ELSEVIER

Journal of Clinical Virology 40 (2007) 38–42

JOURNAL OF CLINICAL VIROLOGY

www.elsevier.com/locate/jcv

Probabilities in norovirus outbreak diagnosis

Erwin Duizer*, Annemarie Pielaat, Harry Vennema, Annelies Kroneman, Marion Koopmans

Center for Infectious Diseases Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Received 24 August 2006; received in revised form 25 May 2007; accepted 30 May 2007

Abstract

Background: Noroviruses are recognized as the most common cause of outbreaks of acute gastroenteritis. Yet, diagnostic testing for norovirus is based mostly on RNA detection by RT-PCR, which is not widely available. While antigen detection tests (ELISAs) are easier to perform, they are in general less sensitive.
Objectives: Our aim was to provide a scientific basis for declaring norovirus as the causative agent of an outbreak of acute gastroenteritis.
Study design: Statistical analysis used binomial distribution to determine the minimal number of positive samples, and the probability of detecting the required number of positive samples, for different tests, required to assign norovirus as the causative agent of an outbreak of acute gastroenteritis.
Results: For either a standard RT-PCR or a commercially available ELISA, finding only 1 sample positive out of 2, 3 or 4 samples is sufficient to assign norovirus as the causative agent of an outbreak of acute gastroenteritis. However, when ELISA is used, the probability of detecting this required minimum number of positive samples is low when small numbers of samples are tested (57% when 2 samples are tested; 72% when 3 samples are tested). In order to reach a 90% probability of detecting a norovirus outbreak (false negativity at outbreak level <10%), at least 3 samples should be tested using RT-PCR, and 6 samples when using an ELISA.
Conclusions: The sensitivity for NoV outbreak diagnosis will increase from 57% to 92%, or from 84% to 96%, for ELISA or RT-PCR respectively, when sample size increases from 2 to 6. Thus, using ELISA instead of RT-PCR for the detection of norovirus in stool samples will result in considerable numbers of false negative outbreaks unless a minimum of 6 samples are tested per outbreak.

Keywords: Norovirus; Outbreaks; Diagnosis

1. Introduction

Noroviruses (NoV) are the most common cause of acute gastroenteritis, not only in sporadic cases, but especially in outbreaks. However, diagnostic tests are still not widely available to clinicians and public health practitioners (Lopman, 2006; Turcios et al., 2006). Currently, several ELISAs are becoming commercially available and one might expect that diagnosis of NoV in gastroenteritis outbreaks will become more common.

The high diversity of these small RNA viruses causes difficulties for diagnosis. The antigenic diversity poses great challenges for the development of generic and sensitive EIA protocols, whereas the high genetic diversity challenges the development of generic RT-PCR protocols. As a result, currently available methods vary greatly in sensitivity, specificity and scope for the detection of NoV (Atmar and Estes, 2001; de Bruin et al., 2006).

RT-PCR protocols are generally considered the gold standard in outbreak diagnosis; still, using an array of several RT-PCR protocols NoV was detected in only 84% of fecal samples from NoV outbreaks (Vinje et al., 2003). In this study

Abbreviations: NoV, norovirus; RT-PCR, reverse transcription polymerase chain reaction; ELISA, enzyme linked immunosorbent assay; se, sensitivity; sp, specificity

* Corresponding author at: Centre for Infectious Diseases Control, National Institute for Public Health and the Environment, P.O. Box 1, 3720 BA Bilthoven, The Netherlands. Tel.: +31 30 274 4142; fax: +31 30 274 4418.
E-mail address: Erwin.Duizer@RIVM.NL (E. Duizer).

1386-6532/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jcv.2007.05.015

we use this detectability of 84% of NoV by the most sensitive method (RT-PCR) to define the prevalence of NoV in fecal samples from NoV outbreaks.

A practical problem is that sometimes limited numbers of samples per outbreak are available. In combination with the relatively low sensitivity of currently available EIA tests, the definition of an outbreak is becoming an issue. Different definitions have been used in NoV outbreak diagnosis: (1) if 2 or more samples are found positive (Richards et al., 2003); (2) 50% or more of samples test positive (Lopman et al., 2003); (3) clinical appearance of a nonbacterial gastroenteritis with at least 1 sample positive for NoV and other laboratory tests negative for bacterial and parasitic agents (Fankhauser et al., 1998). In this study we use statistical analysis in order to define the minimal number of samples to be tested, in combination with the number to be found positive, for different tests and sample sizes, to assign NoV as the causative agent of an outbreak. Furthermore, we calculated the probability of detecting the required number of positives as a function of the number of samples tested. These results can guide the interpretation of diagnosing NoV as the cause of an outbreak of gastroenteritis.

2. Materials and methods

2.1. Definition of an outbreak and norovirus prevalence

NoV are common pathogens and a random fecal sample from the general population might contain the virus without being associated with an outbreak.

In the Netherlands, 35 (5.2%) fecal samples of 669 healthy people were found positive for NoV (de Wit et al., 2001). Based upon this result we determined the upper limit of the prevalence of NoV in feces in the general population at 8% (based on a 97.5% CI about 35/669). We defined an outbreak to be caused by NoV if the prevalence of NoV in the outbreak samples is higher than this upper limit of 8%.

2.2. Sensitivity and specificity of the diagnostic tests

In this study we used literature data on sensitivity (se) and specificity (sp) of several diagnostic tests. Test 1 is a hypothetical "gold standard" test of which the specificity and sensitivity are per definition set at 100% (test 1: se = 1, sp = 1). Test 2 is a RT-PCR method. Even though RT-PCR methods are considered the gold standard for NoV detection, in an international collaborative study to compare RT-PCR assays for the detection of NoV, an average sensitivity for single-round RT-PCR of 72% was found (Vinje et al., 2003). For most RT-PCR protocols virtually no false positive samples are reported, i.e. the specificity of these protocols is very high. In our calculations we therefore used a specificity for the RT-PCR of 99% (test 2: se = 0.72, sp = 0.99). The third test was the IDEIA Norwalk Like virus (Dakocytomation Ltd., Ely, United Kingdom). The average reported sensitivity and specificity of this test is 41% and 98%, respectively (Burton-MacLeod et al., 2004; de Bruin et al., 2006; Richards et al., 2003) (test 3; se = 0.41, sp = 0.98).

2.3. Calculation of the minimal number of positive samples required for outbreak diagnosis

Determinations of uncertainty distributions about p (true prevalence of NoV in an outbreak) for separate outbreak data sets were performed following Vose (Vose, 2000). The number of NoV-positive samples (s) in the sample size (n) is assumed to be binomially (n, p) distributed. Subsequently, the uncertainty distribution about p is calculated using hypothetical data on n ($2 < n < 21$) and s, in combination with the previously defined se and sp. We used the BINOMDIST (s; n; (p*se + (1 − p)*(1 − sp); FALSE) function in Excell (Vose, 2000) to define the uncertainty distribution about p. In order to use this function we assumed that the number of tests is fixed (n), the outcome of each test is positive or negative (where s is the number of positive tests), the tests are independent, and the probability of success (i.e., a positive test result, p) is constant throughout the experiment. This last assumption applies only for stool samples obtained during, or shortly after, the symptomatic phase. Based on the resulting distribution, the minimal number of positive samples needed to diagnose a NoV outbreak was calculated for different sample sizes for the three previously defined tests. The minimal number of positive samples is defined as the number of positive samples for which the probability of the prevalence ≥ 8% becomes >95%.

An example for this reasoning is illustrated in Fig. 1a and b. Fig. 1a shows the confidence distribution of an outbreak tested with test 2 (se = 0.72, sp = 0.99) for which 1 out of 3 samples was tested positive. The striped area under the curve (auc) is the sum of confidences for the true prevalence to be >8%, in this case 97% (i.e., >95%). Fig. 1b shows

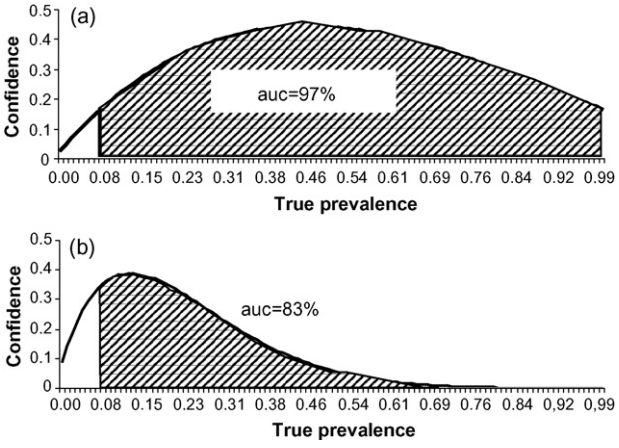

Fig. 1. Confidence distribution of an outbreak tested with test 2 (se = 072, sp = 0.99) for which 1 out of 3 (a) or 1 out of 10 (b) samples was tested positive. The striped area under the curve (auc) is the sum of confidences for the true prevalence to be >8%.

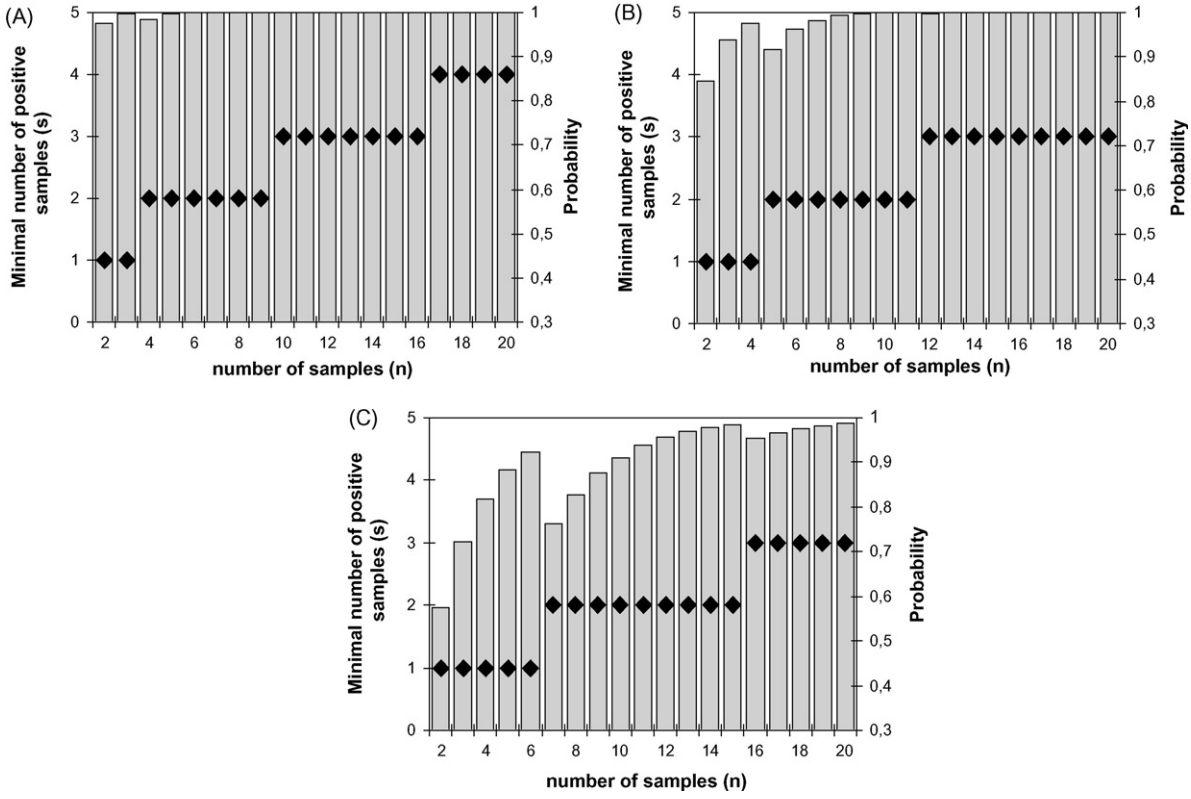

Fig. 2. The minimal number of positive samples to indicate NoV as causative agent of an outbreak (diamonds) and the corresponding probability of actually detecting the required number of positive samples (bars) for test 1 (se = 1, sp = 1) (A), test 2 (se = 0.72, sp = 0.99) (B), and test 3 (se = 0.41, sp = 0.98) (C).

the confidence distribution of an outbreak tested with test 2 (se = 0.72, sp = 0.99) for which 1 out of 10 samples was tested positive. The striped auc is in this case 83% (i.e., <95%) indicating that our confidence on this outbreak being caused by NoV is not high enough.

2.4. Calculation of the probability of detecting the required number of positive samples to diagnose an outbreak

In order to develop further insight in the impact of the sensitivity of the tests for outbreak diagnosis, we calculated the probability of actually detecting the required number of positive samples (resulting from the previous paragraph). For all three tests, specificity and sensitivity were used to calculate the probability of detecting the required number of positive samples to diagnose an outbreak. The probability of detecting NoV per sample was calculated using the BINOMDIST (s; n; (p*se + (1 − p)*(1 − sp); FALSE) function in Excell. In this study we use the detectability of NoV by the most sensitive method (RT-PCR) as measure for the prevalence of NoV in fecal outbreaks samples (i.e., $p = 0.84$, Vinje et al., 2003). This calculates the probability of detecting sufficient NoV positive samples to diagnose NoV as causative agent for that outbreak (i.e., the probability that the prevalence of NoV ≥ 0.08 is >95%).

3. Results

3.1. Minimal numbers of positive samples required for outbreak diagnosis

The minimal number of positive samples to be detected to assign NoV as causative agent of an outbreak with at least 95% confidence, using 1 of a series of 3 tests with defined sensitivities and specificities is shown in Fig. 2A–C (⧫), which shows that the required number of positive samples to identify an outbreak as caused by NoV increases as the number of samples available for testing increases. Furthermore, variations in sensitivity and specificity of the tests affect the required number of positive samples. Moreover, in most cases the required number of positive samples to be found by RT-PCR (test 2; Fig. 2B) and ELISA (test 3; Fig. 2C) is the same, only when the sample size is 5,6,12,13,14, or 15 one extra positive sample is required for the more sensitive RT-PCR to achieve a confidence of >95% for outbreak diagnosis.

For tests 2 and 3, representing the RT-PCR and ELISA respectively, finding only 1 sample positive out of 2, 3 or 4 samples is sufficient to positively diagnose an outbreak as caused by NoV (i.e., $p \geq 0.08$ is >95%), and for the 3 tests finding 2 positive samples in up to 9 samples was sufficient. Note, that if the limit of probability is set to 99% instead of 95% (i.e., the probability of the prevalence ≥8%

Table 1
Maximal number of samples per outbreak for which detection of 1 positive sample is sufficient to indicate NoV as causative agent of an outbreak, for different upper limits of the prevalence of NoV in feces in the general population

Test	Upper limit of the prevalence of NoV in feces in the general population (%)			
	8	4	2	1
se = 1, sp = 1	3	7	14	26
se = 0.72, sp = 0.99	4	8	13	18
se = 0.41, sp = 0.98	6	10	15	20

If more samples than the numbers indicated are tested, 2 or more positives are required to indicate NoV as causative agent.

becomes >99%), detecting only 1 sample positive will never be sufficient.

The number of samples for which only 1 positive sample is required to indicate NoV as causative agent of an outbreak increases with decreasing prevalence of NoV in feces in the general population (Table 1). For example, when the upper limit of NoV prevalence in the population is assumed to be 1%, the transition for requiring 2 positive samples instead of 1 occurred at $n = 27$, 19, and 21, for tests 1, 2, and 3, respectively (Table 1).

3.2. Probability of detecting the required number of positive samples

Based on the sensitivity, specificity and the required number of positive samples per test we calculated the probability of actually detecting at least the required number of positive samples as a measure of the likelihood to confirm the etiology of outbreaks (bars in Fig. 2).

The probability of detecting the required number of positive samples is highly dependent on the sensitivity of the test, and on the number of samples tested. There is a sudden decrease in the probability when there is an increase in number of positive samples required from 1 to 2 (at $n = 4$, 5, and 7 for tests 1, 2, and 3, respectively), this decrease is much less when the required number of positive samples increases from 2 to 3. Fig. 2 shows that when a limited number of 2, 3, or 4 samples is available for testing, the probability of diagnosing an outbreak using RT-PCR (test 2; Fig. 2B) is much higher than if a less sensitive ELISA (test 3: Fig. 2C) is used. When 5 samples are available, the probabilities of detecting the required number of positive samples are 0.92 and 0.88 for RT-PCR and ELISA, respectively. When 6 samples are available, the probabilities of detecting the required number of positive samples are 0.96 and 0.92 for RT-PCR and ELISA, respectively. So, when 5 or 6 samples are tested both tests are almost equally likely to diagnose an outbreak.

We also found that when an RT-PCR method, with the defined specificity and sensitivity, is used to study a candidate NoV outbreak, there is a 99% probability of finding the required number of positive samples to assign a NoV etiology when 8 samples are tested. This indicates that the increase in sensitivity gained by testing more than 8 samples is very limited. If a test with a sensitivity as low as that defined for the study ELISA is used, there is a gradual increase in the probability of correctly identifying a NoV outbreak when the number of samples tested is increased from 8 to 15; however, 99% likelihood is not reached at $n < 21$.

4. Discussion

This study shows that the required number of positive samples needed to assign an outbreak to a pathogen under suspicion is more dependent on the prevalence of the pathogen in the general population than on the sensitivity and specificity of the test used. The sensitivity and specificity greatly influence the probability of detecting the required number of positives, and thus the reliability of the assay for outbreak diagnosis. Considering existing tests, such as the RT-PCR and ELISA with the properties defined as assumption, we can conclude that for 2, 3, and 4 samples tested, detecting 1 positive is sufficient to assign an outbreak to NoV. When only 2, 3, or 4 samples are tested, the probability of finding the required numbers of positive samples increases with the sample number. However, due to the underlying assumed binomial process 1 is more likely to detect the 1 required positive sample when $n = 4$ than the required 2 positive samples when $n = 5$ or 6, using RT-PCR as the diagnostic test. The same holds for the ELISA, with the study-defined specificity and sensitivity, in that the probability for detecting the required number of positive samples is higher for $n = 6$, than for $7 \leq n \leq 10$. The probability of detecting 1 positive sample out of 6 is, of course, higher than detecting 1 out of 4; however, one should realize that detecting less than the required number of positives means that one cannot be confident that the NoV prevalence found in the outbreak is higher than the population background prevalence (8% in this study).

If we consider a limit of 90% probability to detect an outbreak (false negativity at outbreak level <10%), we show that at least 3 samples should be tested using RT-PCR, but 6 (or more than 10) samples should be tested when the study-specified ELISA is used. Since for some outbreaks the number of samples available is often limited, the use of an ELISA detection method for NoV will result in a considerable numbers of false negative outbreaks. If an outbreak is tested negative by an ELISA method, one should consider retesting the samples by a more sensitive method such as RT-PCR. Furthermore, we conclude that in up to 4 samples tested, either by ELISA or RT-PCR, detecting 1 positive sample will be sufficient, even for a pathogen as prevalent as NoV.

Acknowledgements

This work was supported by the European Commission, DG Research Quality of Life Program 6th Framework (EVENT, SP22-CT-2004-502571) and DG SANCO (DIVINE-Net, 2003213).

References

Atmar RL, Estes MK. Diagnosis of noncultivatable gastroenteritis viruses, the human caliciviruses. Clin Microbiol Rev 2001;14(1):15–37.

Burton-MacLeod JA, Kane EM, Beard RS, Hadley LA, Glass RI, Ando T. Evaluation and comparison of two commercial enzyme-linked immunosorbent assay kits for detection of antigenically diverse human noroviruses in stool samples. J Clin Microbiol 2004;42(6): 2587–95.

de Bruin E, Duizer E, Vennema H, Koopmans MPG. Diagnosis of Norovirus outbreaks by commercial ELISA or RT-PCR. J Virol Meth 2006;137(2):259–64.

de Wit MA, Koopmans MP, Kortbeek LM, Wannet WJ, Vinje J, van Leusden F, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol 2001;154(7):666–74.

Fankhauser RL, Noel JS, Monroe SS, Ando T, Glass RI. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. J Infect Dis 1998;178(6):1571–8.

Lopman B. Noroviruses: simple detection for complex epidemiology. Clin Infect Dis 2006;42(7):970–1.

Lopman BA, Reacher MH, Van Duijnhoven Y, Hanon FX, Brown D, Koopmans M. Viral gastroenteritis outbreaks in Europe, 1995–2000. Emerg Infect Dis 2003;9(1):90–6.

Richards AF, Lopman B, Gunn A, Curry A, Ellis D, Cotterill H, et al. Evaluation of a commercial ELISA for detecting Norwalk-like virus antigen in faeces. J Clin Virol 2003;26(1):109–15.

Turcios RM, Widdowson MA, Sulka AC, Mead PS, Glass RI. Reevaluation of epidemiological criteria for identifying outbreaks of acute gastroenteritis due to norovirus: United States, 1998–2000. Clin Infect Dis 2006;42(7):964–9.

Vinje J, Vennema H, Maunula L, von Bonsdorff CH, Hoehne M, Schreier E, et al. International collaborative study to compare reverse transcriptase PCR assays for detection and genotyping of noroviruses. J Clin Microbiol 2003;41(4):1423–33.

Vose D. Animal import and food safety risk assessment. In: Vose D, editor. Risk analysis: a quantitative guide. Chichester: John Wiley & Sons; 2000. p. 349–72.