FOOD-RELATED VIRAL INFECTIONS:

INVESTIGATING THE POSSIBILITIES FOR EFFECTIVE CONTROL

July, 2008

L.P.B. Verhoef, W. van Pelt, M. Koopmans
National Institute for Public Health and the Environment (RIVM)
Center for Infectious Disease Control (CIb)
Laboratory for Infectious Diseases and Perinatal Screening (LIS)
Epidemiology and Surveillance (EPI)
PO Box 1, Postbak 22
3720 BA Bilthoven
the Netherlands
phone: +31-30-274 2605 fax: +31-30-274 4418

Note: this report contains unpublished information. Therefore, this report cannot be made publicly available in its present form. The contents are for use of VWA officials only, to guide interventions. A copy will be provided to the members of the FBVE network.

TABLE OF CONTENTS

SUMMARY

Norovirus outbreaks, including food-borne outbreaks, implicate considerable costs for society. The epidemiology of norovirus outbreaks may differ between genotypes. We investigated whether norovirus genotypes with an increased risk of being ascribed to contaminated food early in the food-chain can be identified, which may be useful information for Food Health Authorities for undertaking prevention measures.

Two norovirus strain collections, which were systematically collected from 1999 through 2004, were compared. We calculated Pierson's correlation coefficients (ρ) for genotype patterns as detected in bivalve mollusks (n=295) and in outbreaks (n=2760) caused by food, food handlers, food consumption, person-to-person transmission and by unknown transmission. For each genotype, proportions of food-related and person-to-person outbreaks including 95% confidence intervals were used to estimate the proportion of outbreaks attributable to food among food handlers, unknowns and bivalve mollusks.

Genotype patterns detected in bivalve mollusks and food highly correlated ($\rho = 0.91$, $p=0.000$) and significantly differed from all other genotype patterns, showing a higher proportion of genogroup I strains. Using the patterns to estimate food-relatedness of outbreaks resulted in an estimated 21% instead of the reported 16%. Some genotypes (I.1, I.2 and I.4) may indicate whether outbreaks are food-related. This information can be used for primary and secondary prevention of food-borne outbreaks with a high risk of international consequences. Still, international information exchange beyond genotype, i.e. sequences, is needed for detection of outbreaks with no clear link in space and time: the 'diffuse outbreaks'. Further research into the information that can be derived from clustering sequences is needed and will be performed this year.

SAMENVATTING

Norovirus uitbraken, waaronder voedselgerelateerde uitbraken, gaan samen met een hoge kostenlast voor de maatschappij. De epidemiologische eigenschappen van norovirussen lijken per genotype te verschillen. In ons onderzoek hebben wij gekeken of bepaalde norovirus genotypen wijzen op contaminatie in het begin van de voedselketen. Dit zou nuttige informatie zijn voor het nemen van maatregelen door de Voedsel en Waren Autoriteit. We hebben twee systematische virusstammencollecties, verzameld tussen 1999 en 2004, vergeleken. De correlatie van genotypenpatronen in 2-kleppige schelpdieren (n=295) en uitbraken (n=2760) door voedsel, voedselbereiders, voedselconsumptie, persoon-op-persoon contact of onbekende transmissie werd berekend met behulp van Pierson's correlatie coëfficiënt rho (ρ). Voor ieder genotype werden de proporties voedselgerelateerde en persoon-op-persoon uitbraken inclusief betrouwbaarheidsinterval berekend. Deze werden vervolgens gebruikt om een schatting te maken van het voedselgerelateerde deel van de uitbraken in voedselbereiders, uitbraken door onbekende oorzaak en in schelpdieren.

Het genotypenpatroon in 2-kleppige schelpdieren komt het meest overeen met het patroon in voedsel ($\rho = 0.91$, $p=0.000$) en dit overeenkomende patroon was significant verschillend van alle andere genotypenpatronen. Het aandeel genogroup I stammen in schelpdieren is hoger dan in uitbraken. Wanneer de genotypenpatronen gebruikt worden om het aandeel voedselgerelateerde uitbraken te schatten, komen we op een hoger percentage (21%) dan de gerapporteerde 16% voedselgerelateerde uitbraken. Sommige genotypen (I.1, I.2, I.4) lijken voedselgerelateerdheid van uitbraken aan te geven. Deze informatie kan gebruikt worden voor primaire en secundaire preventie van voedselgerelateerde uitbraken met verhoogd risico op internationale consequenties.

Ondanks deze bevindingen dient informatie-uitwisseling tussen landen verder te gaan dan het uitwisselen van genotypen. Sequenties zijn nodig voor de detectie van uitbraken waarbij geen duidelijke relatie in tijd of ruimte aanwezig is: de zogeheten 'diffuse uitbraken'. Meer onderzoek is nodig naar de informatie die gehaald kan worden uit clusters van stammen op sequentieniveau, en dit onderzoek zal komend jaar uitgevoerd worden.

1. INTRODUCTION

The Dutch Food and Consumer Safety Authority (VWA) has a mandate to inform the European Union about the prevalence of food-borne infections. Consequently, it is of interest to know the relative importance of different transmission routes in the total amount of gastroenteritis outbreaks. In the case of food-borne infections with a viral cause, identification is complicated by the absence of systematic surveillance. This is particular true for what can be termed 'diffuse food-borne outbreaks', where, as a result of globalization of the food-industry and wide geographic distribution of products, primary introduction of viruses through food may occur in several countries or continents simultaneously. The VWA initiated a study related to the Food-borne Viruses in Europe (FBVE) network to determine which information from a combined epidemiological and virological data-collection of norovirus outbreaks in 13 European countries is relevant to risk assessment by VWA and the European Food Safety Authority (EFSA). During the first year, 2007, the FBVE data-collection was investigated to identify the data needed to provide support interventions by the VWA, and to assess the current surveillance infrastructure for availability of these data. During that first year, an *epidemiological* approach was chosen, of which the details are described (1), and of which the main conclusions are given in the next paragraph. Here we describe the findings of the ongoing investigation using a *virological* approach at clustering strains at genotype level. Results of comparison of clusters of strains at sequence level will be reported in a final report in January 2009.

The main conclusions from the epidemiological approach

As a minimum dataset required supporting VWA action is often lacking for viral outbreaks of gastroenteritis, a strategy was chosen to help prioritize which outbreaks require more complete follow-up. From 5 years of norovirus outbreak surveillance data the following minimum dataset for reporting food-borne outbreaks was epidemiologically determined: setting of the outbreak (i.e. household, restaurant, health care institution, school, hotel, daycare center, other), the number of cases involved, and norovirus genotype. The analysis indicated a clear difference between genogroup (GG) II.4 strains and other viruses, where non-GGII.4 strains were found more frequently in outbreaks with a food-borne mode transmission, and GGII.4 strains were found more frequently in healthcare settings with person-to-person

outbreaks (Verhoef et al. submitted 2008). The epidemiological analysis resulted in a validated predictive multivariate logistic regression model comparing food-borne outbreaks to person-to-person outbreaks, which was used to create a practical web-based selection tool. Incompleteness of surveillance data limited our investigation to 14% of all outbreak reports in the FBVE database. Use of the selection tool is likely to improve surveillance data quality, as the number of outbreaks needing follow-up was reduced to 30% while accepting that one out of five food-borne outbreaks will be missed.

The virological approach

As the epidemiological approach showed that the epidemiology of noroviruses may vary between genogroups, a virological approach was chosen to determine whether further specification into genogroups and genotypes reveals additional differences in the epidemiology of norovirus outbreaks. Noroviruses are highly diverse, with the five genogroups recognized further classified into a growing number of genotypes, and with new genetic clusters continuously evolving. Viruses belonging to different genotypes may differ in virulence. The genotype classification is based on an agreed definition, 80% amino acid homology across the capsid gene. The diversity of strains has consequences for the sensitivity and specificity of molecular detection methods, and has complicated the harmonization of methods for international comparison of strains. In this report we will focus on the norovirus sequences reported by 13 countries to the FBVE database but all genotyped according to the same method described elsewhere (Vennema et al. manuscript in preparation). We will investigate how the FBVE norovirus typing information can be applied to prove or support an epidemiological link in food-related outbreaks.

2 DEFINITIONS

Food-borne outbreak: A food-borne outbreak was defined as follows according to the EFSA guidelines: two or more human cases of the same disease and/or infection or a situation in which the observed number of human cases exceeds the expected number and where the cases are linked, or are probably linked, to the same food source.

Confirmed food-borne outbreak: Where there is laboratory evidence of norovirus in food or analytical epidemiologic evidence of a food source through a case-control or cohort study, the outbreak is defined as confirmed food-borne.

Probable food-borne outbreak: Where there is descriptive epidemiologic data indicating a link to food, the outbreak is defined as probably food-borne.

Diffuse food-borne outbreak: A food-borne outbreak where, as a result of globalization of the food-industry and wide geographic distribution of products, primary introduction of viruses through food may occur in several countries or continents simultaneously.

Norovirus outbreak: Outbreaks are reported when they satisfy the agreed case definition of a cluster of at least 2 patients within 2 days showing signs of acute gastroenteritis indicative of norovirus (2, 3). A gastroenteritis outbreak is ascribed to norovirus based on compatible descriptive epidemiology and laboratory confirmation according to agreed criteria (4).

Confirmed norovirus outbreak: A confirmed outbreak of norovirus is based on compatible descriptive epidemiology and laboratory confirmation of norovirus in at least 1 positive sample of 4 patient samples tested, or in 2 of 5 to 8, or in 3 of 8 to 16.

Probable norovirus outbreak: A probable outbreak is based on compatible descriptive epidemiology and laboratory confirmation of norovirus in 1 of 1 patient sample tested, or 1 in 5 to 8, or 2 in 9 to 16.

3 QUESTION

Do norovirus genotypes differ in behavior and food-relatedness?

In bacterial food-related outbreaks, and specifically in *Salmonella* outbreaks, strains are found to be source-specific (5, 6). Certain *Salmonella* strains are most commonly found either in swine, poultry, eggs or cattle, and are therefore indicative of the source of origin during a *Salmonella* outbreak. This knowledge generated the idea that a comparative situation may be the case for noroviruses. However, noroviruses strains may not be source-specific but mode-specific, as differences at molecular level may result in differences in survivability through the food-chain or transmission-chain. For example, the highly diverse GGII.4 strain dominating norovirus surveillance is most frequently seen in outbreaks occurring from person-to-person transmission (7). Furthermore, GGI.4 strains are often seen in outbreaks involving consumption of shellfish (8, 9). This led to the hypothesis that certain strains may be specifically found in outbreaks originating from contaminated food in general, and specifically in food contaminated early in the food-chain. If this is the case, diffuse food-borne norovirus outbreaks may be detected using strain typing information. We therefore compared the FBVE strain collection, representative for human illness caused by norovirus, to noroviruses identified in bivalve mollusk monitoring. In chapter 5 we demonstrate that norovirus genotypes reveal additional differences in the epidemiology of norovirus outbreaks.

4 CURRENT PRACTICE FOR VIROLOGICAL DATA

Virological data in national surveillance systems

Since surveillance systems are different in each country, surveillance of (viral) gastroenteritis is not harmonized across Europe. Furthermore, as surveillance systems are known to vary in terms of design, effectiveness, and priorities (10, 11), a telephone survey was performed among epidemiologists and data-entry-persons participating in the FBVE network to categorize the national surveillance systems. Of 13 participating countries, 2 (Great Britain, Germany) were responsible for the majority of the outbreak reports (8,276 of 11,239, 74%), but were unable to meet the FBVE network's reporting criterion of providing linked laboratory and epidemiological norovirus outbreak data. The remaining 11 countries (Denmark, Finland, France, Hungary, Ireland, Italy, the Netherlands, Norway, Sweden, Slovenia, and Spain) may be categorized as follows:

A. Intensive surveillance (i.e. at least one reported typed outbreak per million inhabitants) capable of recognizing food-borne outbreaks and with focus on food (Denmark, Finland, Sweden)
B. Intensive surveillance capable of recognizing food-borne outbreaks but without focus on food (Hungary, the Netherlands, Slovenia)
C. No intensive surveillance (i.e. less than one reported typed outbreak per million inhabitants) capable of recognizing food-borne outbreaks (France, Spain, Italy)
D. Surveillance system not capable of recognizing food-borne outbreaks or tagging 'person-to-person transmission' as a diagnosis of exclusion when other modes cannot be clearly identified (Norway, Ireland)

Sequencing and genotyping of outbreak strains

In the FBVE network, collection of sequence results focused on region A of the norovirus genome but allowed other entries (regions B, C, D) due to lack of standardization between the cooperating laboratories in 13 countries. The strains are all genotyped using a previously described method for sequence analysis of a fragment of the ribonucleic acid (RNA) dependent RNA polymerase gene (12). Genotypes were classified on the basis of similarities to reference strains representing known genotypes using the norovirus typing library (www.rivm.nl/bnwww). The

polymerase sequence was applied for assigning genotypes, where possible. If the polymerase genotype could not be based on the capsid type, strains were aligned and assigned genotypes according to their cluster. Strains for which this alignment was necessary were treated separately as recombinant variants. For classification of norovirus strains into genotypes see Vennema et al. (2008, manuscript in preparation). Between January 1999 and December 2004, the FBVE network collected a total of 2760 norovirus strains, which caused outbreaks in Denmark, Finland, France, Germany, Great Britain, Hungary, Ireland, Italy, the Netherlands, Norway, Sweden, Slovenia, and Spain. Sequence lengths varied between 69 nucleotides and full capsid or polymerase genes.

Of these strains, 87 (3%) could not be genotyped due to low frequencies, short sequences or the incapacity of the method applied to type the detected strain beyond its genogroup. An additional 40 (1%) strains could not be linked to epidemiological data, and therefore their origin remained unknown, leaving 2633 norovirus sequences to determine potential food-relatedness of genotypes. The origin and assigned genotypes of outbreak strains are presented in Table 1.

Monitoring of potential sources of infection

Unlike bacteria, viruses do not replicate in food and the infective dose for viral infection can be low. Moreover, detection of viruses in food is complicated, due to the variety of food matrices, for which a variety of validated and standardized RNA extraction methods would be needed. Therefore, detection of pathogens in the food chain is currently mainly limited to bacteria and does not (yet) include viral testing on a regular base with the consequent absence of systematic strain collections representative of presence of viruses in food or water. However, bivalve mollusks are an exception here, since these filter feeders accumulate viruses, and thereby enable virus detection during routine testing. The European Community Reference Laboratory for Bacterial and Viral Contamination of Bivalve Mollusks (ECRL) collected a total 295 norovirus strains between January 1999 and December 2004 during the routine testing of bivalve mollusks. This strain collection may be interpreted as a reflection of strains circulating in the community, either as a cause or as a consequence of illness. The assigned genotypes of these routinely detected strains are shown in Table 1.

Table 1: Number of norovirus strains detected in samples from humans, bivalve mollusks and food, specified by norovirus genotypes.

	Genotypes		Outbreaks					Routine testing
	Pol-based	Cap-based	Food*	Food handler*	Food consumer*	Person-to-person	Unknown	Bivalve mollusks
Genogroups	I.1	I.1	1	0	8	5	18	0
	I.2		0	0	6	1	39	8
	I.3B	I.3	0	3	8	16	83	13
	I.4	I.4	9	1	8	8	51	86
	I.5		0	0	0	1	5	3
	I.6	I.6	2	1	3	21	19	25
	I.7		0	0	1	0	7	2
	I.HK2K	I.Mie2001	0	0	1	0	4	0
	II.1		0	2	5	12	105	7
	II.2	II.2	0	1	13	27	66	0
	II.3		0	0	1	1	44	11
	II.3R	II.3	0	0	1	1	42	2
	II.4	II.4	5	9	47	674	612	63
	II.5		0	0	3	6	15	0
	II.8		1	1	0	1	15	0
	II.NA1		0	0	0	2	7	0
	II.c		0	0	2	8	31	1
	II.d	II.00G260/DE	0	0	1	3	8	0
	IV.1		0	0	2	1	8	0
Recombinants	II.b	II.1, II.2, II.3	3	1	23	100	200	63
	II.1W	II.10	0	0	0	8	24	11
	II.7	II.6, II.7	2	1	18	19	121	0
Total			23	20	151	915	1524	295

* Strains detected in food, food handlers and persons infected by food were considered food-related.

5 THE ADDED VALUE OF VIROLOGICAL DATA

The information derived from genotypes

Table 1 compares the strains detected in bivalve mollusks and outbreaks among humans with the proportion of genogroup I significantly higher in bivalve mollusks (137/295, 46%) than in humans (330/1633, 13%). Figure 1 shows the distribution of strains for outbreaks ascribed to several sources, and for unexplained outbreak strains of human origin. As can be derived from the correlation coefficients in Table 2, the genotype patterns vary between potential sources. Two genotype patterns can be distinguished (p-value for significantly differing ρ's = 0.02): one genotype pattern as detected in humans, and one genotype pattern as detected in food and bivalve shellfish. The significant correlation coefficient for strains detected in food handlers suggested person-to-person rather than a food-borne mode of transmission in these outbreaks.

Table 2: Pearson coefficients and p-values for correlation of norovirus genotype patterns as detected in routine tested bivalve shellfish, and norovirus outbreaks caused by food, food handlers, person-to-person contacts, as well as norovirus outbreaks of unexplained origin.

	Pearson Correlation Coefficients, N = 22					
	Food	Food handler	Food consumer	Person	Unknown	Bivalves
	ρ (p-value)	ρ (p-value)	ρ (p-value)	ρ (p-value)	ρ (p-value)	ρ (p-value)
Food	1.00	0.45 (0.03)	0.54 (0.01)	0.44(0.04)	0.47(0.03)	0.91(0.000)
Food handler		1.00	0.85(0.000)	0.93(0.000)	0.94(0.000)	0.52(0.01)
Food consumer			1.00	0.88(0.000)	0.95(0.000)	0.60(0.01)
Person				1.00	0.97(0.000)	0.52(0.01)
Unknown					1.00	0.57(0.01)
Bivalves						1.00

Figure 1: The distribution of outbreak strains ascribed to food, food-handlers, person-to-person transmission and the distribution of strains in bivalve mollusks. Strain types are sorted for their relevance in outbreaks among unexplained outbreak strains of human origin.

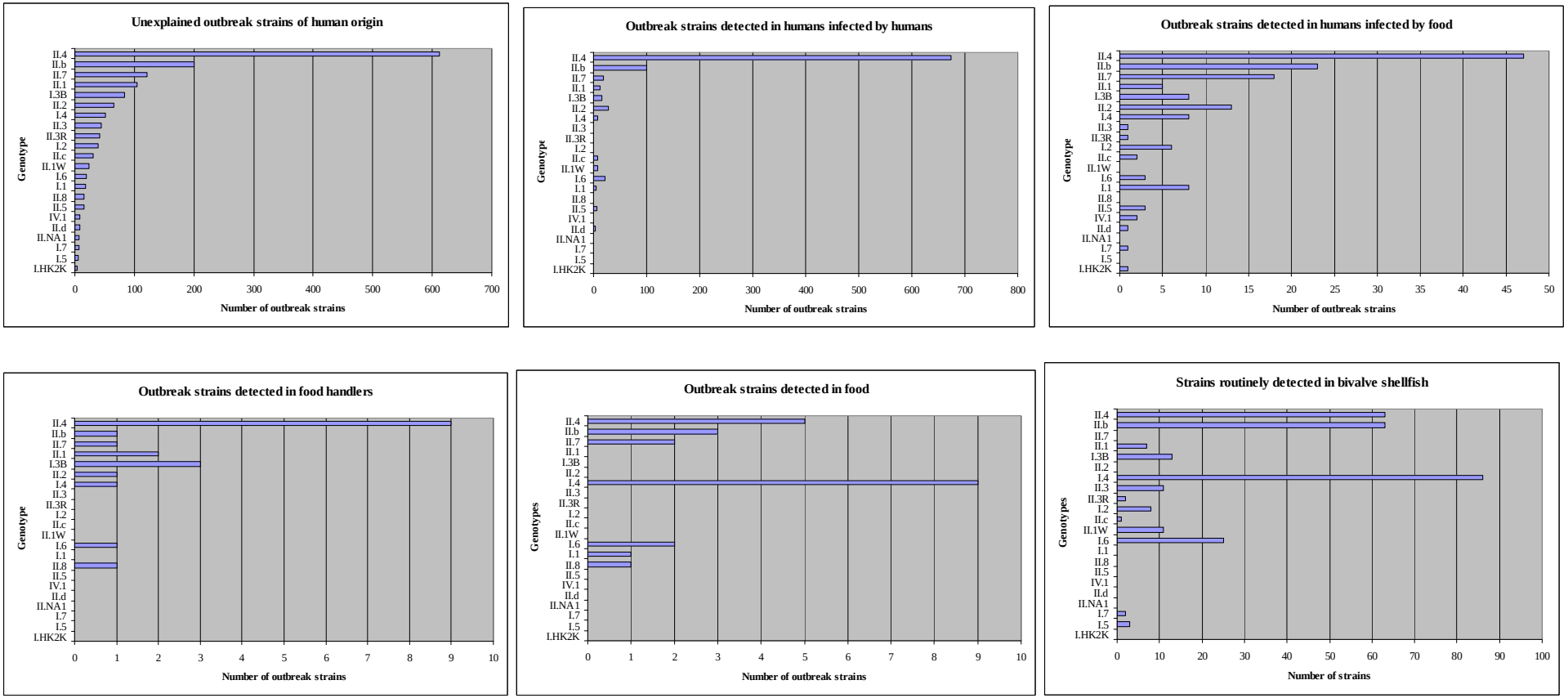
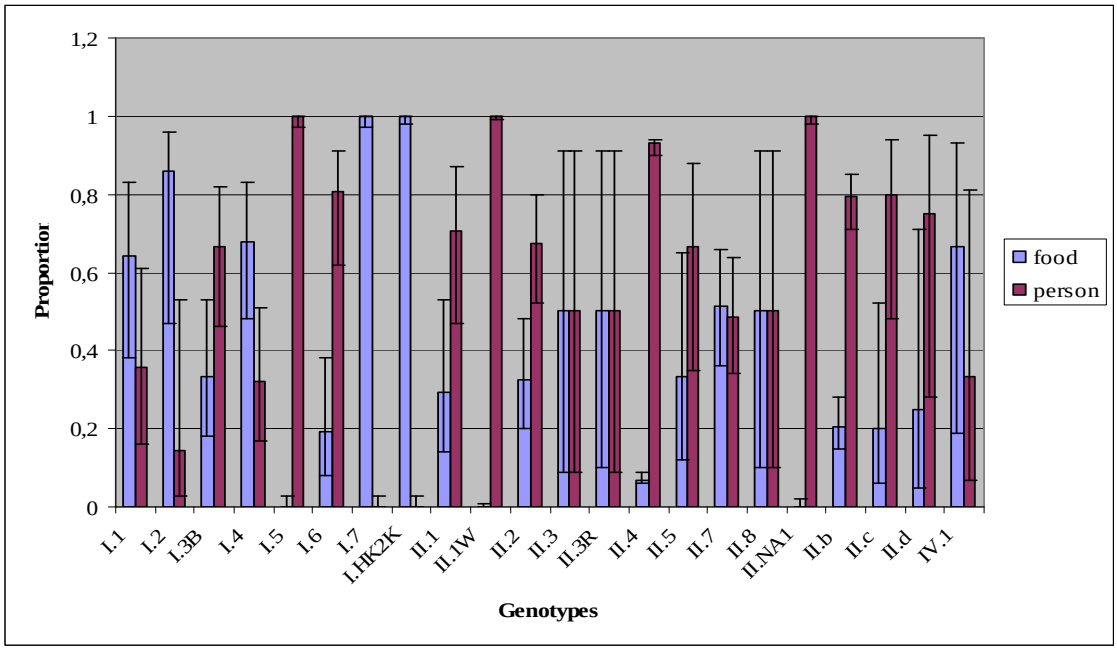


Figure 2 compares the proportions of genotypes detected in food-related outbreaks to those in person-to-person outbreaks, showing that genotypes I.7 and I.HK2K were exclusively, and I.1, I.2 and I.4 more frequently detected in food-related outbreaks. When using the proportions of genotypes and their confidence intervals to distinguish between food-borne and person-to-person outbreaks among outbreaks caused by food-handlers, 5 (95%CI 4-6) are ascribed to food and 15 (95%CI 14-16) to person-to-person transmission.

Figure 2: The proportion of norovirus genotypes in food-related outbreaks (as detected in food or food consumers), and in outbreaks resulting from person-to-person transmission. A 95% confidence interval around this proportion was calculated by Monte Carlo simulation using 10,000 random draws from a beta-distribution.



DETECTION OF GEOGRAPHICALLY DISPERSED OUTBREAKS

For detection of geographically dispersed outbreaks, more data are needed about the diversity of noroviruses belonging to rare genotypes in order to reliably use the data when identifying a probable source of infection. For the more common types, the virological information beyond genotype is needed. Proof of a link between food-borne outbreaks requires identical full-length sequences detected in food and humans, accompanied by compatible epidemiological information. Unfortunately, full-length

sequences are rarely available, with consequent lower evidence for linkage based on identical but shorter sequences. The evidence for linkage differs between regions of the genome, as some regions of the norovirus genome are more conservative than others. Moreover, norovirus strains can mutate when in the human body, resulting in small differences between strains during several transmission cycles in a diffuse outbreak. Such mutations can either be immune-driven or coincidental. Small differences can also exist due to errors in sequence reading. Consequently, small differences between strains do not disprove linkage of outbreaks. The level of concordance between sequences needed for proven linkage needs to be determined, which will be the next step in this project. In the final project-report, the comparison of clusters at sequence level, and consequent linking conditions, will be further elaborated.

6 CONCLUSION AND RECCOMMENDATION

We analyzed a systematically collected outbreak strain collection, as well as routinely collected strains detected in bivalve mollusks, showing a significant difference in GGI/GGII ratio, with GGI strains being more prevalent in bivalve mollusks. Since strains detected in food related to outbreaks show a distribution similar to in bivalve mollusks, this may indicate comparable survivability outside the host rather than transmission mode specific characteristics of genotypes. Based on these 5-year covering systematic strain collections, some genotypes have found to be potentially specific to food (I.HK2K), or to survival in an external host (I.7), while others suggest food-relatedness (I.1, I.2 and I.4). This information can be used for primary and secondary prevention of food-borne outbreaks with a high risk of international consequences, as the identification of a rare norovirus genotype may be a first indicator. To detect geographically dispersed outbreaks, matching genotypes may not provide sufficient information for undertaking prevention measures. Further research is needed into the information that can be derived from clusters of strains at sequence level.

1. Verhoef L, van Pelt W, Koopmans M. Identification and monitoring of viral food-borne outbreaks. Bilthoven: RIVM; 2007.
2. Kaplan JE, Feldman R, Campbell DS, Lookabaugh C, Gary GW. The frequency of a Norwalk-like pattern of illness in outbreaks of acute gastroenteritis. Am J Public Health. 1982 Dec;72(12):1329-32.
3. Koopmans M, Vennema H, Heersma H, van Strien E, van Duynhoven Y, Brown D, et al. Early identification of common-source foodborne virus outbreaks in Europe. Emerg Infect Dis. 2003 Sep;9(9):1136-42.
4. Duizer E, Pielaat A, Vennema H, Kroneman A, Koopmans M. Probabilities in norovirus outbreak diagnosis. J Clin Virol. 2007 Sep;40(1):38-42.
5. Hald T, Vose D, Wegener HC, Koupeev T. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. Risk Anal. 2004 Feb;24(1):255-69.
6. van Pelt W, van de Giessen AW, van Leeuwen WJ, Wannet WJ, Henken AM, Evers EG, et al. Oorsprong, omvang en kosten van humane salmonellose. Deel 1. Oorsprong van humane salmonellose met betrekking tot varken, rund, kip, ei en overige bronnen. Infectieziekten Bulletin. 1999;10(12):240-3.
7. Verhoef L, Depoortere E, Boxman I, Duizer E, van Duynhoven Y, Harris J, et al. Emergence of new norovirus variants on spring cruise ships and prediction of winter epidemics. Emerg Infect Dis. 2008 Feb;14(2):238-43.
8. Le Guyader FS, Bon F, DeMedici D, Parnaudeau S, Bertone A, Crudeli S, et al. Detection of multiple noroviruses associated with an international gastroenteritis outbreak linked to oyster consumption. J Clin Microbiol. 2006 Nov;44(11):3878-82.
9. Boxman IL, Tilburg JJ, Te Loeke NA, Vennema H, Jonker K, de Boer E, et al. Detection of noroviruses in shellfish in the Netherlands. Int J Food Microbiol. 2006 May 1;108(3):391-6.
10. Amato-Gauci A, Ammon A. ECDC to launch first report on communicable diseases epidemiology in the European Union. Euro Surveill. 2007 Jun;12(6):E070607 2.
11. Kroneman A, Harris J, Vennema H, Duizer E, van Duynhoven Y, Gray J, et al. Data quality of 5 years of central norovirus outbreak reporting in the European Network for food-borne viruses. J Public Health (Oxf). 2008 Mar;30(1):82-90.
12. Vinje J, Koopmans MP. Molecular detection and epidemiology of small round-structured viruses in outbreaks of gastroenteritis in the Netherlands. J Infect Dis. 1996 Sep;174(3):610-5.