

Undersökning av sambandet mellan nederbörd och campylobactersmitta hos människor

Linda Kjellberg



UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
<http://www.teknat.uu.se/student>

Abstract

Undersökning av sambandet mellan nederbörd och
campylobactersmitta hos människor

The correlation between precipitation and human Campylobacter infection

Linda Kjellberg

Campylobacter is the most common cause of gastroenteritis in Sweden as well as in many other countries in the world. As many as 7 000 cases are reported yearly in Sweden only. The source of the Campylobacter infection is far from fully known, especially in sporadic human cases. Chickens have long been thought of as the leading source of infection, but they can't explain all cases that have occurred. Recent studies have shown that contaminated drinking water can be a more important source of infection than what was earlier believed.

The purpose of this study is to test the hypothesis that a great amount of precipitation increases the risk of drinking water getting contaminated by Campylobacter, which in turn leads to an increased number of occurred cases of infection. The focus of this study has been to investigate the correlation between precipitation and reported cases of human Campylobacter infection.

Handledare: Yvonne Andersson
Ämnesgranskare: Erik Söderbäck
Examinator: Elisabet Andréadóttir
ISSN: 1650-8319, UPTec STS07 018

Populärvetenskaplig beskrivning

Campylobacter är den vanligaste orsaken bakom mag- och tarminfektion i Sverige och i stora delar av västvärlden. Varje år rapporteras bara i Sverige omkring 7000 fall av campylobacterios och det verkliga antalet insjuknade uppskattas vara tiodubbelt fler. Kostnaden för samhället i form av sjukvård och sjukfrånvaro beräknas till mångmiljonbelopp varje år. Dessutom är det en stor kostnad i form av de drabbades lidande. Trots att så många människor drabbas varje år är orsaken bakom flertalet av alla sporadiska campylobacterfall fortfarande okänd. Kyckling har länge ansetts vara den största, men långt ifrån enda källan till smitta. Indikationer visar att dricksvatten kan vara en betydligt viktigare orsaksskälla än man tidigare trott.

Syftet med det här examensarbetet är att pröva, den på Smittskyddsinstitutet (SMI) framtagna hypotesen, att stora nederbördsmängder ökar risken för dricksvattnet att kontamineras med Campylobacter, vilket i sin tur leder till fler inhemska fall av campylobacterios. Hypotesen har prövats genom att utreda det statistiska sambandet mellan nederbördsmängder och antalet rapporterade campylobacterfall. Den statistiska undersökningen bygger på datamaterial över rapporterade campylobacterfall från SMI och meteorologisk data från SMHI. Undersökningsperioden sträcker sig över åren 1998 – 2003.

För utförande av undersökningen har programvarorna ArcGIS och STATA använts. ArcGIS är ett geografiskt informationssystem och har gjort det möjligt att bearbeta, visualisera och analysera både datamaterial över smittfall och nederbörd med hänsyn till dess geografiska information. STATA är det statistikprogram som har använts för bearbetning av data och genomförande av de statistiska beräkningarna.

För att maximera möjligheten att kunna dra bra slutsatser utifrån den statistiska analysen har flera olika typer av statistiska test genomförts i arbetet. Resultaten från undersökningarna visar att det finns ett visst samband mellan nederbördsmängder och antalet rapporterade smittfall i tiden, men att detta snarare beror på bakomliggande årstidsmässiga faktorer än att nederbördsmängden direkt påverkar antalet inträffade smittfall. Resultaten utesluter inte att det finns någon form av kausalt samband mellan variablerna, men i så fall är det troligen varken särskilt starkt eller linjärt.

Det samband som förväntats under den formulerade hypotesen gick i den här undersökningen inte att påvisa. Hypotesen kan dock inte förkastas enbart utifrån det resultatet. Det är mycket möjligt att hypotesen gäller för vissa specifika fall, men inte generellt och att sambandet därför inte syns i analysen. Trots att det är svårt att dra några generella slutsatser visar den här undersökningen trots allt, att kontaminerat dricksvatten, som resultat av mycket nederbörd, i alla fall inte är den största orsaken till inhemska sporadiska fall av campylobacterios

Förord

Detta examensarbete har genomförts som ett avslutande projekt under min civilingenjörsutbildning i System i Teknik och Samhälle på Uppsala Universitet. Arbetet genomfördes i samarbete med Smittskyddsinstitutets avdelning för epidemiologi. Jag vill passa på att tacka min handledare Yvonne Andersson, samt statistikerna Sharon Kühlmann-Berenzon, Jan-Olov Persson och Helena Pettersson från Smittskyddsinstitutet som varit till ovärderlig hjälp under arbetets gång. Dessutom vill jag tacka min ämnesgranskare Erik Söderbäck från institutionen för biologisk grundutbildning, Dag Jonsson från matematiska institutionen samt Per-Olof Hårdén från Institutionen för geovetenskaper, luft-, vatten- och landskapslära på Uppsala Universitet för den hjälp jag fått. Slutligen vill jag tacka min rumskompis Anna Blom som gjort att arbetet både blivit lättare och roligare att genomföra.

Stockholm 6 mars, 2007

Linda Kjellberg

Innehållsförteckning

1. Inledning	3
1.1 Bakgrund och problemformulering	3
1.2 Syfte och frågeställningar	4
1.3 Tillvägagångssätt	5
1.4 Avgränsningar	5
1.5 Disposition	6
2. Fakta om Campylobacter	7
2.1 Bakterien och dess ekologi	7
2.2 Sjukdomen	8
2.3 Det epidemiologiska läget i Sverige	8
3. Dricksvatten i Sverige	10
3.1 Vattnets kretslopp	10
3.2 Vattenförsörjningssystemet	10
3.2.1 Kommunal vattenförsörjning	10
3.2.2 Enskild vattenförsörjning	11
3.3 Kontaminering av dricksvattnet	11
3.4 Mått på dricksvattenkvalité	12
3.5 Vattenrelaterade utbrott av campylobacterios	12
4. Material och metod	14
4.1 Data över smittfall	14
4.1.1 Bearbetning av smittdata	14
4.1.1.1 Skattning av smittdatum	14
4.1.1.2 Skattning av geografisk plats	16
4.1.2 Begränsning av undersökningen till sommarhalvåret	17
4.1.3 Visualisering av smittdata	17
4.2 Meteorologisk data	19
4.2.1 Bearbetning av meteorologisk data	19
4.2.2 Visualisering av meteorologisk data	19
5. Statistisk undersökning	21
5.1 Testmetod 1: Analys i tiden	21
5.1.1 Tidsserianalys teori	21
5.1.2 Genomförandet på datamaterial	22
5.1.3 Resultat	22
5.1.3.1 Summerat veckovis	22
5.1.3.2 Summerat månadsvis	24
5.2 Testmetod 2: Analys i rummet	25
5.2.1 Bearbetning av datamaterial	25
5.2.2 Variansanalys	27
5.2.2.1 Teori	27
5.2.2.2 Indelning i kategorier baserat på nederbördsmängd	28
5.2.2.3 Genomförande och resultat variansanalys	30
5.2.3 Regressionsanalys	33
5.2.3.1 Poisson regression teori	33
5.2.3.2 Negativ Binominal Regression teori	33
5.2.3.3 Genomförande regressionsanalys	34

5.2.4 Summering	34
5.3 Testmetod 3: Analys i tid och rum	35
5.3.1 Val av region	35
5.3.2 Utförande av test	38
5.3.3 Resultat	38
6. Diskussion	40
6.1 Samband mellan campylobacterfall och nederbörd	40
6.2 Resultatet kopplat till hypotesen	40
6.3 Felkällor och hur dessa skulle kunna undvikas	41
6.4 Framtida forskning	41
6.5 Slutsatser	42
7. Litteraturförteckning	43
7.1 Tryckta källor	43
7.1.1 Publicerade artiklar	43
7.1.2 Publicerade böcker	44
7.1.3 Kompendium	44
7.3 Muntliga Källor	45
7.2 Internetkällor	45
8. Bilagor	46
8.1 Bilaga 1: Information om Campylobacterfall i SmiNet	46
8.2 Bilaga 2: Borttagningsstatistik	48

1. Inledning

1.1 Bakgrund och problemformulering

Alla har vi någon gång drabbats av magsjuka. De flesta av gångerna handlar det om ett relativt snabbt övergående sjukdomsförlopp. I vissa fall är det dock besvärliga bakterier som ger upphov till mag- och tarminfektionen. Symptomen blir då allvarligare, sjukdomsförloppet längre och i sällsynta fall kan infektionen till och med ge upphov till allvarliga komplikationer. Trots att Salmonella är den mest allmänt kända av de magsjukframkallande bakterierna är det Campylobacter som idag är den vanligaste orsaken till mag- och tarminfektion i Sverige och i stora delar av västvärlden.¹ Varje år rapporteras bara i Sverige omkring 7000 fall av campylobacterios, som sjukdomen orsakad av bakterien kallas. Det verkliga antalet insjuknade uppskattas dock vara tiodubbelt fler. Kostnaden för samhället i form av sjukvård och sjukfrånvaro beräknas till mångmiljonbelopp varje år. Dessutom är det en stor kostnad i form av de drabbades lidande.²

Campylobacterfall förekommer både i form av stora utbrott, där många människor blir smittade från en och samma källa, och som sporadiska fall där varje fall smittats av olika smittoförmedlare. Orsaken bakom de stora utbrotten är vanligen relativt lätt att spåra, medan det är betydligt svårare att urskilja källan till de sporadiska fallen. Majoriteten av alla fall sker sporadiskt och orsaken till flertalet av alla campylobacterfall är därför fortfarande okänd.³

Många studier har gjorts för att utreda orsaken till spridningen av Campylobacter. Genom så kallade fall- kontrollstudier har ett antal riskfaktorer identifierats som potentiella källor till sjukdomen. Tillagning och konsumtion av rå kyckling, konsumtion av opastöriserad mjölk och odesinfekterat vatten, nära kontakt med husdjur och längre vistelse på bondgård är några av de vanligaste identifierade riskfaktorena.⁴ Hittills har tillagning och konsumtion av kyckling som köpts hem rå ansetts vara orsaken till de flesta av campylobacterfallen i Sverige. Stora resurser har därför satsats både på att minska utbredningen av Campylobacter hos kyckling och att öka allmänhetens kunskap om hur tillagningen bör gå till. Trots dessa åtgärder ligger antalet campylobacterfall i Sverige på jämn nivå med tidigare år, vilket tyder på att det finns andra minst lika viktiga smittkällor som kyckling.⁵

Med bakgrund av detta har ett flertal studier genomförts för att hitta ytterligare orsaker till spridning av Campylobacter. Resultat från flera av dessa har visat indikationer på att kontaminerat dricksvatten är en betydligt vanligare orsakskälla än man tidigare trott.⁶ Ett flertal studier har visat att enskild vattenförsörjning signifikant ökar risken för sporadiska fall av campylobacterios, medan vattenförsörjning från kommunala anläggningar visat på motsatt effekt, vilket har förklarats med att enskilt vatten är mer känslig för kontaminering av bakterier. Det här tyder på att dricksvatten och i

¹ Hawker et al., 2002

² Livsmedelsverket, 2002

³ Nylen et al., 2002

⁴ Kapperud et al., 2003

⁵ Studahl et al., 2000

⁶ Nygård et al., 2003

synnerlighet dricksvatten från enskilda brunnar kan vara en viktig källa till sporadiska fall av campylobacterios.⁷ Hur vanligt förekommande problem det är med Campylobacter i dricksvattnet är dock mycket svårt att reda ut. Om bakterien förekommer är det vanligtvis i en väldigt låg koncentration, vilket gör det svårt att isolera och påvisa bakterien även om den finns i vattnet. Undersökningar som baserats på provtagningar i vatten har visat varierande resultat gällande påvisning av Campylobacter beroende på vilken provtagningsmetod som använts och antalet prov som tagits. Negativt resultat från kontroll av Campylobacter i dricksvatten, behöver nödvändigtvis inte vara lika med att vattnet är fritt från bakterien och inte kan ge upphov till smitta.⁸

Något som observerats i många studier är att antalet campylobacterfall uppvisar en tydlig årstidsvariation. Sambandet mellan antalet fall och tidpunkten på året har dock endast kunnat påvisas i tempererade länder som till exempel i Norden. I dessa länder kan man se en tydlig topp i antalet rapporterade fall under sommaren. Den exakta tiden för toppen varierar mellan olika länder, men är anmärkningsvärt lika från år till år. Många forskare tror att förståelse för de bakomliggande faktorerna till detta årstidsmässiga mönster är grunden till att lösa gåtan om orsaken till spridningen av Campylobacter. De bakomliggande faktorerna till mönstret är dock ännu inte fastställda. Troligt är att meteorologiska faktorer antingen direkt eller indirekt spelar en stor roll i organismens epidemiologi, vilket således skulle kunna förklara det observerade årstidsmässiga mönstret.⁹ Ett fåtal studier har genomförts för att utreda sambandet mellan klimatmässiga faktorer och campylobactersmitta. Dessa studier har dock främst fokuserat på påverkan av temperatur.¹⁰

Om man tittar på studier som har gjorts för andra vattenburna sjukdomar än Campylobacter kan man se att det inte är ovanligt att dessa visar på ett starkt samband mellan klimatmässiga faktorer och antalet sjukdomsfall. Framförallt har vattenburna utbrott av mikrobiologiska sjukdomar visat tydlig korrelation med kraftig nederbörd. Slutsatser från dessa undersökningar har varit att stora nederbördsmängder ökar risken för kontaminering av bakterier i dricksvattnet, vilket gör att risken för stora vattenburna utbrott ökar när nederbördsmängden har varit stor.¹¹

1.2 Syfte och frågeställningar

Detta examensarbete har genomförts på uppdrag av och i samarbete med Smittskyddsinstitutets (SMI:s) avdelning för epidemiologi. Avdelningen arbetar mycket med att spåra orsaken till de sjukdomsfall som uppstår och har tagit en aktiv del i att försöka lösa mysteriet kring orsaken till sporadiska fall av campylobacterios. Det var med målet att öka kunskaperna i frågan och med utgångspunkt från resultat som framkommit ur tidigare studier som uppslaget till det här examensarbetet växte fram.

Syftet med examensarbetet har varit att pröva den på SMI framtagna hypotesen, att stora nederbördsmängder ökar risken för att Campylobacter på olika sätt förorenar dricksvattnet, vilket i sin tur leder till fler inhemska sporadiska fall av

⁷ Schönberg-Norio et al., 2004

⁸ Livsmedelsverket, 2002

⁹ Tam et al., 2005

¹⁰ Patrick et al., 2004

¹¹ Lindberg & Lindqvist, 2005

campylobacterios. Hypotesen kommer att prövas genom att utreda om det råder något statistiskt samband mellan nederbördsmängder och antalet smittfall av sjukdomen och hur formen på detta samband ser ut.

De konkreta frågeställningarna för arbetet var:

- Finns något statistiskt samband mellan antalet rapporterade campylobacterfall och nederbördsmängder?
- På vilket sätt kan resultatet från den statistiska undersökningen i sin tur säga något om sanningshalten i den formulerade hypotesen?

1.3 Tillvägagångssätt

Undersökningen har baserats på datamaterial från Smittskyddsinstitutet (SMI) över smittfall orsakade av Campylobacter och datamaterial från Sveriges Meteorologiska och Hydrologiska Institut (SMHI) över nederbördsmängder. Datamaterialet från SMI har bearbetats så att varje fall har knutits till ett smittfall både i tid och rum. Geografiska Informations System (GIS) gör det möjligt att lagra, visualisera, bearbeta och analysera lägesbunden information. I det här arbetet har mjukvaran ArcGIS använts för det ändamålet.

För att testa om något samband råder mellan smittfall och nederbördsmängder har tre olika typer av test genomförts. I det första testet har en tidsserieanalys genomförts, där korrelationen mellan antal fall och nederbörd i tiden har undersökts. I det andra testet har en rumslig analys genomförts, där variationer i nederbördsmängd och incidens mellan olika regioner i Sverige studerats och det geografiska sambandet mellan variablerna undersökts. I det tredje testet har en enskild region valts ut och sambandet mellan variablerna har studerats för det valda området. Tre olika test har genomförts för att få större möjlighet att fånga upp ett eventuellt samband. Varje test har sina för- och nackdelar och förhoppningen har varit att summan av testen skulle ge en god grund att basera slutsatserna på.

Den största delen av det här examensarbetet har omfattats av den statistiska analysen och bearbetning av data. Utöver detta har en del litteraturstudier gjorts över tidigare forskning och resultat kring orsaken till campylobacterios. Dessutom har litteraturstudier gjorts kring dricksvattenssystemet i Sverige för att skapa en teoretisk bas för hypotesen som arbetet syftar att pröva.

1.4 Avgränsningar

Undersökningsperioden begränsar sig till tidsperioden mellan åren 1998 och 2003. Datamaterial över smittfall finns tillgängligt sedan sjukdomen blev anmälningspliktig i juli 1989. Informationen kring smittfall tidigare än 1998 är dock mycket ofullständig och saknar i allra flesta fall uppgifter om smittort och smittdatum, vilket behövs för den här undersökningen. Jag har därför valt att avgränsa perioden från och med 1998. Datamaterialet över nederbördsmängder beställdes från SMHI redan i slutet av 2003. Det datamaterialet sträcker sig därför bara fram till 31 oktober det året. Förklaringen till den för arbetet valda tidsperioden mellan 1998 och 2003 är alltså att det är under dessa år som tillfredställande data finns tillgänglig för både smittfall och nederbördsmängder.

1.5 Disposition

I det första kapitlet presenteras bakgrund till ämnet, problemformulering och syfte samt en kortare beskrivning om upplägg och genomförande för studien.

I kapitel två beskrivs fakta om bakterien *Campylobacter* och sjukdomen som den orsakar.

I följande kapitel beskrivs hur dricksvattenssystemet i Sverige ser ut. Detta kapitel syftar till att ge en teoretisk kunskapsgrund för den formulerade hypotesen.

I kapitel fyra presenteras information om de datamaterial undersökningen grundar sig på och hur bearbetningen av data har gått till.

I det femte kapitlet presenteras den statistiska undersökningen. Kapitlet är uppdelat i tre delar där de tre testmetoderna beskrivs i detalj. Resultatet från varje test beskrivs under respektive delavsnitt.

I det sista kapitlet sker en diskussion kring de resultat som kommit fram från den statistiska undersökningen och vilka slutsatser som kan dras av dessa i förhållande till den formulerade hypotesen.

2. Fakta om Campylobacter

2.1 Bakterien och dess ekologi

Campylobacter är spiralformade stavbakterier mellan 1,5 och 3,5µm långa. Bakterien är rörlig och förflyttar sig relativt fort med korkskruvliknande rörelser.¹² Inom släktet Campylobacter finns minst 13 olika arter. De två vanligaste som orsakar infektion hos människa är *C. jejuni* och *C. coli*. *C. jejuni* står för 90 % av alla campylobacterfall i Sverige och *C. coli* står för större delen av de resterande. Flertalet arter inom campylobactersläktet är termofila bakterier som trivs bäst vid en temperatur omkring +43°C. Dit räknas alla arter som ger upphov till infektion hos människa.¹³ Det är dessa jag syftar till när jag skriver om campylobacter i resten av arbetet.

Termofila Campylobacter behöver en temperatur på över +40°C och låg halt fritt syre för att kunna föröka sig. Syrehalten är till exempel alldeles för hög för att bakterien ska kunna tillväxa i luft. Villkoren är däremot uppfyllda i tarmen hos fåglar och varmblodiga djur, där bakterien har sitt naturliga habitat. Campylobacter är vanligt förekommande hos både människor och djur runt om i världen. Sjukdomen som den orsakar är således en zoonos, dvs. en sjukdom som kan spridas mellan människor och djur. Djur kan vara infekterade av bakterien utan att visa några symptom, vilket gör dem till utmärkta bärare och spridare av bakterien.¹⁴

Campylobacter kan spridas till människan via förorenade livsmedel, förorenat vatten eller genom smitta från husdjur. Smitta från människa till människa kan förekomma, men sker troligen inte i någon större utsträckning.¹⁵ Bakterien kan som tidigare nämnts normalt inte tillväxa i mat och vatten, eftersom villkoren på syrehalt och temperatur där inte är uppfyllda. Bakterien kan dock överleva i miljöer utanför värdorganismen i upp till flera veckor, beroende på temperatur. Miljöer kring +4°C ger en lång livslängd, medan bakterien inte klarar sig längre än ett par dagar i temperaturer på mer än +25°C. Normalt krävs en mycket liten infektionsdos på bara omkring 500 bakterier för människan att bli smittad, vilket gör spridning från livsmedel och vatten möjligt trots frånvaro av tillväxt.¹⁶

Trots att Campylobacterernas naturliga habitat är i tarmen på varmblodiga djur förekommer bakterien frekvent även i många andra miljöer. Bakterierna sprids effektivt till nya miljöer via avföringen hos fåglar och andra djur. Fekal kontamination kan ske både direkt eller indirekt, via nederbörd, avrinning och dränage. Tidigare studier har visat att vatten verkar vara det vanligaste mediet som länkar spridningen av Campylobacter från djur till människa.¹⁷

¹² Smibert, 1978

¹³ Hunter, 1997

¹⁴ Studahl & Andersson, 2000

¹⁵ Iwarson, 1985

¹⁶ Studahl & Andersson, 2000

¹⁷ Kapperud et al., 2003

2.2 Sjukdomen

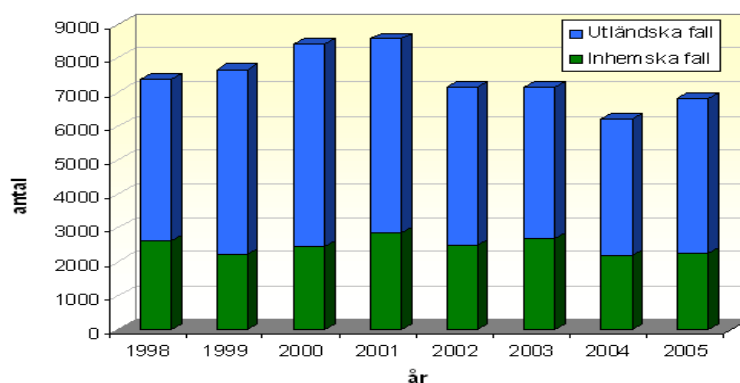
Det var först på 1970-talet som det blev känt att *Campylobacter* kan ge upphov till sjukdom hos människor. Sedan dess har sjukdomsfallen orsakade av bakterien kraftigt ökat i antal och räknas numera som den vanligaste bakteriella orsaken till mag- och tarminfektion i västvärlden.¹⁸

Campylobacteriosis är en akut sjukdom, med varierande svårighetsgrad, där de vanligaste symtomen är diarré, buksmärter och feber. I vissa fall förekommer även illamående och kräkningar. Inkubationstiden är vanligen mellan två till fyra dagar, men tiden från smitta till sjukdomsutbrott kan variera från en dag upp till en vecka. Sjukdomens varaktighet är vanligtvis två, tre dagar, men det är inte ovanligt att symptomen håller i sig upp till en vecka. Det förekommer även fall med betydligt längre sjukdomstid, som kan sträcka sig upp till flera månader eller till och med år. Vissa av de drabbade blir så dåliga att de måste behandlas på sjukhus. Så mycket som omkring 10 % av alla smittade får återfall.¹⁹

De medicinska konsekvenserna av sjukdomen är sällan livshotande, men komplikationer i form av reaktiv artrit (ledinflammation), meningit (hjärnhinneinflammation), Guillain-Barré (organisk nervsjukdom som leder till förlamning) och missfall kan förekomma.²⁰ Trots att komplikationer kan räknas som undantag och sjukdomen i annat fall är relativt ofarlig utgör *campylobacteriosis* ett stort allmänt hälsoproblem i Sverige. Så många som 7000 smittfall rapporteras i Sverige varje år. Antalet inträffade fall är dessutom sannolikt avsevärt högre än de registrerade fallen, eftersom man kan räkna med att endast omkring en tiondel av de insjuknade söker vård.²¹

2.3 Det epidemiologiska läget i Sverige

Varje år rapporteras omkring 7000 *campylobacter*fall i Sverige. Omkring 2000-3000 av dessa har smittats inom landets gränser medan resterande smittats i utlandet. För det här arbetet kommer jag endast att fokusera på de inhemskt smittade fallen.



Figur 2.1: Antal *campylobacter*fall åren 1998-2005²²

¹⁸ Hunter, 1997

¹⁹ Hunter, 1997

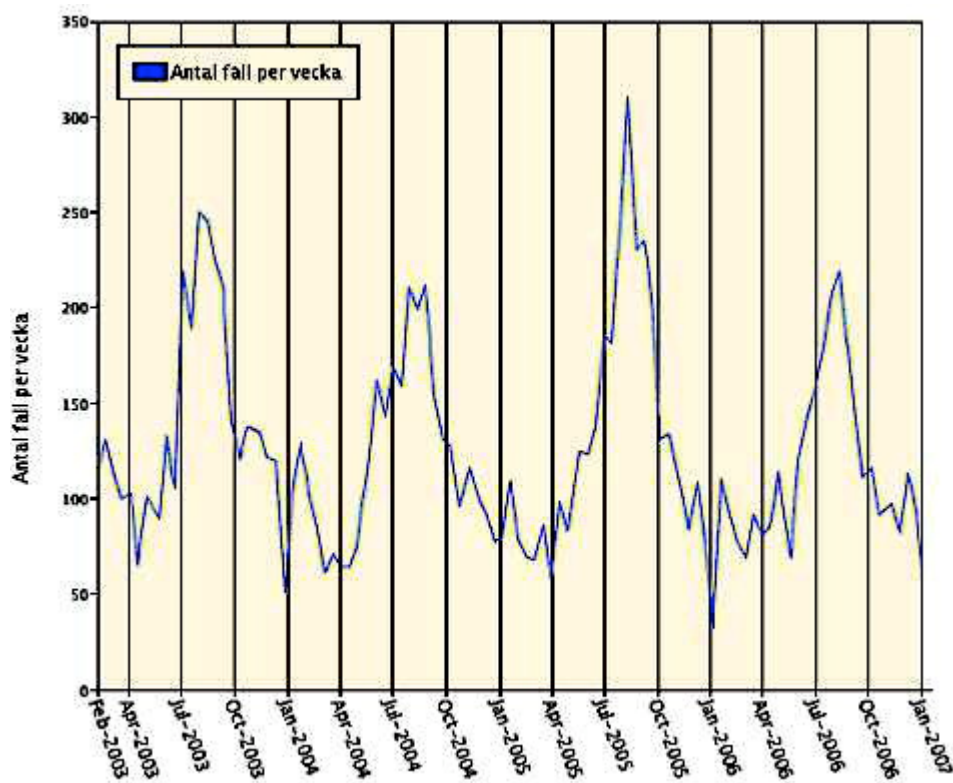
²⁰ Studahl & Andersson, 2000

²¹ Livsmedelsverket, 2002

²² www.smittskyddsinstitutet.se

Könsfördelningen över de smittade är relativt jämn, med en något högre frekvens bland männen. Statistik visar att campylobacterios drabbar alla åldrar, men att flest smittfall inträffar bland barn under fem år och vuxna i lägre medelåldern. Antalet fall varierar stort mellan olika geografiska områden i Sverige. Flest smittfall sker i Stockholms län, men Skåne och Västergötlands län har också ett stort antal smittfall jämfört med resten av landet. Det stora smittantalet i dessa geografiska områden kan förklaras med ett högt invånarantal. För att få en bättre bild av i vilka områden det är störst risk att drabbas kan man istället titta på incidensen, dvs. antalet smittfall per 100 000 invånare. Incidensen för inhemskt smittade fall är som störst på Gotland och som minst i Norrland, men varierar relativt mycket från år till år.²³

Antalet smittfall av campylobacterios varierar kraftigt över året, med störst antal smittade under sommarhalvåret. Det årstidsmässiga mönster som antalet campylobacterfall visar är relativt lika från år till år. Nedan visas en graf över antalet smittfall per vecka fördelat över fyra år.²⁴



Figur 2.2: Den tidsmässiga fördelningen av registrerade campylobacterfall i Sverige år 2003-2007.²⁵

²³ www.smittskyddsinstitutet.se

²⁴ www.smittskyddsinstitutet.se

²⁵ www.smittskyddsinstitutet.se

3. Dricksvatten i Sverige

I detta kapitel presenteras hur dricksvattensystemet i Sverige ser ut. Genom att ha den kunskapen skapas teoretisk förståelse för hypotesen som arbetet syftar att pröva. Trots att detta arbete utreder orsaken till sporadiska fall av campylobacterios presenteras i det här kapitlet även de största vattenrelaterade utbrotten för sjukdomen. Att bakterien kunnat komma in i dricksvattensystemet och smitta människor vid dessa utbrott visar att även sporadiska fall borde kunna uppstå som följd av kontaminerat dricksvatten.

3.1 Vattnets kretslopp

Den vattenmängd som finns på jorden är konstant. Trots detta är vattnet långt ifrån statiskt utan är ständigt i rörelse och under omvandling mellan olika former. När vatten når marken, genom nederbörd i form av snö eller regn, infiltreras en del genom marken, en del sugs upp av växter och en del avdunstar. När vatten infiltreras genom marken, fungerar jorden som ett filter som både renar och tillför mineraler och salter till vattnet. Beroende på vilken sorts mark vattnet passerar får det olika innehåll, kvalitet och egenskaper. Det vatten som silas genom marken tränger antingen djupt ner i jordytan och blir till grundvatten, eller blir genom dränering till ytvatten i åar och sjöar.²⁶

Från vattnets kretslopp i naturen lånar vi människor vårt dricksvatten. Dricksvatten skapas genom rening av råvatten, som tas från yt- eller grundvatten. Det är stor skillnad på råvattnets kvalitet beroende om det är hämtat från yt- eller grundvatten. Till skillnad från grundvatten som redan renats genom infiltrering genom marken kräver råvatten från ytvattentäkter betydligt mer behandling för att bli dricksvatten med bra kvalitet.²⁷

3.2 Vattenförsörjningssystemet

3.2.1 Kommunal vattenförsörjning

Runt 85 % av Sveriges befolkning, dvs. knappt 8 miljoner människor får sin dricksvattenförsörjning från allmänna anläggningar som oftast innefattas av kommunala vattenverk²⁸. Kvaliteten på det kommunala vattnet måste följa de författningar och lagar som finns stiftade inom området. Det är politiker och tillsynsmyndigheter som kontrollerar och ansvarar för att dessa författningar följs.²⁹

Det finns drygt 2000 kommunala vattenverk i Sverige. Det kommunala dricksvattnet kommer från många olika vattenkällor. Drygt hälften av råvatten tas från ytvatten från sjöar och vattendrag. En fjärdedel tas från naturligt grundvatten och resterande tas från konstgjort grundvatten. Generellt kan sägas att de större kommunala vattenanläggningarna är ytvattentäkter och de mindre baseras på grundvatten.³⁰

Vatten från ytvattentäkter renas och desinfekteras alltid innan det används till dricksvatten. Omfattningen av reningen är anpassad till råvattnets kvalitet och

²⁶ Hjärt-Lungfonden, 1993

²⁷ Hjärt-Lungfonden, 1993

²⁸ www.svensktvatten.se

²⁹ www.sgu.se

³⁰ www.vattenportalen.se

variationer angående denna över året.³¹ Vanligt är att ytvatten bearbetas i flera led innan det blir drickbart. Vanliga led i reningen är kemisk fällning, luftning, filtrering, pH-reglering och desinfektering. Desinfektion sker främst genom tillsättning av klor och har som syfte att begränsa tillväxten av mikroorganismer i dricksvattnet.³²

Omkring 80 % av alla kommunala vattenverk utgörs av små grundvattenverk³³. Dessa tar sitt råvatten från de totalt 1800 grundvattentäkterna som finns runtom i landet. Flertalet av grundvattenverken använder inte desinfektion utan enbart pH-justering av vattnet.³⁴ Anledningen till det är att grundvattnet oftast har tillfredställande kvalitet och klorering av vatten är en hälsorisk i sig och därför om möjligt bör undvikas. Eftersom grundvattenverk ofta saknar den mikrobiologiska barriär som klorering innebär är de känsliga för försämringar i råvattenkvaliteten, t.ex. vid kraftig nederbörd³⁵.

3.2.2 Enskild vattenförsörjning

Ett alternativ till kommunal vattenförsörjning är att ordna sin vattenförsörjning på eget sätt. Omkring 1,2 miljoner heltidsboende och lika många fritidsboende får sitt dricksvatten via enskild vattenförsörjning till exempel i form av privata brunnar. I dessa fall är det den enskilde fastighetsägaren eller tomtägarföreningen som ansvarar för vattnets kvalitet.³⁶ Det finns ingen organiserad provtagning eller registrering av den mikrobiologiska kvaliteten på enskilda vattentäkter. Dricksvattenkvaliteten i de enskilda brunnarna är på så sätt inte allmänt känd över landet. En undersökning från Sveriges Geologiska Undersökningar (SGU) har dock visat att kvaliteten på dricksvatten från enskilda vattentäkter varierar kraftigt. Borrade brunnar har generellt visat sig vara av bättre mikrobiologisk kvalitet än grävda. På samma sätt har brunnar i urbergsgrund visat på bättre kvalitet än brunnar i sedimentär berggrund. Vattenprov som tagits i en studie på Gotland var av sämst mikrobiologisk kvalitet under perioden juli till september.³⁷

Ungefär 400 000 enskilda brunnar finns för permanentboende runt om i landet och ungefär lika många finns för fritidsboende. Runt hälften av alla brunnar är borrade och resten är grävda. Andelen boende med kranvatten från enskild brunn är som störst på Gotland (34 %) och som minst i Stockholms län (3 %).³⁸

Vid enskild vattenförsörjning tas dricksvattnet oftast direkt från grundvattnet och dricksvattnet håller därför vanligtvis samma status som grundvattnet gör. Möjligheten finns dock för dem med enskild vattenförsörjning till olika former av rening av vatten. Som det ser ut idag är det dock snarare ett undantag än en regel.³⁹

3.3 Kontaminering av dricksvattnet

Största problemet med kontaminering av dricksvattnet, av sjukdomsframkallande mikroorganismer så som till exempel *Campylobacter*, är vid kommunala grundvattenverk och enskilda brunnar, där dricksvattnet distribueras helt utan

³¹ Socialstyrelsen, Institutet för Miljömedicin & Miljömedicin (Stockholms läns landsting), 2001

³² Hjärt-Lungfonden, 1993

³³ Lindberg & Lindqvist, 2005

³⁴ Socialstyrelsen, Institutet för Miljömedicin & Miljömedicin (Stockholms läns landsting), 2001

³⁵ Lindberg & Lindqvist, 2005

³⁶ www.sgu.se

³⁷ Andersson et al., 1998

³⁸ Socialstyrelsen, Institutet för Miljömedicin & Miljömedicin (Stockholms läns landsting), 2001

³⁹ www.sgu.se

desinfektion. I de fall bakterier på ett eller annat sätt har trängt in i vattentäkten, kommer dessa att kunna spridas direkt med dricksvattnet. De sjukdomsalstrande mikroorganismerna kan tillföras mark och vattentäkter till exempel via spillvatten från läckande avloppsledningar, dagvatteninfiltration, avfallsdeponier, stallgödsel och från ytavrinning från naturgödslad och betad åkermark där fåglar och djur betat och lämnat ifrån sig avföring.⁴⁰

Allra största risken för kontaminering av dricksvattnet gäller troligen enskilda brunnar. Dessa kan förorenas av *Campylobacter* genom inträngning av fekalt förorenat ytvatten. Orsaken kan vara dålig brunnskonstruktion, felaktig placering på brunnen, felkoppling av ledningar, att brunnen inte är tät, eller har trasigt lock. Framförallt vid kraftigt regn och översvämningar kan förorenat markvatten på så sätt tränga in i brunnen.⁴¹

Kontamination av dricksvattnet kan även uppstå i själva distributionsanläggningen, dvs. i transporten mellan vattenverk och användare. För denna typ av kontamination hjälper inte ens säkerhetsbarriärer som filtrering och desinfektion.

3.4 Mått på dricksvattenkvalité

Vid kontroll av dricksvattenkvalitet genomförs vattenanalyser där det undersöks om vattnet innehåller sjukdomsalstrande bakterier. Eftersom sjukdomsframkallande bakterier ofta förekommer i väldigt liten mängd kan de vara svåra att påvisa i dricksvattnet även om de i själva verket finns där. För att på ett enkelt sätt kunna värdera dricksvattnets kvalitet har standardiserade tester skapats som istället undersöker förekomsten av ett antal så kallade indikatorbakterier i vattnet.⁴² Dessa indikatorbakterier är relativt lätta att påvisa och innefattar ett antal bakterier som människor och andra varmblodiga djur kan ha i sin avföring och som kan orsaka sjukdom⁴³. *E. coli*, koliforma och heterofoba bakterier är exempel på indikatorbakterier. Vid påvisning av dessa vid analys, indikeras att dricksvattenkvaliteten inte är av tillfredställande kvalitet och att åtgärder bör genomföras.⁴⁴

Campylobacter räknas inte till indikatorbakterierna och är en bakterie som kan vara mycket svår att påvisa i dricksvatten. Tidigare utredningar har visat att förekomsten av *Campylobacter* inte nödvändigtvis korrelerar med förekomsten av indikatorbakterierna. Analyser som visat på god vattenkvalitet har i flera fall visat sig innehålla *Campylobacter* som givit upphov till smitta.⁴⁵

3.5 Vattenrelaterade utbrott av campylobacterios

Vattenburna sjukdomsutbrott orsakade av mikroorganismer har förekommit under alla tider och inträffar stundtals än idag. Vid flertalet av utbrotten i Sverige kan sjukdomens smittämne inte fastställas. Av de fastställda utbrotten har *Campylobacter* visat sig vara den enskilt vanligaste mikrobiologiska orsaken. Vid vattenburna utbrott av *Campylobacter* insjuknar erfarenhetsmässigt 17 – 30 % av dem som utsatts för

⁴⁰ Naturvårdsverket, 1991

⁴¹ Livsmedelsverket, 2002

⁴² Andersson et al., 1998

⁴³ Hjärt-Lungfonden, 1993

⁴⁴ Andersson et al., 1998

⁴⁵ Brennhovd et al., 1992

smittämnet. Det är en något lägre andel än vad som gäller generellt vid vattenburen smitta.⁴⁶

I Sverige har sedan 1980 fyra stora vattenburna utbrott av *Campylobacter* inträffat, där antalet insjuknade överstigit mer än 1000 individer. Det första inträffade i oktober 1980 i Grums då omkring 1700 av de 10 000 invånare i kommunen drabbades. Liknande utbrott skedde i Kramfors i maj 1994 då omkring 25 % av de 10 000 invånare som erhöll vatten från det kommunala vattenverket drabbades. Vid ett vattenburet utbrott i Marks kommun året efter insjuknade 3000 – 4000 personer i campylobacterios, varav två avled.⁴⁷ År 2003 skedde ytterligare ett utbrott där omkring 3000 personer insjuknade.⁴⁸ Gemensamt för dessa fyra stora utbrott är att de drabbat kommunalt vatten, som saknat eller haft en låg halt av klorering av dricksvattnet. I samtliga fall har dricksvattnet fastställts som den epidemiologiska orsaken till utbrotten. Det exakta händelseförloppet för smittspridningen har dock aldrig kunnat fastställas i detalj.⁴⁹

Av de vattenburna utbrott som dokumenterats genom åren är det de mindre kommunala grundvattenverken som stått för flest av fallen⁵⁰. Sjukdomsfall från enskilda vattenanläggningar har endast rapporterats i ett fåtal fall. Detta betyder dock inte att den enskilda vattenkvaliteten generellt sätt är bättre än den kommunala. Statistiken över sjukdomsfall orsakade av dricksvatten från enskilda brunnar förväntas innehålla ett stort mörkertal. Sjuklighet och problem med dålig vattenkvalitet speciellt vid mindre enskilda brunnar är troligen betydligt vanligare än något rapporteringssystem hittills lyckats fånga. Eftersom det är den enskilda fastighetsägaren som ansvarar för kvaliteten vid de enskilda brunnarna och som får betala för eventuella förbättringar, finns en ovilja till rapportering av bristande dricksvattenkvalitet. Detta kan leda till en ökad smittspridning av till exempel *Campylobacter*.⁵¹

⁴⁶ Andersson & Gustavsson, 1998

⁴⁷ Andersson & Gustavsson, 1998

⁴⁸ Martin et al., 2006

⁴⁹ Andersson & Gustavsson, 1998

⁵⁰ Lindberg & Lindqvist, 2005

⁵¹ Andersson et al., 1998

4. Material och metod

En kortare beskrivning av material och tillvägagångssätt introducerades redan i rapportens inledande kapitel. Här presenteras en utförligare version av de material undersökningen grundar sig på, hur bearbetningen av data har gått till, samt vilka metoder och datorprogram som använts för genomförandet av undersökningen. Resultat och slutsatser baseras på de material och de metoder som använts. Syftet med det här kapitlet är att ge läsarna en tydlig bild av detta, så att de får möjlighet att själva värdera trovärdigheten i de resultat som undersökningen senare kommer fram till.

4.1 Data över smittfall

Smittskyddsinstitutets avdelning för epidemiologi ansvarar för att följa det epidemiologiska läget i Sverige gällande infektionssjukdomar som omfattas av smittskyddslagen. I landet finns det idag 60 smittsamma sjukdomar som är anmälningspliktiga. En av dessa är campylobacterios som blev anmälningspliktig i Sverige den 1 juli 1989.⁵² En sjukdom kan klassas som anmälningspliktig om den ger upphov till svår sjukdom eller andra allvarliga konsekvenser under förutsättning att det är möjligt att förebygga smittspridning.⁵³ Att en sjukdom är anmälningspliktig innebär att läkare och laboratorium enligt smittskyddslagen är skyldiga att anmäla alla fastställda sjukdomsfall till Smittskyddsinstitutet. Anmälan sker genom att en blankett fylls i och skickas dels till smittskyddsläkaren i respektive landsting och dels till SMI.⁵⁴

Smittskyddsinstitutets material över smittfall från anmälningspliktiga sjukdomar lagras sedan 1997 i en databas som heter SmiNet. Där finns information om person- och sjukdomsrelaterade uppgifter knutna till smittfallet. För en komplett lista över informationsposterna i SmiNet, se bilaga 1. Eftersom campylobacterios är en anmälningspliktig sjukdom finns alla fastställda smittfall sedan 1997 registrerade i databasen. Det är tillgängligheten till denna datamängd som skapat förutsättningen för genomförandet av det här examensarbetet.

4.1.1 Bearbetning av smittdata

Eftersom syftet med examensarbetet är att studera variationer i antalet inträffade campylobacterfall både i tid och rum har det varit nödvändigt att relatera varje fall både till ett datum och till en geografisk plats. Det datum och den plats varje fall skulle knytas till har inte alltid gått att hämta direkt från den information som finns i SmiNet utan har ibland behövts uppskattas utifrån tillgänglig data.

4.1.1.1 Skattning av smittdatum

Det datum som varje smittfall i första hand skulle kopplas till var när personen ifråga blev smittad. Det finns uppgifter om smittdatum i SmiNet som läkaren tillsammans med patienten har fyllt i. Smittdatumet finns dock bara angivet i knappt hälften av fallen. Totalt finns 15 305 inhemska fall av campylobacterios registrerade under tidsperioden 1998 till och med 2003. Smittdatumet är dock endast angivet i 7 504 av fallen. Det är ofta svårt för en patient att veta exakt när smittotillfället inträffade, vilket syns tydligt

⁵² www.smittskyddsinstitutet.se

⁵³ Smittskyddslagen (2004:168) 3§ (www.socialstyrelsen.se)

⁵⁴ Lindbäck, 1999

eftersom smittdatumet är angivet i så pass få fall. Även i de fall smittdatumet är angivet kan tillförlitligheten hos datan ifrågasättas. Det angivna smittdatumet är oftast en uppskattning från patienten och läkarens sida. Tidigare studier har visat att det av patienten uppskattade smittillfället ofta skiljer sig mycket från det faktiska smittillfället i tid⁵⁵. Eftersom information om smittdatum ofta saknas och tillförlitligheten kan förväntas vara låg ansåg jag det olämpligt att använda detta datum i studien.

Det datum som ligger närmast smittdatumet i tid är insjukningsdatumet. Det är betydligt lättare för en patient att svara på frågan när man blev sjuk än när man blev smittad, vilket syns tydligt då insjukningsdatumet finns angivet i nästan 88 %, dvs. 13 412 av fallen. Insjukningsdatumet kan dessutom antas vara betydligt säkrare än det angivna smittdatumet. Att insjukningsdatumet både är relativt säkert och frekvent angivet gjorde det lämpligt att använda sig av det, för att uppskatta det datum som varje fall skulle kopplas till.

Smittdatumet kan uppskattas utifrån insjukningsdatumet genom kunskap om inkubationstiden för sjukdomen. Enligt litteraturen är inkubationstiden för *Campylobacter* vanligtvis mellan 2 - 4 dagar⁵⁶. Med hjälp av de ca 15 000 fall av campylobacterios som finns angivet i SmiNet från 1998 till 2003 kunde inkubationstiden skattas genom att beräkna medianen för skillnaden mellan smittdatum och insjukningsdatum. Den beräknade medianen på tillgänglig data blev 2 dagar. Eftersom det angivna smittdatumet antagits vara av så pass låg kvalitet ville jag dock helst undvika att basera inkubationstiden på den datan. Inkubationstiden uppskattades istället utifrån litteraturstudier som talar för en genomsnittlig inkubationstid på 3 dagar⁵⁷. Uppskattning av smittdatum gjordes genom att dra bort 3 dagar från insjukningsdatumet.

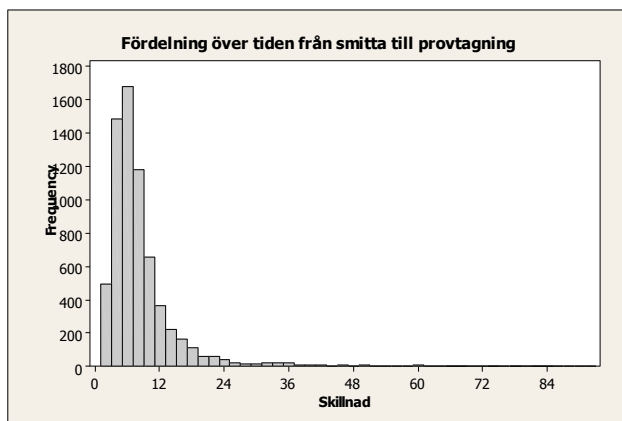
Genom att uppskatta smittdatum utifrån information om insjukningsdatum och genomsnittlig inkubationstid har 88 % av alla fall kunnat knytas till ett datum. För att kunna uppskatta smittdatumet för resterande fall har provtagningsdatumet använts. Eftersom jag inte hittat några tidigare uppskattningar om tiden det tar från smitta till provtagning har jag tvingats uppskatta denna tid utifrån det tillgängliga datamaterialet. Med hjälp av de cirka 15 000 fall som finns registrerade mellan 1998 till 2003 har förskjutningen från smitta till provtagning kunnat uppskattas genom att beräkna medianen för skillnaden mellan provtagningsdatum och smittdatum. Skillnaden kunde endast beräknas för fall där både provtagnings- och smittdatum fanns angiven, varpå medianen endast kunde beräknas utifrån knappt 7000 av fallen. Tydligt är att det på många ställen finns fel i SMI:s databas. Tiden mellan smittdatum och provtagningsdatum varierar enligt datamaterialet från minus 1 304, vilket är en fysisk omöjlighet, till 391 dagar. Ett provtagningsdatum senare än en månad efter smittdatumet, eller samma dag eller före smittdatumet, räknas som fel i databasen och uteslöts ur analysen där medianen beräknades. I figur 4.1 åskådliggörs fördelningen över tiden från smitta till provtagning.

Den uppskattade förskjutningen mellan smitta och provtagning beräknades till 6 dagar. Smittdatum uppskattas alltså genom att dra bort 6 dagar från provtagningsdatumet.

⁵⁵ Andersson, SMI (personlig kommunikation)

⁵⁶ Hunter, 1997

⁵⁷ Andersson, SMI (personlig kommunikation)



Figur 4.1: Fördelningen över tiden från smitta till provtagning. Y-axeln exv. "antal smittfall", X-axeln exv. "tidsskillnad smitta/provtagning i dagar mätt".

Provtagningsdatum finns angivet för 1 728 av de fall som saknade insjukningsdatum, vilket motsvarar 11,3 % av alla fall. Genom att i första hand skatta smittdatum utifrån insjukningsdatum och i andra hand utifrån provtagningsdatum kunde ett smittdatum kopplas till 98,9 % av alla smittfall. Resterande fall uteslöts ur undersökningen på grund av bristande information.

4.1.1.2 Skattning av geografisk plats

Vid registrering av campylobacterios ska information om smittort, insjukningsort och boendeort fyllas i av läkaren i samråd med patienten. Den geografiska plats som jag varit intresserad av i denna studie är platsen där individen i fråga blev smittad, dvs. smittorten. På samma sätt som för smittdatum saknas ibland uppgifter om smittort i databasen. Det kan vara svårt och ibland helt omöjligt för patienten att veta var smittan inträffat, varpå informationen i dessa fall uteblir. I de fall som smittort saknas uteslöts fallet ur undersökningen. Ibland finns smittorten angiven som flera olika alternativ. I dessa fall uteslöts fallet om de olika alternativen har haft en stor geografisk spridning. I annat fall valdes ett av alternativen som smittort. I första hand valdes boendeorten om den fanns med som ett av alternativen. Angiven smittplats som inte gick att knyta till en specifik geografisk plats, som till exempel flygplan, eller "fjällen" uteslöts också ur undersökningen. I Tabell 1 finns information om antalet smittfall som totalt sett utesluts ur undersökningen. Se bilaga 2 för mer detaljerad information kring antalet fall som utesluts och av vilken orsak.

Årtal	Användbara fall	Borttaget	Alla fall
1998	1959	660 (25%)	2619
1999	1717	493 (22%)	2210
2000	1889	577 (23%)	2466
2001	2244	595 (21%)	2839
2002	1989	492 (20%)	2481
2003	2135	555 (21%)	2690
Totalt	11933	3372 (22%)	15305

Tabell 4.1: Smittfall som utesluts ur analysen pga. ofullständiga uppgifter gällande smittdatum och smittort.

För drygt 20 % av alla fall har smittdatum eller smittort inte kunnat skattas och dessa fall har därför uteslutits ur undersökningen.

4.1.2 Begränsning av undersökningen till sommarhalvåret

Undersökningsperioden i detta examensarbete sträcker sig mellan åren 1998 och 2003. I det hittills bearbetade materialet har inträffade smittfall under hela tidsperioden inneslutits. Den hypotes som arbetet syftar att pröva är att höga vattennivåer ger upphov till stort antal smittfall. I undersökningen likställs stora nederbördsmängder med höga vattennivåer, vilket bara är någorlunda sant så länge nederbörden faller i form av regn. Vinterhalvåret är i själva verket en torrperiod, så länge som det är minusgrader, oavsett mängden nederbörd som faller.

För att kunna testa den för arbetet uppställda hypotesen är det bara intressant att ta med smittfall och nederbördsdata från de månader när nederbördsmängden kan likställas med vattennivån, dvs. under sommarhalvåret. Eftersom Sverige är ett avlångt land med stora klimatskillnader mellan norr och syd, finns ingen självklar gräns för de månader som bör inneslutas respektive uteslutas ur undersökningen. Jag har valt att innefatta månaderna maj till oktober och utesluta de resterande. Norra delen av Sverige kan självfallet vara snötäckt även under vissa av dessa månader. De allra flesta smittfallen inträffar dock i mellersta och södra delen av Sverige, så eventuell nederbörd i form av snö borde ge marginell effekt.

De allra flesta smittfallen inträffar under sommarhalvåret. Av de 11 933 fall som inträffat mellan 1998 och 2003 och som kunnat knytas både till smittplats och till smittdatum är det 9 474 som inträffat under månaderna maj till oktober. De 2 459 fall som inträffat under vinterhalvåret har uteslutits ur undersökningen. I enskilda test har dock variationen över hela året studerats, och i dessa fall har alla fall tagits med.

4.1.3 Visualisering av smittdata

Geographical Information System (GIS) är ett verktyg för att bearbeta, visualisera och analysera spatial information, dvs. datamaterial som har en geografisk utbredning. Eftersom både smittfall och nederbörd har en geografisk position har GIS varit ett viktigt redskap för genomförandet av undersökningen i det här examensarbetet. Utöver datamaterialet över smittfall och nederbördsstationer har karor över Sveriges län, kommuner, postnummer och orter använts.

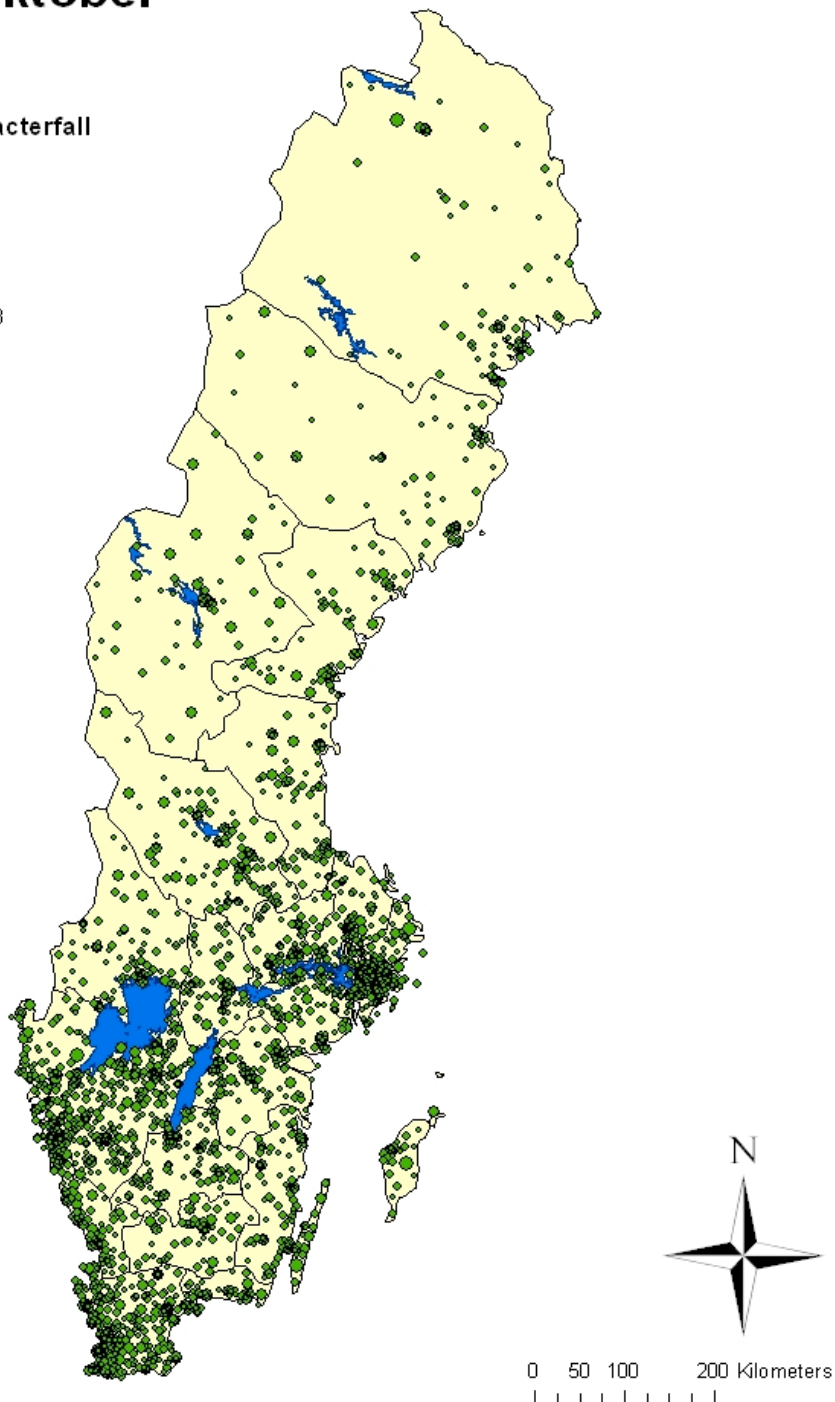
Vid bearbetningen av data knöts alla campylobacterfall till en smittort. Med hjälp av dataprogrammet ArcGIS markerades fallen därefter ut på en Sverigekarta. Se figur 2 nedan för en visuell bild av smittfallens geografiska utbredning. För att underlätta kartläggningen av smittfall lokaliserades postnumret för varje smittplats. Postnumren för alla smittfall matchades mot kartan över Sveriges samtliga postnummer, varpå smittfallen markerades ut på en karta. Anledningen till att postnummer valdes som matchningsnyckel istället för ortnamn är för att varje postnummer är unikt medan samma ortnamn finns på många olika ställen runt om i landet och därmed försvårar automatisk inläggning av data. I 58 % av alla fall var smittorten lika med individens boendeort. I dessa fall användes postnumret för individens bostad. I övriga fall valdes ett godtyckligt postnummer för den angivna smittorten.

Inhemska Campylobacterfall 1998 - 2003 maj - oktober

Campylobacterfall

ANTAL

- 1
- 2 - 4
- 5 - 11
- 12 - 238



Figur 4.2: Alla inhemska Campylobacterfall under månaderna maj till oktober 1998 - 2003 som innesluts i undersökningen

4.2 Meteorologisk data

Väderdata samlas kontinuerligt in av meteorologiska institut i flera länder runt om i världen. I Sverige är det Sveriges Meteorologiska och Hydrologiska Institut (SMHI) som ansvarar för insamling och lagring av väderdata. Runt om i landet finns flera hundra väderstationer, där dagliga mätningar av temperatur och nederbörd görs. Vid de större stationerna mäts även andra klimatmässiga faktorer såsom till exempel vind och lufttryck. Antalet väderstationer i Sverige varierar en del över tiden. Vissa stationer läggs ner medan andra stationer kommer till. Ibland inträffar dessutom driftstopp som resulterar i utebliven nederbördsdata för en kortare eller längre period. Antalet väderstationer som är i drift kan därför variera kraftigt från en tidpunkt till annan.⁵⁸

4.2.1 Bearbetning av meteorologisk data

För att denna undersökning skulle kunna genomföras beställdes nederbördsdata från SMHI. I det beställda datamaterialet finns information om dagliga nederbördsmängder från totalt 456 väderstationer runt om i landet inom tidsperioden januari 1986 till oktober 2003. Totalt innebär det en datamängd på nästan 2 miljoner observationer. Eftersom undersökningsperioden för det här arbetet endast sträcker sig mellan åren 1998 och 2003 kunde nederbördsdata från åren innan 1998 uteslutas. Under den valda tidsperioden finns nederbördsdata tillgänglig från 321 av de totalt 456 väderstationerna. Nederbördsdata från 66 av de 321 stationerna är ofullständig och täcker bara en del av tidsintervallet medan resterande täcker hela den valda tidsperioden. 299 stationer har nederbördsdata som täcker minst hälften av undersökningsperioden. Det är data från dessa 299 stationer som den statistiska undersökningen baseras på.

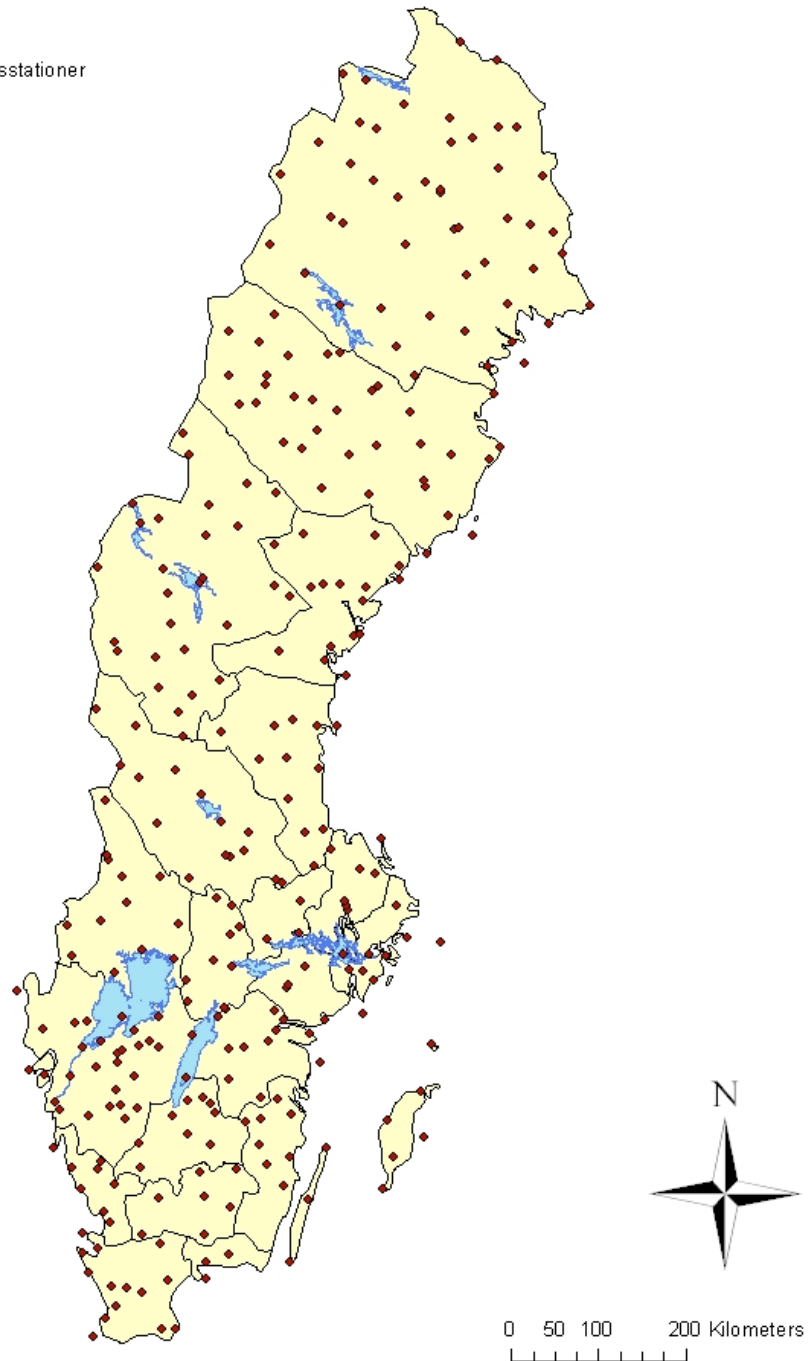
4.2.2 Visualisering av meteorologisk data

I det erhållna datamaterialet från SMHI finns information om longitud och latitud för varje nederbördsstation. Med hjälp av dessa koordinater markerades de 299, för undersökningen relevanta nederbördsstationerna, ut på en karta med hjälp av ArcGIS.

⁵⁸ www.smhi.se

Nederbördsstationer i drift 1998 - 2003

◆ nederbördsstationer



Figur 4.3: Nederbördsstationer i drift under åren 1998 - 2003

5. Statistisk undersökning

Orsaken till spridning av *Campylobacter* grundar sig i bakteriens ekologi, som är ett otroligt komplext system. Även om det här arbetet syftar att utreda sambandet mellan just nederbördsmängder och antal campylobacterfall, går det inte att bortse ifrån att verkligheten är betydligt mer komplicerad än så. Förhoppningen var att det eventuella sambandet mellan nederbörd och smittfall skulle kunna hittas även om det i verkligheten är ett stort antal faktorer utöver nederbördsmängd som påverkar spridningen av bakterien. Upplägget på den statistiska undersökningen avgör vilka bakomliggande faktorer som indirekt kommer med i analysen. Genom att formulera de statistiska undersökningarna på olika sätt kunde åtminstone en del eventuella bakomliggande faktorer minimeras och sambandet mellan just de faktorer jag avsåg att pröva framstå tydligare. Det finns dock ingen enkel metod som eliminerar alla andra faktorer än nederbörd. Genom att formulera olika typer av test, med olika för- och nackdelar, hoppades jag på att kunna få ett resultat som så bra som möjligt skulle säga något om sambandet mellan nederbörd och smittfall i det tillgängliga datamaterialet.

I det här kapitlet utreds det statistiska sambandet mellan antalet smittfall och nederbörd. Eftersom både smittfall och nederbörd varierar både i tid och rum ville jag inkludera både geografiskt läge och tid i beräkningarna. Att dela upp smittfallen både i tid och i rum medförde dock att varje enskild enhet som studerades innefattar väldigt få fall, vilket ger upphov till stor risk för slumpvariationer. Upplägget på undersökningen blev en avvägning mellan en strävan efter att få så noggrann information om smittfall och nederbörd som möjligt och att få tillräckligt stort datamaterial att studera för att undvika slumpvariationer. För att komma till rätta med det här problemet valde jag att utreda sambandet mellan smittfall och nederbörd med tre olika testmetoder. Det här kapitlet delas alltså upp i tre olika avsnitt, där tre olika test med olika syften och mål presenteras. I det första avsnittet undersöks bara variationer mellan variablerna i tiden, i det andra bara i rummet och i det tredje avsnittet läggs fokus på en utvald region där förändringar både i tid och i rum analyseras. Resultatet från varje test presenteras och diskuteras i respektive delkapitel. Slutsatser har dragits med utgångspunkt från summan av resultaten från de tre avsnitten.

I fortsättningen kommer jag att använda beteckningen $\{y\}$ för observerat antal inträffade smittfall, den beroende variabeln och $\{x\}$ för observerad nederbördsmängd, den oberoende variabeln.

5.1 Testmetod 1: Analys i tiden

5.1.1 Tidsserianalys teori

En tidsserie är en mängd observationer $\{x_t\}$, där varje observation är ett mått på variabeln x vid tiden t . Den observerade datamängden består av totalt n observationer. Tidsserien kan uttryckas som $\{x_1, x_2, \dots, x_n\}$. Vid tidsserieanalys kan förändringar i tiden hos en eller flera variabler studeras. Genom att grafiskt åskådliggöra tidsserien som en kurva får man en visuell uppfattning om eventuella trender eller avvikande beteenden hos variabeln över tiden.⁵⁹

⁵⁹ Brockwell & Davis, 2002

I en tidsserieanalys kan alltifrån en enskild till ett stort antal variabler studeras. För en bi- eller multivariat tidsserie kan det ömsesidiga beroendet mellan de inneboende variablerna undersökas. En bivariat tidsserie bygger på en tvådimensionell vektor (x_t, y_t) där värden för både variabel $\{x\}$ och $\{y\}$ observeras vid tiden t . Ett sätt att undersöka sambandet mellan variablerna i en bivariat tidsserie är att beräkna korskorrelationen. En korrelationskoefficient beräknas då för sambandet mellan $\{x\}$ och $\{y\}$.⁶⁰ Korrelationskoefficienten anger styrkan och riktningen på ett linjärt samband mellan variablerna. Korrelationen anges som ett värde inom intervallet $[-1, 1]$ där ett stort negativt värde tyder på ett starkt negativt samband, ett stort positivt värde tyder på ett starkt positivt samband och ett värde nära noll inte tyder på något samband alls.⁶¹

Vid korskorrelation beräknas utöver sambandet mellan $\{x_t\}$ och $\{y_t\}$ även sambandet för ett valt antal förskjutningar i tiden mellan de två variablerna, dvs. korrelationskoefficienter mäts för $(x_t, y_{t \pm k})$, där k är förskjutningen i tiden.⁶²

5.1.2 Genomförandet på datamaterial

Som ett första steg att utreda om det finns något samband mellan smittfall och nederbörd genomfördes en bivariat tidsserieanalys baserat på datamaterialet. Smittfall och nederbörds mängder summerades vecko- och månadsvis för hela Sverige. För varje vecka/månad gavs ett observerat värde för $\{x_t\}$ = nederbörds mängden och $\{y_t\}$ = antalet smittfall och observationerna bildade en bivariat tidsserie. Kurvorna för hur x och y varierar med tiden plottades upp i samma graf och korskorrelationen beräknades för upp till 15 tidsenheters förskjutning.

Syftet med testet var att undersöka det tidsmässiga sambandet mellan smittfall och nederbörd. Utifrån det här testet förväntade jag mig att få svar på frågan om en förändring i nederbörds mängd från en tidpunkt till en annan kan associeras med en förändring i antalet rapporterade fall. Genom att även beräkna korrelationen för förskjutningar i tiden kunde jag få en uppfattning av om den ena variabeln verkar föregå den andra i tiden.

Den för undersökningen valda tidsperioden är år 1998 till 2003. Enligt diskussionen i föregående kapitel skulle undersökningen fokusera på sommarhalvåret, eftersom det är under den tidsperioden störst samband mellan nederbörd och smittfall kan förväntas under den formulerade hypotesen. I det här testet genomfördes dock både en tidsserieanalys på hela året och en baserad på de valda månaderna.

5.1.3 Resultat

5.1.3.1 Summerat veckovis

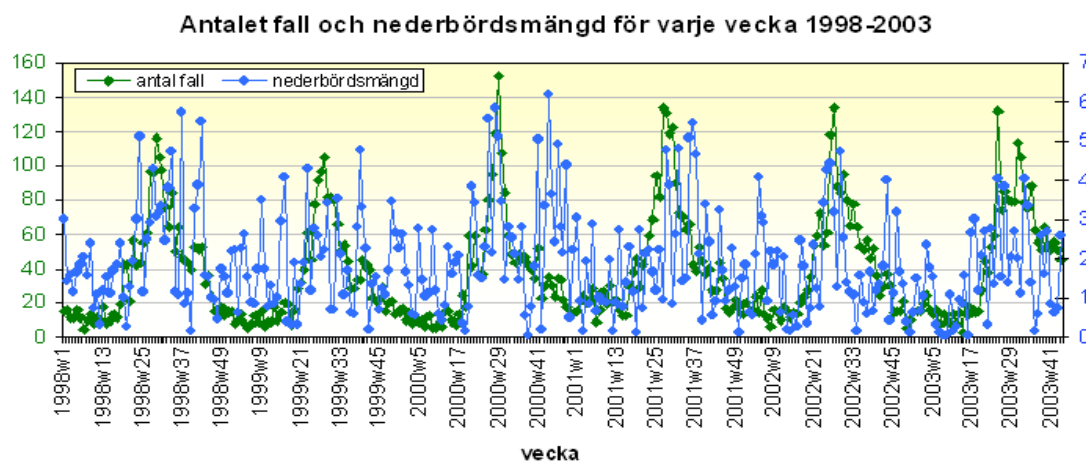
Nedan visas två grafer över hur antalet smittfall och nederbörds mängder förändrats över tiden. Den första figuren beskriver förändringen för hela tidsperioden 1998 – 2003, medan den andra figuren endast innefattar sommarhalvåret. För båda graferna är antalet fall en summa av antalet inträffade smittfall per vecka i hela landet. För nederbörds mängden ges ett medelvärde per dag för alla observationer av nederbörd för varje vecka. Nederbördsdata hämtas från alla aktiva stationer under den veckan. Antalet

⁶⁰ Brockwell & Davis, 2002

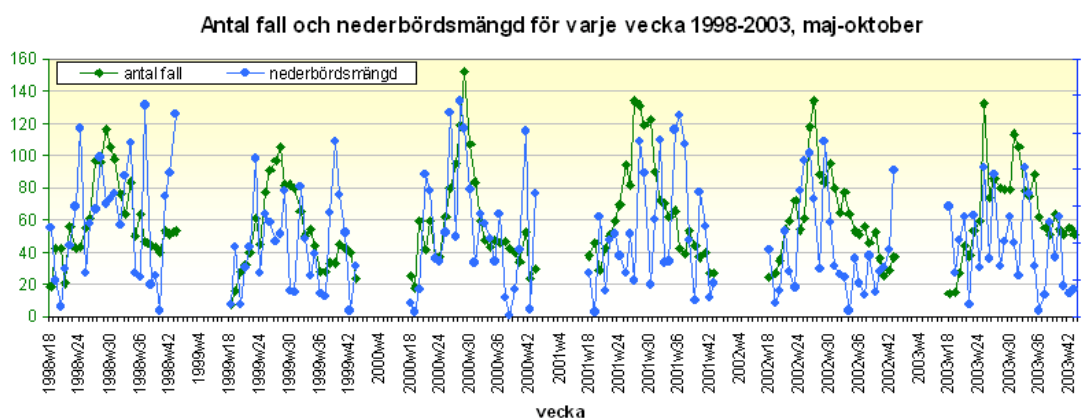
⁶¹ Blom, 1984

⁶² Brockwell & Davis, 2002

nederbördsstationer som är i drift har ingen betydelse, eftersom det är ett medelvärde och inte en summa av nederbörden som används.



Figur 5.1: Antalet campylobacterfall och medelvärde för nederbördsmängd i Sverige för varje vecka under tidsperioden 1998 till 2003. Den vänstra Y-axeln exv. antalet rapporterade campylobacterfall för given vecka, den högra Y-axeln exv. ett medeltal för nederbördsmängd samma vecka. X-axeln exv. veckorna i tidsserien.



Figur 5.2: Antalet campylobacterfall och medelvärde för nederbördsmängd i Sverige för varje vecka under sommarhalvåret 1998 till 2003. Den vänstra Y-axeln exv. antalet rapporterade campylobacterfall för given vecka, den högra Y-axeln exv. ett medeltal för nederbördsmängd samma vecka. X-axeln exv. veckorna i tidsserien.

Som man kan se i graferna följer de två kurvorna varandra under vissa veckor, medan de under andra tider verkar gå åt motsatt håll. Vid beräkning av korskorrelationen gavs den högsta korrelationskoefficienten i båda fallen vid ingen förskjutning i tiden mellan kurvorna.

Korrelationskoefficienten blev i båda fallen relativt svag, med ett värde på 0,39 sett över hela året och 0,34 när bara sommarhalvåret innefattades. Som man kan se i grafen varierar värdet på nederbördsmängden ganska kraftigt från vecka till vecka, något som inte kan observeras på samma sätt för antalet smittfall. Genom att istället summera

antalet smittfall och nederbördsmängd per månad jämnas nederbördskurvan ut lite och grafen blir lättare att tolka.

5.1.3.2 Summerat månadsvis

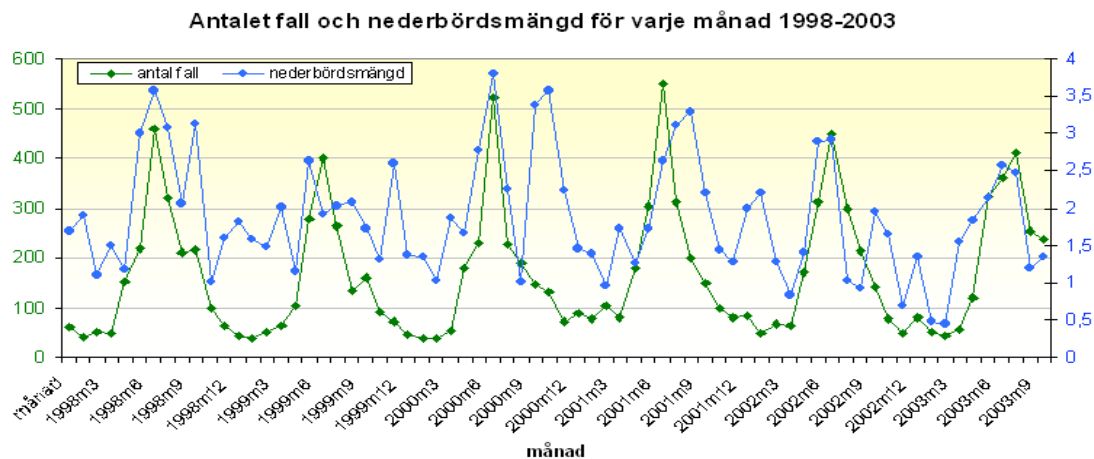


Fig.5.3: Antalet campylobacterfall och medelvärde för nederbördsmängd i Sverige för varje månad under åren 1998 till 2003. Den vänstra Y-axeln exv. antalet rapporterade campylobacterfall för given vecka, den högra Y-axeln exv. ett medeltal för nederbördsmängd samma vecka. X-axeln exv. veckorna i tidsserien.

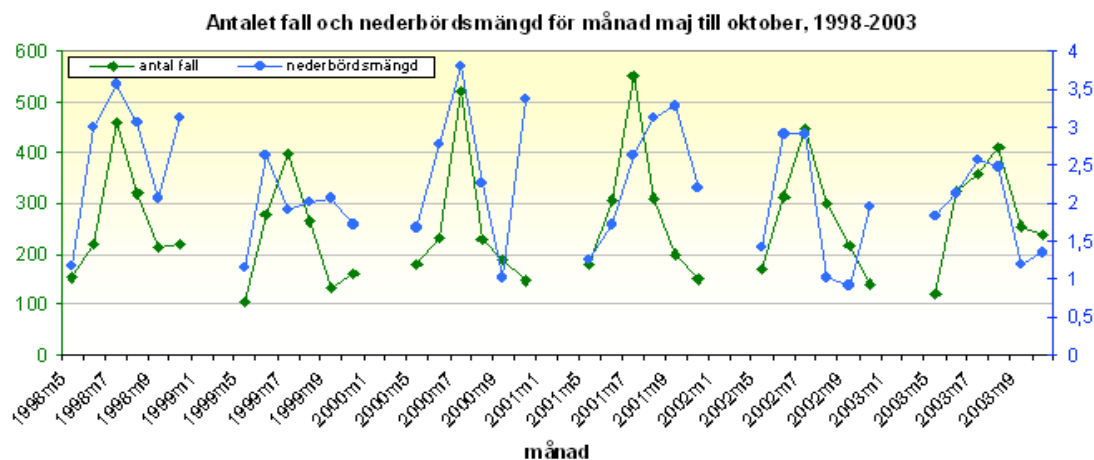


Fig.5.4: Antalet campylobacterfall och medelvärde för nederbördsmängd i Sverige för månad maj till oktober 1998 till 2003. Den vänstra Y-axeln exv. antalet rapporterade campylobacterfall för given vecka, den högra Y-axeln exv. ett medeltal för nederbördsmängd samma vecka. X-axeln exv. veckorna i tidsserien.

När smittfall och nederbördsmängd summerades månadsvis blev det lättare att se hur kurvorna följer varandra i tiden och med blotta ögat kan man se att det tycks finnas ett samband mellan de två variablerna i tiden.

Vid korskorrelation gavs ett högsta värde vid noll månaders tidsförskjutning för båda tidsperioderna. Korrelationskoefficienten blev 0,58 för hela året och 0,48 då bara sommarhalvåret inkluderas. Sambandet blev alltså starkare då variablerna summeras månadsvis än veckovis. Även om det här testet som väntat visade på ett visst positivt samband mellan antalet smittfall och nederbördsmängder säger det egentligen ingenting

om kausaliteten, dvs. om nederbördsmängden verkligen påverkar antalet smittfall. Det enda som det här testet säger är att vid samma tidpunkt som det är höga nederbördsmängder i Sverige sker det stort antal smittfall. Om det beror på att nederbördsmängd påverkar smittfallantalet eller om det är andra bakomliggande faktorer som påverkar är dock osäkert.

Resultatet visar både vid summering vecko- och månadsvis på en högre korrelationskoefficient från testet över hela året än när bara sommarhalvåret var inkluderat. Detta skulle kunna tolkas som ett tecken på att det samband som finns mellan kurvorna snarare tyder på ett årstidsmässigt samband än att smittfall uppstår som följd av nederbördsmängd enligt den formulerade hypotesen. Annars borde sambandet vara starkare när bara sommarhalvåret inkluderades och inte tvärt om.

Det är för tidigt att dra några generella slutsatser om sambandet mellan smittfall och nederbörd. Att ett visst tidsmässigt samband kunde utskildas från det här resultatet motiverar dock ytterligare test att genomföras för att om möjligt finna svaret på frågan.

5.2 Testmetod 2: Analys i rummet

Det första testet visade att ett svagt samband föreligger mellan antal smittfall och nederbördsmängd i tiden. Om det var nederbördsmängden som påverkade antalet smittfall var dock mycket osäkert och en del resultat pekade till och med på att så inte var fallet. I det här testet försökte jag minimera de bakomliggande faktorer som kan ha med årstidsmässighet att göra. Genom att inte titta på vad som har hänt från en tidpunkt till en annan och istället undersöka vad som har hänt under den totala undersökningsperioden och se skillnader mellan olika regioner i Sverige, hoppades jag att få ytterligare information om sambandet mellan smittfall och nederbörd. För det här avsnittet fokuserade jag på undersökningsperioden maj till oktober, 1998-2003.

Eftersom testet jämför olika regioner i Sverige med varandra gick det inte att räkna med antalet smittfall som i tidigare test, eftersom antalet inträffade smittfall i en region till stor del beror på populationsstorleken. I det här testet fick därför incidensen ersätta måttet på antalet inträffade smittfall. Incidensen räknas som antalet fall per 100 000 invånare.

5.2.1 Bearbetning av datamaterial

Av praktiska skäl valdes regionsindelningen för analysen till samma som Sveriges kommunindelning. I Sverige fanns under undersökningsperioden 289 kommuner som var och en här betraktas som en observationsenhet. För varje kommun beräknades ett mått på nederbördsmängd och incidens för regionen. Nederbördsmängden för varje kommun beräknades som ett medelvärde av observationerna från varje väderstation innanför kommunens gränser. I de kommuner som saknade väderstation har ett viktat medelvärde beräknats baserat på observationer från de fem närmsta nederbördsstationerna från kommunens centrum. Ett viktat medelvärde användes för att information från nederbördsstationer i närheten av kommunen skulle bidra mer relativt nederbördsstationer längre bort till medelvärdet. Varje datavärde (X_i) gavs en vikt (W_i) och medelvärdet kunde beräknas enligt följande formel.

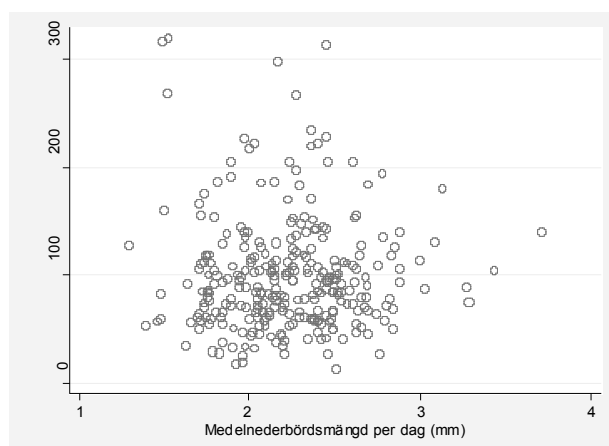
$$\bar{X}_w = \frac{\sum W_i X_i}{\sum W_i}$$

Vikten för varje datavärde bygger på det inverterade avståndet mellan kommunens mittpunkt och nederbördsstationen.

För att få fram information om vilka stationer som ligger i varje kommun och information om de fem närmsta nederbördsstationerna och avstånden till dessa användes Geographical Information System (GIS). Genom geografisk bearbetning i ArcGIS och matematiska beräkningar i STATA kunde en nederbördsmängd beräknas för varje kommun. Nederbördsmängden anges som medelnederbörden per dag under undersökningsperioden, som i det här testet innefattar månaderna maj till oktober, 1998-2003.

Med hjälp av ArcGIS summerades det totala antalet smittfall som inträffat i varje kommun. Från statistiska centralbyrån (SCB) hämtades information om befolkningsmängden för varje kommun. Statistik valdes från slutet av år 2000, eftersom den tidpunkten låg i mitten av undersökningsperioden och därmed bäst kunde förväntas representera befolkningsmängden i kommunen för undersökningen. Genom att dividera det totala antalet fall som inträffat i varje kommun med befolkningsmängden för kommunen och sedan multiplicera 100 000 beräknades ett värde för incidenten för varje kommun.

Efter att incidens och nederbördsmängd för den totala undersökningsperioden beräknats för varje kommun plottades en graf för de 289 observationerna. Se figur 5 för scatterplot över sambandet mellan nederbördsmängd och incidens i Sveriges kommuner.



Figur 5.5: Incidens och nederbördsmängd för Sveriges 289 kommuner. Y-axeln exv. ett medelvärde för incidensen maj till oktober 1998-2003 i Sveriges kommuner. X-axeln exv. ett medelvärde för nederbördsmängd under tidsperioden.

Spridningen för punkterna är väldigt stor och det går inte med blotta ögat se om det finns något samband mellan incidens och nederbörd för Sveriges kommuner. Den slutsats som dock kan dras enbart genom att titta på grafen är att det inte finns något tydligt linjärt samband. Beräkning av Pearsons korrelation mellan nederbörd och incidens stödjer tolkningen om bristen på linjärt samband, då korrelationskoefficienten endast gav ett värde på $-0,0074$.

Hittills hade jag utgått från att sambandet mellan incidens och nederbörd ska vara linjärt, eller åtminstone att en högre nederbördsmängd ger upphov till fler smittfall,

eftersom det är det som förväntas utifrån den formulerade hypotesen. Trots att inget linjärt samband verkar finnas är det fortfarande intressant att reda ut hur relationen mellan variablerna ser ut och om det finns något eventuellt icke-linjärt samband. Styrka mellan två variabler i ett icke-linjärt samband kan testas med Spearmans korrelationskoefficient, som för det här materialet gav ett värde på 0,0472. Det låga värdet tyder på ett svagt samband mellan variablerna och att man inte kan förkasta att de till och med är oberoende.

För att öka förutsättningen för att kunna påvisa ett eventuellt icke-linjärt samband i fortsatt analys delas kommunerna in i kategorier baserade på nederbördsmängd varpå variansanalys genomförs på materialet. Variansanalysen förväntas tala om incidensen överhuvudtaget verkar bero av nederbördsmängden och ge en fingervisning på vilken form ett eventuellt samband har.

5.2.2 Variansanalys

5.2.2.1 Teori

Variansanalys är en samling statistiska modeller som går ut på att ta reda på om stickprovsmedelvärden från olika klasser skiljer sig signifikant från varandra. Grundidén går ut på att jämföra variationen mellan observationer inom grupper med variationen mellan grupper och utifrån det beräkna om det finns någon signifikant skillnad mellan gruppernas medelvärden. Med hjälp av variansanalys kan man testa nollhypotesen att medelvärdet för alla grupper är lika. Om nollhypotesen kan förkastas har man på vald signifikansnivå visat att medelvärdet skiljer sig åt mellan de valda klasserna.⁶³

Variansanalys kan användas på datamaterial där den oberoende variabeln är på nominal- eller ordinalskalanivå. Den beroende variabeln kan då delas in i flera olika kategorier som antingen har eller inte har någon inbördes ordning. Den beroende variabeln måste vara på intervall- eller kvotskalanivå. Det som mäts ska alltså kunna tilldelas ett numeriskt värde och medelvärde och median ska vara möjligt att beräkna.⁶⁴

Envägs variansanalys bygger på den linjära modellen:

$$X_{ij} = \mu + \alpha_j + e_{ij}$$

där X_{ij} = ett godtyckligt observerat värde inom kategori j , μ = totalt medelvärde, α_j = effekten av kategori j och e_{ij} = ”felet” i modellen.⁶⁵

Det finns flera antaganden som ska vara uppfyllda för att resultatet från variansanalysen ska kunna säga något om materialet det utgått från.

- Observationerna ska vara oberoende
- Observationerna från varje grupp ska vara normalfördelade
- Homogena varianser mellan grupperna⁶⁶

⁶³ Montgomery, 2001

⁶⁴ www.wikipedia.org (nominalskala)

⁶⁵ Johnson & Wichern, 2002

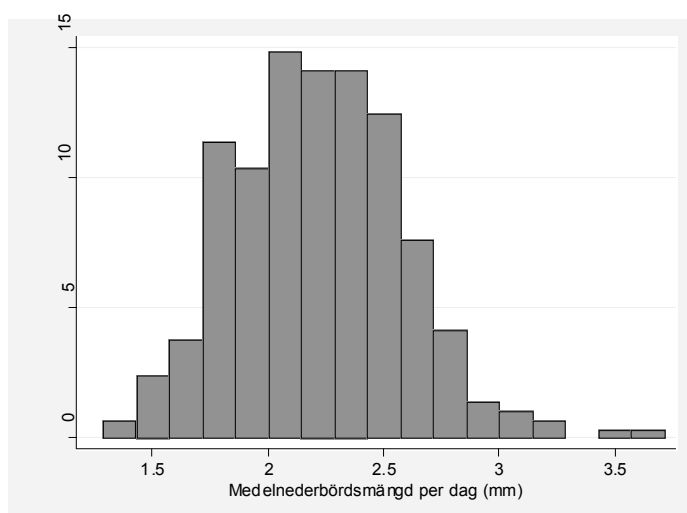
⁶⁶ Johnson & Wichern, 2002

Det finns flera olika typer av test som går inom ramen för variansanalys. I det här arbetet använde jag mig av ”envägs variansanalys” som är ett test för att pröva skillnaden mellan tre eller flera oberoende gruppers medelvärden.

5.2.2.2 Indelning i kategorier baserat på nederbördsmängd

Syftet med variansanalysen är att dela upp Sveriges 289 kommuner i olika kategorier som baseras på hur mycket nederbörd det har fallit i kommunen under undersökningsperioden och sedan undersöka om medelincidensen skiljer sig åt för de olika kategorierna.

För att dela upp kommunerna i lämpliga grupper är det bra att veta hur fördelningen hos kommunernas nederbördsmängd ser ut. Se figur 5 nedan för histogram över kommunernas nederbördsmängd.



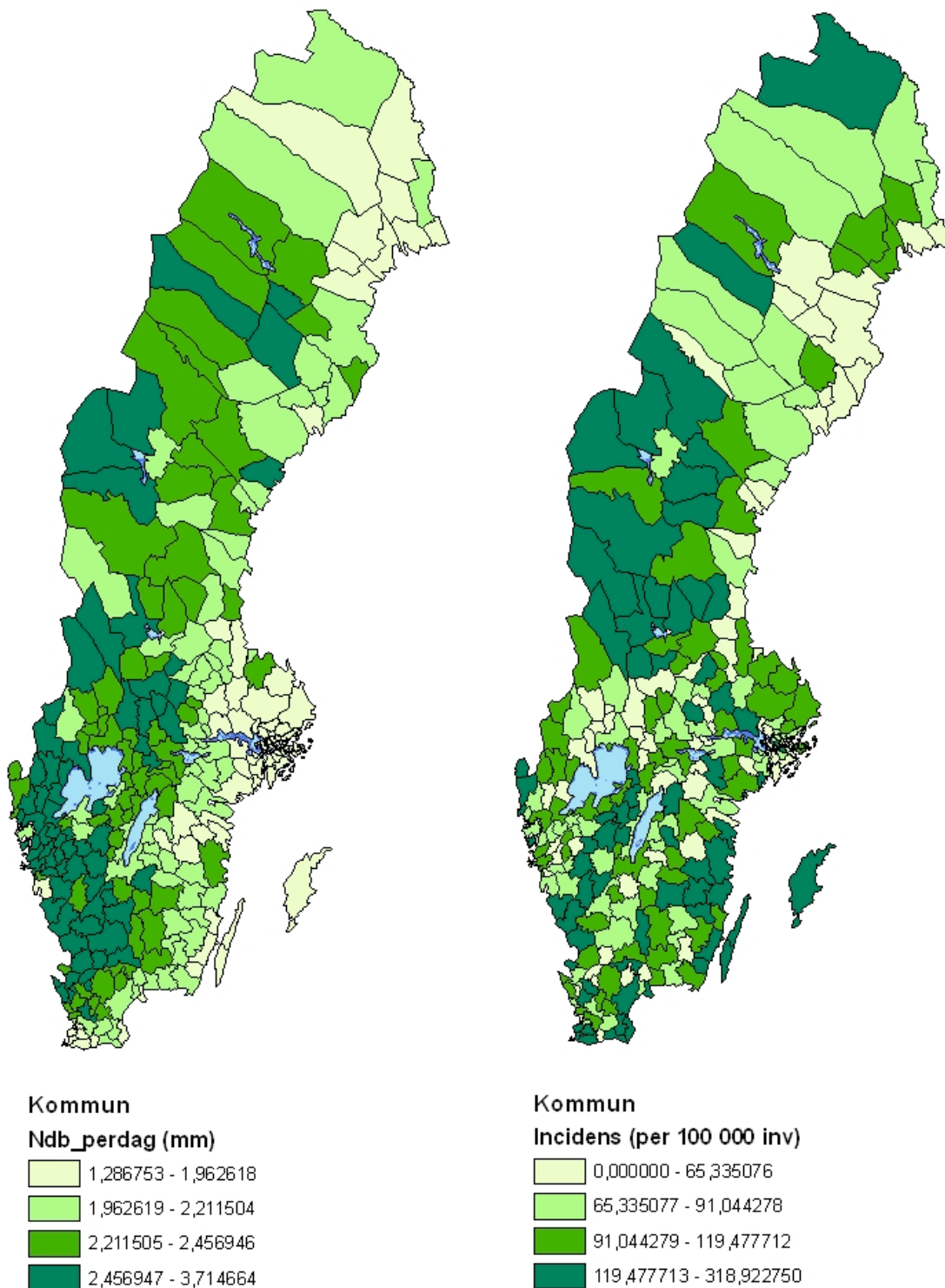
Figur 5.6: Fördelningen för alla kommuners nederbördsmängd

Histogrammet visar inga givna brytpunkter för kategorisering. Jag har därför valt att gruppera efter brytpunkter som sammanfaller med fördelningens kvantiler, dvs. lika många kommuner i varje grupp, fördelade över fyra grupper. Jag har även genomfört testet för annan kategoriindelning, där kommunerna delats upp i tre grupper med lika många i varje. Eftersom det gav liknande resultat nöjer jag mig med att presentera tillvägagångssätt och resultat från den första beskrivna kategoriindelningen.

De fyra grupperna kommer i fortsättningen att kallas för kategori 1, 2, 3 och 4 där kategori 1 är den grupp av kommuner med lägst nederbördsmängd och kategori 4 den med högst. Inom varje kategori finns 72 eller 73 observationer av incidens, eftersom det är 72 eller 73 kommuner i varje kategori som var och en har ett mått på incidens.

Nedan visas två kartor över Sverige där den ena visar Sveriges kommuner uppdelade i fyra kategorier baserade på nederbördsmängder och den andra visar Sveriges kommuner uppdelade i fyra kategorier baserade på incidens. Indelningen av kommuner med avseende på incidens gjordes även den utifrån fördelningens kvantiler. Syftet med kartorna är att visuellt ge en uppfattning av om kommuner med stor nederbördsmängd även har stor incidens och tvärt om.

Karta över Sverige, en jämförelse mellan nederbördsmängd och incidens för varje kommun



Figur 5.7: Karta över Sverige med information om nederbördsmängd och incidens för alla kommuner under perioden 1998-2003, maj-oktober

5.2.2.3 Genomförande och resultat variansanalys

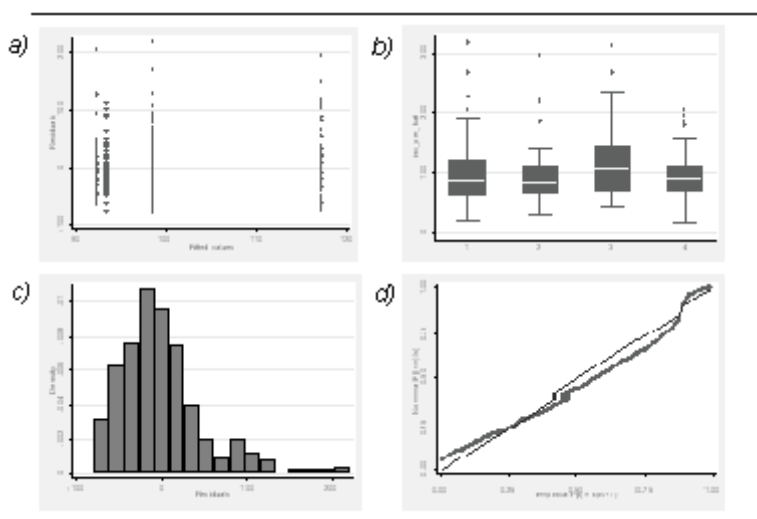
Nollhypotesen som testas i variansanalysen är att väntevärdet för incidensen för varje kategori är lika stor. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. Om denna hypotes kan förkastas vet man att minst ett väntevärde signifikant skiljer sig från de andra och att det då troligen finns någon form av samband mellan variablerna. Den inbördes ordningen mellan kategoriernas väntevärden för incidens kan ge en fingervisning på det eventuella sambandets form.

Modellen som testas i variansanalysen är:

$$Incidens_{ij} = medelincidens + effekt(kategori)_j + e_{ij}$$

för observationsnummer i inom kategori j .

Vid genomförande av envägs variansanalys på datamaterialet visade det sig att antagandet om normalfördelning och lika varianser inte var uppfyllt. Det kan man se i graferna i figur 5.8 som är baserade på resultatet utifrån variansanalysen. Graferna a) och b) tyder på att kriterierna för lika varianser inte är uppfyllda, eftersom spridningen på punkterna i vertikalt led skiljer sig mellan kategorierna i både figur a) och b). Graferna c) och d) talar om att antagandet för normalfördelning inte är uppfyllt, eftersom histogrammet i c) är aningen skevt och kurvan i d) inte sammanfaller med diagonalen. Om antaganden som variansanalysen baseras på inte är uppfyllda måste man vara försiktig med tolkningen av resultaten eftersom de inte riktigt visar vad de avser att visa.



Figur 5.8: Test om antagande för variansanalys är uppfyllda

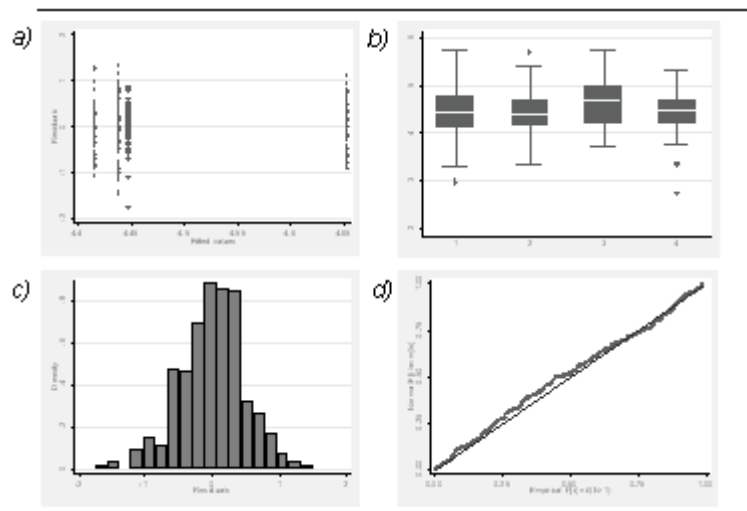
a) Residuals versus fitted values

b) Box-Plot

c) Histogram över residualer

d) Normal QQ plot av residualer

För att skapa bättre förutsättningar för variansanalysen logaritmerades värdena för incidensvariabeln. Villkoren för lika varianser och normalfördelning blev då bättre uppfyllda (se graferna i figur 5.9). Det resultat som presenteras från variansanalysen är alltså gjort för ett logaritmerat värde för incidensen.



Figur 5.9: Test om antaganden för variansanalys är uppfyllda för logaritmerade värden för incidensen.

a) Residuals versus fitted values

b) Box-Plot

c) Histogram över residualer

d) Normal QQ plot av residualer

Sammanfattning av incidens			
kategori	medelvärde	standardavvikelse	frekvens
1	4,439	0,567	73
2	4,417	0,466	72
3	4,654	0,476	72
4	4,448	0,450	72
Total	4,489	0,499	289

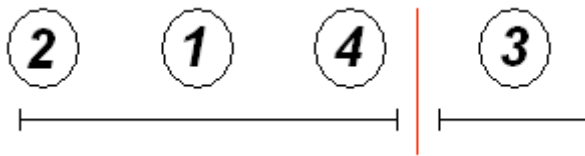
Tabell 5.1: Sammanfattningstabell från variansanalysen med information om medelvärde och standardavvikelse för incidensen för varje kategori.

Källa	SS	df	MS	F	Prob > F
Mellan grupper	2,63	3	0,877	3,62	0,0136
Inom grupper	69,024	285	0,242		
Totalt	71,654	288	0,249		

Tabell 5.2: Resultat från variansanalysen.

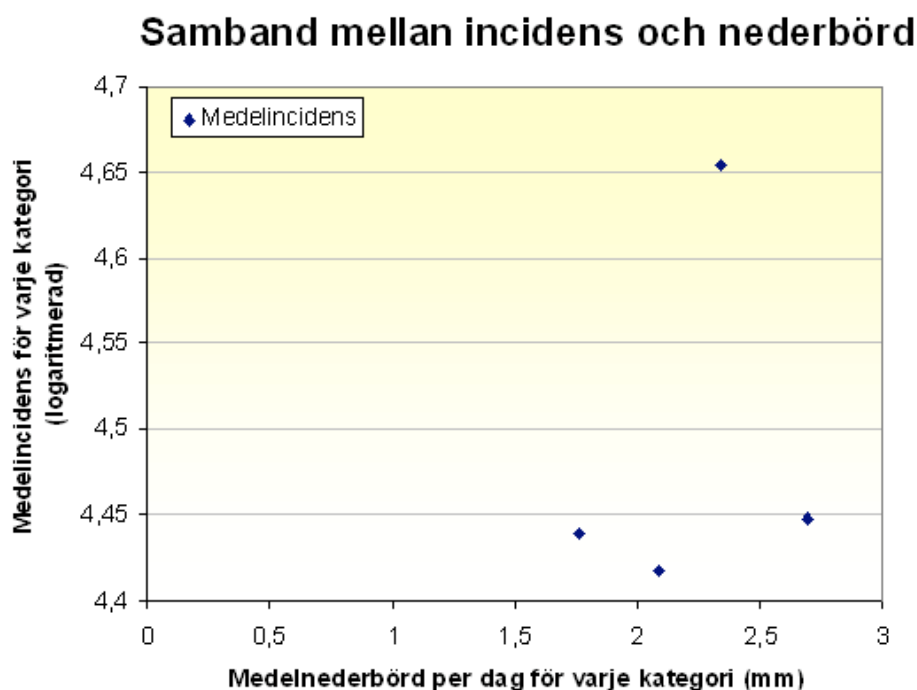
Resultatet från variansanalysen gav ett p-värde på 0,0136, vilket innebär att nollhypotesen kan förkastas på 95 % signifikansnivå. Minst en av kategorierna har alltså väntevärden för incidensen som signifikant skiljer sig från de andra.

Med hjälp av parvisa test med Tukeys metod⁶⁷, undersöktes vilka kategorier vars väntevärden signifikant skilde sig från varandra.



Figur 5.10: Resultat från Tukeys parvisa test.

Resultatet blev, som synes i figur 5.10, att väntevärdet från kategori 3 signifikant skilde sig från väntevärdena i kategori 1, 2 och 4. I figur 5.11 visas sambandet mellan incidens och medelnederbörds mängd för varje kategori.



Figur 5.11: Medelincidensen från varje kategori plottad mot medelnederbörds mängden för varje kategori.

Under den hypotes som arbetet syftar att pröva förväntades att den inbördes relationen för incidensen för varje grupp skulle se ut på följande sätt:

$$\mu_1 < \mu_2 < \mu_3 < \mu_4$$

Resultatet från variansanalysen visade dock på en annan typ av inbördes relation mellan de olika kategorierna. Att resultatet visar signifikant skillnad mellan väntevärdena från de olika kategorierna tyder på att det finns någon form av samband mellan variablerna. Det är därför intressant att gå vidare i analysen och utreda det eventuella icke linjära samband som råder mellan variablerna som variansanalysen tyder på. Sambandet utreddes genom att sätta upp kvadratiske icke linjära modeller som testas med hjälp av regressionsanalys.

⁶⁷ Montgomery, 2001 (s.96)

5.2.3 Regressionsanalys

5.2.3.1 Poisson regression teori

En variabel som mäter antalet händelser, som till exempel antalet inträffade smittfall, har vissa specifika egenskaper att ta hänsyn till vid regressionsanalys. En sådan variabel kan till exempel bara anta positiva heltal. Vid analys av den här typen av data är det lämpligt att anta att den beroende variabeln tillhör en Poisson fördelning, som är den enklaste fördelningen för modellering av variabler som mäter antal. Det finns dock flera andra mer komplicerade fördelningar som kan vara lämpliga för variabler av dessa slag.⁶⁸

Den beroende variabeln Y är lika med antal inträffade händelser under given tidsperiod. $Y = y_1, y_2, y_3, \dots, y_n$ där n är det totala antalet observationer. Karaktäristiskt för en variabel som är Poisson fördelad är att dess väntevärde och varians antas vara lika stor.

$$E[Y] = V[Y] = \mu$$

För beroende variabler som är Poissonfördelade används i första hand Poisson regression för att utreda sambandet mellan Y och förklarande variabler X . Grundmodellen för Poisson regression kan skrivas som:

$$E[y_i | x_{ij}] = \exp(\beta' x_{ij})$$

där antalet inträffade händelser är proportionell mot exponentialen av linjärkombinationen av alla förklarande variabler som är inkluderade i modellen. Exponentialfunktionen garanterar att värdet för Y alltid blir större än noll, såsom önskat. Den generella modellen för Poisson regression skriv dock oftast som

$$\log(E[y_i | x_{ij}]) = \beta_0 + \sum \beta_j x_{ij}$$

där β_0 är skärningen för y-axeln och β_j är parametrarna för de förklarande variablerna x_j .

Den stora begränsningen med Poisson regression är att antagandet om $E[Y] = V[Y] = \mu$ måste vara uppfyllt för datamaterialet för att resultatet från analysen ska kunna bli tolkningsbar. I verkligheten är dock detta antagande långt ifrån alltid uppfyllt. Om $E[Y] > V[Y]$ eller om $E[Y] < V[Y]$ talar man om under- respektive överdispersion. En Negativ Binominal fördelning är i dessa fall bättre att anta för den beroende variabeln.⁶⁹

5.2.3.2 Negativ Binominal Regression teori

För variabler som mäter ett antal inträffade händelser, men som inte uppfyller villkoret $E[Y] = V[Y]$, kan en negativ binominal fördelning antas istället för en Poisson fördelning. Om väntevärdet för Y , $E[Y] = \lambda_i$ antas variansen vara $V[Y] = \lambda_i(1 + \eta^2 \lambda_i)$.

⁶⁸ Armitage & Colton, 2005

⁶⁹ Armitage & Colton, 2005

Vid negativ binominal regression skattas förutom parametrarna β_j även ett värde för η som kompenserar för att väntevärdet och variansen för Y inte är lika.⁷⁰

Det går att testa både om Poisson Regression eller Negativ Binominal Regression är lämplig att anta för det datamaterial analysen bygger på. Deviansen är ett exempel på ett sådant mått som mäter ”goodness of fitness” för modellen. Deviansen talar inte om eventuellt samband mellan variablerna utan snarare om det är lämpligt att ställa upp modellen på det sätt som gjorts i form av om antagandet för variansen är lämplig. Deviansen dividerat med antalet frihetsgrader ska vara ungefär lika med 1 om antagandena för modellen är korrekta. Om detta inte är uppfyllt är resultatet från regressionsanalysen inte tolkningsbar och modellen för regressionsanalysen bör ersättas med en annan.⁷¹

5.2.3.3 Genomförande regressionsanalys

Regressionsanalysen utgår från observerade värden i varje kommun över incidens (antal fall under undersökningsperioden per 100 000 invånare) och medelnederbördsmängd per dag. Kommunerna är inte längre indelade i kategorier utan ses var och en för sig som en undersökningsenhet. Den modell som ställdes upp som grund för regressionsanalysen var:

$$Y = \beta_0 + \beta_1(ndb) + \beta_2(ndb)^2$$

Som ett första steg i analysen antas Y = incidens tillhöra en Poissonfördelning med väntevärde och variansen μ . Vid beräkning av väntevärde och varians för incidensen ser man dock att $E[Y] = 100,8717$ och $V[Y] = 2738,65183$, vilket ger en första indikation på överdispersion. Vid genomförande av Poisson regression på materialet fås visserligen ett signifikant samband, men deviansen dividerat med antalet frihetsgrader ger ett värde på lite över 5, vilket även det tyder på överdispersion. Antagandet om Poissonfördelning är inte uppfyllt och resultatet blir omöjligt att tolka.

Nästa steg blir att anta att Y tillhör en negativ binominal fördelning med väntevärde λ_i och varians $\lambda_i(1 + \eta^2 \lambda_i)$. Även detta test gav en devians dividerat med frihetsgrader som var skilt från ett. Varken med Poisson eller med Negativ Binominal regression har jag kunnat hitta en modell vars antaganden om varians stämmer överens med mitt datamaterial. Detta betyder inte att det inte finns något icke-linjärt samband mellan smittfall och nederbörd. Dock kommer det att krävas mycket arbete för att hitta en lämplig modell för att testa detta med hjälp av regressionsanalys. Jag lämnar därför ytterligare försök till detta och fokuserar istället på att summera de resultat jag fått fram.

5.2.4 Summering

Genom att studera skillnader mellan olika regioner i Sverige gällande incidens och nederbörd hoppades jag finna information om sambandet mellan dessa variabler. Fördelen med analys i rummet är att eventuella bakomliggande årstidsmässiga faktorer inte kommer med i analysen på samma sätt som vid tidsserieanalys. Nackdelen med att jämföra olika regioner med varandra är dock att eventuella regionala skillnader bakas in

⁷⁰ Armitage & Colton, 2005

⁷¹ Stata corporation, 2003

i variabeln för incidens, vilket kan försvåra tolkningen av resultaten. Regionala skillnader som kan påverka incidensen för varje kommun kan till exempel vara administrativa olikheter i rapporteringssystemet för olika län och olika riskbeteenden i olika delar av landet. Detta är faktorer som mycket väl kan påverka incidensen, men som inte alls har någonting med nederbördsmängden att göra. De regionala skillnaderna skulle rent teoretiskt kunna vara förklaringen både på ett visat samband eller på att inget samband kunnat hittas, utan att det behöver stämma med den verkliga relationen mellan smittfall och nederbörd.

De slutsatser som kan dras utifrån analysen i det här avsnittet är att det inte finns något tydligt samband mellan variablerna incidens och nederbörd och om det finns något samband över huvud taget så är det troligen inte linjärt. Resultaten från variansanalysen visade på att hypotesen om att ökad nederbördsmängd ger ökat antal smittfall inte verkar stämma. Tidigare ekologiska studier⁷² som har undersökt hur klimatfaktorer påverkar spridning av sjukdomar har visat på kvadratiske samband, vilket variansanalysen i det här avsnittet även gav stöd för gällande sambandet mellan campylobacterfall och nederbörd. Eftersom jag inte kunde hitta någon lämplig modell att testa sambandet mellan variablerna på i en regressionsanalys, låter jag det exakta utseendet på sambandet fortsätta vara oklart.

Regionala skillnader som inte bara beror av nederbördsmängd kan vara förklaringen till den stora varians som gjorde det svårt att hitta någon lämplig modell att basera regressionsanalysen på. I nästa avsnitt kommer jag försöka eliminera dessa regionala skillnader att komma med i modellen genom att bara titta på ett utvalt område och studera variation över tiden på det utvalda området.

5.3 Testmetod 3: Analys i tid och rum

5.3.1 Val av region

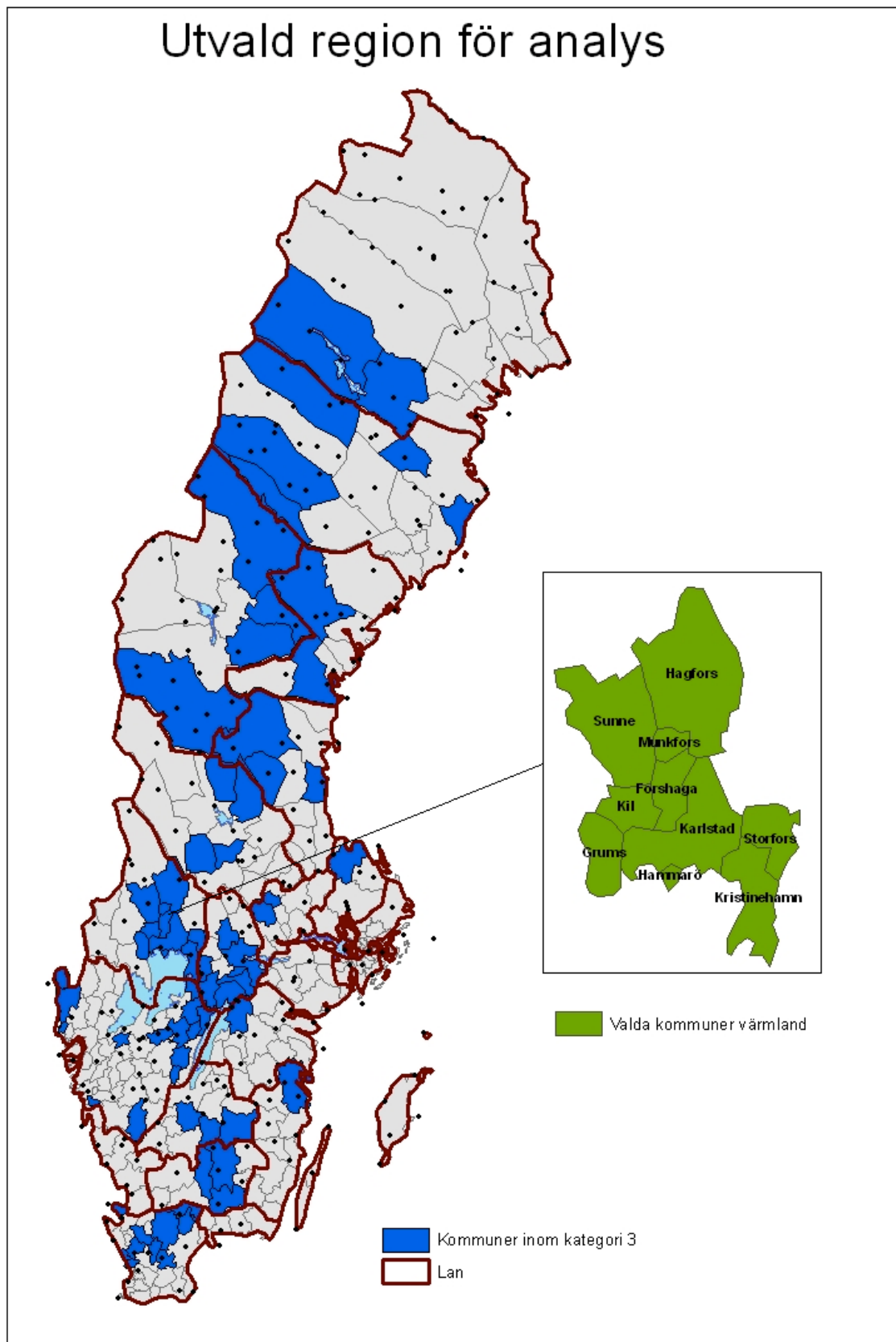
I det här avsnittet analyserades en utvald region. Området valdes så att om ett samband mellan smittfall och nederbörd finns, ska det framstå extra tydligt inom det området. Följande krav har jag därför ställt upp för den region jag ska välja att undersöka:

- Regionen måste vara så pass stor och tätbefolkad att så stort antal smittfall som möjligt har inträffat där. Vid alltför få inträffade fall spelar slumpvariationer stor roll i analysen och tolkningen av resultat blir svår.
- Regionen måste vara tillräckligt liten för att samma nederbördsmängd ska kunna representera hela regionen samtidigt.
- De kommuner som ska ingå i den utvalda regionen ska tillhöra kategori 3 ur kommunindelningen till variansanalysen i föregående avsnitt, eftersom det var kommuner med den nederbördsmängden som uppgav signifikant högre incidens än övriga kategorier och som därför är mest intressant att undersöka.
- Regionen bör ha ungefär samma befolkningsmängd över året och från år till år, så att variationer i antal fall mäter samma sak som incidensen.
- Regionen ska tillhöra samma län så att samma förutsättningar ur anmälnings synpunkt gäller för hela valda regionen.

⁷² Thomson et al., 2006

- Regionen ska ha en stor andel enskilda brunnar, eftersom det är där risken för kontaminering av dricksvatten är som störst och som hypotesen borde framträda som tydligast.

I kartan nedan har gränserna för länsindelningen i Sverige markerats ut som röda linjer. Alla kommuner som tillhörde kategori tre i grupperingen i avsnitt två har markerats blå. Som synes finns det många alternativa regioner att välja bland. Jag har valt att närmare undersöka tio utvalda kommuner i Värmlands län. Enligt opublicerade uppgifter från Svenskt Vatten har dessa tio utvalda kommuner åtminstone större andel enskild vattenförsörjning än medelvärdet för övriga landet. 144 smittfall inträffade i den valda regionen under undersökningsperioden maj till oktober, 1998-2003.



Figur 5.12: Karta över Sverige som motiverar bakgrunden till val av region samt visar placering och utseende på den utvalda regionen.

Kommun	Antal fall	Kategori	Andel kommunalt vatten	Län
Hagfors	11	3	82%	Värmland
Sunne	8	3	49%	Värmland
Munkors	2	3	94%	Värmland
Forshaga	7	3	88%	Värmland
Kil	7	3	75%	Värmland
Grums	4	3	78%	Värmland
Karlstad	63	3	91%	Värmland
Hammarö	16	3	89%	Värmland
Storfors	3	3	75%	Värmland
Kristinehamn	25	3	67%	Värmland

Tabell 5.3: Visar egenskaper för de kommuner som ingår i den region som ska analyseras.

5.3.2 Utförande av test

Antalet smittfall i regionen summerades för varje månad under undersökningsperioden. På samma sätt beräknades ett medelvärde för nederbördsmängd per dag för varje månad. Dessa värden ligger till grund för kommande analys.

ANTAL SMITTFALL							
År	maj	jun	jul	aug	sep	okt	Totalt
1998	1	4	5	2	5	1	18
1999	4	5	6	5	2	0	22
2000	1	8	5	3	1	3	21
2001	2	9	5	2	3	2	23
2002	6	8	8	2	2	2	28
2003	0	6	6	9	7	4	32

Tabell 5.4: Antalet smittfall i inträffade i den utvalda regionen varje månad för undersökningsperioden.

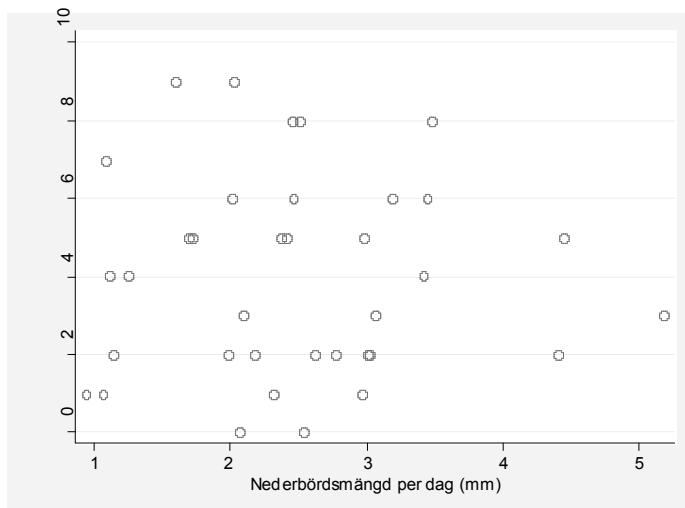
NEDERBÖRDSMÄNGD							
År	maj	jun	jul	aug	sep	okt	Totalt
1998	0,936	3,418	2,376	3,007	2,408	2,966	2,5185
1999	1,245	2,983	3,449	1,697	4,413	2,065	2,642
2000	2,318	2,507	4,456	2,095	1,061	5,181	2,936333
2001	1,138	2,024	1,723	3,021	3,062	2,177	2,190833
2002	2,46	3,486	2,454	2,765	1,983	2,623	2,6285
2003	2,54	3,185	2,01	1,598	1,084	1,112	1,9215

Tabell 5.5: Medelvärde för nederbördsmängden per dag i den utvalda regionen varje månad för undersökningsperioden.

Samband mellan smittfall och nederbörd testades genom att plotta variablerna mot varandra och genomföra regressionsanalys.

5.3.3 Resultat

Genom att plotta variablerna X och Y mot varandra genererades grafen nedan. Trots att eventuella regionala skillnader utesluts ur modellen jämfört med föregående avsnitt syns inget tydligare samband mellan smittfall och nederbörd.



Figur 5.13: Antal smittfall och nederbörds mängd för 10 utvalda kommuner i Värmland. Y-axeln exv. "antalet rapporterade campylobacterfall per månad" och X-axeln exv. "medelnederbörds mängd per dag" för respektive månad.

Genom att bara undersöka en utvald region elimineras regionala skillnader ur modellen. Genom att lägga till årtal och månad som enskilda variabler i regressionsanalysen elimineras även årstidsmässiga faktorer ur modellen. Förutsättningarna borde därför vara bra för att testa sambandet mellan smittfall och nederbörd i den valda regionen på det här sättet. Nackdelen är dock att vid uppdelning av datamängden både i tid och rum blir antalet fall för varje undersökt enhet väldigt få och risken är stor att slumpen spelar stor roll. Att resultaten från detta test ser ut som det gör kan alltså inte uteslutas bero på just slumpen.

På samma sätt som i avsnitt två testades till att börja med Spearmans korrelation, som ett mått på olinjärt samband och Pearsons korrelation som ett mått på datamaterialet.

Spearman's rho = 0,0515

Pearson's korrelationskoefficient = 0,0016

Resultaten tyder på att det i stort sett inte finns något samband alls mellan det två variablerna. Dock blev korrelationskoefficienten för ett eventuellt icke-linjärsamband något högre. Trots att korrelationskoefficienterna blev så små skulle samband mellan variablerna kunna finnas om andra variabler lades till i modellen, såsom till exempel år och nederbörd. Denna modell testas med regressionsanalys.

$$Y = \beta_0 + \beta_1(\text{år}) + \beta_2(\text{månad}) + \beta_3(\text{ndb}) + \beta_4(\text{ndb})^2$$

Modellen testas först med Poisson regressionsanalys och sedan med Negativ Binominal regressionsanalys. På samma sätt som i avsnitt 2 var det svårt att hitta en modell för variansen som beskriver datamaterialet på ett bra sätt. Deviansen dividerat med frihetsgraderna blir dock närmre 1 för det här datamaterialet. Nederbördsvariabeln bidrar dock inte signifikant till modellen, vilket tyder på att det inte finns något samband mellan smittfall och nederbörd för datamaterialet.

6. Diskussion

6.1 Samband mellan campylobacterfall och nederbörd

Syftet med det här examensarbetet har varit att utreda sambandet mellan campylobacterfall och nederbörd. De slutsatser som kan dras utifrån de resultat som kommit fram är att kurvorna för nederbördsmängder och campylobacterfall till viss del följer samma tidsmässiga mönster, men att det snarare beror på bakomliggande årstidsmässiga faktorer än på att nederbördsmängden direkt påverkar antalet smittfall. De geografiska analyser som genomförts i arbetet har inte kunnat påvisa något tydligt samband mellan variablerna. Trots detta resultat är det mycket möjligt att det i verkligheten finns någon form av samband mellan antalet inträffade campylobacterfall och nederbörd. Om det finns är det dock troligen varken särskilt starkt eller linjärt.

Spridningen av Campylobacter påverkas av otroligt många olika faktorer. Det är ett oerhört komplext system och det vore nästan orimligt att tro att detta skulle kunna modelleras med enbart variabler för smittfall och nederbörd. Den något naiva önskan att kunna utreda sambandet mellan smittfall och nederbörd genom en så enkel modell som antagits i det här arbetet byggde på att om sambandet mellan variablerna var tillräckligt stark så skulle det framträda trots försummande av andra bakomliggande variabler. Så var dock inte fallet för det här materialet, vilket är ett resultat i sig som säger något viktigt om relationen mellan smittfall och nederbörd.

6.2 Resultatet kopplat till hypotesen

Den hypotes som arbetet syftade att pröva var att höga vattennivåer leder till ökad risk för dricksvattnet att kontamineras av Campylobacter, vilket i sin tur leder till fler inträffade smittfall. Redan på tidigt stadium väcktes en insikt om svårigheten att pröva denna hypotes genom att utreda sambandet mellan smittfall och nederbörd. Om ett samband mellan smitta och nederbörd skulle kunna fastställas, skulle det visserligen säga en del om sanningshalten i den formulerade hypotesen. Problemet är dock att även om hypotesen verkligen stämmer så är det inte säkert att detta skulle ge utslag i ett samband mellan variablerna eftersom det finns så många fler orsaker till campylobacterios än just dricksvatten. Fall orsakade av andra faktorer än dricksvatten går inte att utesluta ur undersökningen, vilket skapar en stor utspädning av det eventuella sambandet mellan smittfall och nederbörd.

Det som försvårar undersökningen ytterligare är att Sveriges invånare får sin dricksvattenförsörjning från många olika typer av anläggningar och för varje anläggningstyp är risken för kontaminering olika stor och kan uppstå på olika sätt. Som tidigare undersökningar visat är enskild vattenförsörjning förknippat med större risk för smitta, vilket tyder på att dricksvattnet är en viktig smittoorsak för dessa hushåll. Det kan mycket väl vara så att hypotesen endast gäller för enskild vattenförsörjning och att ett samband skulle uppträda tydligt om bara individer med enskild vattenförsörjning inkluderats i undersökningen. Om hypotesen endast gäller för enskild vattenförsörjning är det inte så konstigt att ett samband inte kunnat påvisas i den här undersökningen eftersom enskild vattenförsörjning bara står för en mycket liten del av den totala dricksvattenförsörjningen. Hypotesen skulle troligen ha testats bättre genom att bara

inkludera individer med enskild vattenförsörjning i undersökningen. Vad jag vet finns det dock inget system för att kunna göra den utsortering bland fallen.

Det är alltså mycket svårt att dra några generella slutsatser kring den formulerade hypotesen. Resultat från variansanalysen visade dock att mer nederbörd inte alltid var förenat med fler smittfall, vilket förväntningen var om hypotesen skulle stämma. Den slutsats som kan dras är att om den vanligaste orsaken till inhemska sporadiska fall av *Campylobacter* hade varit dricksvatten, orsakad av kontaminering på grund av mycket nederbörd så hade detta synts i analysen. Detta förfarande kan med hjälp av resultaten från det här arbetet alltså uteslutas.

6.3 Felkällor och hur dessa skulle kunna undvikas

För att kunna genomföra denna studie har många antaganden, uppskattningar och förenklingar behövt göras. Undersökningen har även baserats på ett befintligt datamaterial som skapats för helt andra syften än för den här studien. Felkällor kan därför inte uteslutas och bör tas med i beräkningarna vid analys av resultaten. En av de största källorna till fel, som gör att variabeln för incidens inte säkert beskriver verkligheten, är att flertalet av alla *campylobacter*-fall inte registreras. Det antagandet som har gjorts för den här studien är att skillnader i registreringsfrekvens är lika över hela Sverige, över hela året och mellan olika år. Sanningshalten i detta antagande är dock mycket svår att uppskatta.

Felkällor kommer även in i datamaterialet då varje smittfall knutits till ett smittdatum och en smittort och resterande utesluts. Det antagande som denna uteslutning baserades på var att noggrannheten att fylla i registreringsuppgifter inte skiljde sig åt över tid och rum. I annat fall kan oproportionerligt många fall för en viss ort eller tidpunkt ha uteslutits ur undersökningen. Både vid skattning av smittdatum och av smittort kan felaktigheter även ha kommit in i materialet. Antagandet är att dessa i så fall var så pass små att de inte skulle påverka resultatet nämnvärt och att förskjutningar i datum och plats skulle ta ut varandra eftersom samma förfarande är gjort på så pass många fall.

Eftersom undersökningen har varit tvungen att anpassas efter existerande data har bästa möjliga undersökning försökt skapas, utifrån de uppgifter som funnits tillgängliga. Bättre förutsättningar för analysen hade varit om det på något enkelt sätt hade gått att utesluta kända, icke- dricksvattenrelaterade orsaker till sjukdomen, om smittort och smittotillfälle funnits angivet för samtliga smittfall och om det hade varit möjligt att bara innesluta individer med enskild dricksvattenförsörjning i undersökningen.

6.4 Framtida forskning

Det finns mycket mer arbete som skulle kunna göras för att utreda sambandet mellan nederbörd och *campylobacter*-smitta ytterligare. I och med genomförandet av den här studien har en del funderingar om framtida forskning växt fram. Dels skulle man kunna innefatta variabler om år och månad även i den modell som ställdes upp för regressionsanalys i testmetod 2. Kanske skulle variansen då bli lättare att skatta så att tolkningar skulle kunna göras utifrån analysen. Ett alternativ är att behandla variabeln för nederbördsmängd på annorlunda sätt. Såsom jag har gjort i det här arbetet har ett medelvärde för nederbördsmängder per dag under en vald tidsperiod använts. Vissa andra ekologiska studier som undersökt sambandet mellan smitta och nederbörd har istället använt ett ackumulerat värde för nederbördsvariabeln. Variabeln bygger i dessa

fall på summan av regnmängden för ett visst antal dagar, vilket fångar upp stora regnmängder eller torka flera dagar i rad på ett bättre sätt.

Jag tror dock att det krävs mer än små justeringar för variabler och metoder för att kunna utreda sambandet mellan campylobactersmitta och nederbörd i vidare bemärkelser än jag har gjort. Jag tror att det kommer att krävas att man lägger till andra variabler i modellen så att modellen mer beskriver det komplexa system som spridning av bakterien i verkligheten är, för att få fram ett tillfredställande resultat.

För att kunna svara på frågan huruvida den formulerade hypotesen för arbetet är sann, tror jag att det kommer att krävas en hel del ytterligare undersökningar för att få ett utförligare svar än det här arbetet lyckats generera. Datamaterialet som undersökningen baseras på bör sorteras ut mer noggrant så att utspädningen från andra orsakskällor till smittan minskar. Först då kan ett eventuellt samband mellan smittfall och nederbörd säga något om hypotesen. Det är dock praktiskt svårt att göra denna sortering och dessutom minskar antalet fall för varje utsortering som görs, vilket försvårar genomförandet av den statistiska undersökningen.

För att kunna testa om hypotesen stämmer för enskilda vattentäkter krävs en metod för att kunna sortera bland fallen så att endast dem med enskild vattenförsörjning innefattas i undersökningen. Ett förslag skulle kunna vara att vid registreringen av smittfall i SmiNet lägga till information om dricksvattenförsörjningen, alternativt ställa frågan om vatten druckits från enskild vattentäkt. Risker är dock att denna information, på samma sätt som många andra informationsfält i SmiNet i många fall skulle utebli och på så sätt inte vara till någon nytta.

Inför framtida forskning kan det vara värt att lägga tid på att utreda möjligheten till alternativa metoder för att undersöka hypotesens sanningshalt. Sveriges Geologiska Undersökningar håller på att registrera alla enskilda dricksvattenbrunnar i Sverige. Hittills är omkring 25 % av alla brunnar registrerade, men omkring 8000 - 10 000 brunnar läggs till varje år. I takt med att registret blir mer omfattande kommer det att kunna erbjuda nya möjligheter till analys av frågan. Kanske kan man få en del svar genom att undersöka om kluster av campylobacterfall uppstår kring dricksvattenbrunnarnas läge. Hur nederbördsmängder dessutom skulle kunna komma in i en sådan alternativ analys tål att tänkas på framöver.

6.5 Slutsatser

Några definitiva svar på sambandet mellan campylobactersmitta och nederbörd har det här arbetet inte lyckats generera. Inte heller ett exakt svar på den formulerade hypotesen. Undersökningen har ändå visat på många viktiga resultat, som jag hoppas fortsatta studier kan ha nytta av i framtiden. Arbetet har visat hur GIS kan användas för att lösa epidemiologiska frågeställningar och har även presenterat en rad statistiska metoder, med olika för och nackdelar för att lösa den här sortens problem. Resultatmässigt har undersökningen visat att nederbördsmängd åtminstone inte hör till de starkaste bakomliggande faktorerna för spridning av Campylobacter. Att stora nederbördsmängder leder till ökad risk för kontaminering av dricksvattnet, vilket resulterar i fler inträffade campylobacterfall, kan visserligen inte förkastas, men resultaten i den här undersökningen visar att detta troligen inte är den största förklaringen bakom orsaken till smittan.

7. Litteraturförteckning

7.1 Tryckta källor

7.1.1 Publicerade artiklar

Brennhovd, O., Kapperud, G. & Langeland, G., (1992). „*Survey of thermotolerant Campylobacter spp. And Yersinia spp. In three surface water sources in Norway.*” *International Journal of Food Microbiology*, Ser., 15, p.327-338.

Kapperud, G., Espeland, G., Wahl, E., Walde, A., Herikstad, H., Gustavsen, S., Tveit, I., Natås, O., Bevanger, L. & Digranes, A., (2003). “*Factors Associated with Increased and Decreased Risk of Campylobacter Infection: A Prospective Case-Control Study in Norway.*” *Am J Epidemiol.* 2003 Aug 1; 158(3): p.234-242.

Lindberg, T. & Lindqvist, R., (2005). “*Riskprofil, Dricksvatten och mikrobiologiska risker.*” Livsmedelsverkets rapport nr 28/2005

Lindbäck, J., (1999). “*Geografisk spridning av Campylobacterinfektioner – statistisk analys med MCMC*”

Livsmedelsverket, (2002). “*Riskprojekt 1 – 2000, Campylobacter i kött och vatten – kartläggning av Campylobacter i rått kött och råvatten till dricksvatten i Sverige år 2000.*” Rapport 10 – 2002, Livsmedelsverkets repro.

Martin, S., Penttinen, P., Hedin, G., Ljungström, M., Allestam, G., Andersson, Y. & Giesecke, J., (2006). “*A case-cohort study to investigate concomitant waterborne outbreaks of Campylobacter and gastroenteritis in Söderhamn, Sweden, 2002-3.*” *Journal of Water and Health, IWA Publishing*, 04.4, p.417-424.

Nygård, K., Andersson, Y., Röttingen, J.A., Svensson, Å., Lindbäck, J., Kistemann, T. & Giesecke, J., (2003). “*Association between environmental risk factors and campylobacter infections in Sweden.*” *Epidemiol. Infect.* (Cambridge) Ser., 132, 317-325.

Nylen, G., Dunstan, F., Palmer, S. R., Andersson, Y., Bager, F., Cowden, J., Feierl, G., Galloway, Y., Kapperud, G., Megraud, F., Molbak, K., Petersen, L. R. & Ruutu, P., (2002). “*The seasonal distribution of campylobacter infection in nine European countries and New Zealand.*” *Epidemiol. Infect* (Cambridge) Ser., 128, 383-390.

Patrick, M.E., Christiansen, L.E., Wainø, M., Ethelberg, S., Madsen, H. & Wegener, H.C., (2004). “*Climate as a Predictor of Prevalence of Campylobacter spp. In Humans and Broilers in Denmark.*” *Applied and Environmental Microbiology*, Vol. 70, No. 12, p.7474-7480.

Schönberg-Norio, D., Takkinen, J., Hänninen, M-L., Katila, M-L., Kaukoranta, S-S., Mattila, L. & Rautelin, H., (2004). “*Swimming and Campylobacter Infections.*” *Emerg Infect Dis.* 2004 Aug; 10(8): p.1474-1477.

Smibert R.M., (1978). "*The Genus Campylobacter*" Rev. Microbiol. (Virginia) 32:673-709.

Studahl, A. & Andersson, Y., (2000). "*Risk factors for indigenous campylobacter infection: a Swedish case-control study.*" Epidemiol. Infect. (Cambridge) Ser., 125, 269-275.

Tam, C.C., Rodrigues, L.C., O'Brien, S.J. & Hajat, S., (2005). "*Temperature dependence of reported Campylobacter infection in England, 1989-1999.*" Epidemiol. Infect. (Cambridge) Ser., 134, 119-125.

Thomson, M.C., Doblas-Reyes, F.J., Mason, S.J., Hagedorn, R., Connor, S.J., Phindela, T., Morse, A.P. & Palmer, T.N., (2006). "*Malaria early warnings based on seasonal climate forecasts from multi-model ensembles.*" Nature, Vol. 439, doi:10.1038.

7.1.2 Publicerade böcker

Armitage, P. & Colton, T., (2005). "*Encyclopedia of Biostatistics*", 2nd ed., John Wiley Sons Ltd., Chichester.

Blom, G., (1984). "*Sannolikhetsteori med tillämpningar*", andra upplagan, Studentlitteratur, Lund.

Brockwell, P.J. & Davis R.A., (2002). "*Introduction to Time Series and Forecasting*", 2nd ed., Springer, New York.

Hawker, J., Begg, N., Blair, I., Reintjes, R. & Weinberg, J., (2002). "*Communicable Disease Control Handbook*", MPG Books Ltd., Cornwall.

Heymann, D., (2004) "*Control of Communicable Diseases Manual*", 18th ed., American Public Health Association, Washington.

Hunter, P.R., (1997). "*Waterborne Disease*", 1st ed., John Wiley and Sons.

Iwarson, S., (1985). "*Infektions sjukdomar*", Kristianstads Boktryckeri AB, Kristianstad.

Johnson R.A. & Wichern D.W., (2002). "*Applied multivariate statistical analysis*", 5th ed., Prentice Hall, Upper Saddle River.

Montgomery, D.C., (2001). "*Design and Analysis of Experiments*", 5th ed., John Wiley & Sons Ltd., United States of America.

Stata corporation, (2003). "*Stata base reference manual, volume 3, release eight, N-R*", Stata Press Texas.

7.1.3 Kompendium

Andersson, Y., Gustavsson, P., Hammarquist, C., Kulander, L., Nilsson, P.E.W. & Olsson, A-M., (1998). "*Spring och ring, en studie av självrapportering vid mag-*

och tarmsjukdom” Godrings Tryckeri AB, utgiven i samarbete mellan Gotlands kommun, Länsstyrelsen i Gotlands län och Smittskyddsinstitutet.

Andersson, Y. & Gustavsson, O., (1998). *“Campylobacter – en vattenburen smitta att beakta vid planläggning av vattentäkt och ytvattentyp”* Modin-Tryck

Hjärt&Lungfonden, (1993). *”Vattnet vi dricker”*

Lindbäck, J., (1999). *”Geografisk spridning av Campylobacterinfektioner – statistisk analys med MCMC”* Forskningsrapport, Stockholms Universitet, Matematisk Statistik.

Naturvårdsverket, (1991). *”Grundvattentäkter”* Norstedts Tryckeri AB

Socialstyrelsen, Institutet för miljömedicin & Miljömedicinska enheten Stockholm Läns Landsting, (2001). *”Miljöhälsorapport 2001”* Modin-Tryck, Stockholm

7.3 Muntliga Källor

Andersson Y., Sektionschef Zoonoser SMI, Handledare för examensarbetet, hösten 2006

7.2 Internetkällor

Hemsida: Smittskyddsinstitutet, 2006-09-04

www.smittskyddsinstitutet.se

Hemsida: Socialstyrelsen, 2007-01-15

Smittskyddslag (2004:168)

www.socialstyrelsen.se

Hemsida: Svenskt Vatten, 2006-09-19

www.svensktvatten.se

Hemsida: Sveriges Geologiska Undersökningar, 2006-09-18

www.sgu.se

Hemsida: Sveriges Meteorologiska och Hydrologiska Institut, 2006-09-04

www.smhi.se

Hemsida: Vattenportalen, 2006-09-19

www.vattenportalen.se

8. Bilagor

8.1 Bilaga 1: Information om Campylobacterfall i SmiNet

Statistikdatum	Lokal utbrottsbeteckning
Ansvarig smittskyddsenhet	Nationell utbrottsbeteckning
Person-ID	Grund för diagnos
Namn	Övriga länder där smitta kan ha inträffat
Födelseår	Information av betydelse för s..
Ålder	Födelseland
Kön	Tidpunkt för ankomst, om ej föd...
Hemortslandsting	Givna förhållningsregler
Typ av Patient-ID	Namn på ansvarig smittspårare
Gatuadress	Annan smittväg
Postnummer	Annat yrke/sysselsättning
Ort	Annan orsak till diagnos
Telefonnummer	Vaccinerad
Landsting	Antal doser
Typ av klinik	Laboratoriediagnos
Sjukhus/klinik/mm	Speices
Adress	Typ av anmälan
Postadress	Laboratorium
Tel. kliniken/mottagningen	Lab nr
Klinikens kod	Undersökningsmaterial
Smittskyddsregion	Diagnosisk metod
Läkarens namn	Provtagningsdatum
Typ av läkare	Ankomst för prov
Datum för ansvarig läkares godkännande	Ansvarig läkares namn
Tel. egen anknytning	Remitterande klinik-adress
Registreringsdatum	Remitterande klinik
Smittdatum	Kommentar analysresultat
Insjukningsdatum	Kommentar remiss
Diagnosdatum	Kommentar prov
Klinisk ankomstdatum	Typningsresultat
Labsvardatum	Avresedatum
Avliden datum	Hemkomstdatum
Datum för senaste vaccindos	Hotel där patienten bott under resan
Smittland	Researrangör
Smittort	Annan relevant information
Insjuknandeort	Anger infektionens stadium
Typ av infektion	LGV
Orsak till diagnosen	AIDS kriterier
Smittväg	HIV typ/subtyp
Anmälan avser	Om tidigare HIV-anmäld ange d...
Finns kända/misstänkta fall i området	Är diagnosen TBE
Yrke/Sysselsättning	Är diagnosen West Nile
Tillhör utbrott	Ange identifierat virus
	Gammaglobulin

Har patienten tagit profylax enligt ...
Läkemedel
Om ej tagit profylax enligt ordin...

Typ av skada
Åldersgrupp

8.2 Bilaga 2: Borttagningsstatistik

År	Totalt antal fall	Borttaget pga datum	pga plats totalt
1998	2619	40	620
1999	2210	25	468
2000	2466	12	565
2001	2839	13	582
2002	2481	12	480
2003	2690	13	542

Totalt	15305	115	3257
--------	-------	-----	------

År	pga flera alternativ	pga ej kunde knytas till plats	pga ort fanns på flera ställen
1998	25	47	15
1999	24	32	13
2000	25	37	13
2001	29	49	22
2002	36	27	4
2003	33	48	11

Totalt	172	240	78
--------	-----	-----	----

År	Totalt borttaget
1998	660
1999	493
2000	577
2001	595
2002	492
2003	555

Totalt	3372
--------	------