch, 1995).

While regression analyses may also be used to test for association, one would tend to use regression to predict one variable from the other (SigmaStat, 1997) and often the purpose of testing for associations fade from focus when doing so.

For this study, correlation-specific tests were used.

- The parametric **Pearson Product Moment** to measure the strength of the association between pairs of variables without regard to which variable is dependent or independent and the residuals are normally distributed with constant variance. The Pearson Product Moment Correlation coefficient is the most commonly used correlation coefficient (SigmaStat, 1997).
- The non-parametric **Spearman Rank Order Correlation** was also used to measure the strength of association between pairs of variables without specifying which variable is dependent or independent but for data sets not normally distributed with constant variance.

## 7.2 Regression analysis

Regression procedures use the values of the independent variable (e.g., age in years) to predict the value of a dependent variable (e.g., daily ingestion volumes). For this study, regression analyses are used to:

- Estimate (predict) values of one variable based on the knowledge of the other variable.

Independent variables are the known, or predictor (explanatory), variables. These are most often plotted as X-axis values (for this study age in years). When the independent variables are plotted, they result in corresponding values for the dependent (response) variables (ingestion volumes), assigned to the Y-axis (Glantz, 1997; Pearson and Turton, 1993; Wadsworth, 1989). Regression procedures find associations between independent and dependent variables that, when graphed on Cartesian co-ordinate systems, produce straight lines, planes or curves.

A cubic polynomial regression analysis (detailed discussion in Appendix D) was applied to predict modified ingestion volumes for this study.

# 8. MINIMUM SAMPLE SIZE

It would be ideal in any study to test the whole target population (or for this study all the waters in a volume). However, this is physically and financially impossible. Instead, a representative subset of the potential data, called a sample, is extended to the entire population (Helsel and Hirsch, 1995). The larger the sample size, the greater the accuracy (power) (Pearson and Turton, 1993). The power, or sensitivity, of a test is the probability that the test will detect a difference or effect if there really is a difference or effect.

One should therefore determine approximately how big the sample size has to be, crude or not, in order to detect an effect or difference at a specified level of statistical significance or power (Helsel and Hirsch, 1995; SigmaStat, 1997).

## 8.1 Sample size for ANOVA

When testing for ANOVA during this study, minimum sample size estimates were based on a sample size of 15 prescribed by Standard Methods (1998) for an intra-laboratory proficiency programme. After assessing the first 15 samples, the mean differences of each (n = 15) data set was used to estimate the final minimum sample size in SigmaStat Version 2.03, and to confirm whether the initial sample sizes were large enough.

ANOVA testing procedures (parametric or non-parametric) depend on whether the comparative data is normally distributed with equal variance. However, to determine the minimum sample size, the normality of data is generally ignored and the size determined according to the following parameters (Helsel and Hirsch, 1995; SigmaStat, 1997):

- The size of the minimum expected differences in the group means is entered. Based on typical zero-hypotheses or data reliability theory, no differences should be encountered between the means of data groups. However, standard statistical packages used to calculate the estimated sample sizes, do not accept a 0 entry, as this is seen as statistically unrealistic (SigmaStat, 1997). Literature is also not very clear on how to approach the selection of minimum expected differences in the group means. The size of the minimum differences in the group means for the lognormal data were therefore calculated for each comparison group individually, based on the mean differences encountered after using data from the initial 15 samples.
- The size of the standard deviation of the data is entered. The size of the standard deviation could be the size expected (an estimate) or can be derived from previous experiments. Again, literature was unclear about what could be expected. It was decided to use studies by Jagals (1994, 1997, 2000), Griesel (2001), as well as data from this study, to calculate an overall mean standard deviation for each microorganism group used for each water- and other category during this study. These were then entered as the "expected" standard deviations and the calculated sample size suggested by the programme was then used as a minimum sample size.

- Desired power (sensitivity) set to 0.80.
- Alpha (∝) was set at 0.05.

## 8.2 Sample size to test associations (correlation coefficients)

To determine the sample size necessary to detect a specified correlation coefficient with SigmaStat, 1997, the following was specified:

- Expected value of the correlation coefficient (correlation expected). As the correlation between environmental variables such as two microorganism groups are rarely 1, a more realistic coefficient of 0.99 was used wherever correlation coefficients were applied in this study (Helsel and Hirsch, 1995).
- Desired power (sensitivity) of the test. Power is the probability that the correlation coefficient quantifies an actual association. The closer the power is to 1, the more sensitive the test. According to Helsel and Hirsch (1995), sensitivity in water resources testing is traditionally set to achieve a power of 0.80, which means that there is an 80% chance of detecting an association with 1-∝ confidence (i.e. a 95% confidence when ∝ = 0.05).
- Alpha (∝) is used to determine the sample size. The desired ∝ level is the acceptable probability of incorrectly concluding that there is an association. This indicates that a 1 in 20 chance of being wrong is acceptable (willing to conclude that there is an association when $P < 0.05$).

## 8.3 The confirmed minimum sample sizes were all 15

Based on these approaches, the statistical programme SigmaStat Version 2.03 (1997) was used to calculate the minimum sample size needed for such statistical significance. The result of the calculations confirmed that minimum sample sizes of 15 were sufficient for all the statistical tests done in this study.

# MODIFIED DAILY WATER INGESTION VOLUMES (Domestic use only)

Chapter 2 discussed the rationale behind extrapolating daily intake volumes from locally sourced water based on three race-based data sets from previous studies reported by Bourne et al. (1987, 1992), to derive at hypothetical intake volumes for this study. It was decided to apply the international age-grouping (infants, children, adolescents, adults and the elderly) reported by Roseberry and Burmaster (1992). Regression curves were fitted to the data in order to derive hypothetical daily ingestion volumes for children up to the elderly. Daily ingestion volumes for infants was instead calculated from the (yellow) Weight / Age (kg) chart (Robinson et al., 1982; WHO, 1994), used for all race groups at municipal clinics all over South Africa. This is shown in Appendix D2, while Appendix D1 discusses the statistical approaches followed.

## Appendix D1

## 1 CURVE FITTING

There are two reasons for fitting curves to data (Glantz, 1997; Helsel and Hirsch, 1995; SigmaPlot, 2002; SigmaStat, 1997; Wadsworth, 1989):

- To improve the display of the data with functions such as pointing out tendencies or trends in data i.e. whether there was an increase / decrease in the occurrence of a dependent variable when plotted against an independent variable.
- To mathematically describe the data points e.g. for future use in modelling other similar situations.

According to Glantz (1997), Helsel and Hirsch (1995), as well as SigmaStat® (1997), polynomial regression is a suitable test to use if the data passed normality and equal variance. Polynomial regression assumes an association between the independent (usually

on the X-axis) and the dependent  axis) variables that fit the general equation for a 1st order polynomial of $y = y_0 + m * x$ (linear shape).

However, the data generally did not appear to be in a straight-line for this component of the study. Non-linear polynomial smoothed curve fitting in this component of the study, returned equations for the polynomial cubic order of $y = y_0 + a_1 * x + b_2 * x^2 + c_3 * x^3$. "Cubic" means that the $+ c_3 * x^3$ was the 1st exponent of the independent variable that would return the smoothed spline curve required for interpolation and prediction depending on the correlation coefficient.

Regression, in other words, finds the equation that most closely describes, or fits the actual data in a set (Appendix D2). It then plots the curve that best describes the shape, as well as behaviour of data, and returns an equation that may be used as a modelling function. This resulting equation therefore, when plotted over the original data, produces a mathematical answer about the behaviour of the independent variable (age in years – Appendix D2) (SigmaPlot, 2002) that may be used to predict the behaviour of the dependent variable (modified ingestion volumes in mℓ/hd.d) – often referred to as trend analysis (Glantz, 1997; Helsel and Hirsch, 1995).

## 2. TREND ANALYSES

The purpose of using trend analysis in this study was to determine modified water ingestion volumes from a series of observations for a response variable that have been collected for various age groups. Where the X variable is age as a test for trend, it is directly analogous to regression. Trend analyses was also used in this study to describe (model) the rate of that change, as well as to estimate and / or predict - with a regression line - how much the response variable increased or decreased on the average as the explanatory variable (on the X-axis) changed.

To display whether there was dependence between the two continuous variables, age as the independent variable on the X-axis) and water intake volumes (as the dependent variable on

the Y-axis), polynomial regression [boilerplate] ting curves (with resulting equation) to the data. The computer programr [boilerplate] (2002) was used for this function, commonly also known as curve-fitting. The strength of these associations was then quantified with a correlation coefficient (Glantz, 1997; Helsel and Hirsch, 1995; Wadsworth, 1989).

Central University of Technology, Free State

# Modified South African ingestion volumes for this study

**Appendix D2**

## South African Studies

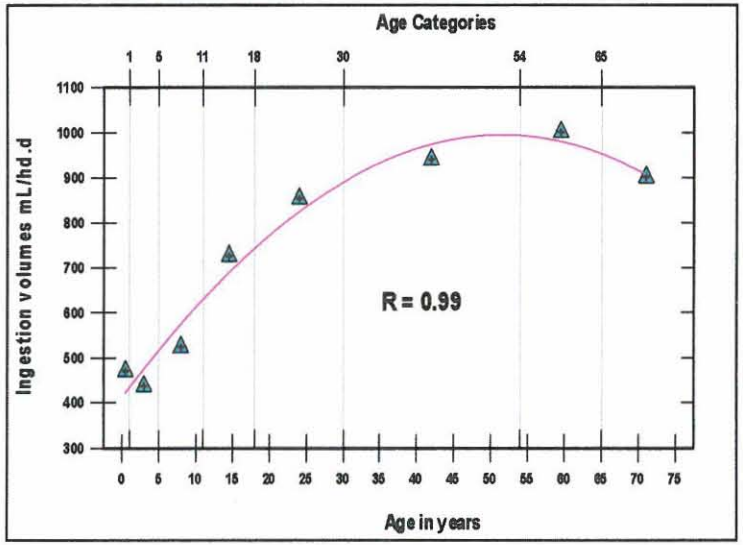
| South African age grouping | White people (Bourne et al., 1987) | | | Coloured people (Bourne et al., 1987) | | | Black people (Bourne et al., 1992) | | | Group means |
|---|---|---|---|---|---|---|---|---|---|---|
| | M | F | Mean | M | F | Mean | M | F | Mean | |
| 0≤age<1 | 508 | 261 | 385 | 595 | 518 | 557 | Not Available | | | 471 |
| 1≤age<5 | 490 | 484 | 487 | 388 | 385 | 387 | | | | 437 |
| 5≤age<11 | 709 | 599 | 654 | 526 | 503 | 515 | 402 | 402 | 402 | 524 |
| 11≤age<18 | 986 | 866 | 926 | 797 | 692 | 745 | 494 | 515 | 505 | 725 |
| 18≤age<30 | 1089 | 1211 | 1150 | 734 | 764 | 749 | 719 | 602 | 661 | 853 |
| 30≤age<54 | 1037 | 1426 | 1232 | 838 | 985 | 912 | 707 | 646 | 677 | 940 |
| 54≤age | 1368 | 1377 | 1373 | 848 | 928 | 888 | 830 | 652 | 741 | 1001 |
| 65≤age | Not Available | | | Not Available | | | M+F | | 901 | 901 |

## Modified intentional daily ingestion volumes for South Africans

| Age categories for this study Roseberry and Burmaster (1992) | Infants 6mths ≤age<1 | Children 1≤age<11 | Teens 11≤age<20 | Adults 20≤age<65 | Elderly 65≤age |
|---|---|---|---|---|---|
| Locally sourced water intake (mℓ/hd.d) | 1,318 | 630 | 773 | 952 | 865 |

| Extrapolation curve | |
|---|---|
| $y=y0+a*x+b*x^2+c*x^3$ | Coefficient |
| y0 | 411.4257 |
| a | 22.1202 |
| b | -0.1993 |
| c | -0.0002 |

| Extrapolated ingestion values | |
|---|---|
| Independent variable AGE in Years | y = mℓ/hd.d |
| 1 | 433 |
| 11 | 630 |
| 20 | 773 |
| 65 | 952 |
| 75 | 865 |




| Age / weight (kg) Robinson et al., 1982 WHO, 1994 | | | | Ingestion (150mℓ/kg) |
|---|---|---|---|---|
| Age | Boys | Girls | Mean weight | Volume |
| 6 months | 7.8 | 7.2 | 7.5 | 1,125 |
| 7 months | 8.3 | 7.7 | 8 | 1,200 |
| 8 months | 8.8 | 8.2 | 8.5 | 1,275 |
| 9 months | 9.2 | 8.6 | 8.9 | 1,335 |
| 10 months | 9.5 | 8.9 | 9.2 | 1,380 |
| 11 months | 9.8 | 9.2 | 9.5 | 1,425 |
| 12 months | 10.2 | 9.6 | 9.9 | 1,485 |
| Mean | 9.09 | 8.49 | 8.79 | 1,318 |

# Microbiological Water Quality Data: RSQC (Chapter 3)

## Appendix E

| Sampling Dates | Sampling points | *E. coli* | *E. coli* Mean | Log *E. coli* | Log *E. coli* Mean |
|---|---|---|---|---|---|
| 2001-09-10 | RS2 | 3,170 | | 3.50 | |
| | RS1 | 4,040 | | 3.61 | |
| | BS | 10,000 | 5,737 | 4.00 | 3.70 |
| 2001-10-01 | RS2 | 80,000 | | 4.90 | |
| | RS1 | 100,000 | | 5.00 | |
| | BS | 1,200,000 | 460,000 | 6.08 | 5.33 |
| 2001-10-08 | RS2 | 1,370 | | 3.14 | |
| | RS1 | 2,670 | | 3.43 | |
| | BS | 13,600 | 5,880 | 4.13 | 3.57 |
| 2001-10-22 | RS2 | 13,300 | | 4.12 | |
| | RS1 | 43,000 | | 4.63 | |
| | BS | 43,100 | 33,133 | 4.63 | 4.46 |
| 2001-11-05 | RS2 | 1,000 | | 3.00 | |
| | RS1 | 7,000 | | 3.85 | |
| | BS | 9,000 | 5,667 | 3.95 | 3.60 |
| 2001-11-12 | RS2 | 17,000 | | 4.23 | |
| | RS1 | 40,000 | | 4.60 | |
| | BS | 361,000 | 139,333 | 5.56 | 4.80 |
| 2001-11-26 | RS2 | 5,200 | | 3.72 | |
| | RS1 | 23,000 | | 4.36 | |
| | BS | 28,700 | 18,967 | 4.46 | 4.18 |
| 2001-12-03 | RS2 | 2,670 | | 3.43 | |
| | RS1 | 22,000 | | 4.34 | |
| | BS | 88,200 | 37,623 | 4.95 | 4.24 |
| 2002-01-07 | RS2 | 6,300 | | 3.80 | |
| | RS1 | 18,000 | | 4.26 | |
| | BS | 325,500 | 116,600 | 5.51 | 4.52 |
| 2002-01-14 | RS2 | 43,000 | | 4.63 | |
| | RS1 | 290,900 | | 5.46 | |
| | BS | 920,800 | 418,233 | 5.96 | 5.35 |
| 2002-02-04 | RS2 | 3,970 | | 3.60 | |
| | RS1 | 13,400 | | 4.13 | |
| | BS | 15,000 | 10,790 | 4.18 | 3.97 |
| 2002-02-11 | RS2 | 12,200 | | 4.09 | |
| | RS1 | 46,110 | | 4.66 | |
| | BS | 166,666 | 74,992 | 5.22 | 4.66 |
| 2002-03-04 | RS2 | 11,000 | | 4.04 | |
| | RS1 | 25,900 | | 4.41 | |
| | BS | 227,900 | 88,267 | 5.36 | 4.60 |
| 2002-03-13 | RS2 | 14,300 | | 4.16 | |
| | RS1 | 143,000 | | 5.16 | |
| | BS | 209,800 | 122,367 | 5.32 | 4.88 |
| 2002-03-25 | RS2 | 3,000 | | 3.48 | |
| | RS1 | 21,800 | | 4.34 | |
| | BS | 75,400 | 33,400 | 4.88 | 4.23 |
| 2002-04-08 | RS2 | 27,230 | | 4.44 | |
| | RS1 | 172,600 | | 5.24 | |
| | BS | 727,000 | 308,943 | 5.86 | 5.18 |
| n | | 48 | 16 | 48 | 16 | | |
| Average | | 117,496 | 117,496 | 4.45 | 4.45 | 28,444 | AntiLog E52 |
| Geomean | | 28,444 | 49,439 | | | | |
| Median | | 22,500 | 56,308 | 4.35 | 4.49 | | |
| Std | | 239,401 | 147,577 | 0.75 | 0.57 | | |
| 95% CI | | 67,726 | 72,311 | 0.21 | 0.28 | | |
| UL | | 185,221 | 189,807 | 4.67 | 4.74 | 46,431 | AntiLog E57 |
| LL | | 49,770 | 45,185 | 4.24 | 4.17 | 17,425 | AntiLog E58 |
| Max | | 1,200,000 | 460,000 | 6.08 | 5.35 | 1,200,000 | AntiLog E59 |
| Min | | 1,000 | 5,667 | 3.00 | 3.57 | 1,000 | AntiLog E60 |
| 95th Percentile | | 598,900 | 428,675 | 5.76 | 5.33 | 569,017 | AntiLog E61 |

| Sampling Dates | Sampling points | *Salmonellae* | *Salmonellae* Mean | Log *Salmonellae* | Log *Salmonellae* Mean |
|---|---|---|---|---|---|
| 2001-09-10 | RS2 | 95 | | 1.98 | |
| | RS1 | 73 | | 1.86 | |
| | BS | 95 | 88 | 1.98 | 1.94 |
| 2001-10-01 | RS2 | 150 | | 2.18 | |
| | RS1 | 733 | | 2.87 | |
| | BS | 4,383 | 1,755 | 3.64 | 2.89 |
| 2001-10-08 | RS2 | 286 | | 2.46 | |
| | RS1 | 36 | | 1.56 | |
| | BS | 286 | 203 | 2.46 | 2.16 |
| 2001-10-22 | RS2 | 110 | | 2.04 | |
| | RS1 | 206 | | 2.31 | |
| | BS | 584 | 300 | 2.77 | 2.37 |
| 2001-11-05 | RS2 | 36 | | 1.56 | |
| | RS1 | 286 | | 2.46 | |
| | BS | 110 | 144 | 2.04 | 2.02 |
| 2001-11-12 | RS2 | 95 | | 1.98 | |
| | RS1 | 949 | | 2.98 | |
| | BS | 760 | 601 | 2.88 | 2.61 |
| 2001-11-26 | RS2 | 457 | | 2.66 | |
| | RS1 | 286 | | 2.46 | |
| | BS | 1,001 | 581 | 3.00 | 2.71 |
| 2001-12-03 | RS2 | 36 | | 1.56 | |
| | RS1 | 36 | | 1.56 | |
| | BS | 287 | 120 | 2.46 | 1.86 |
| 2002-01-07 | RS2 | 95 | | 1.98 | |
| | RS1 | 150 | | 2.18 | |
| | BS | 110 | 118 | 2.04 | 2.07 |
| 2002-01-14 | RS2 | 95 | | 1.98 | |
| | RS1 | 387 | | 2.59 | |
| | BS | 110 | 197 | 2.04 | 2.20 |
| 2002-02-04 | RS2 | 36 | | 1.56 | |
| | RS1 | 95 | | 1.98 | |
| | BS | 148 | 93 | 2.17 | 1.90 |
| 2002-02-11 | RS2 | 95 | | 1.98 | |
| | RS1 | 387 | | 2.59 | |
| | BS | 36 | 173 | 1.56 | 2.04 |
| 2002-03-04 | RS2 | 36 | | 1.56 | |
| | RS1 | 286 | | 2.46 | |
| | BS | 718 | 347 | 2.86 | 2.29 |
| 2002-03-13 | RS2 | 286 | | 2.46 | |
| | RS1 | 148 | | 2.17 | |
| | BS | 760 | 398 | 2.88 | 2.50 |
| 2002-03-25 | RS2 | 36 | | 1.56 | |
| | RS1 | 286 | | 2.46 | |
| | BS | 36 | 119 | 1.56 | 1.86 |
| 2002-04-08 | RS2 | 286 | | 2.46 | |
| | RS1 | 36 | | 1.56 | |
| | BS | 286 | 203 | 2.46 | 2.16 |
| n | | 48 | 16 | 48 | 16 | | |
| Average | | 340 | 340 | 2.22 | 2.22 | 167 | AntiLog M52 |
| Geomean | | 167 | 232 | | | | |
| Median | | 149 | 200 | 2.17 | 2.16 | | |
| Std | | 647 | 411 | 0.49 | 0.32 | | |
| 95% CI | | 183 | 201 | 0.14 | 0.16 | | |
| UL | | 523 | 541 | 2.36 | 2.38 | 230 | AntiLog M57 |
| LL | | -16 | 139 | 2.08 | 2.07 | 121 | AntiLog M58 |
| Max | | 4,383 | 1,755 | 3.64 | 2.89 | 4,383 | AntiLog M59 |
| Min | | 36 | 88 | 1.56 | 1.86 | 36 | AntiLog M60 |
| 95th Percentile | | 883 | 890 | 2.94 | 2.75 | 878 | AntiLog M61 |

| C1 | *E. coli* | | *Salmonellae* | |
|---|---|---|---|---|
| Sample date | C1 *E. coli* | Log C1 *E. coli* | C1 *Salmonellae* | Log C1 *Salmonellae* |
| 2001-09-10 | 1.0 | 0.00 | 0.1 | -1.00 |
| 2001-10-01 | 1.0 | 0.00 | 0.1 | -1.00 |
| 2001-10-08 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2001-10-22 | 30 | 1.48 | 30 | 1.48 |
| 2001-11-05 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2001-11-12 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2001-11-26 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2001-12-03 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2002-01-07 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2002-01-14 | 13,300 | 4.12 | 1,898 | 3.28 |
| 2002-02-04 | 4.7 | 0.67 | 0.1 | -1.00 |
| 2002-02-11 | 6.7 | 0.82 | 0.1 | -1.00 |
| 2002-03-04 | 0.1 | -1.00 | 0.1 | -1.00 |
| 2002-03-13 | 5.0 | 0.70 | 0.1 | -1.00 |
| 2002-03-25 | 1.3 | 0.12 | 0.1 | -1.00 |
| 2002-04-08 | 1.3 | 0.12 | 0.1 | -1.00 |
| n | 16 | 16 | 16 | 16 |
| Mean | 834 | 0.07 | 121 | -0.58 |
| Geomean | 1.16 | | 0.26 | |
| Minimum | 0.10 | -1.00 | 0.10 | -1.00 |
| Maximum | 13,300 | 4.12 | 1,898 | 3.28 |
| Std Dev | 3,324 | 1.36 | 474 | 1.20 |
| 95 % Ci | 1,629 | 0.67 | 232 | 0.59 |
| UL | 2463.28 | 0.73 | 352.95 | 0.01 |
| LL | -794.32 | -0.60 | -111.67 | -1.17 |
| 95th Percentile | 3,348 | 2.14 | 497 | 1.93 |

| F1 | *E. coli* | | *Salmonellae* | |
|---|---|---|---|---|
| Sample date | F1 *E. coli* | Log F1 *E. coli* | F1 *Salmonellae* | Log F1 *Salmonellae* |
| 2001-09-10 | 10 | 1.0 | 0.1 | -1.0 |
| 2001-10-01 | 7 | 0.8 | 0.1 | -1.0 |
| 2001-10-08 | 20 | 1.3 | 0.1 | -1.0 |
| 2001-10-22 | 150 | 2.2 | 259 | 2.4 |
| 2001-11-05 | 1,500 | 3.2 | 584 | 2.8 |
| 2001-11-12 | 237 | 2.4 | 0.1 | -1.0 |
| 2001-11-26 | 67 | 1.8 | 0.1 | -1.0 |
| 2001-12-03 | 50 | 1.7 | 0.1 | -1.0 |
| 2002-01-07 | 107 | 2.0 | 0.1 | -1.0 |
| 2002-01-14 | 133 | 2.1 | 0.1 | -1.0 |
| 2002-02-04 | 20 | 1.3 | 0.1 | -1.0 |
| 2002-02-11 | 20 | 1.3 | 0.1 | -1.0 |
| 2002-03-04 | 25,300 | 4.4 | 0.1 | -1.0 |
| 2002-03-13 | 3,000 | 3.5 | 0.1 | -1.0 |
| 2002-03-25 | 100 | 2.0 | 0.1 | -1.0 |
| 2002-04-08 | 20 | 1.3 | 0.1 | -1.0 |
| n | 16 | 16 | 16 | 16 |
| Mean | 1,921 | 2 | 53 | -1 |
| Geomean | 105 | | 0.28 | |
| Minimum | 7 | 1 | 0.10 | -1 |
| Maximum | 25,300 | 4 | 584 | 3 |
| Std Dev | 6,284 | 1 | 156 | 1 |
| 95 % Ci | 3,079 | 0.47 | 76 | 1 |
| UL | 5,001 | 2 | 129 | 0 |
| LL | -1,158 | 2 | -24 | -1 |
| 95th Percentile | 8,575 | 4 | 340 | 3 |