Analysis of dissemination of waterborne pathogens from raw water and related health issues in Sweden using GIS

Nikolaos Tsalikis

Master of Science Thesis in Geoinformatics
TRITA-GIT EX 15-007

School of Architecture and the Built Environment
Royal Institute of Technology (KTH)
Stockholm, Sweden

June 2015
Abstract

Water-related diseases are a major issue of public health worldwide. Raw drinking water retrieved from surface and groundwater undergoes treatment at water treatment plants in order to disable certain microorganisms that if were to be released in drinking water, would cause a variety of unwanted gastrointestinal symptoms. As such, it is imperative that research goes in the direction of preventing the entry of such microorganisms in drinking water by identifying potential risk situations involving high concentrations in raw water.

The objective of this thesis research is to examine variations and peaks in concentrations of notable pathogens found in raw water samples retrieved from the influx of five different water treatment facilities in Sweden, along with visualisations of related health symptoms that might have been the result of pathogen contamination in areas that were supplied by those waterworks. The primary source for the raw water data is Folkhälsomyndigheten (Public Health Agency of Sweden) and covers the period of March 2014 to 2015 in an ongoing project, however data from before this project started have also been retrieved from Vattentäktsarkivet covering the period of 2009-2013 for the representative waterworks. The data used to visualise the related health symptoms were retrieved from Healthcare Guide 1177, based on telephone triage. Along with that data, SCB data on population were used along with a shapefile of postal areas of Sweden, provided by Postnummerservice.

In order to visualise symptoms across time and to examine possible rises of symptoms related to gastrointestinal diseases across the period of analysis, Excel and Geographical Information System (GIS) toolboxes Spatial Analyst and Data Management were used in order to process the data from 1177, create time-animated outputs that depicted the symptoms monthly for every area and to identify areas that reported the most symptoms using density maps and hot-spot analysis. Statistical analysis and graphs for the data regarding the raw water dataset were performed in Excel.

The results implied minor correlations between high concentrations of notable pathogens in raw water and a rise of symptoms reported to 1177 in some months and areas during the period of analysis. The suspect months and areas were identified through a visualisation of clusters of symptoms and increases in comparison to normal seasonal effects. The winter months along with March and April generally showed the highest number of reported symptoms in all counties and a comparison of incidence rates revealed that the municipalities of Östersund and Trollhättan had the highest incidence rates in regards to their population.

A large source of error in the analysis is the raw water dataset itself. Since raw water always undergoes treatment, it is typically assumed that pathogens are removed prior to entering drinking water. It is possible that pathogens might elude this process, though it is difficult to exclude other sources of infection, such as contaminated food or seasonal contagious effects like winter-vomiting disease without access to patient records.

Keywords: GIS, Raw Water, Pathogens
Acknowledgements

First of all, I would like to thank my supervisor Professor Yifang Ban for the opportunity to work on a project of such scale and for her guidance and comments during the course of this thesis research. I would also like to thank my co-supervisor Lars Skog from ESRI for his active interest and his invaluable comments on GIS and ideas on how to best approach the project.

Additionally I would like to thank both Anette Hansen and Pär Bjelkmar from Folkhälsomyndigheten for promptly providing me with the needed data, ideas and their expertise on their respective fields of knowledge.

Jan Haas from KTH also has my gratitude for his assistance with ArcMap related issues. Finally I would like to thank both Katarina Skärlund and Sharon Kuhlmann-Berenzon from Folkhälsomyndigheten for promptly providing me with the required shapefile of Sweden.
# Table of contents

Abstract .............................................................................................................................. I  
Acknowledgements ............................................................................................................. III  
Table of contents ................................................................................................................. 1  
List of figures ..................................................................................................................... 3  
List of abbreviations .......................................................................................................... 6  
1. Introduction ................................................................................................................... 7  
   1.1 Background ............................................................................................................. 7  
   1.2 Objectives ............................................................................................................. 7  
2. Literature review .......................................................................................................... 8  
   2.1 Indicator organisms and pathogens ........................................................................ 8  
      2.1.1 Escherichia Coli .............................................................................................. 8  
      2.1.2 Verotoxigenic Escherichia Coli ...................................................................... 8  
      2.1.3 Coliforms ....................................................................................................... 9  
      2.1.4 Enterococci .................................................................................................... 9  
      2.1.5 Clostridium .................................................................................................... 9  
      2.1.6 Cryptosporidium ............................................................................................ 10  
      2.1.7 Giardia .......................................................................................................... 10  
      2.1.8 Campylobacter .............................................................................................. 11  
      2.1.9 Salmonella .................................................................................................... 11  
      2.1.10 Phages ......................................................................................................... 11  
      2.1.11 Norovirus ..................................................................................................... 12  
   2.2 Pathogen related attributes ..................................................................................... 12  
      2.2.1 Chemical Oxygen Demand .......................................................................... 12  
      2.2.2 Turbidity ....................................................................................................... 12  
      2.2.3 pH ................................................................................................................ 13  
      2.2.4 Temperature .................................................................................................. 13  
2.3 Treatment process ..................................................................................................... 13  
   2.3.1 Factors affecting process and regulations ......................................................... 13  
   2.3.2 Operational monitoring ...................................................................................... 14  
   2.3.3 Methods of disinfection ....................................................................................... 14  
2.4 Known outbreaks in Sweden and internationally attributed to waterborne pathogens ................................................................................................................. 16  
3. Study areas and data description .................................................................................. 17
3.1 Study areas ......................................................................................................................... 17
  3.1.1 Minnesgärde waterworks network ................................................................................. 17
  3.1.2 Råssnäs waterworks network ....................................................................................... 18
  3.1.3 Tallvägen waterworks network .................................................................................... 19
  3.1.4 Sjöbo waterworks network .......................................................................................... 20
  3.1.5 Överby waterworks network ......................................................................................... 21
3.2 Data collection ..................................................................................................................... 22
  3.2.1 Raw water data ............................................................................................................. 22
  3.2.2 Healthcare guide 1177 data ......................................................................................... 23
  3.2.3 Maps ............................................................................................................................. 23
  3.2.4 SCB data ....................................................................................................................... 23
4. Methodology .......................................................................................................................... 24
  4.1 Tools used ......................................................................................................................... 24
  4.2 Raw water dataset analysis ............................................................................................... 24
  4.3 Symptom dataset analysis ............................................................................................... 24
5. Results and Discussion ........................................................................................................... 27
  5.1 Minnesgärdet waterworks and Östersund ................................................................. 27
  5.2 Råssnäs waterworks and Motala .................................................................................... 40
  5.3 Tallvägen waterworks and Härnösand ........................................................................... 50
  5.4 Sjöbo waterworks and Borås .......................................................................................... 59
  5.5 Överby waterworks and Trollhättan ................................................................................. 70
  5.6 Comparison across the five counties ............................................................................... 81
  5.7 General error sources and limitations .......................................................................... 87
6. Conclusions and future research ......................................................................................... 88
References ................................................................................................................................. 89
Appendix I – Other maps of interest ...................................................................................... 93
List of figures

Figure 1: Waterworks locations in Sweden ................................................................. 17
Figure 2: Minnesgården waterworks location and distribution areas .......................... 18
Figure 3: Råssnäs waterworks location and distribution areas .................................. 19
Figure 4: Tallvägen waterworks location and distribution areas ............................... 20
Figure 5: Sjöbo waterworks location and distribution areas ..................................... 21
Figure 6: Överby waterworks location and distribution areas ................................... 22
Figure 7: Query editor example used in the join process ......................................... 25
Figure 8: Hot spot analysis example (source: ESRI) .................................................. 26
Figure 9: Time-series data for 2009 through 2015 from raw water retrieved from Minnesgården influx A: Coliforms, B: E.Coli .......................................................... 28
Figure 10: Time series data for 2014 through 2015 from raw water retrieved from Minnesgården influx A: Enterococci B: Phages ..................................................... 29
Figure 11: Clinical characteristics examined in the Minnesgården waterworks distribution network ........................................................................................................ 30
Figure 12: Total symptoms reported sorted in age groups in areas supplied by Minnesgården waterworks .......................................................... 30
Figure 13: Total symptoms reported per month in the Minnesgården water distribution network ........................................................................................................ 31
Figure 14: Symptom allocation during January of 2009 in Minnesgården supply network .... 32
Figure 15: Symptom distribution during March 2009 in the Minnesgården waterworks network ........................................................................................................ 33
Figure 16: Symptoms reported to 1177 during the Cryptosporidium outbreak of November 2010 .......................................................... 34
Figure 17: Symptoms reported to 1177 during the Cryptosporidium outbreak in December 2010 .......................................................... 35
Figure 18: Symptoms reported to 1177 during May of 2010 in the Minnesgården supply network ........................................................................................................ 36
Figure 19: Symptom distribution during March 2014 in Östersund and surrounding areas ... 37
Figure 20: Density of total symptoms reported per square km in Östersund and surrounding areas .......................................................................................... 38
Figure 21: Areas identified as statistically significant based on symptom counts in the Minnesgården network ............................................................................. 39
Figure 22: Incidence proportions in areas supplied by Minnesgården waterworks between 2009 and 2015 ..................................................................................... 40
Figure 23: Time series data from raw water samples retrieved from the influx of Råssnäs waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli, C: Turbidity ............... 42
Figure 24: Time series data from raw water samples retrieved from the influx of Råssnäs waterworks through 2014 to 2015 for A: Clostridium, B: Phages ......................... 43
Figure 25: Clinical characteristics examined in Råssnäs waterworks supply areas through 2009 to 2015 ..................................................................................... 44
Figure 26: Distribution of reported symptoms by age groups in Råssnäs distribution areas ... 44
Figure 27: Symptom count per month in Råssnäs water distribution network .......... 45
Figure 28: Symptom count allocation in Råssnäs waterworks supply network during March 2014 .......................................................... 46
Figure 29: Symptom count in areas supplied by Råssnäs waterworks during April of 2014 .. 47
Figure 30: Density map of total symptoms per square km for areas supplied by Råssnäs waterworks during 2009-2015 .......................................................... 48
Figure 31: Areas identified as statistically significant based on total symptom counts in the Råssnäs distribution network ................................................................. 49
Figure 32: Symptom incidence in areas supplied by Råssnäs waterworks through 2009 to 2015 ........................................................................... 50
Figure 33: Time series data from raw water recovered from the influx of Tallvägen waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli .......................................................... 52
Figure 34: Time series data from raw water recovered from the influx of Tallvägen waterworks through 2014 to 2015 for A: Enterococci, B: Clostridium, C: Phages .......................... 53
Figure 35: Clinical characteristics examined in the areas supplied by Tallvägen waterworks through 2012 to 2015 .................................................................................. 54
Figure 36: Distribution of reported symptoms by age groups in areas supplied by Tallvägen waterworks .................................................................................. 54
Figure 37: Symptom counts aggregated by month in areas supplied by the Tallvägen waterworks .............................................................................. 55
Figure 38: Symptom allocation in areas supplied by the Tallvägen waterworks during January 2015 .................................................................................. 55
Figure 39: Symptom allocation in areas supplied by the Tallvägen waterworks during April 2013 .................................................................................. 56
Figure 40: Density map of symptoms per square km in the Tallvägen waterworks network.. 57
Figure 41: Areas identified as statistically significant based on total number of symptoms inside the Tallvägen waterworks supply network .................................................................. 58
Figure 42: Incidence proportions in areas supplied by Tallvägen waterworks through 2012 to 2015 .................................................................................. 59
Figure 43: Time series concentrations of Phages detected in raw water retrieved from Sjöbo waterworks through 2014 to 2015 .............................. 60
Figure 44: Time series concentrations of indicator organisms detected in raw water retrieved from Sjöbo waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli .................. 61
Figure 45: Time series data of pathogens detected in raw water retrieved from Sjöbo waterworks influx through 2013 to 2015 for A: Enterococci, B: Clostridium .............................. 62
Figure 46: Clinical characteristics reported to 1177 in areas supplied by Sjöbo waterworks through 2009 to 2015 .................................................................................. 63
Figure 47: Distribution of reported symptoms by age groups in areas supplied by Sjöbo waterworks .............................................................................. 63
Figure 48: Symptoms reported in areas supplied by Sjöbo waterworks aggregated by month 64
Figure 49: Symptom spatial allocation in areas supplied by Sjöbo waterworks during March 2013 .................................................................................. 65
Figure 50: Symptom spatial allocation in areas supplied by Sjöbo waterworks during April 2013 .................................................................................. 66
Figure 51: Symptom spatial distribution in areas supplied by Sjöbo waterworks during February of 2012 .................................................................................. 67
Figure 52: Symptom density per square km in areas supplied by Sjöbo waterworks .................................................................................. 68
Figure 53: Areas identified as statistically significant based on their total symptoms between 2009 and 2015 in the Sjöbo waterworks distribution network .................................................................. 69
Figure 54: Incidence proportions in areas supplied by Sjöbo waterworks through 2009 to 2015 .................................................................................. 70
Figure 55: Time series data from raw water samples retrieved from Överby waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli .......................................................... 72
Figure 56: Time series data from raw water samples retrieved from Överby waterworks through 2014 to 2015 for A: Enterococci, B: Phages, C: Clostridium .............................. 73
Figure 57: Clinical characteristics reported to 1177 from areas supplied by Överby waterworks through 2009 to 2015 .......................................................... 74
Figure 58: Aggregation of symptoms to age groups in areas supplied by Överby waterworks .......................................................... 74
Figure 59: Aggregation of symptoms by month in areas supplied by Överby waterworks .... 75
Figure 60: Symptom count allocation in areas supplied by the Överby water treatment plant during April 2013 .......................................................... 76
Figure 61: Symptom count allocation in areas supplied by Överby waterworks during November 2012 .......................................................... 77
Figure 62: Symptom count allocation during December 2012 in areas supplied by Överby waterworks .......................................................... 78
Figure 63: Total symptom density map for areas supplied by Överby waterworks .......... 79
Figure 64: Areas identified as statistically significant based on total number of symptoms in the Överby waterworks network .......................................................... 80
Figure 65: Incidence proportions in areas supplied by Överby waterworks through 2009 to 2015 .......................................................... 81
Figure 66: Impact of symptom count per month on total population residing in the examined postal areas from each county .......................................................... 82
Figure 67: Symptoms in Råssnäs distribution network and percentage of population affected in those areas during January 2013 .......................................................... 83
Figure 68: Symptoms in Råssnäs distribution network and percentage of population affected in those areas during August 2014 .......................................................... 84
Figure 69: Symptoms in Råssnäs distribution network and percentage of population affected in those areas during September 2014 .......................................................... 85
Figure 70: Symptoms in Råssnäs distribution network and percentage of residents affected in those areas during October 2014 .......................................................... 86
Figure 71: Yearly comparison of incidence rates across municipalities .......................... 87
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>km(^2)</td>
<td>square kilometer</td>
</tr>
<tr>
<td>ml</td>
<td>millilitres</td>
</tr>
<tr>
<td>l</td>
<td>litres</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet Light</td>
</tr>
<tr>
<td>VTEC</td>
<td>Verotoxigenic Escherichia Coli</td>
</tr>
<tr>
<td>MID</td>
<td>Minimum Infectious Dose</td>
</tr>
<tr>
<td>MSB</td>
<td>Myndigheten för Samhällsskydd och Beredskap (Swedish Civil Contingencies Agency)</td>
</tr>
<tr>
<td>COD</td>
<td>Chemical Oxygen Demand</td>
</tr>
<tr>
<td>HUS</td>
<td>Haemolytic Uremic Syndrome</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony Forming Unit</td>
</tr>
<tr>
<td>MPN</td>
<td>Most Probable Number</td>
</tr>
<tr>
<td>EPA</td>
<td>Environment Protection Agency</td>
</tr>
<tr>
<td>NTU</td>
<td>Nephelometric Turbidity Units</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographical Information System</td>
</tr>
<tr>
<td>SLU</td>
<td>Sveriges Lantbruksuniversitet (Swedish University of Agricultural Sciences)</td>
</tr>
<tr>
<td>SCB</td>
<td>Statistiska Centralbyrån (Statistics Sweden)</td>
</tr>
<tr>
<td>FoHM</td>
<td>Folkhälsomyndigheten (Public Health Agency of Sweden)</td>
</tr>
<tr>
<td>SGU</td>
<td>Sveriges Geologiska Undersökningar (Geological Survey of Sweden)</td>
</tr>
<tr>
<td>SMI</td>
<td>Smittskyddsinstitutet (Swedish Institute for Infectious Disease Control)</td>
</tr>
</tbody>
</table>
1. Introduction
1.1 Background

Water-related diseases are a major cause of health issues and mortality worldwide. Newly-recognised pathogens as in disease causing organisms and new strains of existing pathogens are discovered that present significant and constant challenges to the public health sectors. Drinking water contaminated with pathogens or using it in food preparation will typically cause a variety of unwanted symptoms in the human or animal population that consumes it. Those at greatest risk of contracting an infection are usually infants and young children or people whose immune system is compromised and lastly people that suffer from an existing sickness (WHO, 2008).

The supply of safe drinking-water involves the use of several mechanisms or barriers that prevent the entry and transmission of dangerous pathogens from raw water to the one that finally arrives to the consumer’s taps. It is the responsibility of water producers to provide safe drinking water in the population of its respective supply network as a component of effective policy for health protection. Microbiological analyses are carried out throughout the year in labs to validate the suitability of water for drinking purposes and to determine the most sufficient treatment techniques. Those techniques generally differ between counties depending on the water characteristics and how contaminated it is expected to be (Drinan, 2001).

Studies presented by research teams in known past outbreaks of waterborne pathogens established connections between irregularities of water quality characteristics and increased number of gastrointestinal contacts to hospitals or health agencies that provide support and advice via telephone triage (Tornevi, 2014). Further research is required in order to establish a link between pathogens and unwanted gastrointestinal symptoms that could be observed in the general population as a direct result of their entry in drinking water supplies. As such, connecting datasets concerning pathogens and reported symptoms is a necessity that can be achieved through visualisation and identification of irregularities that might show possible routes of infection.

1.2 Objectives

The purpose of this Master thesis is to examine variations and peaks in notable waterborne pathogen concentrations found in the raw water sources of five surface waterworks in different parts of Sweden and to analyse symptoms that have been reported in their regions of supply that might be attributed to a presence of unwanted microorganisms in drinking water. As such the project aims to identify risk situations of such pathogen concentration peaks in raw water that might be the subject to possible outbreaks of diseases related to those pathogens.

In an ongoing study Microbiological drinking water risks: Risk Classification of Swedish surface raw water (in Swedish “Mikrobiologiska dricksvattenrisker – Riskklassning av svenska ytravatten”), that is financed by the Swedish Civil Contingences Agency (MSB), samples are taken and analysed weekly from six surface waterworks. In this thesis parts of the results from these analyses will be used in order to determine a possible correlation between the pathogen distributions of concentrations and public health deterioration in terms of gastrointestinal symptoms that were reported by telephone triage in surrounding populated areas. Visualization is the key to display such complex phenomena in terms of graphs, statistical tables and maps.
2. Literature review

2.1 Indicator organisms and pathogens

Drinking water can be an important source of infectious agents, particularly of those that cause enteric infections. While person-to-person contact is just as important it is common to recognise water as a source of disease. Not all individual members of any population sample will be susceptible to a pathogenic organism in the water in the same way. Waterborne infections will typically depend on the following factors: (WHO, 2003b).

- the amount of any pathogenic organism in drinking water
- how virulent and infectious the pathogen strain is
- the amount of water ingested by individuals which has not been adequately disinfected
- the minimum infectious dose (MID) of the ingested pathogen
- the immune capability or susceptibility of individuals

Below are some short descriptions and characteristics of typical microorganisms, either pathogens or indicators and chemical measurements that are included in the aforementioned project analysis to be examined for the purposes of this thesis and are typically related to gastrointestinal symptoms in the human population if found in drinking water.

2.1.1 Escherichia Coli

Escherichia Coli or most commonly E. Coli is one of the most common species of coliform bacteria, that being a large assemblance of various bacterial species cultured together as a single group. It is a normal component of the intestines in humans and most warm-blooded animals and are usually found in sewage in high numbers (WHO, 2004).

In a typical bacteriological examination of water, it is not possible to test for each possible type of disease causing organisms. As such, Escherichia Coli is routinely used as an indicator of presence of other malicious pathogens on a statistical basis, showing the likelihood of other pathogens being present and contracting a disease by consuming infected water. The reason for that is because it is easy to culture and to be examined in the laboratory (Ashbolt, 2001).

The safe levels of E. Coli vary in recreational water such as beaches suitable for swimming and lakes but drinking water itself should have no E. Coli present after proper treatment. If E.Coli is found in drinking water then typically other malicious organisms might be present (SLVFS, 2001).

2.1.2 Verotoxigenic Escherichia Coli

The term Verotoxigenic E. coli (VTEC) refers to special strains of E. Coli that are able to produce toxin and are particularly virulent and infectious. Those strains have emerged as important gastrointestinal pathogens for individuals of all ages, but with an increased incidence and severity of illness in young children and the elderly (Griffin, 1991).

Most outbreaks of VTEC related gastroenteritis have occurred after consumption of under cooked food and non-pasteurized milk products or water contaminated with the germs. The germs are also contagious from one infected person to another. VTEC infection is more common in the spring and summer months than during the winter (Rowe, 1995). VTEC strains
have been shown to survive similarly to regular E. Coli strains in routine drinking water conditions (Duncan, 2003).

People infected by VTEC typically manifest the following symptoms:
- Severe stomach pain and bloody diarrhoea that occur one to eight days after consumption of contaminated food or water. Dehydration due to loss of fluid is also common.
- Fever, if present, is usually mild.
- People usually recover without problems after 10 days but the disease can have serious effects especially for young children and the elderly. A very serious complication in young children is a type of kidney failure called haemolytic uremic syndrome (HUS) (Karmali, 1983).

2.1.3 Coliforms

Total coliforms include organisms that can survive and grow in water. They are not as useful as an index of faecal pathogens, but they can be used as an indicator of water treatment effectiveness. However, there are better indicators for such purposes, since as a disinfection indicator, the test for total coliforms is far slower and less reliable than direct measurement of disinfectant residual. In addition, total coliforms would be destroyed much easier than some more persistent species of enteric viruses, as they are more sensitive to disinfection (Ashbolt, 2001).

Total coliforms should be absent immediately after disinfection, and the presence of these organisms would indicate that the treatment process was not utterly successful. Their presence in distribution systems and stored water supplies can reveal microorganism regrowth and possible existence of other pathogens as well (WHO, 2004).

2.1.4 Enterococci

Intestinal enterococci typically occur in the faeces of humans and other warm-blooded animals. They can be present in large numbers in sewage and water environments polluted by sewage or wastes from humans and animals (WHO, 2004).

They tend to survive longer in water environments than E. Coli and are generally more resistant to common disinfectants. Intestinal enterococci have been used in testing of raw water as an index of faecal pathogens that have a higher survival rate than E. Coli. In addition, they have been used to test water quality after repairs to distribution systems have been made (Ashbolt, 2001).

Enterococci are typically associated with gastrointestinal diseases and as such should have no presence in drinking water after treatment (SLVFS, 2001).

2.1.5 Clostridium

The presence of Clostridium or C. perfingens as otherwise noted, in drinking-water is usually an indicator of intermittent faecal contamination. Potential sources of contamination should be investigated at any occurrence as the gastrointestinal symptoms related to its presence can be quite severe. Typical filtration processes that have been designed to remove enteric viruses should also remove the presence of Clostridium. However, detection in water immediately after
treatment should lead to an immediate investigation on the performance of the water treatment facility’s filters (Franco, 1993).

Clostridium is virtually always present in sewage. It produces spores that are exceptionally resistant to unfavourable conditions in water environments, including Ultraviolet Light irradiation, temperature and pH extremes, and disinfection processes, such as chlorination (Franco, 1993). Because of the spores’ exceptional resistance to disinfection processes and other unfavourable environmental conditions, C. perfringens has been proposed as an index of enteric viruses in treated drinking-water supplies (Ashbolt, 2001).

At any rate, Clostridium should have no presence in drinking water after decontamination (SLVFS, 2001).

### 2.1.6 Cryptosporidium

Cryptosporidium is well recognised as a cause of severe watery diarrhoea that can last several days to a whole week. (Dubey et al., 1990). Reported cases of Cryptosporidium infections in the population range from 0.6 to 20% depending on the geographic locale and while it has greater prevalence in lands where water treatment systems are inadequate it has been reported in both USA and UK as the most significant source of waterborne diseases throughout the world (Duncan, 2003).

Cryptosporidium oocysts have been detected in 11.5% of surface waters in Sweden (SMI, 2011). Cryptosporidium is a challenge to identify at laboratory conditions because of low detection limits and because it can easily be confused with other artifacts similar in appearance (Casemore, 1985). At any rate, Cryptosporidium should have no presence in drinking water of any kind (SLVFS, 2001). According to the website of Folkhälsomyndigheten, in 2014 a total number of 414 infections was reported in Sweden.

### 2.1.7 Giardia

Giardia has been known as a human parasite for 200 years. Symptoms generally include diarrhoea and abdominal cramps that range in severity. Giardiasis is usually self-limiting, but it may be chronic in some patients, lasting more than 1 year, depending on their immune system. Studies on human volunteers have shown that fewer than 10 cysts would constitute a meaningful risk of infection (Stuart, 2003).

Waterborne outbreaks of giardiasis have been associated with drinking-water supplies for over 30 years. Giardia cysts are more resistant than enteric bacteria to disinfectants such as chlorine, but they are not as resistant as Cryptosporidium oocysts. Control measures that can be applied to manage potential risk from Giardia include prevention of source water contamination by human and animal waste, followed by proper treatment, disinfection and protection of water during distribution. Due to the resistance of the cysts to various disinfectants, E. Coli cannot be relied upon as an index of the presence of Giardia in drinking-water supplies (WHO, 2004).

The website of Folkhälsomyndigheten reports a total number of 1260 Giardia infections in Sweden during 2014. According to SMI, Giardia was found in 4% of Swedish surface waters (SMI, 2011).
2.1.8 Campylobacter

Campylobacter exist in a variety of environments. Wild and domestic animals, especially poultry, wild birds and cattle, are important reservoirs along with pets and other animals. Food, including meat and unpasteurized milk, are important sources of Campylobacter infections. Water is also a significant source. The occurrence of the organisms in surface waters has proven to be strongly dependent on rainfall and water temperature (Frost, 2001). Campylobacter are faecally borne pathogens and are not particularly resistant to disinfection. Hence, E.coli is an appropriate indicator for the presence/absence of Campylobacter in drinking-water supplies (WHO, 2004).

Contaminated drinking-water supplies have been identified as a source of outbreaks. The number of cases in these outbreaks ranged from a few to several thousand, with sources including inadequately chlorinated surface water supplies and faecal contamination of water storage reservoirs by wild birds (Koenraad, 1997).

Amongst the diseases attributed to waterborne pathogens the most common bacterium responsible is Campylobacter. In USA 45% of the people reporting diarrhoea are found to have infections caused by Campylobacter. These infections manifest quite fast, typically within 24 hours of ingestion, however the illness itself is usually self-limited and lasts up to 3 days. The disease occurs sporadically and large outbreaks are unusual, however the Centers for Disease Control and Prevention typically receive 10000 reports of the disease each year, which would roughly translate to 1 infected person in 100000 (Madigan et al., 2000).

In Sweden there were a total of 8288 cases of Campylobacter infections reported throughout 2014 according to the website of Folkhälsomyndigheten.

2.1.9 Salmonella

Salmonella is typically associated with poultry, wild birds and reptiles. The pathogens typically gain entry into water systems through faecal contamination from sewage discharges, livestock and wild animals. Contamination has been detected in a wide variety of foods and milk as well (WHO, 2004).

Salmonella is relatively sensitive to disinfection. Control measures that can be applied to manage risk include protection of raw water supplies from animal and human waste, adequate treatment and protection of water during distribution. Escherichia Coli is a generally reliable index for Salmonella in drinking water supplies (WHO, 2004).

Salmonella infections are relatively mild and similar to characteristics with Campylobacteriosis. Approximately 30% of the people with diarrhoea who visit a doctor in the USA are found to have Salmonellosis. In Sweden there was a total number of 2211 Salmonella infections reported in 2014 according to the Folkhälsomyndigheten website.

2.1.10 Phages

Coliphages are viruses that use E. Coli and closely related bacteria as hosts for replication and as such can be released by those hosts into the faeces of humans and other animals (WHO, 2004).
They are typically used in water quality assessment as indicators to determine the presence or absence of viruses since they share many characteristics such as structure and mode of reproduction. However, they cannot be absolutely relied upon to determine the existence of viruses since there is no direct correlation between their numbers and the numbers of enteric viruses. They are typically used in lab testing to determine water quality as they are easy to grow within a day, but are unsuitable for operational monitoring. If they exist in drinking water then their presence would most likely mean that disinfection was not effective for enteric viruses (Ashbolt, 2001).

2.1.11 Norovirus

Noroviruses are a major cause of acute viral gastroenteritis in all age groups. Symptoms typically include nausea, vomiting and abdominal pain. Close to 40% of infected individuals manifest diarrhoea; some might have fever, chills and headache. Since a lot of cases present with vomiting only and no diarrhoea, the associated condition is also known as “winter vomiting disease.” The symptoms are usually relatively mild and rarely last for more than 3 days. High attack rates in outbreaks indicate that the infecting dose is low and that the disease is highly contagious. Its source being infected water, food or contact with other people that are already affected by it. Since viruses typically present high resistance to disinfection, E. Coli cannot be used as an indicator of presence or absence (WHO, 2004).

2.2 Pathogen related attributes

2.2.1 Chemical Oxygen Demand

In environmental chemistry, the chemical oxygen demand (COD) test is commonly used to indirectly measure the combined effect of substances and conditions in water. Most applications of COD determine the amount of organic pollutants found in surface water, making COD a useful measure of water quality. It is expressed in milligrams per liter (mg/L) also referred to as ppm (parts per million) and indicates the mass of oxygen consumed per liter of solution (Boyles, 1997).

2.2.2 Turbidity

In simple terms turbidity refers to how clear the water is. The clarity of water is the first characteristic people would notice when looking at water. Natural water that is very clear (in this case water of low turbidity) will allow us to see through it at considerable depths, while high turbidity water on the other hand appears cloudy (Spellman, 2014).

High levels of turbidity can protect microorganisms from the effects of disinfection, stimulate and enhance the growth of bacteria and quite possibly give rise to a significant chlorine demand required to destroy pathogens (WHO, 2003b).

Turbidity adversely affects the efficiency of disinfection. It is also measured to determine what type and level of treatment are needed. Turbidity tests can be carried out with a tube that allows a direct reading in nephelometric turbidity units (NTU) and is routinely used by water treatment plants to determine the characteristics of water that is about to undergo treatment (WHO, 2004).
Following an investigation after a large Cryptosporidium outbreak that occurred in Milwaukee, Morris et al. (1996) presented an analysis of hospital records and readings of water turbidity over a period of 16 months before the outbreak itself. The team discovered a strong correlation between rises in turbidity values and an increased attendance of children presenting gastrointestinal symptoms at hospitals.

According to Livsmedelsverket guidelines, water should not exceed turbidity values of 0.5 NTU when it enters the distribution system and any abnormal changes in values should be investigated and dealt with instantly (SLVFS, 2001).

2.2.3 pH

Most bacteria grow adequately at neutral pH levels close to a value of 7. Extreme acidic or basic conditions that are far from the value of 7 generally inhibit growth, although for some bacteria these conditions might actually improve their growth rate (Spellman, 2014).

It is necessary to know the pH of water, because more alkaline water requires a longer contact time with chlorine for adequate disinfection with chlorination being possibly ineffective above pH 9. Suggested levels according to USEPA are between 6.5 and 8.5. Low Ph is associated with a bitter metallic taste while high Ph with a slippery feel and soda taste (USEPA, 1994).

2.2.4 Temperature

Temperature is one of the most important parameters in surface water systems because such systems are subject to great temperature variations, as they are directly affected by the different weather conditions.

Generally, the warmer the environment, the faster the rate of growth for various bacteria. For most bacteria, each increase of 10°C would double their growth rate. Heat can also be used to kill bacteria. This is not an exclusive rule however since some pathogens such as Cryptosporidium and Giardia can survive better in low temperatures (Spellman, 2014).

Water temperature can affect the efficiency of water treatment processes. High water temperatures can increase chlorine demand because of increased reactivity, as well as increased levels of algae and other organic matter in the raw water (Drinan, 2001).

2.3 Treatment process

2.3.1 Factors affecting process and regulations

Surface waters typically contain waterborne pathogens that can be largely responsible for infectious epidemic diseases. Bacterial numbers increase significantly during storm events which increase contamination by washing material from the ground surface into surface waters. As a result the highest bacterial numbers in surface waters are typically observed during rainy periods (Spellman, 2014). Tornevi et.al (2014) showed that generally the lowest raw water quality occurs 2 days after rainfall and remains as such for several more days. Heavy rainfall was linked with highly increased quantities of E.Coli and increases in turbidity levels along with increased levels of gastrointestinal contacts to 1177 (Tornevi, 2014).

Water treatment operators are primarily concerned with the elimination of bacterial pathogens and parasites that can be harmful for human health. These pathogens enter potential drinking
water supplies through faecal contamination and are ingested by humans if the water is not properly treated and disinfected. It is generally recommended that owners of all public water supplies collect raw water samples and deliver them to a certified laboratory for bacteriological examination at least monthly. The number of samples required typically varies according to the number of people served by the waterworks. The general guideline indicates that at least one sample per month should be collected for each 1000 persons served by the waterworks (WHO, 2004).

2.3.2 Operational monitoring

In most cases, operational monitoring will be based on simple and rapid observations or tests, such as measuring turbidity and pH, inspecting water colour and the structural integrity of the filters and machinery involved in the treatment processes, rather than complex microbial or chemical tests that typically require a lot of time. The complex tests that are carried out in labs are generally applied as part of validation and verification activities rather than as part of operational monitoring (WHO, 2004).

When designing a new water treatment system, all raw water quality factors should be taken into account in selecting technologies for the disinfection treatment. Variations in turbidity and other parameters of raw surface waters can be significant, and allowance must be made for such deviations. Treatment plants should be designed to take account of variations that are known or expected to occur with significant frequency rather than for average water quality; otherwise, filters may rapidly become blocked or sedimentation tanks overloaded with adverse effects for both the treatment plant and public health itself (WHO, 2004).

2.3.3 Methods of disinfection

Water treatment unit processes include (1) storage of raw water, (2) prechlorination, (3) coagulation–floculation, (4) water softening, (5) filtration, and (6) disinfection. Filtration and disinfection are the primary means of removing contaminants and pathogens from drinking water supplies. In each of these unit processes, the reduction or elimination of pathogens is variable and influenced by a number of factors such as sunlight, sedimentation, and temperature (Spellman, 2014).

The most commonly used disinfection processes are Chlorination, Ozonation, and UV irradiation. These methods are very effective in killing bacteria and can be reasonably effective in inactivating viruses (depending on their type and strain) and many parasites, including Giardia and Cryptosporidium. For effective removal of unwanted bacteria, filtration with the aid of coagulation/floculation (which reduces particles and turbidity) followed by disinfection (usually with a combination of disinfectants) is the most practical method (WHO, 2004).

Ozonation refers to Ozone gas that enters the water supply in order to oxidize pathogens and other chemicals, neutralizing their health impact. It imparts no flavour or odour, but unlike chlorine, it does not continue to provide sanitization throughout the distribution system. It is however an expensive system to install and operate (Madeline, 2013).
Chlorine is known to have no effect against some organisms such as Cryptosporidium, however Ultraviolet Light has been demonstrated to kill these oocysts. UV light should be required as the main disinfectant along with chlorine used for residual action against bacteria in distribution. It is used to break down organic molecules in the water supply, effectively eliminating most pathogens. It is safe to operate, and relatively inexpensive for small drinking water systems. UV light however does not provide continual protection within the distribution system (Harrington, 2001).

The World Health Organisation has established a guideline value of 5mg/litre for free chlorine in drinking-water. Most individuals are able to taste or smell chlorine in drinking-water at concentrations well below 5mg/litre, and some at levels as low as 0.3 mg/litre. At a residual free chlorine concentration of between 0.6 and 1.0 mg/litre, there is an increasing likelihood that some consumers may object to the taste. The taste threshold for chlorine is below the health-based guideline value (WHO, 2003).

It is possible to reduce the concentration of chlorine to less than 0.1 mg/litre. However, it is normal practice to supply water with a chlorine residual of a few tenths of a milligram per litre to act as a preservative during distribution. According to Livsmedelsverket guidelines, the value of chlorine in drinking water should not go over 0.4 mg/l (SLVFS, 2001).
2.4 Known outbreaks in Sweden and internationally attributed to waterborne pathogens

In November 2010 about 27,000 inhabitants of Östersund (almost 50% of the population) in Sweden, were affected by a waterborne outbreak of Cryptosporidiosis. The outbreak had a high attack rate, especially among young and middle-aged people. Cryptosporidium was identified as the pathogen responsible for the outbreak as it was found in drinking water showing that perhaps the disinfection process was unsuccessful and a general recommendation to boil the water before consumption was issued for the following three months. Although the cause of the outbreak was not made totally clear, the suspected source was sewage water being discharged directly into a stream, which led to the lake from where the drinking water was taken, instead of being sent to a local sewage treatment plant. This outbreak was documented as the largest known outbreak in Europe (Widerström, 2014).

A second city, Skellefteå was also affected by the same parasite in April 2011, with around 20000 people becoming ill and forcing residents to boil tap water for five months (Skellefteå Municipality, 2011).

The largest known documented waterborne outbreak in the world happened in Milwaukee in 1993. Cryptosporidium found in drinking water was identified as the source of the infection of approximately 400000 people out of which 100 died (MacKenzie et al., 1994). Treated water showed turbidity levels well above normal values, while the investigations of MacKenzie and the Centers for Disease Control and Prevention revealed that the outbreak was caused by Cryptosporidium oocysts that had successfully passed through the filtration system of one of the city's water-treatment plants. This abnormal condition at the water purification plant lasted for two weeks, after which, the plant was shut down.

While the incidents described above are not every day phenomena, they prove that insufficient microbiological barriers or possible malfunctions at the filtration systems of water treatment facilities can have dire consequences for public health if pathogens successfully pass in drinking water.
3. Study areas and data description
3.1 Study areas

The area of study for this thesis includes municipalities supplied by five waterworks located in various parts of Sweden. The locations of these waterworks are shown in Figure 1 below:

i. Östersund/Minnesgärde waterworks
ii. Motala/Råssnäs waterworks
iii. Härnösand/Tallvägen waterworks
iv. Borås/Sjöbo waterworks
v. Trollhättan/Överby waterworks

Figure 1: Waterworks locations in Sweden

3.1.1 Minnesgärde waterworks network

The Minnesgärde water treatment plant supplies Östersund, Frösön, Orrviken, Brunflo and Lockne with drinking water, an amount that corresponds to 16-17 million litres per day for consumption. The raw water is taken from the lake Storsjön which is next to the city of Östersund (Vatten Östersund, 2014). The water passes through two microbiological barriers (ozonation and chloramination) as recommended by the drinking water regulations in Sweden for surface waterworks. A UV water disinfection system was also installed but only after the outbreak of November 2010 (Widerström, 2014). The water distribution area is shown in Figure 2 below. It must be noted that all the maps shown in these sections are based on postal area polygons and as such are not entirely accurate.
3.1.2 Råssnäs waterworks network

The Råssnäs water treatment plant supplies Motala, Fågelstad, Varv, Österstad, Norrsten, Fornåsa, Klockrike, Nykyrka, Fivelstad, Lönsås, Ekebyborna, Älvan and Bona with close to 7.5 million litres of drinking water daily (Motala kommun, 2015). The raw water is taken from Motalaviken and passes through the standard microbiological barriers. A UV light water disinfection system is also present since 1999 (Peterson, 2009). Figure 3 below displays the location of the waterworks along with the areas it supplies with drinking water:
3.1.3 Tallvägen waterworks network

The Tallvägen treatment plant supplies the county of Härnösand with drinking water along with six smaller treatment plants that are located in Viksjö, Häggdånger, Rö, Brunne, Starred, Smitingen (Löfnäs, 2015). Since the raw water dataset only includes samples taken from the influx of Tallvägen waterworks, the smaller treatment plants and their regions of supply will be excluded from the analysis of this thesis.

The raw water to be processed by the Tallvägen treatment plant is taken from the lakes of Långsjö and Bondsjö and goes through the standard microbiological barriers including a UV light disinfection system (Sjöstrand, 2015). The treatment plant location along with its respective distribution area is shown in Figure 4 that follows:
3.1.4 Sjöbo waterworks network

The Sjöbo waterworks treatment plant supplies Borås, Fristad, Viskafors, Borgstena, Kinnarumma, Svaneholm, Sjömarken, Sandared, Sandhult, Olsfors and Hultafors with drinking water. The raw water is taken from the lake Öresjö and goes through the standard microbiological barriers including a UV light disinfection system (Borås Energi och Miljö, 2015). The water treatment plant location and areas that are supplied with drinking water are shown in Figure 5 below:
3.1.5 Överby waterworks network

The Överby waterworks treatment plant supplies close to 50 000 people in the areas Trollhättan, Sjuntorp, Velanda, Åsaka, Norra Björke (Åström, 2013). The raw water is collected from Göta Älv river and is processed with the standard disinfection systems including a UV light system (Trollhättan Energi, 2015). The treatment plant location and areas supplied by it are shown in Figure 6 below:
3.2 Data collection

The dataset used in this thesis were kindly provided by a number of sources and are outlined below:

3.2.1 Raw water data

A dataset of sample measurements at different times of year that were collected from the influx of the five water treatment plants shown in Figure 1. The dataset displays the most notable pathogens and indicators that were briefly analysed above and is based on samples that were analysed on a weekly or monthly basis. It must be noted that the entire dataset is based on samples that were taken before the water treatment.

The source for this data is Folkhälsomyndigheten (Public Health Agency of Sweden) and the project “Mikrobiologiska dricksvattenrisken – Riskklassning av svenska ytråvatten” with project coordinator Livsmedelsverket (National Food Agency of Sweden) and covers the period of March 2014 to 2015 in an ongoing project. Data from before these agencies started this project have been retrieved from Vattentäktsarkivet in order to expand the period of study from 2009 to March 2014 in the waterworks of interest. As stated earlier, Vattentäktsarkivet is a database updated by field labs maintaining it in their respective areas and is maintained by SGU.
The period of 2009 to March 2014 is less consistent and samples are not always posted on a weekly basis. In addition not all pathogens and indicators described above are analysed consistently. The study period will be explained with more detail in the respective waterworks analysis made at a later part in this thesis. The file was provided in the form of an Excel spreadsheet.

### 3.2.2 Healthcare guide 1177 data

A dataset of symptoms possibly related to the examined waterborne pathogens that were reported in the five counties where the aforementioned water treatment plants are stationed, collected from the Healthcare Guide 1177 (1177 Vårdguiden) based on telephone triage. Healthcare Guide 1177 is a 24-hour nurse-on call service that provides healthcare advice by telephone (1177) or website ([www.1177.se](http://www.1177.se)). Data from this service have been successfully used in previous works as a means to determine signals for local waterborne and foodborne outbreaks that occurred in Sweden from 2007-2011 ([Andersson et.al, 2014](#)). As such this service can be extremely helpful by providing the means to identify possible future outbreaks and to increase awareness.

The data file comes in the form of an Excel spreadsheet and involves reported symptoms and information on age groups and geography based on areal postal codes for the whole respective county. The information is aggregated and anonymised. The study period covers every day between 2009 until March of 2015 for all examined areas except the county of Härnösand which starts from June of 2011 until March of 2015.

### 3.2.3 Maps

For the purposes of correlating pathogens with related symptoms visualisation has been used as a primary tool. Basemaps used for the visualisations in the five study areas were produced by Lantmäteriet and extracted from the national digital map library. The data were delivered in the RT90 2.5 gov V in raster format.

Postnummerservice (Addressmaster Norden) has supplied the map of Sweden which is divided in polygons based on postal code areas that are used in the analysis of this thesis to geolocate the incident. This map was also delivered in RT90 2.5 gov V coordinate system in the form of a shapefile and is the latest version of January 2015.

### 3.2.4 SCB data

SCB data (Statistics Sweden) used in this research include demographic information about the study period of 2009 to 2015 primarily the number of residents that have been registered in the respective postal codes that are used in the visualisation analysis.
4. Methodology
4.1 Tools used

For the purposes of the analysis that takes place in this thesis two popular programs have been used: ArcMap and Excel.

ArcMap is a powerful software developed by ESRI and is commonly used for visualisation purposes with specialised tools for different fields of GIS. Amongst the various toolboxes that ArcMap provides, Data Management tools and Spatial Analyst were essential for visualising and examining in detail the data provided by 1177.

Excel is a very popular spreadsheet application developed by Microsoft and is commonly used for analysing statistical data, through a variety of graphs, pivot tables and mathematical formulas.

4.2 Raw water dataset analysis

The first step in the analysis was to create summary statistic tables for every examined pathogen with water quality indicators such as turbidity, pH and temperature that are routinely used by the treatment plants for operational monitoring.

The data were aggregated by year and line graphs were created to smooth the visualisation process and identify peaks in the concentrations of pathogens. Those concentrations when identified were noted down along with the respective months presented. As stated earlier high pathogen concentrations or abnormalities in related chemical attributes would indicate that the treatment processes would have to be adjusted accordingly in order to effectively neutralise pathogens before entering drinking water and released in the distribution network. Water that might be contaminated with even a small presence of such organisms could be associated with a rise of gastrointestinal symptoms in the same or possibly the next months. As such the raw water dataset is used for validation purposes.

4.3 Symptom dataset analysis

The dataset includes information such as date of phone call, age of the person that made the phone call or had the symptoms, the symptom or symptoms that were reported in the categories of diarrhoea, blood in stool, abdominal pain, fever and nausea and finally the zip code to geolocate the incident for the whole respective municipality.

The first step was to filter the databases to exclude areas in every county that are not supplied by the respective water treatment facility. Using pivot tables and mathematical formulas in the spreadsheets, the data was aggregated per month to determine number of symptoms and graphs were produced in order to display the percentages of clinical characteristics presented and distribution of symptoms sorted by age groups.

Comparative charts were also created using the demographical information retrieved from SCB in order to display the amount of people affected in regards to the total population residing in the equivalent postal code area of interest. This data was used as a means to normalise the symptoms and compare it across the five counties since the population sizes were different. This was done in order to determine which area had the highest reported symptoms at a given
time period and determine patterns of possible infection. In addition to that, yearly incidence rates per 10000 residents were estimated for all municipalities in the form of a bar chart.

The symptom data was then processed accordingly in a new table that was imported in geodatabases in ArcMap in order to visualise them per month. The objective was to create time-animated rasters that would display the symptoms per month in every affected area, along with density maps that would show the number of symptoms reported per square kilometre and a hot spot analysis would be carried out that would determine which areas had the most symptoms and as such identify clusters of symptoms. In addition to that, maps of incidence proportions of the population residing in the examined areas were created.

Once the data was imported in ArcMap and a proper geodatabase was established, Make Query Table tool was used in order to match postal codes from the zip code shapefile of Sweden and the processed table that contained the postal codes that showed symptoms in every affected area along with the count of those symptoms and the respective date in the form of year and month (YYYYMM). The results were returned in a layer that displays only the matched postal areas in the form of polygons which would be only those supplied with drinking water from the respective waterworks. This allowed the creation of a joined shapefile that could then easily be visualised by month. The resulting maps that show the symptom count per month were used to determine possible sudden rises of symptoms that could be attributed to water borne pathogens. The expression used in Query Editor may be viewed in Figure 7 below as an example taken from the process followed for Borås county.

![Query Builder](image)

Figure 7: Query editor example used in the join process
In order to identify areas that showed the biggest count of symptoms, Hot Spot Analysis tool was used with Inverse distance spatial relationship. This tool identifies statistically significant spatial clusters of high values (hot spots) and low values (cold spots). It creates a new Output Feature Class with a z-score which is a measure of statistical significance indicating whether the observed spatial clustering of high or low values is more pronounced than one would expect in a random distribution of those same values. A high z-score for a feature indicates a spatial clustering of high values and the higher that score is, the more intense the clustering. A z-score near zero indicates no apparent spatial clustering. A visual example as recovered from ESRI ArcGIS online help is shown in Figure 8 below:

Figure 8: Hot spot analysis example (source: ESRI)

In order to create density maps that would display the symptom count spatial distribution, Kernel Density tool was used. It must be noted that in order to properly use this tool exact coordinates of incidents are typically required, however the symptoms database uses zip codes in order to geolocate the incidents which are shown as polygons on the map instead of points. In order to overcome this difficulty Feature to Point tool was used which essentially converts the zip code polygons to points using their centroid locations. Even though this might cause some spatial location errors in larger polygons, the centroids still contain the symptom count attributes of their respective sources and as such may be used to display an approximate spatial distribution of symptoms. After this was accomplished Kernel Density was applied with bilinear interpolation as a means to visualise symptom counts per square kilometre.
5. Results and Discussion

5.1 Minnesgärdet waterworks and Östersund

The analysis of raw water data taken from the influx of Minnesgärdet waterworks is generally more consistent after March of 2014 when Livsmedelsverket and FoHM started their project and contains measurements of all pathogens that were briefly described above even though it is not always on a weekly basis. Before this particular date, complementary data were retrieved from Vattentäktsarkivet that primarily lists E. Coli and Coliforms as indicators for the existence of other pathogens. Measurements of water quality with the exception of temperature also lack a lot in posted measurements. It must also be noted that the year of 2013 is completely absent from the database. A summary table of the pathogen analysis and water quality attributes is listed below:

| Table 1: Summary statistics of raw water samples during 2009-2015 for Minnesgärdet waterworks |
|---------------------------------|-----------------|----------------|-------------|------|-------|-------|
|                                | Samples | Missing | Mean | St.Dev | Minimum | Maximum |
| Coliforms MPN/100 ml           | 302     | 4        | 26,19 | 96,03 | 0       | 1100    |
| E. Coli MPN/100 ml             | 302     | 4        | 2,53  | 14,62 | 0       | 236     |
| Enterococci MPN/100 ml         | 135     | 171      | 0,99  | 1,77  | 0       | 20      |
| Clostridium CFU/100 ml         | 253     | 55       | 1,09  | 3,82  | 1       | 50      |
| Cryptosporidium oocysts/100 l  | 20      | 286      | 1,05  | 0,22  | 1       | 2       |
| Giardia cysts/100 l            | 20      | 286      | 1,40  | 1,79  | 1       | 9       |
| Phages (per 100 ml)            | 95      | 211      | 2,78  | 3,71  | 1       | 20      |
| Campylobacter MPN/l            | 23      | 283      | 0,10  | 0,08  | 0,06    | 0,3     |
| Salmonella MPN/l               | 23      | 283      | 0,06  | 0,00  | 0,06    | 0,06    |
| Verotoxigenic E.Coli MPN/l     | 23      | 283      | 0,07  | 0,03  | 0,06    | 0,18    |
| Chemical Oxygen Demand         | 22      | 284      | 1,09  | 0,45  | 1,7     | 3,8     |
| Turbidity                      | 191     | 115      | 0,38  | 0,21  | 0,13    | 1,3     |
| Temperature                    | 299     | 7        | 6,53  | 3,20  | 2,2     | 14,5    |
| pH                             | 100     | 206      | 7,44  | 0,12  | 7,2     | 7,7     |

Measurements of water quality such as turbidity, COD and pH did not show any particular variation. Turbidity showed a mean value of 0.38 NTU which is well below the established guideline of 1.5 NTU and never exceeded it, while pH remained at neutral levels close to 7 without showing any extreme basic or acidic conditions that would stimulate bacterial growth. Temperature did not show any abnormal pattern and typically varied between 2.2 degrees during colder months and up to 14.5 degrees in warmer months.

When it comes to pathogens, Giardia, Salmonella and VTEC showed little to non-existent presence. The exception to this is Cryptosporidium which was found in drinking water during the outbreak of November 2010, however this data is not present in the examined dataset. The following graphs in Figure 9 and Figure 10 show the time-series concentrations of E. Coli, Coliforms, Enterococci and Phages.
Figure 9: Time-series data for 2009 through 2015 from raw water retrieved from Minnesgärdet influx A: Coliforms, B: E.Coli
Coliforms showed a big increase in concentrations during May, June and October in 2009 and 2010 and then during April and November of 2014. E. Coli spiked particularly high during April and November of 2014. Enterococci also showed an increase during the same months. Finally Phages had significant presence during October of 2014 and February of 2015.

The months identified above as suspect months are cross-referenced to symptom counts to establish a possible correlation between their increased presence and rise of gastrointestinal symptoms. The database extracted from 1177 includes over 20000 phone calls from areas supplied by the waterworks between 2009 and 2015 and a total of 55502 people reside in those postal areas as of 2014. The examined clinical characteristics are presented in Figure 11 below.
Figure 11: Clinical characteristics examined in the Minnesgärden waterworks distribution network

Figure 12: Total symptoms reported sorted in age groups in areas supplied by Minnesgärden waterworks

Figure 12 shows that the most likely age groups to present these symptoms are 0-9 and 20-29. Young children generally have a weaker immune system and the higher amount of calls can also be attributed to worried parents and as such this age group might be overrepresented. Of course the number of inhabitants are not the same for each age group and different age groups would use the 1177 service differently. Abdominal pain appears to be the most dominant symptom in most phone calls as it could be the most immediate symptom to be noticed and reported as a start of gastrointestinal problems.
The chart in Figure 13 presents the symptom count aggregated by month from 2009 to 2015. While during this study period there is an average of 300 monthly symptoms reported as shown by the trendline, in November 2010 and December this amount is three times the average monthly symptom number. These are the months when the Cryptosporidium outbreak occurred. A total of 1457 symptoms were reported to 1177 during this period of two months. However the total number of infected people were estimated to be over 27000 people according to the article mentioned earlier, which is almost 50% of people living in Östersund. This relatively low number of phone calls can perhaps be attributed to the fact that most people would be most likely to contact the nearest doctor and hospital immediately instead of using 1177 service. That number is still high enough to reveal an outbreak during those months. The number of reported symptoms is still high in the start of the next year as well that follows the outbreak. The figure also shows an increase of symptoms during the first three months of 2009 and then during March and April of 2014.
January, February and March of 2009 showed the highest symptom numbers second to the Cryptosporidium outbreak. However reported concentrations of pathogens were not as high in the relative raw water data during those months. It is possible that the disinfection process was not as effective as it should in the waterworks treatment plant during the winter months of 2008 or the start of 2009 however there is no accurate data to support such a claim. It is also interesting to notice that UV light which is exceptionally effective against Cryptosporidium oocysts was installed only after the outbreak of November of 2010. At any rate the visualisation shows that there was a cluster of symptoms particularly in the southern and more populated areas of Östersund.
The outbreak of Cryptosporidium in Östersund and the consequences it had for the general population in terms of health is shown in Figure 16 and Figure 17 below. The maps clearly show a significant number of symptoms for the most areas that were supplied by the waterworks network. According to a report written by Smittskyddsinstitutet (Swedish Institute for Infectious Disease Control) during the 27th of November and 20th of January of 2011 confirmed oocysts of Cryptosporidium were found in drinking water in concentrations of 0.02 to 1.3 per 10 litre, which proves that even a small amount can cause such severity of symptoms (SMI, 2011).
Figure 16: Symptoms reported to 1177 during the Cryptosporidium outbreak of November 2010
A possible correlation that could indicate a problem with the treatment process may also be detected some months prior to the outbreak itself. During May of 2010 raw water samples showed a high concentration of 370 Coliforms per 100 ml and also 34 Clostridium per 100 ml. Even though the number of symptoms shown in Figure 18 below are not very high in number, clusters of symptoms are still formed in the more populated areas to the south of Östersund.
Another possible correlation between drinking water pathogens and a rise of symptoms may be detected in March of 2014 visualised in Figure 19 below. This month showed the third highest amount of symptoms detected inside the city of Östersund but also in Orrviken and Brunflo. By cross checking the water data set it appears that a small presence of Cryptosporidium oocysts (less than 2 oocysts per 100 l) were indeed detected in raw water during March and 9 cysts of Giardia per 100 l were detected during April according to the measurements provided by FoHM. The symptoms reported in April however were shown to be lower in comparison to the
same month in previous years so a seasonal effect might be taking effect. The map showing the symptom allocation during March is shown in Figure 19 below:

Figure 19: Symptom distribution during March 2014 in Östersund and surrounding areas

Figure 20 and Figure 21 presented below show the areas that held the most symptoms by square kilometer and which areas were identified as statistically significant based on their total count of reported symptoms during the whole study period of 2009 to 2015. Most cases were reported in the main city of Östersund and particularly in the central and southern parts.
Figure 20: Density of total symptoms reported per square km in Östersund and surrounding areas
Figure 21: Areas identified as statistically significant based on symptom counts in the Minnesgårdet network

Finally Figure 22 shows the incidence proportion in regards to the population residing in each postal area. The southern parts of Östersund appear more pronounced in symptom to population ratio.

Additional maps of interest showing the symptom allocation for February 2009, January 2011 and April 2014 are shown in Appendix I.
5.2 Råssnäs waterworks and Motala

The measurements of raw water samples as recovered from the influx of Råssnäs waterworks are very consistent after March of 2014 when FohM and Livsmedelsverket started their project. Before that date however, data recovered from Vattentäktsarkivet are lacking a lot of information as samples have been recorded only once per three months for this particular waterworks. It is possible that all the relevant data were analysed but they have not been recorded in the retrieved dataset, since the sources and purposes of the two datasets are different. The summary statistics table based on the available data is presented below.
Temperature typically varies from 0.28 as a lowest point to 20.9 degrees in the warmer months. No significant rises in COD while pH is totally absent from the dataset. The warmer months generally showed a substantial increase in pathogen concentrations as it may be viewed in Figure 23 and Figure 24 below. This increase could be explained by factors such as heavy rainfall, runoff from land or other sorts of sewage impacts. Turbidity showed some variation that exceeded the recommended levels of 1.5 NTU during March and May of 2014. Coliforms were found in high numbers from July to October of 2014. Clostridium showed an increase during June and July of 2014, while Phages showed a significant presence during April of 2014. Cryptosporidium and Giardia oocysts were also detected during the same month with a number of 3.3 oocysts per 100 l for Cryptosporidium and 1.7 Giardia cysts per 100 l. Campylobacteria, Salmonella and VTEC did not show any particular presence during this period.

| Table 2: Summary statistics of raw water samples between 2009-2015 for Råssnäs waterworks |
|-----------------------------------------------|---|--------|--------|--------|--------|
| Samples                                      | Samples | Missing | Mean   | St.Dev | Minimum | Maximum |
| Coliforms MPN/100 ml                         | 125      | 2       | 40.27  | 110.17 | 1       | 770     |
| E. Coli MPN/100 ml                           | 125      | 2       | 1.26   | 0.96   | 1       | 8       |
| Enterococci MPN/100 ml                       | 85       | 41      | 1.05   | 0.30   | 1       | 3       |
| Phages (per 100 ml)                          | 85       | 42      | 1.44   | 2.56   | 1       | 24      |
| Clostridium CFU/100 ml                       | 85       | 41      | 2.15   | 3.73   | 1       | 27      |
| Cryptosporidium oocysts/100 l                | 15       | 111     | 1.80   | 0.40   | 1.7     | 3.3     |
| Giardia cysts/100 l                          | 15       | 111     | 1.70   | 0.00   | 1.7     | 1.7     |
| Campylobacter MPN/l                          | 20       | 107     | 0.62   | 1.24   | 0.06    | 4.8     |
| Salmonella MPN/l                             | 20       | 107     | 0.06   | 0.00   | 0.06    | 0.06    |
| Verotoxigenic E.Coli MPN/l                   | 20       | 107     | 0.06   | 0.00   | 0.06    | 0.08    |
| Chemical Oxygen Demand                       | 53       | 74      | 1.58   | 0.35   | 0.9     | 2.4     |
| Turbidity                                    | 113      | 9       | 0.80   | 0.67   | 0.18    | 3.9     |
| Temperature                                  | 112      | 15      | 9.12   | 5.63   | 0.28    | 20.9    |
Figure 23: Time series data from raw water samples retrieved from the influx of Råssnäs waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli, C: Turbidity
Figure 24: Time series data from raw water samples retrieved from the influx of Råsnäs waterworks through 2014 to 2015 for A: Clostridium, B: Phages

Following the same analysis as performed for Östersund, and after filtering the areas that are not supplied by Råsnäs waterworks, the symptoms reported to 1177 can be summarised in Figure 25 and Figure 26 according to their characteristics and numbers. The dataset includes close to 14300 calls and a total number of 37164 people reside in the examined postal areas as of 2014.
The results are similar to Östersund. Abdominal pain is the primary reported symptom and a high percentage of symptoms in young children below the 9th year of age. Age group of 20-29 is still the second group in frequency of reports however significantly lower. The total symptom counts aggregated per month is shown in Figure 27 below:
The symptoms show an increasing trend in the average number of symptoms and more seasonal increases in comparison to Östersund. The months that had the highest reported symptoms were the first five months of both 2011 and 2013, followed by March, April and May of 2014. When it comes to symptoms before 2014 the recorded raw water data is not sufficient to draw a conclusion of possible correlations but as shown by Figure 23 and Figure 24 earlier, March and April of 2014 are suspect months due to a higher concentration of Phages, Cryptosporidium and Giardia in raw water and also an increase of turbidity which would however serve as operational monitoring for the waterworks. Figure 28 and Figure 29 below show the symptom spatial allocation in the network in those particular months:
Figure 28: Symptom count allocation in Råssnäs waterworks supply network during March 2014
Both figures show a cluster of symptoms particularly in the main city of Motala but April also shows an increase of symptoms in the smaller cities of Fornåsa and Klockrike. However Figure 27 also proves that a seasonal effect takes place during spring every year with symptoms being high in an equal manner historically so it is a difficult to assume with certainty a correlation to the waterborne pathogens, especially since the database before 2014 is limited.

Figure 30 and Figure 31 below show the areas that held the most symptoms and were deemed as more statistically significant. As it may be seen the most symptoms were recorded in the eastern parts of the city of Motala and also in the cities of Fornåsa and Borensberg.
Figure 30: Density map of total symptoms per square km for areas supplied by Råssnäs waterworks during 2009-2015
Figure 31: Areas identified as statistically significant based on total symptom counts in the Råsnäs distribution network.

Finally in Figure 32 below the incidence proportion and the impact on the population residing in the supplied postal areas is shown. The smaller cities to the north of Motala show very low incidence proportions while inside the main city and to the south of Motala the impact is higher.
5.3 Tallvägen waterworks and Härnösand

In the case of Härnösand municipality 1177 only started to record symptoms from June of 2011 and the data becomes more consistent in late 2011 and start of 2012. The raw water database itself follows the same limited pattern as described for the Råssnäs dataset. A summary table is shown below:
The total number of recorded data is not sufficient to draw certain conclusions about water quality itself. In general it appears that temperature has a lowest point of 1.5 degree and highest of 18.2 degrees. Turbidity shows an equivalent variation that ranges from 0.1 to 4.2 throughout 2014.

Time series concentrations for the data is shown in Figure 33 and Figure 34 below. Coliforms show a substantial increase from July to September and then reduce but are still in significant concentrations during October and November of 2014. E. Coli and Enterococci are only present in small amounts during October and June of the same year. Phages showed higher concentrations during March of 2014 and October of the same year, while Clostridium appeared mostly during June. Campylobacter, VTEC and Salmonella showed no presence at all throughout the study period. Finally Cryptosporidium and Giardia were detected in every monthly sample after March of 2014 with concentrations ranging from 2-4 oocysts per 100 litre.
Figure 33: Time series data from raw water recovered from the influx of Tallvägen waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli
As explained earlier the dataset retrieved from 1177 becomes more consistent in the start of 2012 and as such the period of analysis will be limited from the start of 2012 to March of 2015 for Härnösand. The dataset includes close to 3400 calls and 21586 people live in the examined postal areas according to SCB data of 2014. The clinical characteristics and distribution of symptoms by age group are shown in Figure 35 and Figure 36 below:
The figures show that the dominant symptom was abdominal pain at a higher percentage compared to the other municipalities, while the age group of 20-29 almost reached the age group of 0-9 which was reported highest in all municipalities. Figure 37 below shows the aggregated symptoms per month which occurs after filtering the database properly and removing areas of the county that are not supplied by the waterworks:
Figure 37: Symptom counts aggregated by month in areas supplied by the Tallvägen waterworks

The highest count of symptoms was detected in January of 2015 which amounted to 18.5% increase in comparison to the previous month. The map displaying the symptoms is shown in Figure 38 below. It must be noted that a presence of 3 Cryptosporidium oocysts per 100 litres of raw water was found in the samples.

Figure 38: Symptom allocation in areas supplied by the Tallvägen waterworks during January 2015
The months identified as suspect months in the time series analysis of pathogens found in raw water showed very low count of reported symptoms to 1177 so a correlation could not be established. It is interesting to note however that March and April of 2013 also showed high numbers of symptoms not noticed in the same months in other years. The database of raw water data does not hold sufficient samples however to determine a correlation. A map showing the symptoms during April of 2013 is shown in Figure 39 below. It appears that there was a higher amount of symptoms in areas outside Härnösand to the north-west.

Figure 39: Symptom allocation in areas supplied by the Tallvägen water treatment plant during April 2013

Figure 40 below identifies the areas in the supplied network that held the most symptoms based on their total count per square kilometre presented as a density map. The highest numbers were reported in the west parts of Härnösand which are the most populated as well but also in the cities to the north-west.
Figure 40: Density map of symptoms per square km in the Tallvägen waterworks network.

Figure 41 shown below identifies areas that were deemed statistically significant based on the total number of symptoms reported to 1177 between June of 2011 and March of 2015. It is interesting to notice that the main city of Härnösand was not identified as the only most significant symptom-wise area which is the only exception of the examined counties. Instead the cities to the north west of Härnösand were also identified as hot spots of statistical significance, however that could also be attributed to the fact that they all adhere to same postal code and appear as a large polygon on the map.
Figure 41: Areas identified as statistically significant based on total number of symptoms inside the Tallvägen waterworks supply network

To gain better insight into the incidence proportions in regards to population residing in the examined areas Figure 42 below shows that Härnosand and the surrounding areas had all very low rates of symptoms in comparison to the other municipalities.
Figure 42: Incidence proportions in areas supplied by Tallvägen waterworks through 2012 to 2015

5.4 Sjöbo waterworks and Borås

The data from both sources (SGU and FoHM) are very consistent regarding Sjöbo waterworks. Sample measurements have been recorded on weekly basis for Coliforms and E.Coli since 2009 while for Clostridium and Enterococci since the start of September of 2012. Turbidity and temperature are measured at least monthly before 2014, while Cryptosporidium and the other pathogens are always measured as part of the FoHM project from March of 2014.

The raw water samples recovered from the influx of Sjöbo waterworks in Borås showed very high amounts of pathogens at all times. Especially during September of 2013 the concentration of Coliforms reached a peak of 13000 (MPN/100 ml). It also was very high with 648 Coliforms per 100 ml during August of 2014. Table 4 below shows the summary statistics of the pathogens and chemical characteristics:
Temperature showed a consistent variation with a low point of 0.7 degrees during the colder months and 14 degrees in the warmer ones. PH showed a mean value of 7.2 and never dropped below 6.7 or over 7.7.

E. Coli showed significant presence from October to March during every year after 2011 while Enterococci reached their peak during March and April of 2014. Clostridium highest concentration points were detected during March 2014, January and February of 2015. Phages also showed the highest presence during the same months. Cryptosporidium had 5 oocysts per 100 l during March and August of 2014 followed by Giardia which had 4 cysts per 100 l. The rest of the pathogens showed minimal signs of existence.
Figure 44: Time series concentrations of indicator organisms detected in raw water retrieved from Sjöbo waterworks through 2009 to 2015 for A: Coliforms, B: E.Coli
Figure 45: Time series data of pathogens detected in raw water retrieved from Sjöbo waterworks influx through 2013 to 2015 for A: Enterococci, B: Clostridium

The database retrieved from 1177 after the filtering of areas not supplied by Sjöbo waterworks includes close to 30000 calls. A total population of 85286 people resides in the supplied postal codes as of 2014 according to SCB. After proper aggregation Figure 46 and Figure 47 below show the clinical characteristics examined and the allocation of symptoms in age groups:
Abdominal pain was the dominant symptom followed by fever or nausea. The symptoms were more prevalent in the age groups of 0-9 and 20-29, following the same pattern as the other counties.
The Sjöbo water distribution network is the most expansive studied throughout this thesis. Because of the higher number of residents the average number of symptoms is between 400 and 500 with an increasing trend from 2009 to 2015. As shown in Figure 48 above the highest number of symptoms were spotted in March and April of 2013, however a seasonal effect appears to take place since March and April show the highest number of symptoms every year. Figure 49 and Figure 50 show the spatial distribution of symptoms in Borås and surrounding areas supplied by Sjöbo waterworks during those months.
Figure 49: Symptom spatial allocation in areas supplied by Sjöbo waterworks during March 2013

The majority of symptoms are spotted inside Borås and mainly in the northern and western parts and also in the city of Fristad to the north of Borås.
A possible correlation supported by the water dataset in terms of gastrointestinal symptom increase was spotted in February of 2012 that presented the highest symptoms in comparison to February of any other year, an increase of 28% to their average number of symptoms. Figure 51 shows the spatial distribution of symptoms during that month. The symptoms mostly appeared in the centre and north of Borås and also in the city of Fristad.
Figure 51: Symptom spatial distribution in areas supplied by Sjöbo waterworks during February of 2012

Figure 52 and Figure 53 show the density of symptoms and areas identified as statistically significant based on the total number of reported symptoms to 1177. Both figures indicate that the majority of symptoms appear inside the more populated city of Borås, followed by Fristad.
Figure 52: Symptom density per square km in areas supplied by Sjöbo waterworks
Figure 53: Areas identified as statistically significant based on their total symptoms between 2009 and 2015 in the Sjöbo waterworks distribution network.

Figure 54 below identifies areas that had the highest impact of symptoms in regards to the population residing in them. The city of Borås itself is quite densely populated so the impact was not that high, while Viskafors to the south had a higher incidence proportion in comparison to other small cities that surround Borås.
5.5 Överby waterworks and Trollhättan

Raw water samples taken from the influx of Överby waterworks in Trollhättan were always highly contaminated and most notable pathogens had presence in low or high concentrations. The database was very consistent for E.Coli and Coliforms with weekly samples before March of 2014 and then weekly samples after that date for most notable pathogens. Table 5 shows summary statistics of most notable pathogens and indicators along with chemical characteristics:
Cryptosporidium oocysts showed up in every weekly sample after March of 2014 in concentrations of lower than 1.7 cysts up to 8.8 cysts (per 100 l) which was noted in August of 2014. Giardia cysts showed the same pattern with a peak of 5 cysts (per 100 l) during October and November of 2014 and then during January and February of 2015. Campylobacter showed a peak of 9.2 MPN/l during March and April of 2014 and then during October and November in the same year. Small traces of Salmonella with a peak of 0.46 MPN/l were spotted during August of 2014 and an important concentration of 48 VTEC/l was also detected during that same month.

Turbidity was mostly above the limit of 1.5 NTU with important variations and a peak of 18.4 NTU. Ph showed a mean value of 7.32 without significant changes. Figure 55 and Figure 56 below show the variations and peaks of the other notable pathogens throughout the entire study period:

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Samples</th>
<th>Missing</th>
<th>Mean</th>
<th>St.Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coliforms MPN/100 ml</td>
<td>373</td>
<td>1</td>
<td>135.35</td>
<td>344.93</td>
<td>0</td>
<td>2419</td>
</tr>
<tr>
<td>E. Coli MPN/100 ml</td>
<td>373</td>
<td>1</td>
<td>39.20</td>
<td>83.63</td>
<td>0</td>
<td>727</td>
</tr>
<tr>
<td>Enterococci MPN/100 ml</td>
<td>112</td>
<td>262</td>
<td>14.81</td>
<td>54.16</td>
<td>1</td>
<td>488</td>
</tr>
<tr>
<td>Phages (per 100 ml)</td>
<td>111</td>
<td>263</td>
<td>6.99</td>
<td>6.82</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>Clostridium CFU/100 ml</td>
<td>163</td>
<td>211</td>
<td>8.72</td>
<td>28.97</td>
<td>0</td>
<td>343</td>
</tr>
<tr>
<td>Cryptosporidium oocysts/100 l</td>
<td>24</td>
<td>350</td>
<td>2.89</td>
<td>1.73</td>
<td>1.7</td>
<td>8.8</td>
</tr>
<tr>
<td>Giardia cysts/100 l</td>
<td>24</td>
<td>350</td>
<td>2.66</td>
<td>1.20</td>
<td>1.7</td>
<td>5</td>
</tr>
<tr>
<td>Campylobacter MPN/l</td>
<td>30</td>
<td>344</td>
<td>2.57</td>
<td>3.02</td>
<td>0.06</td>
<td>9.2</td>
</tr>
<tr>
<td>Salmonella MPN/l</td>
<td>32</td>
<td>342</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.46</td>
</tr>
<tr>
<td>Verotoxigenic E.Coli MPN/l</td>
<td>32</td>
<td>342</td>
<td>1.65</td>
<td>8.46</td>
<td>0.06</td>
<td>48</td>
</tr>
<tr>
<td>Chemical Oxygen Demand</td>
<td>216</td>
<td>158</td>
<td>4.31</td>
<td>0.44</td>
<td>3.2</td>
<td>6.1</td>
</tr>
<tr>
<td>Ph</td>
<td>194</td>
<td>180</td>
<td>7.32</td>
<td>0.13</td>
<td>7</td>
<td>7.7</td>
</tr>
<tr>
<td>Turbidity</td>
<td>301</td>
<td>73</td>
<td>4.29</td>
<td>2.55</td>
<td>1.0</td>
<td>18.4</td>
</tr>
</tbody>
</table>
Coliforms had a constant presence but showed the highest concentrations during February, September and October of 2013 and then during April, September, October and December of 2014. E. Coli also showed the highest numbers during the exact same months. Enterococci peak was detected during October 2014, while Clostridium during March of 2014. Finally Phages had a significant presence during both 2014 and 2015 with the highest numbers spotted during April of 2014.
The database retrieved from 1177 after the filtering of areas not supplied by Överby waterworks includes close to 20000 calls. A total of 55811 people reside in those postal codes as of 2014 according to SCB. The characteristics of the reported symptoms and their distribution to age groups are shown in Figure 57 and Figure 58 below:
The clinical characteristics are similar to the other municipalities with abdominal pain as the primary symptom of each phone call, followed by fever or nausea. The most phone calls regarded children between 0-9 and adults of the age 20-29.
April of 2013 appears to hold the peak of symptoms, with a total number of 446 symptoms, corresponding to 0.8% of the population in those postal areas being sick. That number is at least 0.2% higher in comparison to April of any other year. A seasonal effect appears to take place with March every year to be the dominant month on symptom counts. It is interesting to notice that Coliforms had a very significant presence of 900 MPN/100 ml in late February 2013 so it is possible that a rise of symptoms would occur in March and April during which a presence of 12 Clostridium (cfu/100 ml) was detected in the raw water samples.

November and December of 2012 appear to have significantly high numbers as well in comparison with November and December of the other respective years. It must be noted that an average Coliform presence of 88 (MPN/100 ml) and 70 E.Coli (MPN/100 ml) was detected during these months along with a presence of 36 Clostridium/100 ml in late October.

The maps showing the spatial distribution of symptoms in Överby supplied areas for April of 2013 and November and December of 2012 appear in Figure 60 and Figure 61 below:
During April 2013 the most symptoms seem occur in the majority of the city of Trollhättan along with Sjuntorp and smaller cities to the south east of Trollhättan. Not all areas appear to be equally affected.
During November 2012 the most symptoms appear in the southern parts of the city of Trollhättan and less so in Sjuntorp and cities to the south east of Trollhättan. Finally during December of 2012 it appears that most of the city of Trollhättan shows high symptom counts. December of 2012 with the symptom allocation is shown in Figure 62 below:
Figure 62: Symptom count allocation during December 2012 in areas supplied by Överby waterworks

Figure 63 and Figure 64 identify the areas that had the biggest and most significant total count of symptoms and they appear to be the western, eastern and southern parts of Trollhättan.
Symptom density in the Överby water distribution network through 2009 to 2015

Figure 63: Total symptom density map for areas supplied by Överby waterworks
Figure 64: Areas identified as statistically significant based on total number of symptoms in the Överby waterworks network.

Figure 65 shows the impact of symptoms on the population residing in each postal area. The city of Trollhättan looks less affected with the exception of suburban areas to the south and north. Sjuntorp shows a high impact of symptoms relative to its population.

Additional maps showing seasonal effects on March 2011 and March 2013 are shown in Appendix I.
5.6 Comparison across the five counties

In order to compare the symptoms between municipalities of different population sizes and determine possible patterns of which municipality had the highest number of symptoms and on which months, it is necessary to normalise the number of symptoms to show the impact on the total number of people living the examined postal areas. Those statistics were recovered from SCB for each corresponding year. The graph below shows the impact on four out of five municipalities’ population based on the number of symptoms reported in the examined postal codes. Härnösand has been omitted from this graph for ease of viewing process as its total number of symptoms starts to count after June of 2011 and was always significantly lower than the other four counties.
Upon examining Figure 66 it must be noted that the more the lines deviate from each other, the easier it is to detect an abnormality as in a higher amount of symptoms in said municipality for specific months. For example in the case of Östersund it can easily be noticed that during November and December of 2010 when the Cryptosporidium outbreak occurred the total impact on the population residing in the supplied postal areas percentage wise has been way above 1%.

In a similar manner after examining the peaks it may be noticed that Östersund had a very elevated number of symptoms during the start of 2009 which supports what was shown earlier in Figure 14 and Figure 15. Borås showed a high number of symptoms during February 2012, which was also identified in Figure 51 as a month with a very increased number of symptoms.

Motala seems to deviate from the other municipalities in this comparison during January 2013 and then during March, April and August, September, October of 2014. March and April have already been shown in Figure 28 and Figure 29, so the rest of those suspect months will be examined:
Figure 67: Symptoms in Råssnäs distribution network and percentage of population affected in those areas during January 2013

Figure 67 definitely shows a cluster of symptoms in the main city of Motala and then high counts in smaller cities to the north and west of Motala. However the raw water dataset does not hold sufficient data in this particular period to show any interesting correlation. It is interesting though to notice that January of 2013 that is depicted here had 94 more reported symptoms in comparison to January of the previous year.

Finally Figure 68, Figure 69 and Figure 70 below show the allocation of symptoms during August, September and October of 2014. It must be noted that Coliforms showed significant presence in raw water samples as it may be viewed in Figure 23 in the Motala section above, so a possible correlation might be detected in these months. The symptoms were reported from the main city of Motala for the most part. There were 30 more symptoms on an average basis compared to the equivalent months of 2012 and 2013.
Figure 68: Symptoms in Råssnäs distribution network and percentage of population affected in those areas during August 2014
Figure 69: Symptoms in Rässnäs distribution network and percentage of population affected in those areas during September 2014.
Symptoms in areas supplied by the Råssnäs waterworks during October 2014

Figure 70: Symptoms in Råssnäs distribution network and percentage of residents affected in those areas during October 2014

Figure 71 below shows the incidence rates per 10000 residents across municipalities per year. The municipality with the lowest rate was Härnösand with an average yearly incidence rate of 440 symptoms per 10000 residents and has been omitted from the graph for ease of viewing. Trollhättan held the highest average incidence rate with 639 symptoms per 10000 residents, followed closely by Östersund with 636 symptoms per 10000 residents.

It is interesting to notice that after the outbreak of 2010 that occurred in Östersund the incidence rate became the lowest amongst the four compared municipalities. It must also be noticed that after 2010 the differences in yearly incidence rates between the compared municipalities are small, with Borås and Trollhättan having similar incidence rates and Motala having a smaller incidence rate with the exception of 2014 where it held the peak of symptoms.
5.7 General error sources and limitations

The shapefile of Sweden used for the visualisations displays postal areas as polygons on the map and many small cities or villages that do not belong to the waterworks distribution network are sometimes assigned to the same five digit postal code. That means they will also be assigned in the same polygon on the map, thus preventing their removal. As such the water distribution network that is displayed in all maps is not entirely accurate. However since most of these villages and cities usually have very small populations and the majority of symptoms was detected in the bigger cities, the margin of error seems to be small.

Another limitation is attributed to the dataset from 1177. It uses postal codes to geolocate each incident, instead of exact coordinates. As such, density maps that have been presented in this thesis are based on the centroid attributes of each polygon instead of exact coordinates, making their location on the maps inaccurate, even if the total density count is correct. In addition, no patient records have been examined and as such this thesis can only investigate for pathogen related symptoms instead of confirmed cases of infections. It must also be noted that the 1177 service is used differently between various groups of people in respect to different ages, rural areas and between counties.

The final limitation has to do with the raw water dataset. Asides the fact that the data refers to raw, unprocessed water and not drinking water itself, the samples are mostly consistent for indicators such as E.Coli and Coliforms. Other notable pathogens are only examined after March of 2014 during the project started by FoHM and Livsmedelsverket. Because of those limitations the process of correlating symptoms to pathogens cannot be accurate, since the thesis is not able to examine the treatment process that takes place.
6. Conclusions and future research

In general, the results seem credible. Trollhättan was identified as the municipality with the highest mean yearly incidence rate followed closely by Östersund which had a Cryptosporidium outbreak in 2010 but afterwards had the lowest incidence rate amongst the five examined municipalities. Härnösand had the lowest incidence rate per year, while Motala held the highest position only in 2014.

A lot of seasonal effects take place and some months like March, April and the winter months in general always showed more symptoms in each respective municipality. However, aggregating the symptoms by month and comparing them across the municipalities with the aid of graphs and visualisations performed in GIS, allowed the detection of suspect months with respective rises in gastrointestinal contacts to 1177. The results however, were not always supported by the raw water dataset and a high number of symptoms in an area did not necessarily imply that a high concentration of pathogens was discovered in raw water samples. High levels of contacts to 1177 can also be attributed to common seasonal epidemics such as winter-vomiting disease, not necessarily related to drinking water.

During some months and areas some patterns could be established based on the visualisations performed in GIS, as clusters of symptoms were found in the bigger cities and significant presence of pathogens in the raw water dataset could imply the possibility that some microorganisms successfully passed in drinking water. However, without access to specific patient records that show an exact number of confirmed cases of infections due to specific pathogens, the limitations to the attempted correlations are obvious, and the purpose of this thesis is not to evaluate the treatment process, but rather to identify risk situations and patterns.

GIS has proven to be an invaluable tool that allows the conversion of entire databases into time-animated outputs that are able to portray in a simple and understandable manner the spatial allocation of any chosen symptom over any period of time. Along with Healthcare Guide 1177, GIS analysis can be a good indicator able to provide increased awareness in the detection of a possible future waterborne or foodborne outbreak. With this combination, the usage of Healthcare Guide 1177 can be improved to allow the studying of seasonal and historical patterns across different months of the year and in different counties. One problem when comparing counties is that the usage of the Healthcare Guide 1177 service is different between counties and depends on the different age groups that use the service along with an overrepresentation of younger ages and people living in more central and populated areas.

Potential future research could include further insight into the microbiological analysis and control of pathogen indicators. The data received were mostly consistent for indicators such as E. Coli and Coliforms but lacked information on other notable pathogens and indicators from before March of 2014. While it is difficult to allocate funding for the analysis of all potential pathogens, it is important to maintain proper control and consistent analysis for at least the most common indicators found in raw water.
References


Peter C Rowe MD FRCPC Department of Paediatrics, Johns Hopkins University School of Medicine Baltimore, Maryland, USA (1995) Escherichia coli O157:H7, other verotoxin-producing E coli and the haemolytic uremic syndrome in childhood.


Appendix I – Other maps of interest

Symptom count during February of 2009 in the Minnesgårdet supply network

Symptom count during January of 2011 in the Minnesgårdet supply network
Symptom count during April of 2014 in the Minnesgården supply network

Total symptoms: 318
Impact on population: 0.58%

Symptom allocation during March 2010 in the Sjöbo water distribution network

Total symptoms: 633
Impact on population: 0.73%


