SUPPLEMENT ARTICLE

Laboratory Diagnostic Challenges in Case/Control Studies of Diarrhea in Developing Countries

Roy M. Robins-Browne[1,2] and Myron M. Levine[3]

[1]Department of Microbiology and Immunology, The University of Melbourne, [2]Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia; and [3]Center for Vaccine Development, University of Maryland School of Medicine, Baltimore

Case/control studies of acute infectious diarrhea require accurate and dependable laboratory tests to detect pathogens in samples from both symptomatic patients and healthy control subjects. The methods used to detect these pathogens have usually been evaluated on patient samples only, and their performance on samples from control subjects is mostly unknown. Because many pathogens occur at a high overall frequency in developing countries and thus may be present in a notable proportion of control subjects as well as patients, the relative ability of a diagnostic test to detect these pathogens in diarrheic and normal stools can have a profound effect on the interpretation of case/control data.

The laboratory procedures used to detect etiologic agents in patients with acute infectious diarrhea are constantly evolving. Nevertheless, the principles underlying the performance and interpretation of these procedures are well established. As with clinical diagnostic microbiology in general, the choice of the tests used to detect a possible etiologic agent is determined by clinical relevance, practicability, and cost. For the most part, this approach is satisfactory, although in some cases no etiologic agent is identified. This may be because the diarrhea is not infectious in origin, or because a particular agent is not identified either because it was not sought or because the procedures used to detect it were not sufficiently sensitive and gave a false-negative result. On the other hand, when a single pathogen is found, interpretation of the results is straightforward, insofar as the pathogen is usually assumed to be responsible for the patient's symptoms.

In case/control studies, in addition to investigating diarrheic samples for pathogens, we undertake the far less familiar task of investigating feces or rectal swabs from subjects without diarrhea. The detection of a pathogen in these individuals indicates asymptomatic carriage, the possible reasons for which are discussed in the accompanying article by Levine and Robins-Browne in this supplement.

Analysis of the outcome of a case/control study involves comparing the frequencies of the detection of pathogens in cases and control subjects, which are used to determine an odds ratio (OR). As the OR indicates the strength of the association between a pathogen and the occurrence of diarrhea, it is used as a measure of the relative pathogenicity of different pathogens. The OR is also one factor in the equation utilized to calculate attributable fraction, which provides an estimate of the relative contribution of the pathogen(s) of interest to the diarrheal disease burden (see the article by Blackwelder et al in this supplement). Because the frequency of detection of pathogens in control subjects can have a profound effect on the interpretation of case/control data, it is essential to understand the performance of laboratory tests in samples from subjects without diarrhea, as well as those from patients.

Correspondence: Roy M. Robins-Browne, MB, BCh, PhD, Dept of Microbiology and Immunology, Grattan St, The University of Melbourne, Victoria 3010, Australia (r.browne@unimelb.edu.au).

Clinical Infectious Diseases 2012;55(S4):S312–16

© The Author 2012. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1093/cid/cis756

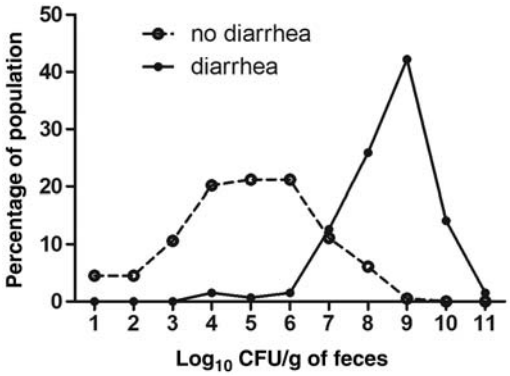

Figure 1. Frequency distribution of the number of the colony-forming units from 198 fecal samples from rabbits without diarrhea, and 135 samples from rabbits with diarrhea caused by infection with 1 of 3 different rabbit-specific enteropathogenic *Escherichia coli* strains. Data were obtained from quantitative cultures of samples on selective agar containing antibiotics to which the infecting strains were resistant. The data show that feces from rabbits without diarrhea contain significantly lower numbers of bacteria than those with diarrhea. Rabbits without diarrhea included similar numbers of animals with subclinical infection and animals sampled during the incubation or convalescent periods of symptomatic infection. Abbreviation: CFU, colon-forming units.

Two key factors that govern the usefulness and reliability of a laboratory test are its sensitivity and specificity. For almost all tests, these parameters are first determined by using "spiked" samples and then evaluated under field conditions, often in comparison with other tests. A test that is highly sensitive and specific will reliably detect a pathogen in cases with few false-negative or false-positive results. However, diagnostic tests are seldom evaluated in control subjects (ie, individuals without symptoms). In industrialized countries this is seldom an issue because the prevalence of most pathogens in healthy subjects is low. In developing countries, however, where sanitation is poor and exposure to contaminated food and water is virtually a daily norm, intestinal pathogens circulate at high frequency and children are liable to become repeatedly infected with them. Accordingly, endemic pathogens will be present at a far higher frequency overall than in industrialized countries. The ability to detect these pathogens in control subjects will differ according to the sensitivity of the diagnostic procedures that are used.

Table 1. Results of Hypothetical Case/Control Studies Where the Frequency of a Pathogen in Patients Is Fixed at 25% and Its Frequency in Control Subjects Ranges From 0% to 25%

Frequency in 100 Patients	Frequency in 100 Control Subjects	Odds Ratio	*P* Value[a]
25	0	Infinity	<.0001
25	5	6.3	.001
25	10	3	.009
25	15	1.9	.11
25	20	1.3	.5
25	25	1	1

[a] Fisher exact test, 2-tailed.

Most comprehensive case/control studies of diarrhea include culture for bacterial pathogens, such as *Salmonella* species, *Shigella* species, and *Escherichia coli*. Culture of fecal samples has an intrinsically low sensitivity to detect pathogens, especially in individuals without diarrhea, because the complex microbiota of healthy individuals makes it difficult to detect a pathogen among the high background "noise." The fact that culture of feces is a useful diagnostic procedure despite its low sensitivity can be explained partly by the fact that in patients with diarrhea, the pathogen is generally excreted in far higher numbers and makes up a much greater proportion of the cultivable microbiota than in healthy subjects who are asymptomatic carriers of the same pathogen. The odds of finding a pathogen in both cases and controls can be considerably improved by using selective media with or without prior enrichment. The use of such media has revolutionized our understanding of the epidemiology of bacterial enteropathogens such as *Campylobacter jejuni* in developing and industrialized countries [1, 2].

To illustrate these points, we will use data from studies we have undertaken with an animal model of diarrhea caused by a subtype of enteropathogenic *E. coli*, known as rabbit-specific enteropathogenic *E. coli* (REPEC). Infection of infant rabbits with REPEC closely parallels infection of human infants with human-specific enteropathogenic *E. coli* (EPEC) in terms of age-related susceptibility, clinical presentation, and associated intestinal pathology [3, 4].While establishing this model at the University of Melbourne and determining the median infectious dose of different REPEC strains, the natural course of infection was charted by observing rabbits for symptoms of diarrhea and correlating this with quantitative culture of REPEC on selective media containing antibiotics to which the challenge strains were resistant. Detailed descriptions of our methods have been published previously [5].

To investigate the hypothesis that bacteria are present in greater numbers (and therefore more easily detected) in cases with diarrhea than in control subjects, we reanalyzed published and unpublished data from experiments in which we infected rabbits with 1 of 3 different wild-type strains of REPEC of differing virulence. Quantitative culture of fecal samples from these animals indicated that rabbits with diarrhea excrete significantly more bacteria (1.1×10^8 colony-forming units [CFU] per gram of feces [mean]; 1.6×10^8 CFU [median]) than rabbits without symptoms (2.8×10^4 CFU

Table 2. Results of Hypothetical Case/Control Studies Where the Frequency of a Pathogen in Patients Ranges From 10% to 30%, and the Difference in Its Frequency in Patients and Control Subjects Is Fixed at 10%

Frequency in 100 Patients	Frequency in 100 Control Subjects	Odds Ratio	*P* Value[a]
10	0	Infinity	.001
15	5	3.4	.03
20	10	2.3	.07
25	15	1.9	.11
30	20	1.7	.14

[a] Fisher exact test, 2-tailed.

[mean]; 3.9×10^4 CFU [median]; $P < .0001$) (Figure 1). Our data indicated that a test with a detection limit of 10^7 CFU per gram would be positive in 96% of cases of diarrhea and in 18% of infected, but asymptomatic, individuals (Figure 1). In contrast, a test with a detection limit of 10^4 CFU per gram would be positive in 100% of cases and 80% of infected controls. This analysis exemplifies how increasing the sensitivity of the test can improve detection limits disproportionately in control samples compared with samples from patients. The influence this may have on the interpretation of hypothetical case/control data is shown in Tables 1 and 2. Our data suggest that as test sensitivity increases, quantitative assays may be useful in distinguishing clinical from subclinical infection. This suggestion has been borne out by a recent report by Barletta et al [6], who found that the use of quantitative polymerase chain reaction (PCR) to diagnose EPEC infection in children in an endemic setting yielded higher values in patients than in subjects without diarrhea.

A striking example of how improved detection of intestinal pathogens can influence data obtained from a case/control study comes from the comprehensive English Infectious Intestinal Disease Study (1993–1996). By using PCR for 8 groups of pathogens to investigate 4627 archived fecal samples from 2422 cases and 2205 controls in the original study (which did not use PCR-based detection of enteropathogens), Amar et al [7] increased the detection rate of at least 1 agent (or toxin) from 53% in the original study to 75% in cases, and from 19% to 42% in controls. Furthermore, the use of PCR-based diagnosis for 8 groups of pathogens increased the number of cases in whom >1 pathogen was detected from 272 to 993 (a 73% increase), and from 32 to 280 (a 89% increase) in controls. The greatest increase in detection rates that resulted from the use of PCR was for rotavirus and norovirus.

For example, in children aged <1 year the detection rate of rotavirus in cases went from 29 of 144 (20%) detected by enzyme immunoassay (EIA) and/or electron microscopy (EM) to 70 of 144 (49%) detected by PCR and EM. In controls the increase in diagnostic yield using PCR was even greater: from 3 of 183 (2%) to 53 of 183 (29%). In this age group, the OR before PCR was 15.1 (95% confidence interval [CI], 4.5–78.8); and with the PCR data included it was 2.3 (95% CI, 1.4–3.7). Despite the fall in OR, the attributable fraction (discussed in the article by Blackwelder et al in this supplement) increased from 19% using the original detection method to 28% using PCR. This can be explained by the significantly increased detection of rotavirus in cases. Also, in the original study, norovirus was detected by using EM in 14 of 144 (10%) cases <1 year old and in 2 of 183 (1%) age-matched controls (OR, 9.7 [95% CI, 2.2–89.3]). Together, however, EM and PCR revealed norovirus in 70 of 144 cases (49%) and 57 of 183 (31%) controls (OR, 2.1 [95% CI, 1.3–3.4]). In this case, the attributable fraction increased from 9% using the original method to 25% using more sensitive detection. These data, albeit from an industrialized country, clearly indicate how using a test with enhanced sensitivity can influence the major outcomes of a case/control study. Amar et al [7] also reported that quantitative PCR for norovirus may permit asymptomatic carriage to be distinguished from symptomatic infection, a finding that was subsequently confirmed by other researchers [8]. Similar findings linking the number of virus particles in feces to disease severity have also been reported for rotavirus [9, 10].

Although mucosal immunity will account for some instances of asymptomatic carriage of particular pathogens, the presence of mucosal antibodies may also interfere with the ability to detect pathogens when using EIA. In the case of rotavirus, for example, the most common method of diagnosis is a type of "sandwich" EIA, in which an immobilized antibody is used to capture a rotavirus antigen from feces, after which the captured antigen is revealed by using a second, labeled antibody. Tests of this type are capable of detecting between 10^5 and 10^6 rotavirus particles per milliliter [11]. However, the sensitivity of this assay may fall during the course of the illness, as patients develop immunity to rotavirus and secrete mucosal antibodies that coat the virus and interfere with its detection by EIA [11, 12]. By contrast, PCR using reverse transcription to amplify rotavirus RNA is able to detect as few as 1000 virus particles per milliliter [11], and is unaffected by mucosal immunity. In a study of children hospitalized for diarrhea with rotavirus, PCR-based diagnosis revealed that 11 of 37 (30%) children were still infected with rotavirus >3 weeks after hospitalization compared with only 2 of 37 (5%) when EIA was used to detect the virus [11]. These data indicate that the EIA for rotavirus is more likely to be positive in patients experiencing their first infection with a particular virus than in children who are convalescing from an acute infection or are reinfected with a strain of rotavirus they have encountered previously. By contrast, the EIA for *Giardia lamblia* is extremely sensitive and can be used to identify asymptomatic

carriers of this pathogen, for example, during the investigation of outbreaks of giardiasis in industrialized countries [13].

Apart from test sensitivity, another factor that may influence the comparison of laboratory data from cases and control subjects in case/control studies of diarrhea is the nature of the samples that are investigated. In an ideal case/control study, the diagnostic samples that are collected from patients and controls should be the same. In case/control studies of diarrhea, however, this is generally not the case, because in patients with diarrhea, especially watery diarrhea, much of the sample will have originated in the small intestine, and the normal microbiota of the large intestine will have been purged or significantly diluted, whereas in controls, the fecal samples or rectal swabs that are investigated will reflect the microbiota of the distal large intestine.

Although this difference may not matter in some instances, in others it could be important. For example, some bacteria only cause disease in the small intestine. An example is enterotoxigenic *E. coli* (ETEC), in which enterotoxins that are responsible for diarrhea act predominantly in the small intestine [14]. Furthermore, studies with EPEC infection in adult volunteers have shown that these bacteria are virulent only when ingested by mouth and not when they are inoculated directly into the large intestine [15]. This may be explained by the fact that environmental signals required to activate virulence gene expression are absent from the large intestine [16]. Given that *E. coli* is well adapted to persist in the large intestine as part of the normal microbiota, it is conceivable that strains of pathogenic *E. coli*, including ETEC, EPEC, and enteroaggregative *E. coli*, may colonize the large intestine of healthy people or convalescent patients and behave as nonpathogens, whereas the same bacteria isolated from the small intestine would be of considerable diagnostic significance. Similar circumstances may apply to other enteric pathogens that also differ in their ability to cause disease depending on their site of intestinal colonization. On the other hand, diarrheic stools may contain commensal microorganisms, which normally reside mainly in the proximal intestine and are not readily detectable in formed stools. In this case, the association of the agent with diarrhea may lead to the false conclusion that it is a causative agent.

As detailed in the article in this supplement by Nataro et al, the methods used to detect and identify pathogens in The Global Enteric Multicenter Study (GEMS) were state-of-the-art. Nevertheless, the relative ability of these methods to detect pathogens in cases and control subjects in developing countries is not known. For a thorough understanding of case/control data, especially when comparing the relative contribution of different pathogens to the overall burden of disease, we need a more thorough understanding of the performance of diagnostic procedures as used on samples from cases and controls. Some possible areas of further study could include the quantitative analysis of patients' samples (particularly when there is >1 pathogen), and examination of virulence gene expression to indicate if a putative etiologic agent is behaving as a pathogen or commensal.

Notes

Acknowledgments. We are grateful to Dr William Blackwelder and Dr Tamer Farag for helpful discussion and advice, and to Louise Adams for providing unpublished data.

Financial support. This research was supported by the Australian National Health and Medical Research Council and the Bill & Melinda Gates Foundation (grant number 38874).

Supplement sponsorship. This article was published as part of the supplement entitled "The Global Enteric Multicenter Study (GEMS)," sponsored by the Bill & Melinda Gates Foundation.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Blaser MJ, Taylor DN, Feldman RA. Epidemiology of *Campylobacter jejuni* infections. Epidemiol Rev **1983**; 5:157–76.
2. Patton CM, Mitchell SW, Potter ME, Kaufmann AF. Comparison of selective media for primary isolation of *Campylobacter fetus* subsp. *jejuni*. J Clin Microbiol **1981**; 13:326–30.
3. Clarke SC, Haigh RD, Freestone PP, Williams PH. Virulence of enteropathogenic *Escherichia coli*, a global pathogen. Clin Microbiol Rev **2003**; 16:365–78.
4. Milon A, Oswald E, De RJ. Rabbit EPEC: a model for the study of enteropathogenic *Escherichia coli*. Vet Res **1999**; 30:203–19.
5. Robins-Browne RM, Tokhi AM, Adams LM, Bennett-Wood V. Host specificity of enteropathogenic *Escherichia coli* from rabbits: lack of correlation between adherence in vitro and pathogenicity for laboratory animals. Infect Immun **1994**; 62:3329–36.
6. Barletta F, Ochoa TJ, Mercado E, et al. Quantitative real-time polymerase chain reaction for enteropathogenic *Escherichia coli*: a tool for investigation of asymptomatic versus symptomatic infections. Clin Infect Dis **2011**; 53:1223–9.
7. Amar CF, East CL, Gray J, Iturriza-Gomara M, Maclure EA, McLauchlin J. Detection by PCR of eight groups of enteric pathogens in 4,627 faecal samples: re-examination of the English case-control Infectious Intestinal Disease Study (1993–1996). Eur J Clin Microbiol Infect Dis **2007**; 26:311–23.
8. Phillips G, Lopman B, Tam CC, Iturriza-Gomara M, Brown D, Gray J. Diagnosing norovirus-associated infectious intestinal disease using viral load. BMC Infect Dis **2009**; 9:63.
9. Kang G, Iturriza-Gomara M, Wheeler JG, et al. Quantitation of group A rotavirus by real-time reverse-transcription-polymerase chain reaction: correlation with clinical severity in children in South India. J Med Virol **2004**; 73:118–22.
10. Phillips G, Lopman B, Tam CC, Iturriza-Gomara M, Brown D, Gray J. Diagnosing rotavirus A associated IID: using ELISA to identify a cut-off for real time RT-PCR. J Clin Virol **2009**; 44:242–5.
11. Richardson S, Grimwood K, Gorrell R, Palombo E, Barnes G, Bishop R. Extended excretion of rotavirus after severe diarrhoea in young children. Lancet **1998**; 351:1844–8.
12. Coulson BS, Grimwood K, Hudson IL, Barnes GL, Bishop RF. Role of coproantibody in clinical protection of children during reinfection with rotavirus. J Clin Microbiol **1992**; 30:1678–84.
13. Garcia LS, Shimizu RY. Evaluation of nine immunoassay kits (enzyme immunoassay and direct fluorescence) for detection of *Giardia lamblia* and *Cryptosporidium parvum* in human fecal specimens. J Clin Microbiol **1997**; 35:1526–9.

14. Turner SM, Scott-Tucker A, Cooper LM, Henderson IR. Weapons of mass destruction: virulence factors of the global killer enterotoxigenic *Escherichia coli*. FEMS Microbiol Lett **2006**; 263:10–20.
15. Koya G, Kosakai N, Kono M, Mori M, Fukasawa Y. Observations on the multiplication of *Escherichia coli* O-111 B4 in the intestinal tract of adult volunteers in feeding experiments: the intubation study with Miller-Abbott's double lumen tube. Jpn J Med Sci Biol **1954**; 7:197–203.
16. Yang J, Tauschek M, Robins-Browne RM. Control of bacterial virulence by AraC-like regulators that respond to chemical signals. Trends Microbiol **2011**; 19:128–35.