

Hochschule für Angewandte Wissenschaften Hamburg

**Fachbereich Ökotrophologie
Studiengang Gesundheit**

**Case-Case-Studie eines Salmonelloseausbruchs in
Hamburg-Altona:**

Ein epidemiologisches Analysewerkzeug für Surveillancedaten

Tag der Abgabe: 28.01.2005

Vorgelegt von:

Ralf Krumkamp



Betreuung durch:

Prof. Dr. Ralf Reintjes

Dr. Martin Dirksen-Fischer

**Case-Case-Studie eines Salmonelloseausbruchs in
Hamburg-Altona:
Ein epidemiologisches Analysewerkzeug für Surveillancedaten**

Diplomarbeit von Ralf Krumkamp

**Betreuung durch:
Prof. Dr. Ralf Reintjes
Dr. Martin Dirksen-Fischer**

Danksagung:

Für die freundliche Kooperation bedanke ich mich bei Fr. Gryzewski, Fr. Schweizer, Fr. Spode, Hr. Dr. Dirksen-Fischer, Hr. Dr. Niessen, Hr. Dr. Paschen und Hr. Prof. Dr. Reintjes.

Abstract

Die Analyse von Infektionsausbruchsausbrüchen, die über ein Surveillancesystem erfasst wurden, ist vielen Fehlerquellen unterlegen. Herkömmlicherweise werden Daten meldepflichtiger Erkrankungen hinsichtlich der Häufigkeit potentieller Risikoquellen der Infizierten beurteilt. Diese Analyse erlaubt eine deskriptive Beurteilung. Für eine analytische Auswertung muss eine epidemiologische Studie, z.B. eine Fall-Kontroll-Studie durchgeführt werden. Eine solche Untersuchung ist nicht nur aufwendig und zeitintensiv, sondern auch anfällig für Bias. Eine Fallgruppe, welche durch ein Surveillancesystem erfasst wird, ist stark selektiert. Dadurch ist es schwierig, die geeigneten Kontrollen zu finden. In einer Case-Case-Studie besteht die Kontrollgruppe ebenfalls aus Erkrankten, die durch dasselbe Surveillancesystem erfasst wurden. So werden Selektions- und Recall-Bias vermieden. Dieses Studiendesign führt jedoch zu Einschränkungen in der Ergebnisinterpretation.

In dieser Arbeit wurde eine Case-Case-Studie mit Meldedaten eines Salmonelloseausbruchs in Hamburg-Altona durchgeführt. Die Kontrollgruppe bestand aus weiteren Salmonelloseerkrankten des Jahres 2003.

Der Konsum von Rindfleisch (OR 11,81; CI 2,35-66,66; $p < 0,001$) und Schweinefleisch (OR 8,44; CI 1,72-46,39; $p < 0,001$) zeigt signifikante Odds Ratios. Die weiteren untersuchten Variablen sind unauffällig.

Die durchgeführte Case-Case-Studie ist in der Lage, die Ursache des Ausbruchs aufzuzeigen. Sie ist nicht geeignet, einen kausalen Zusammenhang zwischen Infektion und Exposition herzustellen. Bei der Analyse handelt es sich um eine einfache und schnelle Methode, interpretierbare Risikoschätzer für Surveillancedaten zu erhalten.

Das Studiendesign ist für Erkrankungen geeignet, die sich in Subgruppen gliedern. So lassen sich einzelne Ausbrüche identifizieren und in Fall- und Kontrollgruppen einteilen. Um das Studiendesign in die Analyse meldepflichtiger Erkrankungen zu initiieren, muss eine standardisierte Befragung jedes Falls in der routinemäßigen Surveillance durchgeführt werden.

Abstract

Analysing reportable disease data without an epidemiologic investigation is difficult. Commonly surveillance outbreak data are interpreted using the frequency of potential sources. This provides only a descriptive overview. For an analytical analysis it is necessary to conduct an epidemiologic study like a case-control-study. This is not only costly and time consuming, it is also susceptible for bias. The case group which is captured from a surveillance system is subject to strong selection. Therefore it is difficult to find appropriate controls. A case-case-study uses infected persons as controls who have been registered by the same surveillance system as the cases. This reduces bias but opens new problems for the interpretation of the results.

A case-case-study of a salmonella outbreak in Hamburg-Altona was executed by using surveillance data. The “control group” consists of all reported salmonella cases in Hamburg-Altona in 2003.

The consume of beef (OR 11.81; CI 2.35-66.66; $p < 0.001$) and pork (OR 8.44; CI 1.72-46.39; $p < 0.001$) shows significant odds ratios. Other inquiry sources were inconspicuous.

The study is able to reveal the cause of the salmonella outbreak, but it is not qualified to put it in a causal relationship. However, it is an easy and quick way to obtain quantified risk rates to estimate surveillance data.

The study design is appropriate for diseases which are split into subgroups. Using sub typing of the germ is needful to identify a potential outbreak. To establish the study design in reportable disease control it is necessary to use standardised survey for every case.

Inhaltsverzeichnis:

1	Hintergrund.....	1
1.1	Ausbruchsuntersuchung durchgeführt mit routinemäßig erhobenen Surveillancedaten	1
1.1.1	Deskriptive Interpretation von Surveillancedaten	1
1.1.2	Ausbruchsuntersuchung anhand einer Fall-Kontroll-Studie	2
1.1.3	Ausbruchsuntersuchung anhand einer Case-Case-Studie	4
1.2	Case-Case-Studien in der epidemiologischen Forschung	5
1.3	Salmonellose.....	6
1.3.1	Mikrobiologie der Salmonellose	7
1.3.2	Epidemiologie der Salmonellose	9
2	Material & Methode	11
2.1	Epidemiologische Analysemethoden der Studie	11
2.1.1	Bivariate Analyse	11
2.1.2	Stratifikation	12
2.1.3	Logistische Regression.....	14
2.2	Datensammlung	15
2.2.1	Fallgruppe.....	15
2.2.2	Kontroll/Fallgruppe	16
2.3	Datenanalyse.....	17
3	Ergebnisse.....	19
3.1	Deskriptive Analyse des Datensatzes	19
3.2	Ergebnisse der epidemiologischen Analyse	21
3.2.1	Bivariate Analyse	21
3.2.2	Stratifizierte Analyse nach Geschlecht.....	21
3.2.3	Stratifizierte Analyse nach Altersgruppen.....	23
3.2.4	Logistische Regression.....	23
4	Diskussion	24
4.1	Interpretation der Ergebnisse.....	24
4.1.1	Bivariate Analyse	24
4.1.2	Stratifizierte Analyse nach Geschlecht.....	25

4.1.3	Stratifizierte Analyse nach Altersgruppen	26
4.1.4	Logistische Regression.....	26
4.2	Unterschiedliche Erhebung der Fall- und Kontroll/Fallgruppe.....	27
4.3	Stärke des Zusammenhangs	28
4.4	Vorteile des Studiendesigns	29
5	Ausblick	30
5.1	Implementierung einer Case-Case-Analyse in ein laufendes Meldesystem	30
	Literaturverzeichnis.....	35
	Eidesstattliche Erklärung.....	39
	Anlage 1: Fragebogen für die Dateneingabe.....	41
	Anlage 2: Analysesyntax der bivariaten und stratifizierten Analyse	45
	Anlage 3: Syntax zur Erstellung des Logistischen-Regressionsdatensatzes	49

Darstellungsverzeichnis

Abbildungen:

Abbildung 1: Deskriptive Auswertung von Häufigkeiten, Gegenüberstellung von zwei vergleichbaren Gruppen	1
Abbildung 2: Schematische Darstellung einer Fall-Kontroll-Studie.....	2
Abbildung 3: Schematische Darstellung einer Case-Case-Studie.....	5
Abbildung 4: Epidemische Kurve der Fallgruppe in Wochen zusammengefasst, gemeldete Salmonellosen aus Hamburg-Altona, 2003.....	20
Abbildung 5: Epidemische Kurve der Kontroll/Fallgruppe in Monaten zusammengefasst, gemeldete Salmonellosen aus Hamburg-Altona, 2003.....	20
Abbildung 6: Implementierung einer Case-Case-Analyse in ein laufendes Meldesystem	30

Tabellen:

Tabelle 1: Verteilung der 10 häufigsten Serovare von <i>S. enterica</i> der übermittelten Salmonellosefälle in Deutschland 2003	8
Tabelle 2: Verteilung der <i>Salmonella</i> Serovare in der Kontroll/Fallgruppe, gemeldete Infektionen aus Hamburg-Altona, 2003	19
Tabelle 3: Ergebnistabelle bivariate Analyse, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003	21
Tabelle 4: Ergebnistabelle stratifizierte Analyse nach Geschlecht, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003	22
Tabelle 5: Ergebnistabelle stratifizierte Analyse nach Alter, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003	22

1 Hintergrund

1.1 Ausbruchsuntersuchung durchgeführt mit routinemäßig erhobenen Surveillancedaten

Deutsche Gesundheitsämter sind nach dem „Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen“ (IfSG) § 11 Abs. 1 verpflichtet, dem Robert Koch Institut mögliche Infektionsquellen von Ausbrüchen meldepflichtiger Erkrankungen zu übermitteln [vgl. Beck Texte 2003, S. 328 ff]. Die Analyse der Informationen infizierter Personen kann nach verschiedenen Methoden erfolgen. Sie unterscheiden sich hinsichtlich des Aufwands der Durchführung und der Validität ihrer Ergebnisse.

1.1.1 Deskriptive Interpretation von Surveillancedaten

Bei einer Hospitation in Hamburger Gesundheitsämtern¹ wurde berichtet, dass für die Ermittlung potentieller Infektionsquellen gemeldete infizierte Personen von Medizinalaufsehern kontaktieren werden. In einem Gespräch werden sie bezüglich verschiedener Expositionen befragt.² Konnte ein Infektionsausbruch identifiziert werden, vergleicht man diese erhobenen Daten innerhalb des Ausbruchs. Werden Häufungen in einzelnen Expositionen beobachtet, kann man Rückschlüsse auf die Infektionsursache schließen. Eine Risikoquelle, die oft genannt wurde, kann für die Ursache eines Ausbruchs

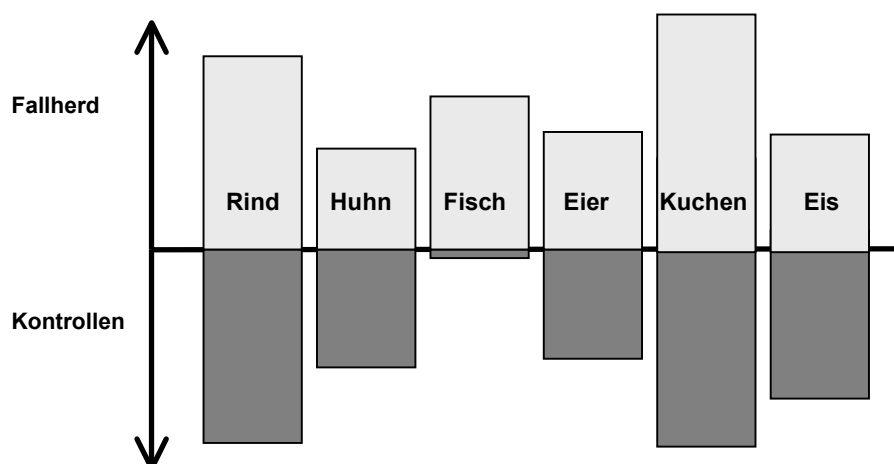


Abbildung 1: Deskriptive Auswertung von Häufigkeiten, Gegenüberstellung von zwei vergleichbaren Gruppen

¹ Die Hospitation erfolgte im April 2004 im Gesundheitsamt Hamburg Altona und Hamburg Eimsbüttel.

² Im weiteren Text wird diese Standardbefragung gemeldeter infizierter Personen als „routinemäßige Surveillance“ bezeichnet.

verantwortlich sein. Diese Analyse ist eine deskriptive Auswertung von Expositionen. Die Ergebnisse können in keinen ursächlichen Zusammenhang mit der Infektion gestellt werden. Hierfür bedarf es einer analytischen epidemiologischen Studie. Um herauszufinden, welches Verhalten zu einer Infektion geführt hat, muss untersucht werden, wie sich gesunde Personen verhalten haben, die unter dem gleichen Risiko standen sich zu infizieren (Abbildung 1). Eine Exposition, die in einer Gruppe erkrankter Personen genauso bestand wie in einer Gruppe vergleichbarer gesunder Personen scheint nicht für die Infektion verantwortlich zu sein. Beobachtet man eine Exposition unter den Infizierten, unter den Gesunden hingegen nicht, kann ein kausalen Zusammenhang zur Infektion bestehen [vgl. Giesecke 2002, S. 27].

1.1.2 Ausbruchuntersuchung anhand einer Fall-Kontroll-Studie

Ein geeignetes Studiendesign, um unterschiedliches Verhalten zwischen gesunden und infizierten Personen zu untersuchen, ist das einer Fall-Kontroll-Studie (Abbildung 2). In ihr werden Expositionen von infizierten Fällen mit denen von gesunden Kontrollen verglichen.

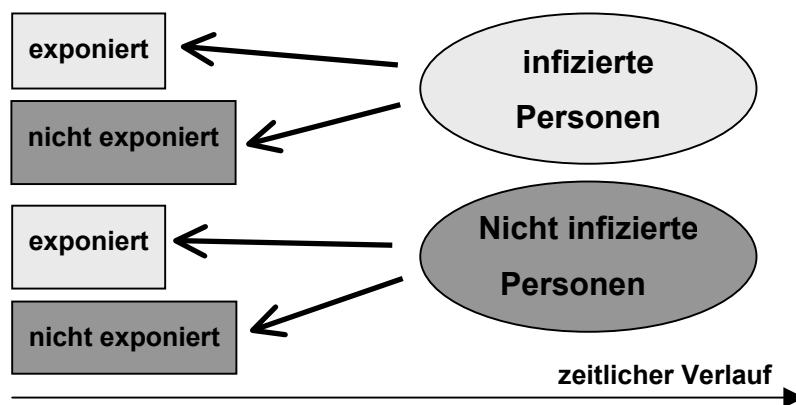


Abbildung 2: Schematische Darstellung einer Fall-Kontroll-Studie

Für eine Fall-Kontroll-Studie werden Fälle und Kontrollen zu vermutlichen Risikofaktoren befragt. In beiden Gruppen berechnet man die Wahrscheinlichkeit, mit diesen Faktoren exponiert gewesen zu sein und setzt sie rechnerisch in ein Verhältnis. Das Ergebnis ist ein Risikoschätzer, die Odds Ratio. Die Odds Ratio und weitere statistische Signifikanzwerte geben Aufschluss über die Wahrscheinlichkeit, erkrankt zu sein, wenn man im Vorfeld einer Exposition ausgesetzt war [vgl. Rothman 2002, S. 73 ff].

Um einen Zusammenhang herzustellen muss die verwendete Kontrollgruppe unter dem selben Risiko einer Infektion gestanden haben wie die infizierten Fälle. Ebenso ist die Struktur innerhalb der Gruppen wichtig. Faktoren wie Geschlecht, Alter und sozialer Status können die Entstehung einer Erkrankung beeinflussen. Solche Ausprägungen sollten daher in beiden Gruppen gleichmäßig verteilt sein. Nur so geben die Studienergebnisse den Unterschied im Verhalten zwischen Gesunden und Infizierten und nicht den Unterschied zwischen der Struktur der zu vergleichenden Gruppen wieder. Systematische Fehler, die aufgrund schlecht zu vergleichender Gruppen entstehen, werden Selektionsfehler oder Selection-Bias genannt [vgl. Rothman 2002, S. 78 ff].

Das Design einer Fall-Kontroll-Studie ist aus mehreren Gründen nicht für die routinemäßige Analyse von Infektionsmeldedaten geeignet. Eine solche epidemiologische Studie ist aufwendig in der Durchführung. Der zeitliche, finanzielle und personelle Einsatz, der für die sachgemäße Durchführung einer Fall-Kontroll-Studie nötig ist [vgl. DAE 2004], schließt eine routinemäßige Durchführung aus.

Ein weiteres Problem einer Fall-Kontroll-Studie, die mit Surveillancedaten durchgeführt wird, ist die Erhebung einer Kontrollgruppe. Die Quote an infizierten Personen, die von einem Meldesystem erfasst werden, liegt weit unter der Inzidenz in der Bevölkerung. Internationale Studien beschreiben z.B., dass nur jeder fünfzigste Salmonellosefall gemeldet wird [vgl. RKI 2004a]. Daher ist davon auszugehen, dass es sich bei den gemeldeten Infizierten um eine Gruppe mit speziellem Verhalten handelt. Das Robert Koch Institut führt folgende Gründe, die die Meldung in ein laufendes Meldesystem beeinflussen, auf:

- „Schwere Erkrankungsbilder führen eher zum Arztbesuch als leichte.
- Asymptomatische Infektionen werden in der Regel nur im Rahmen von Routineuntersuchungen (z.B. betriebsärztliche Untersuchungen) oder Umgebungsuntersuchungen bei Ausbrüchen labordiagnostisch erfasst.
- Berufstätigkeit, Alter, Geschlecht und Sozialstatus wirken sich unterschiedlich auf die Wahrscheinlichkeit aus, einen Arzt aufzusuchen.
- Das Alter eines Patienten wirkt sich auf die Wahrscheinlichkeit aus, mit der eine labordiagnostische Untersuchung veranlasst wird.
- Die Bereitschaft, sich in ärztliche Behandlung zu begeben, unterliegt unter Umständen saisonalen Schwankungen und regionalen Einflüssen.

- Labordiagnostische Untersuchungen werden vom Arzt eventuell deshalb nicht angeordnet, weil befürchtet wird, dass diese das Budget der Arztpraxis belasten. Hierzu ist zu sagen, dass Laboruntersuchungen für meldepflichtige Erreger aus der Budgetierung bewusst herausgenommen wurden und somit das Laborbudget des Arztes nicht belasten. Diese Tatsache scheint vielen Ärzten allerdings nicht bekannt zu sein.“ [RKI 2003, S. 27]

Zudem werden weitere Fälle aus Unwissenheit, fehlender Organisation oder erhöhtem Arbeitsaufwand von Seiten der Labore nicht gemeldet [vgl. RKI 2003, S. 28].

Bei einer analytischen Studie muss die verwendete Kontrollgruppe zu den Fällen passen. Die Erstellung einer Kontrollgruppe für Infizierte, die über ein laufendes Meldesystem erfasst wurden, erweist sich somit als extrem aufwendig. Die gemeldeten Fälle weisen eine spezifische Charakteristik auf, welche sich in den Kontrollen widerspiegeln muss.

1.1.3 Ausbruchsuntersuchung anhand einer Case-Case-Studie

Ein alternatives Studiendesign, die Case-Case-Studie, ist nicht von dem oben beschriebenen Selektionsfehler beeinflusst. In ihr bestehen die Kontrollen³ ebenfalls aus gemeldeten Infizierten (Abbildung 3). In diesem Studiendesign werden Expositionen von Infizierten eines Ausbruchs mit denen weiterer infizierter Personen verglichen. Die Kontroll/Fallgruppe hat denselben Selektionsprozess des Surveillancesystems durchlaufen wie die Fallgruppe. Die in Kapitel 1.1.2 beschriebene Charakteristik, die durch eine Selektion des Meldesystems entsteht, trifft auch für die Personen dieser Kontroll/Fallgruppe zu. Sie passt daher von ihrer Beschaffenheit zu den zu untersuchenden Fällen [vgl. McCarthy & Giesecke 1999].

Weitere systematische Fehler, die in Fall-Kontroll-Studien eine bedeutende Rolle spielen, sind Formen des Erinnerungsfehlers (Recall-Bias). Diese Art eines Informations-Bias entsteht durch eine fehlerhafte Erinnerung bestandener Expositionen. Ein Recall-Bias entsteht durch verschiedene Ursachen. Zum einem kann er durch eine verzerrte Wahrnehmung der Vergangenheit erfolgen. Eine erkrankte Person setzt sich z.B. intensiv mit ihrem Verhalten und den Ereignissen, die die Erkrankung beeinflusst haben könnten, auseinander. Das kann zu einer ausführlicheren oder überzogenen Erinnerung von

³ Im weiteren Text werden die Kontrollen einer Case-Case-Studie als „Kontroll/Fälle“ bezeichnet und ihre Analysegruppe als „Kontroll/Fallgruppe“.

bestandenen Expositionen führen. Werden diese Informationen für eine Studie erfragt, kommt es zu einem unterschiedlichen Antwortverhalten zwischen Fällen und ihren gesunden Kontrollen. Diese unterschiedliche Wahrnehmung verfälscht die Ergebnisse der Studie [vgl. Rothman 2002 S. 98 ff]. Da in einer Case-Case-Studie Fälle und Kontroll/Fälle mit der gleichen Erkrankung infiziert sind, ist diese Fehlerart nicht von Bedeutung. In beiden Gruppen erinnern die infizierten Personen sich unter der gleichen Voraussetzung an ihre Expositionen.

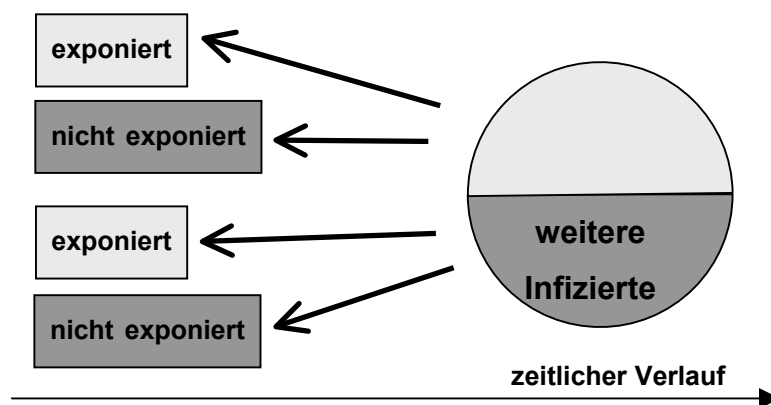


Abbildung 3: Schematische Darstellung einer Case-Case-Studie

Diese Diplomarbeit untersucht, ob eine Case-Case-Studie, durchgeführt mit routinemäßig erhobenen Surveillancedaten, die Quelle eines Infektionsausbruchs identifiziert. Eine geeignete Erkrankung für diese Analyse ist die Salmonellose. Ihre Fälle können aufgrund ihrer Typenvielfalt gut einem Ausbruch zugeordnet werden [vgl. Giesecke 2002, S. 144 ff]. Da es sich um eine meldepflichtige Erkrankung handelt, liegen Ausbruchsdaten von allen gemeldeten Salmonellose-Infizierten in örtlichen Gesundheitsämtern vor. Diese Daten können für eine Analyse in eine Fall und eine Kontroll/Fallgruppe eingeteilt werden und in einer Case-Case-Studie analysiert werden.

1.2 Case-Case-Studien in der epidemiologischen Forschung

Veröffentlichungen von Case-Case-Studien sind vorwiegend in der Krebsepidemiologie zu finden. Mit dem Studiendesign werden unterschiedliche Verläufe einer Krebsart untersucht. In einer Analyse können z.B. Ursachen, die zu verschiedenen Metastasierungs- oder Rezidivierungsprozessen einer Krebsart führen, erforscht werden. Die Fallgruppe besteht hierbei aus Krebserkrankten, die den beschriebenen Krankheitsverlauf aufweisen,

die Kontroll/Fallgruppe aus Erkrankten ohne diese Symptomatik. Unterschiede in den untersuchten Expositionen können Aufschlüsse über einen ursächlichen Zusammenhang geben [vgl. Last 1995 S. 22; Porta et al. 1999; Taylor 1989]. Dieses Studiendesign kann auch aus Fall-Kontroll-Studie mit Restriktion bezüglich der Krebs symptomatik bezeichnet werden. In beiden Gruppen handelt es sich um Krebsfälle, weshalb das Design als Case-Case-Studie bezeichnet wird. Die Zuteilung in die Analysegruppen erfolgt aufgrund des unterschiedlichen Krankheitsverlaufes. Somit gibt es eine klare Differenzierung zwischen Fällen und Kontrollen deren Ursache untersucht wird.

Veröffentlichungen von Case-Case-Studien sind in der Infektionsepidemiologie weniger weit verbreitet. Anwendung findet das Design um Unterschiede zwischen Subtypen eines Keimes zu untersuchen. Einzelne Untergruppen eines Erregers weisen oftmals einen typischen Übertragungsweg auf. Vergleicht man Expositionen von Infizierten verschiedener Subtypen miteinander, kann man diese Unterschiede darstellen. Ein Beispiel hierfür ist eine Studie von Gillespie und O'Brien. In ihr dienten mit *Campylobacter coli* infizierte Personen als Fallgruppe und *Campylobacter jejuni*-Infizierte als Kontroll/Fallgruppe. Die Analyse ergab, dass Infizierte mit *Campylobacter coli* mehr mit Trinkwasser exponiert waren. *Campylobacter jejuni*-Infizierte wiesen einen stärkeren Zusammenhang zu dem Konsum von Pasteten auf [vgl. Gillespie et al. 2002a]. Ähnliche Herangehensweisen sind in weiteren Studien zu finden [vgl. Gillespie et al. 2002b; Kist & Freitag 2000]. In diesen Studien dient das Design ebenfalls zur Identifizierung von Unterschieden zwischen Untergruppen einer Erkrankung.

Für die von McCarthy und Giesecke beschriebene und in dieser Arbeit angewandte Methode zur Analyse von Infektionsausbrüchen [vgl. McCarthy & Giesecke 1999] konnten keine Veröffentlichungen gefunden werden.

1.3 Salmonellose

Zu den Leitsymptomen der Salmonellose gehören Durchfälle, die häufig mit Übelkeit, Erbrechen und Fieber einhergehen. Die Symptome bestehen in der Regel ein bis zwei Tage. Die Inkubationszeit ist abhängig von der Infektionsdosis und liegt zwischen fünf bis 72 Stunden. Die mittlere Inkubationszeit beträgt 20 bis 24 Stunden. Eine Keimausscheidung erfolgt über drei bis sechs Wochen. In seltenen Fällen sind auch Dauerausscheidungen bis zu sechs Monaten möglich. Eine antibakterielle Chemotherapie

ist nur bei immunsuppressiven Infizierten, wie z.B. an Leukämie oder HIV erkrankten Personen, angezeigt [vgl. RKI 2002a].

Eine Ansteckung erfolgt meist über kontaminierte Lebensmittel. Seltener berichtet werden fäkal-orale Übertragungen oder Infektionen über den Kontakt zu ausscheidenden Haustieren. Ein besonderes Risiko geht hierbei von Reptilien aus [RKI 2000]. Bei nahrungsmittelassoziierten Salmonellose-Infektionen steht der Konsum von Fleischwaren, vor allem Rind- und Schweinefleisch, von Geflügel wie Huhn, Ente, Gans und Pute, Fisch und Meeresfrüchten, Milch und Milchwaren sowie von Ei und Eiprodukten im Vordergrund [vgl. Köhler et al. 2001, S. 288 ff; RKI 2002; Pschyrembel 2001, S. 1477]. Vereinzelt wird auch von selteneren Erkrankungsursachen, wie z.B. dem Genuss von Anistee [vgl. RKI 2003] oder Schokolade [vgl. RKI 2002b] berichtet. Eine kurzzeitige Erhitzung der Lebensmittel auf 75 Grad Celsius tötet die Bakterien ab. Es geht also eine erhöhte Gefahr von nicht oder nicht ausreichend erhitzten Nahrungsmitteln aus. Durch unsachgemäßen Umgang mit Lebensmitteln sind Kreuzkontaminationen möglich. Hierbei werden Keime von kontaminierten Produkten auf weitere Nahrungsmittel übertragen, über die nun eine Ansteckung erfolgen kann [vgl. RKI 2002a]. Reisen ins Ausland stehen ebenfalls mit einem erhöhten Salmonellose-Risiko in Zusammenhang. Im Jahr 2003 wurde bei 10% der gemeldeten Salmonellosen ein ausländisches Infektionsland angegeben. Die typischen Urlaubsländer der Deutschen - Türkei, Spanien, Griechenland und Italien - waren hier mit je einem Prozent angegeben [RKI 2004c, S.138].

1.3.1 Mikrobiologie der Salmonellose

Die Salmonellose gehört zu der Familie der Enterobacteriaceae. Bei den *Salmonella*-Keimen handelt es sich um gramnegative Stäbchen. Die Gattung wird in die beiden Arten *Salmonella enterica* und *Salmonella bongori* eingeteilt. Infektionen des Menschen werden durch den *S. enterica* Stamm verursacht. Er splittet sich in sechs weitere Subspezies, die mit Namen oder Nummern bezeichnet werden, auf. Hierbei handelt es sich um die Subspezies *enterica* (I), Subspezies *salamae* (II), Subspezies *arizonae* (IIIa), Subspezies *diarizonae* (IIIb), Subspezies *houtenae* (IV) und Subspezies *indica* (V). Von humanpathologischer Bedeutung sind die Erreger der Subspezies *enterica*. Die Subspezies *salamae*, *arizonae* und *diarizonae* werden nur sporadisch bei Durchfallpatienten isoliert. Die Subspezies *houtenae* und Subspezies *indica* sind von veterinärmedizinischer Bedeutung. Die Subspezies lassen sich weiter in einzelne Serovare einteilen. Die korrekte

Bezeichnung für das *Salmonella*-Isolat Enteritidis lautet somit wie folgt: *Salmonella enterica*, Subspezies *enterica*, Serovar Enteritidis. Eingebürgert hat sich die Kurzform *Salmonella* Enteritidis oder einfach *S. enteritidis*. Zu der Subspezies *enterica* gehören auch die Serovare *S. thypi* und *S. paratyphi*, die Erreger des Typhus und Paratyphus. Hierbei handelt es sich um eine systematische Erkrankung mit Darmbeteiligung [vgl. Köhler et al. 2001; S. 288]. Auf sie wird in dieser Arbeit nicht weiter eingegangen.

Die Einteilung in die sechs *Salmonella enteritica*-Subspezies wird anhand der spezifischen metabolischen Leistung der Erreger vorgenommen. Durch eine genaue Serotypisierung der Keime erfolgt eine Zuordnung, nach dem Kaufmann-White-Schema, in die einzelnen Serovare. Hierfür werden die Körperantigene (O-Antigene) und die Geißelantigene (H-Antigene) des isolierten Keims ermittelt. Über die O-Antigene erfolgt eine Einteilung in neun Gruppen. Sie besitzen die alphabetische Bezeichnung Gruppe A bis I. Für eine weitere Identifizierung bestimmt man die H-Antigene und erhält die genaue Bestimmung des Serovares [vgl. Köhler et al. 2001; S. 288].

Tabelle 1: Verteilung der 10 häufigsten Serovare von *S. enterica* der übermittelten Salmonellosefälle in Deutschland 2003 (n=63.044)⁴

Serovar	Anzahl	Prozent
<i>S. enteritidis</i>	42.180	66,9
<i>S. typhimurium</i>	12.169	19,3
<i>S. infantis</i>	535	0,8
<i>S. virchow</i>	241	0,4
<i>S. derby</i>	191	0,3
<i>S. hadar</i>	181	0,3
<i>S. bovismorbificans</i>	152	0,2
<i>S. antum</i>	149	0,2
<i>S. agona</i>	110	0,2
<i>S. goldcoast</i>	92	0,1

Quelle: RKI. Bakterielle Gastroenteritiden - Situationsbericht 2003. Epidemiologisches Bulletin 2004; 31: 252.

Bisher sind über 2.400 verschiedene *Salmonella*-Serovare bekannt [vgl. RKI 2002a]. In Deutschland waren *S. enteritidis* und *S. typhimurium* die mit Abstand am häufigsten gemeldeten Salmonellose-Serovare im Jahr 2003 [vgl. RKI 2004b]. Eine Auflistung der

⁴ Nach Rücksprache mit dem Robert Koch Institut handelt es sich bei der im Epidemiologischen Bulletin veröffentlichten Zahl von 60.303 Infizierten um einen Druckfehler.

zehn meistgemeldeten Serovare in Deutschland aus dem Jahr 2003 befindet sich in der Tabelle 1.

Neue genetische Typisierungsverfahren gewinnen langsam an Bedeutung. Mit der Puls-Feld-Gel-Elektrophorese (PFGE) ist man in der Lage, den genetischen „Fingerabdruck“ eines Erregers zu identifizieren. Eine solche Bestimmung ermöglicht eine genauere Zuordnung des isolierten Keimes. Dieses erscheint besonders sinnvoll, da ein Großteil der Infektionen durch die beiden Erreger *S. enteritidis* und *S. typhimurium* verursacht werden. Mit der PFGE kann man auch innerhalb dieser Serovare unterscheiden und somit Infizierte eindeutig einem Herd zuordnen. Angewandt wird diese Methode zur Ausbruchbestimmung bereits in einigen Staaten der USA. Das „Center for Disease Control and Prevention“ hat das Netzwerk „PulseNet“ ins Leben gerufen. Durch eine staatenübergreifende, standardisierte Feintypisierung wird somit eine wirksame Surveillance von nahrungsmittelbedingten Infektionskrankheiten ermöglicht [vgl. Swaminathan et al. 2001]. In Deutschland wird nach diesem Vorbild das „German PFGE-Netzwerk“ aufgebaut. Es entsteht unter der Leitung des Salmonellose-Überwachungszentrums in Weringerode [vgl. Tschäpe 2003].

1.3.2 Epidemiologie der Salmonellose

Zur Zeit ist in Deutschland die Salmonellose die häufigste meldepflichtige Erkrankung. Im Jahr 2003 wurden 63.044 Fälle gemeldet. Dies ergibt eine Inzidenz von 76,4 auf 100.000 Personen. Die Salmonellose ist seit Anfang der neunziger Jahre in ihren Meldezahlen rückläufig. Allein im Jahr 2003 reduzierten sich die Meldungen im Vergleich zum Vorjahr um 13 %.

Die epidemische Kurve der Salmonellose zeigt einen saisonalen Jahresverlauf. Infektionen mit dem Erreger häufen sich in den wärmeren Monaten. Ein stetiger Anstieg der Fälle verzeichnet sich im Frühjahr. Die meisten Infektionen (Peak) ereignen sich im August und September eines Jahres. Zu den Wintermonaten nehmen die Meldungen wieder ab.

Geschlechterspezifische Unterschiede sind bei der Salmonellose nicht zu beobachten. Die höchste altersspezifische Inzidenz findet sich bei Säuglingen und Kleinkindern unter zehn Jahren. In 2003 wurden die meisten Infektionen mit 450 Fällen auf 100.000 in der Gruppe der Einjährigen gemeldet. Mit steigendem Alter fielen die Meldezahlen stetig ab.

In den Altersstufen über 15 Jahren bewegte sich die Inzidenz um ca. 50 Meldungen auf 100.000.

In der geografischen Verteilung weisen alle östlichen Bundesländer eine höhere Inzidenz auf. Die meisten Salmonellosen wurden in Thüringen gemeldet, die wenigsten in Bremen. Es ist nicht geklärt, ob hier eine real höhere Inzidenz oder ein unterschiedliches Meldeverhalten zwischen den Bundesländern besteht [vgl. RKI 2004c, S. 137 ff].

2 Material & Methode

Für die Studie wurde das im Kapitel 1.1.3 beschriebene Design einer Case-Case-Studie verwendet. Um die Ergebnisse zu bewerten wurden sie mit denen einer Fall-Kontroll-Studie verglichen, die mit dem gleichen Herd durchgeführt wurde.

2.1 Epidemiologische Analysemethoden der Studie

Für die Case-Case-Studie wurden die Analyseschritte einer Fall-Kontroll-Studie durchgeführt. Im folgenden Abschnitt werden diese Analyseschritte und die Berechnung ihrer Risikoschätzer dargestellt.

2.1.1 Bivariate Analyse

Als epidemiologische Risikoschätzer der bivariaten Analyse wurden die Odds Ratio (OR), ihr Konfidenzintervall (CI) und der dazugehörige p-Wert berechnet.

Um die Odds Ratio zu berechnen, listet man die Häufigkeit einer Variablen für Fälle und Kontrollen auf. Daraus wird der Quotient aus exponiert und nicht-exponiert zu einer Variablen in der Fallgruppe, ins Verhältnis zu exponiert und nicht-exponiert in der Kontrollgruppe gesetzt. Der berechnete Risikoschätzer gibt Aufschluss über die Wahrscheinlichkeit erkrankt zu sein, wenn eine bestimmte Exposition bestanden hat, verglichen mit nicht erkrankten Personen [vgl. Rothman 1986 S. 62 ff]. In einer Case-Case-Studie wird das Verhältnis zwischen dem infizierten Herd und den Infizierten der Kontroll/Fallgruppe ermittelt. So gibt die OR den Vergleich zwischen dem Ausbruchsherd und weiteren infizierten Personen wieder.

Liegt die OR bei eins, ist keine Verbindung zwischen der Erkrankung und der Variablen nachzuweisen. Somit ist die Nullhypothese, die besagt, dass kein Zusammenhang besteht, bestätigt. Ein Wert über eins gibt die Stärke des Risikos und ein Wert unter eins die Stärke eines Schutzfaktors wieder. Die OR, die in der bivariaten Analyse berechnet wird, heißt roher Effektschätzer oder Crude OR [vgl. Rothman 1986, S. 62 ff].

Bei der OR handelt es sich um einen Punktwert. Um den Einfluss des Zufalls darzustellen bedarf es zusätzlicher Größen, die eine weitere Interpretation ermöglichen. Eine kleine Studienpopulation unterliegt z.B. stärkeren Verfälschungen durch den Zufall als eine große Analysegruppe. Diesen Einfluss stellt das Konfidenzintervall dar. Es wird anhand der Standardabweichung des OR berechnet. Das CI gibt die Streuung von OR wieder, die

entstehen würde, wenn die Studie unter gleichen Bedingungen wiederholt werden könnte. Herkömmlich wird ein 95% Konfidenzintervall berechnet. Es besteht aus zwei Werten, zwischen denen 95% der Ergebnisse liegen, würde diese Studie mehrmals wiederholt. Je enger die Werte beieinander liegen, desto präziser ist das Ergebnis der berechneten OR. Um einen Zusammenhang zu bestätigen, darf dieser Wertebereich die eins nicht mit einschließen. So kann eine Bestätigung der Nullhypothese ausgeschlossen werden [vgl. Rothman 1986, S. 119 ff].

Der p-Wert⁵ ist ein weiteres Maß, um die statistische Aussagekraft des Ergebnisses zu beurteilen. Der p-Wert gibt die Wahrscheinlichkeit an, mit der die berechneten Daten zufallsbedingt sind und somit die Nullhypothese bestätigt werden kann. Gewöhnlich wird der p-Wert als signifikant interpretiert, wenn er unter 0,05 liegt. Bei diesem Wert geht man mit einer fünfprozentigen Wahrscheinlichkeit davon aus, dass für die analysierte Variable ein Risikoschätzer berechnet werden kann, der die Nullhypothese bestätigt. Ein p-Wert von 1,00 würde eine 100-prozentige Wahrscheinlichkeit beschreiben, dass die berechnete OR durch den Zufall bedingt ist [vgl. Rothman 1986, S.115 ff]. Herkömmlich wird der p-Wert über den Chi-Quadrat-Test ermittelt. Bei kleineren Stichproben erfolgt die Berechnung über den Fischer-Exakt-Test [Rothman 1986, S. 155 ff].

2.1.2 Stratifikation

In der multivariaten Analyse kontrolliert man nach möglichen Störgrößen. Hierbei handelt es sich um Confounder oder Effektmodifikatoren. Beobachtet man, dass der Faktor A die Krankheit B verursacht, kann eine weitere Risikogröße C ihren Einfluss ausüben. C kann ein Confounder sein, der im Zusammenhang mit dem Faktor A steht und in Wirklichkeit die Erkrankung bedingt. Der Faktor A gibt somit stellvertretend den Einfluss des Confounders C wieder. Ein Confounder steht nicht in einem ursächlichen Zusammenhang mit dem betrachteten Risikofaktor. Er ist somit keine Folge von A [vgl. Rothman 1986, S. 107 ff]. Ein Beispiel für einen Confounder ist der scheinbare Zusammenhang zwischen der Geburtenfolge eines Kindes und dem Auftreten einer Trisomie 21. Eine Studie zeigte, dass sich die Wahrscheinlichkeit, ein Kind mit Trisomie 21 zur Welt zu bringen, mit jeder Geburt erhöhte. In Wirklichkeit bestand der Zusammenhang zu Faktoren, die durch das Alter der schwangeren Frau bedingt waren. Die

⁵ „p“ steht für den englischen Begriff „probability“.

ermittelte OR maß also nicht die Geburtenfolge des Kindes, sondern das Alter der Frauen während der Schwangerschaft, welches sich mit jeder Geburt erhöhte. Das Alter ist in diesem Fall ein Confounder [vgl. Rothman 2002, S. 101 ff].

Risikofaktoren, die eine Erkrankung nicht bedingen, aber das Erkrankungsrisiko positiv oder negativ beeinflussen, werden als Effektmodifikatoren bezeichnet. Hierbei handelt es sich nicht um einen Einfluss, den man nicht beseitigen möchte, sondern den es zu beschreiben gilt [vgl. Rothman 1986, S. 177 ff].

Um die Auswirkungen eines Confounders zu beseitigen oder die Stärke einer Effektmodifikation zu messen, bedient man sich der Stratifikation. In dieser Analyse werden die Risikoschätzer innerhalb der Ausprägungen einer dritten Variable ausgewertet. Die Berechnungen erfolgen z.B. getrennt für Männer und Frauen oder Raucher und Nichtraucher. Dadurch erhält man Werte, die um den Einfluss dieser dritten Variablen bereinigt wurden. Diese Einteilungen in Untergruppen einer Ausprägung werden als Straten bezeichnet. Möchte man, wie im obigen Beispiel untersuchen, ob das Alter einer Mutter bei ihrer Schwangerschaft den Zusammenhang zwischen Geburtsfolge des Kindes und der Entstehung einer Trisomie 21 beeinflusst, berechnet man Risikoschätzer für die einzelnen Altersgruppen. Beobachtet man in den Straten der verschiedenen Altersgruppen dasselbe Risiko wie in der unstratifizierten Analyse, scheint keine Beeinflussung durch diese dritte Variable vorzuliegen. Unabhängig vom Alter besteht das gleiche Risiko ein Kind mit Trisomie 21 zur Welt zu bringen. Berechnet man, im Gegensatz zur bivariaten Analyse, innerhalb der Straten kein oder nur ein geringes Risiko, scheint es sich um einen Confounder zu handeln. Der Einfluss des Alters hat innerhalb der Straten nur einen geringen Effekt. Somit ist nur ein geringes oder kein Risiko zu messen. Weichen die berechneten OR gegensätzlich von der Crude OR ab, liegt eine Effektmodifikation vor. Die dritte Variable besitzt verstärkende oder abschwächende Auswirkungen auf das Ergebnis [vgl. Rothman 1986, S. 177 ff].

Die Epidemiologen Mantel und Haenszel haben eine Formelsammlung zur Berechnung von gewichteten, durchschnittlichen Risikoschätzern für eine stratifizierte Analyse erstellt. Hierbei werden die Risikoschätzer einer stratifizierten Variablen bezüglich der Größenverhältnisse in den einzelnen Untergruppen gewichtet und daraus der Mittelwert gebildet. So erhält man Werte, die um den Einfluss einer Störgröße bereinigt sind. Wie stark der Einfluss einer dritten Variablen herausgerechnet werden kann, hängt von der Einteilung der Straten ab. Je enger die Grenzen eines Stratums sind, desto geringer ist der Einfluss eines Confounders. Eine Stratifikation nach Alter für die beiden Gruppen unter 40

Jahre und über 40 Jahre wird weiterhin eine deutliche Verzerrung durch die Störvariable aufweisen. Feinere Einteilungen der Straten sind oftmals aufgrund der zu erwartenden kleinen Fallzahlen nicht möglich [vgl. Rothman 1986, S. 195 ff].

2.1.3 Logistische Regression

Eine Methode, um den Einfluss mehrerer Variablen gleichzeitig zu analysieren, ist das mathematische Modell der Logistischen Regression (LR). In einer stratifizierten Auswertung ist man selten in der Lage, nach mehreren Confoundern gleichzeitig zu untersuchen. Die einzelnen Straten müssten hierfür wieder in Untergruppen eingeteilt werden und dadurch ergeben sich sehr kleine Fallzahlen. Mit der Logistischen Regression kann der Einfluss von mehreren unabhängigen Faktoren auf eine abhängige Variable berechnet werden. Die Risikoschätzer werden in Abhängigkeit von den weiteren Variablen, die im LR-Modell enthalten sind, ermittelt. Hierfür sind verschiedene Herangehensweisen beschrieben. In der vorliegenden Analyse wurde das „stepwise backward Modell“ angewandt. Expositionen, die einen p-Wert von unter 0,1 besitzen oder die unter Verdacht stehen, einen Einfluss auf das Ergebnis auszuüben, werden in das LR-Modell aufgenommen. Variablen, die im LR-Modell eine niedrige Signifikanz aufweisen, werden schrittweise entfernt. Kommt es zu einer merkbaren Veränderung in den Risikoschätzern der abhängigen Variablen, scheint es sich um eine Exposition, die Auswirkung auf das Ergebnis besitzt, zu handeln. Im letztendlichen LR-Modell bleiben die sich beeinflussenden Faktoren stehen. In ihm wird die Stärke der Beziehung, die sie untereinander besitzen, dargestellt. So erhält man einen Überblick über die Variablen, die sich auf das Ergebnis auswirken [Rothman 2002, S.189 ff].

Die Logistische Regression ist kein Ersatz für eine Stratifikation, sondern eine weitere Methode, um nach Einflussgrößen zu suchen und ihre Beziehungen untereinander darzustellen [vgl. Rothman 1986, S. 285 ff]. Bei der LR handelt es sich um eine komplexe Berechnung, bei der man schnell die Übersicht verliert. Kenneth Rothman schreibt dazu: „In a multivariable analysis, the reader is typically in the dark and the researcher may not be much better off than the reader. For this reason, a multivariable analysis should be used only as a supplement to a stratified analysis, rather than as the primary analytic tool.” [Rothman 2002, S. 192].

Die Multivariate Analyse ist nur in der Lage, nach Einflussgrößen zu analysieren, die in einer Studie erfragt wurden. Es ist davon auszugehen, dass jede Studie durch Störgrößen beeinflusst wird, die nicht dargestellt werden können.

2.2 Datensammlung

Die für die Analyse verwendeten Daten sind Meldedaten zur Salmonellose aus dem Jahr 2003 des Bezirkes Hamburg-Altona. Der Verwaltungsbezirk Hamburg-Altona umfasst die Stadtteile Altona-Altstadt, Altona-Nord, Ottensen, Bahrenfeld, Groß-Flottbek, Othmarschen, Lurup, Osdorf, Nienstedten, Blankenese, Iserbrook, Sülldorf und Rissen. Ende 2003 waren hier insgesamt 245.474 Personen gemeldet [vgl. Finanzbehörde Hamburg 2004, S. 8].

Im Fall einer Salmonelloseinfektion in den oben genannten Stadtteilen erfolgt seitens der untersuchenden Labore eine Meldung an das Gesundheitsamt Altona. Diese umfasst Angaben zur Person und in der Regel eine Spezifizierung des Erregers. Medizinalaufseher des Amtes nehmen zu der gemeldeten Person Kontakt auf und befragen sie bezüglich ihrer Exposition zu bekannten Risikofaktoren. Diese und weitere Informationen zur Person und zum Krankheitsverlauf werden in ein Formular eingetragen, welches vor Ort archiviert wird. Hieraus wurde die Analysedatei erstellt.

In EpiInfo 6.04d [vgl. CDC 2004] wurde eine Datenmaske angefertigt, über die die benötigten Informationen jedes Falles eingegeben werden konnten.⁶ Hierbei handelt es sich um Angaben zur Person, zur Bezeichnung des Salmonellose-Serovares, ggf. eine Zuordnung zu einem Infektionsherd und um Abfragen von verschiedenen Expositionen, insbesondere Lebensmittelexpositionen.

2.2.1 Fallgruppe

Der Fallherd besteht aus elf Personen, die sich mit einer monozyklischen Variante des *Salmonella* Typhimurium infizierten.⁷ Ihre Datensätze stammen aus einer

⁶ Die Programmierung der Datenmaske befindet sich im Anhang als Anlage 1.

⁷ In den Veröffentlichungen des Hamburger Infekt Infos wird der Keim mit seiner Seroformel „*Salmonella* 1,4,5,12:i:-, Subspezies I“ bezeichnet [Institut für Hygiene und Umwelt 2003a; Institut für Hygiene und Umwelt 2003b]. Nach Rücksprache mit dem Hamburger Hygieneinstitut lautet seine Serovarbezeichnung „monozyklische Variante des *Salmonella* Typhimurium“. Er ist antigenetisch eng mit *Salmonella* Typhimurium verwandt. Die H2-Antigenphase ist bei diesem Typ allerdings nicht ausgebildet.

Nachuntersuchung, die mit einer Fall-Kontroll-Studie durchgeführt wurde [vgl. Institut für Hygiene und Umwelt 2003]. Routinemäßig erhobene Surveillancedaten konnten für die Fallgruppe nicht verwendet werden, da erkannte Ausbrüche oft zielgerichtet befragt wurden. Bestand bereits eine verdächtige Exposition, wurde sie im Interview erfragt. Ließ sie sich bestätigen, erfolgte keine weitere Erhebung.

Die Nachuntersuchung des Herdes erfolgte standardisiert über einen dreizehnseitigen Fragebogen. Er diente als Vorlage für die Eingabemaske der Datei. Konnten Fragen mit „ja“ oder „nein“ beantwortet werden⁸, gab es die Abstufungsmöglichkeiten „sicher ja“, „wahrscheinlich ja“, „sicher nein“, „wahrscheinlich nein“ und „weiß nicht“. „Sicher ja“ und „wahrscheinlich ja“ wurde bei der Datenanalyse als „ja“ gewertet, „sicher nein“ und „wahrscheinlich nein“ als „nein“. Die Antwort „weiß nicht“ wurde als „keine Angabe“ behandelt. Es wurden die Expositionen der letzten drei Tage vor der Infektion erfragt. In einigen Kategorien gab es zusätzlich die Möglichkeit zur offenen Nennung.⁹ Diese Angaben wurden für die Analyse nicht verwendet.

Die Interviews wurden telefonisch sechs Wochen nach Meldung des letzten Falls durchgeführt. Dadurch lagen zwischen der Meldung einer Infektion und dem Interview zwischen sechs bis zwölf Wochen.

Von den elf gemeldeten Infizierten lag für acht Personen ein Fragebogen der Nachuntersuchung vor. Drei von diesen Personen teilten sich einen Fragebogen. Hierbei handelte es sich um eine Mutter mit ihren beiden Kindern. Die erfolgten Angaben wurden für die einzelnen Fälle eingegeben. Für zwei weitere Infizierte lag kein Interview der Nachuntersuchung vor. Bei ihnen wurden die Daten aus der routinemäßig erhobenen Surveillance benutzt. Von einer Person waren keine Daten vorhanden. Die Angaben der zehn Personen sind für die Fallgruppe verwendet worden.

2.2.2 Kontroll/Fallgruppe

Für weitere 215 Salmonelloseinfizierte aus dem Jahr 2003 lagen routinemäßige Surveillancedaten vor. Die gemeldeten Fälle wurden durch Medizinalaufseher des Gesundheitsamtes telefonisch befragt. Das Interview erfolgte anhand eines vierseitigen Fragebogens. Er enthielt Fragen zur Person und zum Krankheitsverlauf sowie zu bestehenden Risikofaktoren. Die Erhebung der Nahrungsmittelexpositionen erfolgte über

⁸ z.B.: Haben Sie in den letzten drei Tagen von der Infektion Rindfleisch verzehrt?

⁹ z.B.: Wo haben Sie das konsumierte Rindfleisch gekauft?

offene Fragen.¹⁰ Standardmäßig wurde der Konsum von Rindfleisch, Schweinefleisch, Fisch & Meeresfrüchten, Milch & Milchprodukten, sowie Ei & Eiprodukten erfragt. Bestätigte die infizierte Person eine Exposition, wurde diese aufgenommen. Kein Eintrag erfolgte, wenn das Nahrungsmittel nicht konsumiert wurde. Für die Analyse wurden die aufgeführten Nahrungsmittel mit ja bewertet. Eine nicht aufgelistete Exposition wurde als nicht konsumiert und mit nein gewertet. In den Gesprächen wurden die Expositionen der letzten drei Tage von der Erkrankung erfragt. Die Interviews erfolgten in der Regel zehn Tage nach Auftreten der Infektion.¹¹

Von den 215 gemeldeten Personen konnten sieben nicht erreicht werden, was einem Anteil von 3% entspricht. Für sie lagen keine Fragebögen vor. Für 37 Personen (17%) erfolgte keine genauere Befragung, da sie bereits einem Herd zugeordnet wurden, für den der Infektionsgrund bekannt war. 56 der befragten Personen (26%) hatten keine Angaben zu Nahrungsmittelexpositionen gemacht. Von weiteren 18 Infizierten (8%) waren die Antworten nicht eindeutig und daher nicht zu verwenden.

Die Daten der verbleibenden 97 gemeldeten infizierten Personen (45%) dienten als Kontroll/Fallgruppe. Diese Gruppe wurde verwendet, da sie zeitnah um den Fallherd lag, aus demselben Meldebezirk stammte und eine ausreichende Größe für eine statistische Signifikanz besaß.

2.3 Datenanalyse

Die bivariate und stratifizierte Analyse erfolgte mit dem Programm EpiInfo 6.04d [vgl. CDC 2004]. Hierfür wurde eine Datei mit der Kodierung ja=1 und nein=2 erstellt. Die Analysesyntax befindet sich im Anhang (Anlage 2).

Die stratifizierte Analyse erfolgte für die Variablen Geschlecht und Alter. Die infizierten Personen sind in die Altersgruppen Kinder & Jugendliche und Erwachsene eingeteilt worden. In die erste Kategorie wurden alle Infizierten bis einschließlich 18 Jahre zusammengefasst. Die Gruppe der Erwachsenen umfasst alle Infizierten über 18 Jahre. Für die stratifizierten Variablen wurden Mantel-Haenszel Odds Ratios (MH-OR) berechnet.

Die Logistische Regression wurde mit Logistic 3.11Ef [vgl. Dallal 1988] durchgeführt. Hierfür wurde eine Datei mit 1/0 Kodierung erstellt. In dieser Datei gilt ja=1 und nein=0.

¹⁰ z.B.: Welche Nahrungsmittel haben Sie in den letzten drei Tagen vor der Infektion verzehrt?

¹¹ Informationen aus Gesprächen mit Medizinalaufsehern im Gesundheitsamt Hamburg-Altona im April 2004.

Die Geschlechterausprägung wurde als weiblich=0 und männlich=1 abgespeichert. Die Syntax zur Erstellung der Datei befindet sich im Anhang (Anlage 3).

3 Ergebnisse

3.1 Deskriptive Analyse des Datensatzes

Die zehn Erkrankungen der Fallgruppe wurden durch eine monozyklische Variante der *Salmonella* Typhimurium verursacht.

In der Kontroll/Fallgruppe infizierten sich die meisten Personen mit *S. enteritidis*. 78,4% der Erkrankungen sind diesem Serovar zuzuordnen. Das zweithäufigste Serovar, mit einem Anteil von 5,2%, war *S. typhimurium*. Bei zwölf Erkrankten (12,4%) wurde keine genaue Typisierung nach dem Kaufmann-White-Schema vorgenommen. Hier erfolgte eine Einteilung über die Bestimmung der O-Antigene in die Gruppen A bis I oder eine Zuordnung in die entsprechende Subspeziesgruppe (vgl. Kapitel 1.3.1). Eine Auflistung der *Salmonella*-Serovare aus der Kontroll/Fallgruppe befindet sich in Tabelle 2.

Tabelle 2: Verteilung der *Salmonella*-Serovare in der Kontroll/Fallgruppe, gemeldete Infektionen aus Hamburg-Altona, 2003

Serovar	Anzahl	Prozent
<i>S. enteritidis</i>	76	78,4%
<i>S. typhimurium</i>	5	5,2%
<i>S. infantis</i>	2	2,1%
<i>S. stanly</i>	1	1,0%
<i>S. virchow</i>	1	1,0%
<i>S. Gruppe B</i>	4	4,1%
<i>S. Gruppe D</i>	3	3,1%
<i>S. Gruppe C</i>	2	2,1%
<i>S. Subspezies I</i>	3	3,1%

Die Geschlechterverteilung in beiden Gruppen ist nahezu identisch. In der Fallgruppe wurden 60 % und in der Kontroll/Fallgruppe 60,8% Frauen gemeldet.

In der Fallgruppe betrug das Durchschnittsalter 28,3 Jahre, mit einer Standardabweichung (SD) von 27,0. Der Median lag bei 18 Jahren. Die jüngste gemeldete Person war sieben Jahre alt, die älteste Person hatte ein Alter von 84 Jahren. In der Kontroll/Fallgruppe lag der Mittelwert für das Alter bei 34,7 Jahren (SD 23,8). Der Median betrug 32 Jahre. Die jüngste gemeldete Person hatte das erste Lebensjahr noch nicht vollendet und die älteste gemeldete Person war 91 Jahre alt. In der Fallgruppe sind somit jüngere Personen etwas häufiger gemeldet worden.

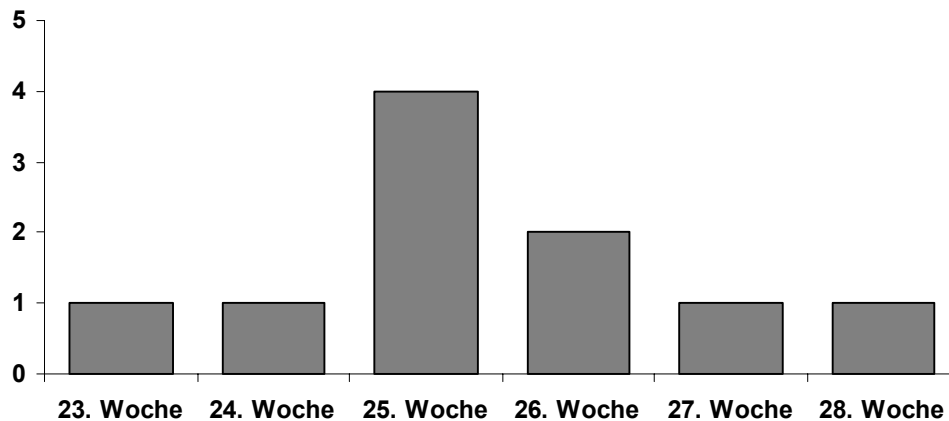


Abbildung 4: Epidemische Kurve der Fallgruppe in Wochen zusammengefasst, gemeldete Salmonellen aus Hamburg-Altona, 2003

Die Salmonelloseinfektionen der Fallgruppe ereigneten sich zwischen Anfang Juni bis Mitte Juli 2003, über einen Zeitraum von sechs Wochen. Ein Peak mit vier Fällen bestand in der 25sten Kalenderwoche. In der Folgewoche wurden zwei Infektionen verzeichnet. In den übrigen Wochen erfolgte jeweils eine Meldung (Abbildung 4).

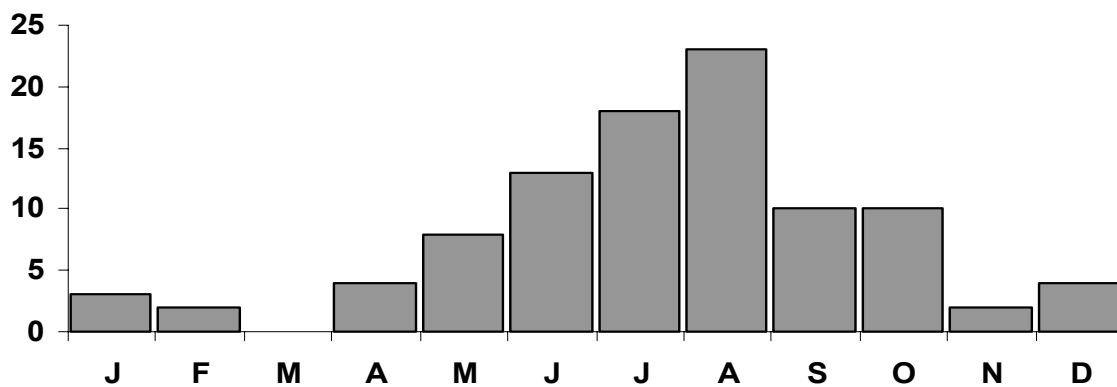


Abbildung 5: Epidemische Kurve der Kontroll/Fallgruppe in Monaten zusammengefasst, gemeldete Salmonellen aus Hamburg-Altona, 2003

Die epidemische Jahreskurve der Kontroll/Fallgruppe zeigt einen saisonalen Verlauf. Die Meldungen zwischen Januar und April lagen auf einem Niveau von null bis vier Fällen. Es folgt eine stetige Steigerung der Meldezahlen von April mit vier Infektionen, bis August mit einem Peak 23 Infektionen. Im September und Oktober gingen die Meldungen auf jeweils zehn Salmonellosen zurück. Die Kurve flacht zum Ende des Jahres weiter ab (Abbildung 5).

Aus datenschutzrechtlichen Gründen geht diese Arbeit auf die geografische Verteilung der Infizierten nicht weiter ein.

3.2 Ergebnisse der epidemiologischen Analyse

3.2.1 Bivariate Analyse

Für die Variablen Rindfleisch und Schweinefleisch wurden signifikante Odds Ratios berechnet. Rindfleisch zeigt, mit einer OR von 11,81, den stärksten Zusammenhang. Das Konfidenzintervall von 2,35 bis 66,66 ist klar im positiven Bereich. Schweinefleisch weist mit einer OR von 8,44 einen geringeren aber ebenfalls deutlichen Zusammenhang auf. Das CI (CI 1,72 - 46,39) bestätigt das Ergebnis. Beide Variablen haben einen stark signifikanten p-Wert.

Tabelle 3: Ergebnistabelle bivariate Analyse, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003

Exposition	OR	CI	p-Wert
Rindfleisch	11,81	2,35 - 66,66	<0,001
Schweinefleisch	8,44	1,72 - 46,39	<0,001
Geflügel	2,17	0,30 - 11,03	0,287
Fisch	2,57	0,54 - 11,88	0,162
Ei	3,33	0,70 - 17,65	0,081
Milch	3,62	0,80 - 16,54	0,047

Für die Exposition mit Milch & Milchprodukten ergab sich ein signifikanter p-Wert von 0,047, das CI kann diesen Zusammenhang allerdings nicht bestätigen. Für die weiteren berechneten Variablen Geflügel, Fisch & Meeresfrüchte und Ei & Eiprodukte¹² wurden keine signifikanten Risikoschätzer berechnet (Tabelle 3).

3.2.2 Stratifizierte Analyse nach Geschlecht

Bei der stratifizierten Analyse kam es in den Fallgruppen zu kleinen Zahlen. Das Stratum der Frauen umfasst sechs und der Männer vier Personen. Im Stratum der

¹² In den Tabellen sind die Variablen Fisch & Meeresfrüchte zu Fisch, Ei & Eiprodukte zu Ei und Milch & Milchprodukte zu Milch zusammengefasst.

Kontroll/Fallgruppe befinden sich 59 Frauen und 38 Männer. Aufgrund der kleinen Fallzahlen wurde der p-Wert über den Fischer-Exakt-Test ermittelt.

Betrachtet man die nach Geschlecht stratifizierte Analyse für die Variable Rindfleisch, liegen die OR in den Stratern auseinander. Die OR für Frauen beträgt 8,73 und die für Männer 19,80. Die MH-OR (MH-OR 11,61) und die Crude OR (OR 11,81) liegen nah beieinander. Alle Risikoschätzer sind signifikant.

Tabelle 4: Ergebnistabelle stratifizierte Analyse nach Geschlecht, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003

Exposition	Frauen			Männer			Mantel-Haenszel		
	OR	CI	p-Wert	OR	CI	p-Wert	OR	CI	p-Wert
Rindfleisch	8,73	1,13-81,80	0,022	19,80	1,31-629,80	0,018	11,61	2,72-49,51	<0,001
Schweinefleisch	19,58	1,85-496,69	0,004	3,22	0,26-40,01	0,277	8,15	1,96-33,92	0,001
Geflügel	2,18	0,24-17,23	0,591	2,20	<0,01-35,86	0,474	2,19	0,51-9,43	0,288
Fisch	9,80	1,25-93,15	0,017	0,00	<0,01-5,43	0,557	2,39	0,65-8,82	0,166
Ei	4,83	0,48-118,24	0,205	2,80	0,23-34,21	0,565	3,80	0,84-17,11	0,072
Milch	3,92	0,53-29,48	0,130	3,22	0,26-40,01	0,277	3,61	0,95-13,67	0,049

Tabelle 5: Ergebnistabelle stratifizierte Analyse nach Alter, Case-Case-Studie eines Salmonelloseausbruchs in Hamburg-Altona, 2003

Exposition	Kinder & Jugendliche			Erwachsene			Mantel-Haenszel		
	OR	CI	p-Wert	OR	CI	p-Wert	OR	CI	p-Wert
Rindfleisch	17,25	1,06-587,43	0,028	9,67	1,27-89,09	0,016	11,87	2,77-50,90	<0,001
Schweinefleisch	24,00	1,35-903,75	0,016	5,78	0,79-51,14	0,055	8,87	2,10-37,51	<0,001
Geflügel	0,00	<0,01-14,23	1,000	4,83	0,66-36,15	0,087	2,31	0,53-10,01	0,253
Fisch	8,00	0,51-151,74	0,112	1,56	0,18-11,52	0,636	2,70	0,70-10,46	0,135
Ei	6,00	0,43-177,67	0,272	2,52	0,35-21,70	0,405	3,43	0,84-14,02	0,073
Milch	5,10	0,37-149,77	0,284	2,68	0,29-21,04	0,271	3,55	0,86-14,67	0,076

Für die Exposition mit Schweinefleisch wurden zwischen den Geschlechtern ebenfalls deutlich unterschiedliche Odds Ratios berechnet. Die OR für Frauen liegt signifikant bei 19,58. Für Männer beträgt sie 3,22 und weist keine Signifikanz auf. Die MH-OR beträgt 8,15 und weicht somit von der Crude OR (OR 8,44) kaum ab.

Im Stratum der Männer erfolgte für die Exposition mit Fisch & Meeresfrüchten keine Nennung. Somit konnten keine Risikoschätzer in dieser Gruppe berechnet werden. Die MH-OR ist in der Lage mit Nullen zu rechnen und kann interpretiert werden. Die OR für Frauen liegt signifikant bei 9,80. Crude OR (OR 2,57) und MH-OR (MH-OR 2,39) sind unsignifikant und weichen kaum voneinander ab.

Die Stratifikation für die weiteren Variablen zeigt keine auffälligen Risikoschätzer. Die berechneten Odds Ratios weichen nur leicht voneinander ab (Tabelle 4).

3.2.3 Stratifizierte Analyse nach Altersgruppen

In der Fallgruppe der Kinder & Jugendlichen befanden sich vier Personen, ihre Kontroll/Fallgruppe umfasste 27 Infizierte. In der Fallgruppe der Erwachsenen befanden sich 6 Personen und 70 Infizierte bildeten ihre Kontroll/Fallgruppe. Aufgrund der kleinen Fallzahlen wurde der p-Wert über den Fischer-Exakt-Test berechnet.

Die OR im Stratum der Kinder & Jugendlichen für die Variable Rindfleisch fällt mit 17,25 höher aus als für Erwachsene (OR 9,67). Beide Werte sind signifikant. Die Crude OR und die MH-OR sind mit 11,81 und 11,87 nahezu identisch.

Für die Variable Schweinefleisch zeigt sich eine signifikante OR für Kinder & Jugendliche, von 24,00. Die OR für Erwachsene liegt nicht signifikant bei 5,78. Die Crude OR (OR 8,44) weicht nur gering von der signifikanten MH-OR, die 8,87 beträgt, ab.

In der Gruppe der Kinder & Jugendlichen erfolgte keine positive Angabe zur Exposition mit Geflügel. Die berechneten Risikoschätzer sind somit nicht zu verwenden. Die Crude OR, mit 2,17 und die MH-OR mit 2,31 weichen kaum von einander ab.

Bei den weiteren stratifizierten Variablen zeigt sich für die Gruppe der Kinder & Jugendlichen eine gering höhere OR, als bei den Erwachsenen. Alle Werte weisen keine Signifikanz auf. Crude OR und MH-OR weichen kaum voneinander ab (Tabelle 5).

3.2.4 Logistische Regression

Die Variablen Rindfleisch und Schweinefleisch sind für das LR-Modell verwendet worden. Der Konsum von Schweinefleisch und Rindfleisch beeinflusst sich stark. Für Rindfleisch ergibt sich eine knapp nicht signifikante OR von 6,12 (CI 0,97-38,78; p 0,055). Die OR für Schweinefleisch liegt nicht signifikant bei 2,80 (CI 0,44-17,87; p 0,277).

4 Diskussion

4.1 Interpretation der Ergebnisse

Die Risikoschätzer einer Case-Case-Studie müssen als Mischung aus einem deskriptiven und einem analytischen Wert interpretiert werden. Die Fallgruppe besteht aus Personen eines infizierten Herdes und die Kontroll/Fallgruppe aus weiteren Infizierten einer definierten Zeitspanne. In dieser Vergleichsgruppe ist somit der durchschnittliche Infektionsweg dieser Zeitspanne enthalten. Die Risikoschätzer geben die Stärke der Exposition zu einer Variablen im Fallherd, verglichen mit den infizierten Personen der Kontroll/Fallgruppe, wieder.

Eine Fall-Kontroll-Studie vergleicht infizierte Personen mit einer Gruppe nicht infizierter Kontrollen. Durch die Erhebung einer „gesunden“ Kontrollgruppe, die unter einem möglichst vergleichbaren Risiko stand sich anzustecken, kann unterschiedliches Verhalten, das zu einer Infektion geführt hat, untersucht werden. Diesen kausalen Zusammenhang stellt eine Case-Case-Studie nicht her. In ihr wird ein signifikanter Unterschied im Verhalten des Herdes zu weiteren Infizierten dargestellt. Die Risikoschätzer sind somit ein Maß für die Abweichung des Infektionsweges in der Fallgruppe, im Gegensatz zum Infektionsweg in der Kontroll/Fallgruppe.

4.1.1 Bivariate Analyse

In der durchgeführten Case-Case-Studie zeigen sich für den Konsum von Rind- und Schweinefleisch signifikante Odds Ratios. Die Risikoschätzer weisen auf einen stärkeren Zusammenhang zwischen den Infizierten des Herdes und der Exposition zu Rindfleisch auf. Eine Fall-Kontroll-Studie, die mit dem gleichen Ausbruch durchgeführt wurde, identifizierte einen Fleischer im Bezirk als Quelle der Infektionen. Die Studienfrage in dieser Untersuchung richtet sich nach der Bezugsquelle der konsumierten Fleischwaren. Es wurde untersucht, ob die Infizierten des Herdes ihre Fleischwaren in einer Fleischerei, die unter Verdacht stand, erwarben. Diese Studie ermittelte hierfür eine OR von 6,2 (CI 1,2-32,3)¹³ [vgl. Institut für Hygiene und Umwelt 2003]. Die Studienfrage der Case-Case-Studie und der Fall-Kontroll-Studie unterscheiden sich. In der Case-Case-Studie wurde nach den konsumierten Lebensmitteln und nicht nach einer Bezugsquelle analysiert. Der in

¹³ Der p-Wert ist in der Veröffentlichung nicht mit aufgeführt.

der Fall-Kontroll-Studie dargestellte Zusammenhang zu dem Fleischer ist allerdings eine Bestätigung des Ergebnisses der Case-Case-Studie. In ihr wurde ebenfalls der Konsum von Fleischwaren mit der Fallgruppe in Verbindung gebracht.

4.1.2 Stratifizierte Analyse nach Geschlecht

In der stratifizierten Analyse nach Geschlecht ergab sich in den Straten für Rindfleisch und Schweinefleisch ein unterschiedlicher Wert. Männer aus der Fallgruppe haben eine höhere Wahrscheinlichkeit Rindfleisch zu konsumieren. Für Frauen berechnet sich ein niedrigerer, ebenfalls signifikanter Zusammenhang. Frauen weisen hingegen eine höhere signifikante OR für den Konsum von Schweinefleisch auf, wobei sich für Männer kein Zusammenhang zeigt. Diese Werte weisen auf eine Effektmodifikation hin. Wahrscheinlicher ist allerdings, dass Männer und Frauen des Herdes sich unterschiedlich infizierten. Um sich über Rind- und Schweinefleisch anzustecken, müssen beide Nahrungsmittel kontaminiert worden sein. Die Verbreitung des Erregers kann über einen infizierten Mitarbeiter der Fleischerei erfolgt sein. Ein Dauerausscheider des Salmonelloseerregers ist in der Lage, durch unhygienisches Arbeiten mehrere Produkte zu verunreinigen. Die epidemische Kurve des Ausbruchs unterstützt diese Vermutung (Abbildung 4, S. 20). Die Infektionen ereigneten sich über einen Zeitraum von sechs Wochen. Das ist ein Hinweis auf eine Gemeinschaftsquelle, die lange gestreut hat. Ein Dauerausscheider, der mehrere Fleischwaren über die Zeitspanne kontaminierte, kann ein solches Ausbruchsmuster verursachen. Über Kreuzkontaminationen, die bei der unsachgemäßen Verarbeitung der Fleischwaren auftreten, könnten ebenfalls mehrere Produkte verunreinigt werden. Somit verbreitet sich der gleiche Erreger über verschiedene Nahrungsmittel. Wenn die so infizierten Nahrungsmittel eingefroren wurden und langsam in den Verkauf geraten sind, kann sich ebenfalls diese epidemische Kurve zeigen.

Eine weitere Erklärung für den Unterschied in den stratifizierten Odds Ratios ist der Einfluss eines Recall-Bias. Hierbei erinnern sich die Befragten zum Zeitpunkt des Interviews fehlerhaft an bestandene Expositionen. Die OR ist in diesem Fall durch eine falsche Erinnerung der konsumierten Fleischsorten beeinflusst. Einige Personen haben im Nachhinein Schwierigkeiten, sich an verzehrte Nahrungsmittel korrekt zu erinnern. Die Unterscheidung einzelner Fleischsorten kann somit fehlerhaft erfolgt sein. Diese Gefahr besteht besonders, da die Befragung des Fallherdes zwischen sechs und zwölf Wochen

nach Meldung der Infektion erfolgte. Zusätzliche Fehlerquellen ergeben sich durch die kleinen Fallzahlen in den Straten. Dadurch ein Einfluss des Zufalls wahrscheinlicher.

Die Variable Fisch & Meeresfrüchte zeigt für Frauen ein signifikant positives Ergebnis. Für Männer konnte aufgrund der kleinen Fallzahlen kein Risikoschätzer ermittelt werden. Die MH-OR lässt allerdings auf keine Störvariable schließen. Da für den Infektionsausbruch der Konsum von Rind- und Schweinefleisch ermittelt wurde, scheint das Ergebnis der Stratifikation durch einen Recall-Bias oder durch den Zufall beeinflusst zu sein.

4.1.3 Stratifizierte Analyse nach Altersgruppen

In der stratifizierten Analyse nach Alter liegen, speziell für Rindfleisch und Schweinefleisch, höhere Risikoschätzer für Kinder & Jugendliche vor. Auch hier kann ein möglicher Recall-Bias das Ergebnis verzerren. Diese Vermutung liegt besonders nahe, da die Gruppe der Kinder & Jugendlichen in allen Expositionen höhere Effektschätzer als die Gruppe der Erwachsenen aufzeigt. Eine korrekte Wiedergabe der Nahrungsmittlexpositionen ist für Kinder als schwierig zu bewerten. Bestandene Expositionen zu unterscheiden und richtig zu erinnern führt speziell in dieser Altersgruppe zu Fehlern [Linnane et al 2002; Geddie 2001]. Dies ist von besonderer Bedeutung, da die Infizierten des Herdes zwischen sechs und zwölf Wochen nach Auftreten der Symptome interviewt wurden.

Auch bei der Stratifikation nach Alter ist, aufgrund der kleinen Fallzahlen, von einer Verzerrung durch den Zufall auszugehen.

4.1.4 Logistische Regression

Die Logistische Regression zeigt, dass Rindfleisch und Schweinefleisch sich deutlich beeinflussen. Beide Variablen sind im Modell nicht signifikant. Bei der Exposition zu Rindfleisch liegen die Signifikanzwerte allerdings im grenznahen Bereich (OR 6,12, CI 0,97-38,78, p 0,055). Die Risikoschätzer für Schweinefleisch weisen keinen Zusammenhang auf. Da in einer Case-Case-Studie kein kausaler Zusammenhang hergestellt wird, muss der CI und der p-Wert spezifischer als in einer Fall-Kontroll-Studie interpretiert werden. Der infizierte Herd wird in dieser Analyse mit dem durchschnittlichen Infektionsweg der Erkrankung verglichen. Somit kann es sein, dass die berechneten Werte sich nicht so deutlich, wie im Vergleich zu gesunden Kontrollen, abheben. Die

Risikoschätzer einer Case-Case-Studie können dadurch eine geringere Signifikanz, als die einer Fall-Kontroll-Studie, aufweisen (vgl. auch Kap. 4.3).

Die Logistische Regression zeigt eine stärkere Assoziation des Fallherdes zum Konsum von Rindfleisch, die sich im Zusammenhang mit dem Konsum von Schweinefleisch stark abschwächt. Dieses spiegelt sich ebenfalls in der nach Geschlecht stratifizierten Analyse wieder. Die Variable Rindfleisch weist für Männer und Frauen einen signifikanten Zusammenhang auf. Für Schweinefleisch bestätigt sich nur eine Assoziation für Frauen. Die LR unterstützt somit die Vermutung, dass Rindfleisch für die Infektion hauptverantwortlich ist und weitere Infektionen durch den Konsum von Schweinefleisch verursacht wurden.

4.2 Unterschiedliche Erhebung der Fall- und Kontroll/Fallgruppe

Die verwendeten Daten der Fallgruppe und der Kontroll/Fallgruppe wurden unterschiedlich erhoben. Für die Fallgruppe konnten keine Daten der routinemäßigen Surveillance verwendet werden, da eine zielgerichtete Erhebung von Ausbrüchen durch einen Interviewer-Bias verfälscht wurde. Die Daten der Fallgruppe stammen aus einer Nachuntersuchung, die zu einem späteren Zeitpunkt erfolgte. In ihr wurden systematisch und ausführlich Nahrungsmittel abgefragt. Dies geschah über ja/nein-Fragen. Die Daten der Kontroll/Fallgruppe stammen aus der routinemäßigen Surveillance. Hierbei kam es zu einer Erfassung von ausgewählten Nahrungsmitteln, deren Konsum bestätigt werden konnte.

Diese unterschiedliche Form der Erhebung kann verschiedene Arten eines Recall-Bias verursachen. Durch die Nachuntersuchung der Fallgruppe zu einem späteren Zeitpunkt sind die konsumierten Nahrungsmittel weniger präsent als in der routinemäßigen Surveillance. Hierdurch können Expositionen fehlerhaft erinnert werden. Die ausführlichere und wiederholte Abfrage kann wiederum die befragten Personen anhalten, sich mehr Mühe mit der Beantwortung ihrer Fragen zu geben. Sie sind sich bewusst, Gegenstand einer Studie zu sein und bewerten somit die Befragung höher als bei der routinemäßigen Erhebung. Es gibt mehrere Möglichkeiten, wie hierdurch das Antwortverhalten beeinflusst wird. Es muss daher von einer Art „diffusen Verzerrung“ ausgegangen werden, die das Ergebnis beeinträchtigt. Um ein möglichst fehlerfreies Bild zu erhalten sollte eine Befragung zeitnah und einheitlich erfolgen.

Ein weiterer systematischer Fehler, der in der Studie besteht, ist der Selektions-Bias. Angaben von 17% der ursprünglichen Kontroll/Fallgruppe wurden nicht erfasst, da man sie bereits einem Herd zuordnen konnte. Es ist davon auszugehen, dass mehrere der Herde durch typische Infektionsquellen verursacht wurden. Expositionen von Infizierten dieser Herde wurden nicht erfasst. Dieser Selektions-Bias kann die Wahrscheinlichkeit einer Exposition zu Infektionsquelle in Kontroll/Fallgruppe verringern und somit die Crude OR erhöhen. Da es sich bei Rind- und Schweinefleisch ebenfalls um häufige Infektionsquellen handelt, kann die ermittelte OR in Wirklichkeit niedriger ausfallen.

Weitere 26% der gemeldeten Personen hatten keine Angaben zu verzehrten Nahrungsmitteln gemacht. Ihre Expositionen sind ebenfalls nicht mit in die Analyse eingeflossen.

Für eine weitere Erhebung ist es wichtig, auf eine einheitliche und zeitnahe Befragung jedes Falles zu achten. So kann der Einfluss von Selektions- und Recall-Bias minimiert werden.

4.3 Stärke des Zusammenhangs

Die Stärke eines gefundenen Zusammenhangs hängt von der Häufigkeit einer Exposition im untersuchten Herd sowie von der Häufigkeit dieser Exposition in der Kontroll/Fallgruppe ab. Geht man davon aus, dass sich ein Herd mit einer oft vorkommenden Infektionsquelle infizierte, wird die OR niedriger ausfallen. Wenn der Quotient aus exponiert zu nicht exponiert in der Fallgruppe und in der Kontroll/Fallgruppe hoch ist, wird die ermittelte Crude OR niedriger ausfallen. Eine seltene Exposition, die in der Kontroll/Fallgruppe kaum genannt wird kann hingegen zu einer hohen Crude-OR führen. Die Stärke eines Zusammenhangs hängt direkt mit dem Übertragungsweg, der in der verwendeten Kontroll/Fallgruppe erfasst wurde, zusammen.

Von einer Abschwächung der ermittelten Risikoschätzer können besonders verbreitete *Salmonella*-Serovare betroffen sein. Einzelne Serovare kommen in bestimmten Nahrungsmitteln häufiger vor. *S. enteritidis* wird z.B. verstärkt über Eiprodukte übertragen [vgl. Klist & Freitag 2000]. Da ca. 80% der gemeldeten Salmonellosen auf diesen Erreger zurückgehen, ist eine Häufung in der Nennung von Ei & Eiprodukten zu erwarten. Liegt die Quelle eines Ausbruches ebenfalls in diesen Nahrungsmitteln, kann es zu einer Unterschätzung des Risikos kommen.

Um dieses Problem zu umgehen, kann man das Erregerserovar des Fallherdes aus der verwendeten Kontroll/Fallgruppe entfernen. Dieses führt aber speziell bei verbreiteten Subtypen zu einer starken Verkleinerung der Kontroll/Fallgruppe und somit zu einer Abschwächung der statistischen Signifikanz.

In der vorliegenden Untersuchung wurde der Erreger ebenfalls über ein verbreitetes Nahrungsmittel übertragen und das Studiendesign hat die Quelle signifikant identifiziert.

4.4 Vorteile des Studiendesigns

Ein wichtiger Vorteil ergibt sich durch den Vergleich von zwei infizierten Gruppen. Wie im Kapitel 1.1.3 beschrieben, werden durch das Design einer Case-Case-Studie Erinnerungs- und Selektions-Bias nahezu ausgeschlossen. Dieses führt zu einer guten Vergleichbarkeit unter den zu analysierenden Gruppen. Fall-Kontroll-Studien sind besonders durch diese systematischen Fehler beeinflusst. Dieses gilt verstärkt wenn sie mit Infektionsmeldedaten, die die wirkliche Inzidenz einer Krankheit unterrepräsentieren, durchgeführt werden. In diesem Fall scheint der Einsatz einer Case-Case-Studie besonders geeignet [McCarthy & Giesecke 1999].

Durch die Berechnung der Risikoschätzer in einer Case-Case-Analyse ist man in der Lage, einen Ausbruch nach wissenschaftlichen Kriterien zu bewerten, ohne eine aufwendige Studie durchführen zu müssen. Ihre Risikoschätzer besitzen mehr Aussagekraft als eine deskriptive Untersuchung der routinemäßigen Surveillance. Die erstellte Ergebnistabelle enthält quantifizierte interpretierbare Risikoschätzer für die Interpretation eines Herdes.

Ein weiterer Vorzug besteht in der einfachen und schnellen Durchführung dieser Untersuchung. Da die Kontroll/Fallgruppe aus bereits gemeldeten Infizierten besteht, kann die Analyse direkt nach der Erhebung des zu untersuchenden Herdes durchgeführt werden. Hierfür müssen die Meldedaten in entsprechender Form verwaltet werden. Die Erhebung mit Hilfe einer Datenbank ist dafür geeignet.

5 Ausblick

5.1 Implementierung einer Case-Case-Analyse in ein laufendes Meldesystem

Die vorliegenden Ergebnisse zeigen, dass sich das Design einer Case-Case-Studie für die Umsetzung in ein laufendes Surveillancesystem eignet. Um das Studiendesign in ein Meldesystem einzusetzen, müssen einige Faktoren berücksichtigt werden (Abbildung 6).

Infizierte Personen werden seitens der untersuchenden Labore an das Gesundheitsamt gemeldet. Hierbei ist es vom Vorteil, wenn Angaben über das Serovar des Erregers vorliegen. Somit ist eine Zuordnung zu einem Ausbruch möglich. Kommt es in einem zeitlichen und räumlichen Kontext zu Häufungen eines Serovars, liegt die Vermutung einer gemeinsamen Infektionsquelle nahe.

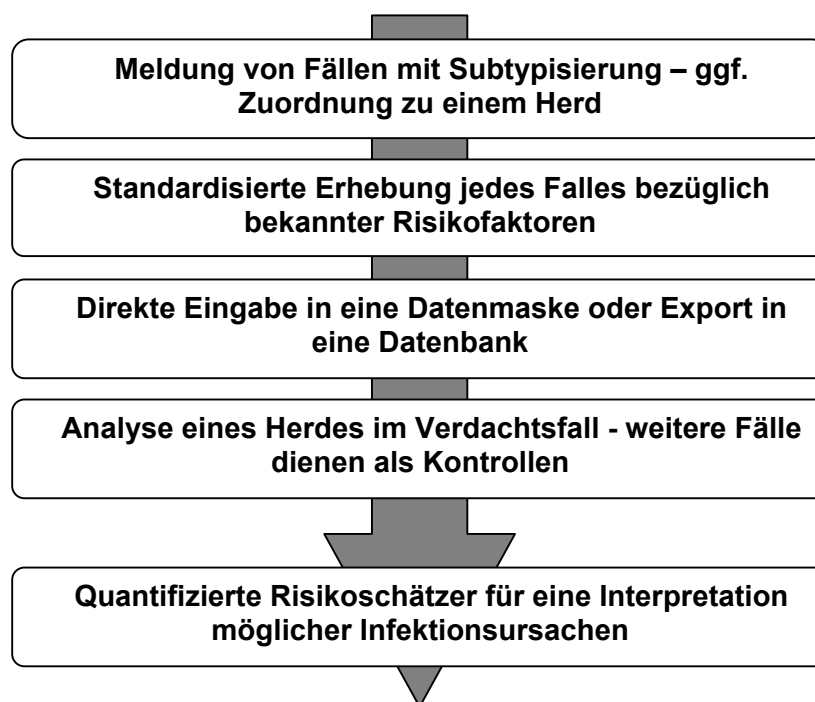


Abbildung 6: Implementierung einer Case-Case-Analyse in ein laufendes Meldesystem

Erfolgt die Meldung eines Salmonellosefalls, wird diese infizierte Person von Mitarbeitern des Gesundheitsamtes kontaktiert. In einem Gespräch werden verschiedene Infektionsursachen erfragt. Um für eine Case-Case-Analyse verwendbare Daten zu erhalten, muss jeder Fall standardisiert erhoben werden. Für Salmonellose-Infektionen gelten als Hauptursache kontaminierte Lebensmittel. Weitere Übertragungswege sind

ebenfalls zu erheben. Von Interesse ist hierbei der Zeitraum 72 Stunden vor Auftreten der Symptome [vgl. RKI 2002a]. Folgende Expositionen sollten in den Interviews erfragt werden:¹⁴

- Rindfleisch
- Schweinefleisch
- sonstige Fleischwaren, z.B. Lamm und Wild
- Geflügel, z.B. Huhn, Ente, Gans und Pute
- Milch und Milchprodukte
- Fisch, Fischprodukte und Meeresfrüchte
- Eier und Eiprodukte, besonders mit Rohei zubereitete Speisen wie Speiseeis, Dessertspeisen, verschiedene Backwaren mit Eischäumen oder Cremebestandteilen sowie Mayonnaise
- Kontakt zu Haustieren
- Reisen ins Ausland
- Kontakt zu weiteren Erkrankten

Die Exposition zu diesen zehn Variablen sollte mit den Antwortmöglichkeiten „ja“, „nein“ und „keine Angabe“ erfragt werden. Hieraus können die statistischen Risikoschätzer ermittelt werden. Zu jeder Variable sollte es die Möglichkeit einer offenen Nennung geben. Somit können z.B. der Ort, an dem ein Nahrungsmittel gekauft wurde oder detaillierte Angaben über Reisen erhoben werden. Aus diesen Angaben können weitere Rückschlüsse für einen identifizierten Herd gemacht werden. Die erhobenen Variablen sollten sich dem aktuellen Wissensstand anpassen, da sich die Erreger in ihrer Ätiologie und die Menschen in ihrem Verhalten weiter verändern. Neue Herstellungsverfahren von Lebensmitteln und neue Konsummuster können zu einer Veränderung von Risikoexpositionen führen. Alle Angaben müssen standardmäßig für jeden Fall erfragt werden. Das ist zum einen wichtig um ein möglichst genaues Bild des Herdes zu erhalten. Weiterhin dienen alle Angaben ebenfalls der Kontroll/Fallgruppe. Jeder erhobene Fall ist zugleich ein potentieller Kontroll/Fall.

¹⁴ Die Expositionen wurden aus Kapitel 1.3.1 zusammengestellt.

Um die Daten für eine Analyse zu verwalten empfiehlt sich eine Datenbank oder ein Analyseprogramm. Die gemeldeten Personen werden herkömmlich per Telefoninterview befragt. Diese Angaben können direkt in ein Programm eingegeben werden. So vermeidet man erhöhten Arbeitsaufwand und Fehler bei der Datenübertragung. Als Programme eignen sich Standardprogramme wie FileMaker, Microsoft Access, Microsoft Excel, OpenOffice Calc oder OpenOffice Datenbank. Spezielle epidemiologische Programme wie EpiInfo, SAS oder SPSS können ebenfalls verwendet werden. Eine sorgfältige Datenverwaltung ermöglicht eine schnelle und einfache Auswertung.

Ereignet sich ein Ausbruch, den man untersuchen möchte, definiert man die Datensätze als Fälle und Kontroll/Fälle. Die Größe der Fall- und Kontroll/Fallgruppe ist mit entscheidend für die statistische Signifikanz der Ergebnisse. In der vorliegenden Studie hat sich eine Fallgruppe von 10 Infizierten und eine Kontroll/Fallgruppe von rund 100 Personen als geeignet erwiesen. In der stratifizierten Analyse wurden jedoch schon signifikante Ergebnisse mit vier Fällen und 28 Kontroll/Fällen erzielt. Bei so kleinen Fallzahlen ist der Einfluss des Zufalls sehr groß und es besteht die Gefahr, einen Ausbruch falsch zu interpretieren. Geeignete Fallzahlen müssen in weiteren Analysen ermittelt werden. Die Kontroll/Fallgruppe sollte aus Infizierten, die zeitnah zum Ausbruch gemeldet wurden, bestehen. So kann die Studie auf Veränderungen, wie saisonale Schwankungen, reagieren.

In einer statistischen Analyse können nun Risikoschätzer für die erhobenen Variablen berechnet werden. Als epidemiologische Risikoschätzer werden die Werte einer Fall-Kontroll-Studie, die Odds Ratio, das Konfidenzintervall und der p-Wert berechnet. Anhand der nun vorliegenden Ergebnisse kann der Ausbruch bezüglich der erhobenen Variablen beurteilt werden.

Das Design einer Case-Case-Studie lässt sich ohne große logistische Veränderungen in ein laufendes Meldesystem umsetzen. Werden die Daten in der oben beschriebenen Form verwaltet, entsteht kein zeitlicher Mehraufwand in der Verwaltung der Meldedaten.

Die hier beschriebene Implementierung einer Case-Case-Analyse in ein laufendes Meldesystem wird im Frühjahr 2005 im Gesundheitsamt Hamburg-Altona und Hamburg-Eimsbüttel umgesetzt. Salmonellose-Ausbrüche werden dann durch eine Microsoft Access Datenbank verwaltet und über eine Microsoft Excel Programmierung analysiert.

Literaturverzeichnis

Beck Texte. Gesundheitsrecht. München: DTV, 2003.

Centers for Disease Control and Prevention (CDC). What is Epi Info? Aktualisiert: Oktober 2004; www.cdc.gov/epiinfo

Deutsche Arbeitsgemeinschaft für Epidemiologie (DAE). Leitlinien und Empfehlungen zur Sicherung von guter Epidemiologischer Praxis, Aktualisierte Version: 2004. www.daepi.de

Dallal GE. Logistic: A Logistic Regression Program for the IBM PC. The American Statistician 1988; 42: 272.

Finanzbehörde Hamburg. Hamburg Handbuch. Hamburg: Dumrath & Fassnacht, 2004.

Geddie LF, Beer J, Bartosik S. The Relationship Between Interview Characteristics and Accuracy of Recall in Young Children: Do Individual Differences Matter? Child Maltreatment. 2001; 6: 59-68.

Gillespie IA, O'Brien J, Frost JA. A Case-Case Comparison of *Campylobacter coli* and *Campylobacter jejuni* Infection: A Toll for Generating Hypotheses. Emerging Infectious Diseases 2002a; 8: 937-942.

Gillespie IA, O'Brien J, Frost JA. Ciprofloxacin resistance in *Campylobacter jejuni*: case-case analysis as a tool for elucidation risks at home and abroad. Journal of Antimicrobial Chemotherapy 2002b; 50: 561-568.

Giesecke J. Modern infectious disease epidemiology. New York: Oxford University Press, 2002.

Institut für Hygiene und Umwelt. Ein Cluster von Infektionen durch *Salmonella* 1,4,5,12:i:-, Subspezies I in Hamburg, Juni/Juli 2003. Infekt-Info 2003a; 15: 1-5.

Institut für Hygiene und Umwelt. Gastroenteritis-Häufung durch *Salmonella* 1,4,5,12:i:-
Ergebnisse einer Fall-Kontroll-Studie. Infekt-Info 2003b; 16: 1-6.

Last JM. A Dictionary of Epidemiology. New York: Oxford University Press, 1995.

Kist MJ, Freitag S. Serovar specific risk factors and clinical features of *Salmonella enterica* ssp. *Enterica* serovar Enteritidis: a study in South-West Germany. Epidemiology and Infection 2000; 124: 383-92.

Köhler W, Eggers HJ, Fleischer B. Medizinische Mikrobiologie. München Jena: Urban und Fischer, 2001.

Linnane E, Roberts RJ, Mannion PT. An outbreak of *Salmonella enteritidis* phage type 34a infection in primary school children: the use of visual aids and food preferences to overcome recall bias in a case control study. Epidemiology and Infection 2002; 129: 35-39.

McCarthy N, Giesecke J. Case-case comparison to study causation of common infectious diseases. International Journal of Epidemiology 1999; 28: 764-768.

Porta M, Malats J, Jarrod M. Serum concentrations of organochlorine compounds and K-ras mutations in exocrine pancreatic cancer. The Lancet 1999; 354: 2125-2129.

Pschyrembel. Klinisches Wörterbuch. Berlin: Walter de Gruyter, 2001.

Robert Koch Institut (RKI). Infektionsgefahr durch Reptilien als Haustiere. Epidemiologisches Bulletin 2000a; 21: 169.

Robert Koch Institut (RKI). Merkblatt für Ärzte – Salmonellose. Aktualisiert: Dezember 2002a; www.rki.de/infektionen

Robert Koch Institut (RKI). *Salmonella* Oranienburg in Schokolade: Internationaler Ausbruch von Oktober bis Dezember 2001. Epidemiologisches Bulletin 2002b; 3: 17-20.

- Robert Koch Institut (RKI). Zu einer Häufung von Salmonella Enteritidis mit erfolgreichem Nachweis des Erregers im Lebensmittel. Epidemiologisches Bulletin 2004a; 18: 149-151.
- Robert Koch Institut (RKI). Bakterielle Gastroenteritiden – Situationsbericht 2003. Epidemiologisches Bulletin 2004b; 31: 251-260.
- Robert Koch Institut (RKI). Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2003. Berlin: Robert Koch Institut, 2004c.
- Rothman KJ. Modern Epidemiology. Boston/Toronto: Little, Brown and Company, 1986.
- Rothman KJ. Epidemiology an Introduction. New York: Oxford University Press, 2002.
- Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV. PulseNet: The Molecular Subtyping Network for Foodborne Bacterial Disease Surveillance, United States. Emerging Infectious Diseases 2001; 7: 382-389.
- Taylor JA. Oncogenes and Their Applications in Epidemiologic Studies. American Journal of Epidemiology 1989; 130: 6-13.
- Tschäpe H. German PulseNet (German PFGE-Netzwerk). Aktualisiert: Juli 2003; www.foodborne-net.de/content/e25/e70/index_ger.html

Eidesstattliche Erklärung

Ich versichere, dass ich vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quelle kenntlich gemacht.

Hamburg, den 28.01.2005

Anlage 1

Fragebogen für die Dateneingabe

Eingabemaske Case-Case-Studie
Diplomarbeit von Ralf Krumkamp

ID: <idnum>

Geschlecht: <A>

Beruf: _____

Alter: ##

gemeldet in Stadtteil: <A> E=Eimsbüttel
A=Altona

Postleitzahl: #####

Ausbruchsnummer: ## (99 wenn ohne Nr.)

Subtyp: # 1=S. enteritidis
2=S. typhimurium
3=S. virchow
4=S. hadar
5=S. heidelberg
6=S. Subspez. I
9=anderer Subtyp

wenn Subtyp=9 bitte Namen hier eintragen _____

Diagnosetyp: # 1=klinisch-epidemiologisch
2=klinisch-labordiagnostisch
3=labordiagnostisch ohne klinisches Bild
4=labordiagnostisch bei unbekanntem klinischen Bild

Beginn der Erkrankung: <mm/dd/yyyy>
mm/dd/yyyy

Dauer der Erkrankung: ##

Aufgetretene Symptome

Uebelkeit: <Y>

Erbrechen: <Y>

Durchfall: <Y>

Fieber: <Y>

Vor Beginn der Erkrankung auf Reise innerhalb Deutschlands? <Y> wo:

Vor Beginn der Erkrankung auf Reise ausserhalb Deutschlands? <Y> wo:

Abgefragte Nahrungsmittel:

Hintergrunddaten (ggf. wo oder was
gegessen):

Fisch und Meeresfruechte <Y> _____
(Fertiggerichte, Div. Fisch, Krabben, Muscheln, Sossen ...)

Gefluegel <Y> _____
(Huehnchen, Pute, Truthahn, Ente, Fasan, Wildgefluegel, Pastete, KFC ...)

Schwein <Y> _____
(Keule, Kotelett, Hackfleisch, Eisbein, Schnitzel ...)

Rindfleisch <Y> _____
(Hackfleisch, Hamburger, Schmorgerichte, Salami, sonstige wurst ...)

Lamm <Y> _____
(Keule, Kotelett, Hackfleisch, Kebabs ...)

Anderes Fleisch <Y> _____
(Wild, Pferd, Kaninchen, Ziege ...)

Eiprodukte <Y> _____
(Eier, Ruehr-/Spiegeleier, Sossen mit Eiern, Mayonnaise, Quiche, Pfannkuchen ...)

Kaese <Y> _____
(div. Kaese)

Milch und Milchprodukte <Y> _____
(Milch, Joghurts, Quark, Creme fraiche, Eis ...)

Pasta <Y> _____
(Spaghetti, Lasagne, Makkaroni, Tagliatelle ...)

Salat, Obst, Gemuese <Y> _____
(gemischer Salat, rohes Gemuese ...)

Pasteten, Kuchen <Y> _____
(Cremebroetchen, Eierkuchen, Biskuitkuchen ...)

Imbiss & Suessigkeiten <Y> _____
(Knabberkram, Trockenfruechte, Bonbons, Schokolade, Gummibaerchen ...)

Fast Food <Y> _____
(Pizza, Wuerste ...)

andere Lebensmittel: _____ <Y> _____

Besondere Freizeitaktivitäten <Y> _____

Erkrankte im Bekanntenkreis <Y> wo: _____

Kontakt mit Erkrankten <Y> wo: _____

Kontakt mit Tieren <Y> welches _____

Sonstiges: _____

Anlage 2

Analysesyntax der bivariaten und stratifizierten Analyse

```

define ausbruch _
if ausbruchsn = 1 then ausbruch = 1 else ausbruch = 2
define cluster ##
let cluster = ausbruchsn
define sasubtyp _____
let sasubtyp = wennsubtyp
recode subtyp to sasubtyp 1 = "S. enteritidis" 2 = "S. typhimurium" 3 =
"S. virchow" 6 = "S. subspez. I"
define onset <mm/dd/yyyy>
let onset = beginndere
define sex _
let sex = geschlecht
define age ##
let age = alter
define agegr string
if age <18 then agegr = 1 else agegr = 2
define agegr2 string
recode age to agegr2 by 10
define PLZ #####
let plz = postleitza
define contact _
if erkranktei = "Y" then contact = 1 else contact = 2
define meat _
if schwein = "Y" then meat = 1 else meat = 2
if rindfleisch = "Y" then meat = 1
if lamm = "Y" then meat = 1
define beef _
if rindfleisch = "Y" then beef = 1 else beef = 2
define pork _
if schwein = "Y" then pork = 1 else pork = 2
define fish _
if fischundme = "Y" then fish = 1 else fish = 2
define egg _
if eiprodukte = "Y" then egg = 1 else egg = 2
if pastetenku = "Y" then egg = 1
define milk _
if milchundmi = "y" then milk = 1 else milk = 2
define fast_food _
if fastfood = "Y" then fast_food = 1 else fast_food = 2
if Imbissues = "Y" then fast_food = 1
define pasta2 _
if pasta = "Y" then pasta2 = 1 else pasta2 = 2

```

```

define salat _
if salatobstg = "Y" then salat = 1 else salat = 2
define s-ent _
if wennsubtyp = 1 then s-ent = 1 else s-ent = 2
route c:\epi6\altona\s-altona.rec
write recfile id ausbruch cluster sasubtyp onset sex age agegr agegr2
plz contact meat beef pork fish egg milk fast_food pasta2 salat s-ent
close
read c:\epi6\altona\s-altona.rec
route c:\epi6\altona\s-des.txt
title 1 -----
title 2 case-case-studie salmonellose/altona
title 3 deskriptive analyse - clusterverteilung allgemein
title 4 -----
freq cluster
title 3 deskriptive analyse fallgruppe
select ausbruch = 1
freq sex
freq age
freq agegr2
freq onset
freq contact
select
title 2 deskriptive analyse kontroll/fall-gruppe
select ausbruch = 2
freq sex
freq age
freq agegr2
freq onset
freq sasubtyp
select
route c:\epi6\altona\s-epi.txt
title 1 -----
title 2 case-case-studie salmonellose/altona
title 3 univariate analyse
title 4 -----
title 5
tables meat ausbruch
tables beef ausbruch
tables pork ausbruch
tables fish ausbruch
tables egg ausbruch
tables milk ausbruch
tables fast_food ausbruch
tables pasta2 ausbruch
tables salat ausbruch

```



```
route C:\epi6\altona\s-str.txt
title 3 stratifizierte analyse-sex
tables meat ausbruch sex
tables beef ausbruch sex
tables pork ausbruch sex
tables fish ausbruch sex
tables egg ausbruch sex
tables milk ausbruch sex
tables fast_food ausbruch sex
tables pasta2 ausbruch sex
title 3 stratifizierte analyse-agegoup (1 < 18; 2 >= 18)
tables meat ausbruch agegr
tables beef ausbruch agegr
tables pork ausbruch agegr
tables fish ausbruch agegr
tables egg ausbruch agegr
tables milk ausbruch agegr
tables fast_food ausbruch agegr
tables pasta2 ausbruch agegr
tables salat ausbruch agegr
route C:\epi6\altona\ausgabe\s-ent.txt
title 1 -----
title 2 case-case-studie salmonellose/altona
title 3 case-case-vergleich s. enteritidis mit weiteren salmonellosen
title 4 -----
title 5
tables meat s-ent
tables beef s-ent
tables pork s-ent
tables poultry s-ent
tables fish s-ent
tables egg s-ent
tables milk s-ent
tables contact s-ent_
close
```


Anlage 3

Syntax zur Erstellung des Logistischen-Regressionsdatensatzes

```

define ausbruch #
if ausbruchsn = "1" then ausbruch = 1 else ausbruch = 0
define contact #
if erkranktei = "Y" then contact = 1 else contact = 0
define meat #
if schwein = "Y" then meat = 1 else meat = 0
if rindfleisch = "Y" then meat = 1
if lamm = "Y" then meat = 1
define beef #
if rindfleisch = "Y" then beef = 1 else beef = 0
define pork #
if schwein = "Y" then pork = 1 else pork = 0
define fish #
if fischundme = "Y" then fish = 1 else fish = 0
define poultry
if gefluegel = "Y" then poultry = 1 else poultry = 0
define egg #
if eiprodukte = "Y" then egg = 1 else egg = 0
if pastetenku = "Y" then egg = 0
define milk #
if milchundmi = "y" then milk = 1 else milk = 0
define fast_food #
if fastfood = "Y" then fast_food = 1 else fast_food = 0
if Imbissues = "Y" then fast_food = 1
define pasta2 #
if pasta = "Y" then pasta2 = 1 else pasta2 = 0
define salat #
if salatobstg = "Y" then salat = 1 else salat = 0
define PLZ #####
let plz = postleitza
define cluster ##
let cluster = ausbruchsn
define onset <mm/dd/yyyy>
let onset = beginndere
define age ##
let age = alter
define sex #
if geschlecht = "M" then sex = 1 else sex = 0
define agegr #
if age < 18 then agegr = 1 else agegr = 0
route c:\epi6\altona\lr\lr-salm.rec

```

```
write recfile ausbruch cluster onset sex age agegr plz contact meat beef  
pork poultry fish egg milk fast_food pasta2 salat  
close
```