Comparison of Nine Commercially Available *Clostridium difficile* Toxin Detection Assays, a Real-Time PCR Assay for *C. difficile* *tcdB*, and a Glutamate Dehydrogenase Detection Assay to Cytotoxin Testing and Cytotoxigenic Culture Methods[▽]

Kerrie Eastwood,¹ Patrick Else,¹ André Charlett,² and Mark Wilcox^{1*}

Microbiology Department, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom,¹ and Statistical Unit, Health Protection Agency, Centre for Infections, Colindale, London, United Kingdom²

Received 3 June 2009/Returned for modification 19 July 2009/Accepted 14 August 2009

The continuing rise in the incidence of *Clostridium difficile* infection is a cause for concern, with implications for patients and health care systems. Laboratory diagnosis largely relies on rapid toxin detection kits, although assays detecting alternative targets, including glutamate dehydrogenase (GDH) and toxin genes, are now available. Six hundred routine diagnostic diarrheal samples were tested prospectively using nine commercial toxin detection assays, cytotoxin assay (CYT), and cytotoxigenic culture (CYTGC) and retrospectively using a GDH detection assay and PCR for the toxin B gene. The mean sensitivity and specificity for toxin detection assays were 82.8% (range, 66.7 to 91.7%) and 95.4% (range, 90.9 to 98.8%), respectively, in comparison with CYT and 75.0% (range, 60.0 to 86.4%) and 96.1% (91.4 to 99.4%), respectively, in comparison with CYTGC. The sensitivity and specificity of the GDH assay were 90.1% and 92.9%, respectively, compared to CYT and 87.6% and 94.3%, respectively, compared to CYTGC. The PCR assay had the highest sensitivity of all the tests in comparison with CYT (92.2%) and CYTGC (88.5%), and the specificities of the PCR assay were 94.0% and 95.4% compared to CYT and CYTGC, respectively. All kits had low positive predictive values (range, 48.6 to 86.8%) compared with CYT, assuming a positive sample prevalence of 10% (representing the hospital setting), which compromises the clinical utility of single tests for the laboratory diagnosis of *C. difficile* infection. The optimum rapid single test was PCR for toxin B gene, as this had the highest negative predictive value. Diagnostic algorithms that optimize test combinations for the laboratory diagnosis of *C. difficile* infection need to be defined.

Clostridium difficile is a major nosocomial pathogen causing a range of symptoms from mild to severe diarrhea and is the etiological agent of pseudomembranous colitis. The incidence of *C. difficile* infection has increased markedly in many countries, notably associated with the epidemic spread of PCR ribotype 027 (NAP1) since its recognition in the United States and Canada (6, 7, 13). It is essential to have accurate laboratory diagnosis of *C. difficile* infection to ensure patients receive appropriate treatment and that correct infection control measures are put in place. Also, inaccurate testing will potentially lead to poor quality surveillance data that may lead to inappropriate infection prevention measures.

The cytotoxin assay (CYT), first described by Chang et al., detects the toxins produced by *C. difficile* in the supernatants of patient feces, using both antitoxin-protected and nonprotected cell monolayers (2). This assay is commonly used as the gold standard method for comparison in toxin kit evaluations, although its use in routine microbiology laboratories has largely been superseded. Cytotoxigenic culture (CYTGC) has been used as an alternative gold standard method to CYT testing, i.e., where CYT testing is performed using culture supernatants instead of directly from the fecal sample (1). These are

lengthy assays, however, with results delayed for 24 to 48 h for the CYT and for more than 72 h for the CYTGC assay.

Rapid, commercially available, toxin detection kits removed the need for laboratories to maintain the cell lines necessary for CYT testing. Although originally designed to detect either toxin A or toxin B, the kits currently available detect both toxins to enable detection of toxin A-negative, toxin B-positive strains. Alternative detection methods have now been developed, including an assay that detects a surface-associated enzyme of *C. difficile*, glutamate dehydrogenase (GDH). Zheng et al. reported that the Techlab *C. diff* Chek-60 GDH assay had good sensitivity compared to CYT testing of 92%, but it had a low specificity of 89.1% and poor positive predictive value (PPV) of 57.7% (21). Commercial molecular diagnostic tests, such as the BD GeneOhm *C. difficile* PCR assay, which detects the *tcdB* toxin gene of *C. difficile*, are now available. A recent study compared this assay to CYT testing and found a sensitivity and specificity of 90.9% and 95.2%, respectively (15). The PPVs of the BD GeneOhm *C. difficile* PCR assay were only 70.2% compared with CYT testing and 89.5% compared with CYTGC (15), with a prevalence of toxin-positive fecal samples of 15.2%.

Despite numerous evaluations of *C. difficile* testing methods, no evaluation has compared all methods on the same sample set. This study compared six commercially available enzyme immunoassays (EIAs) and three lateral-flow assays for detection of *C. difficile* toxins A and B, a PCR assay for detection of the *tcdB* gene of *C. difficile*, and an assay for detection of *C. difficile*-specific GDH, with CYT testing and CYTGC.

* Corresponding author. Mailing address: Microbiology Department, Old Medical School, Leeds General Infirmary, Leeds LS1 3EX, W. Yorks., United Kingdom. Phone: 44 113 392 6818. Fax: 44 113 343 5649. E-mail: mark.wilcox@leedsth.nhs.uk.

[▽] Published ahead of print on 26 August 2009.

TABLE 1. *C. difficile* detection assays included in the evaluation in this study

Assay type	Assay	Target(s)	Supplier
Well-type EIA	Premier toxin A+B	Toxins A and B	Meridian
	GA <i>Clostridium difficile</i> antigen	Toxins A and B	The Binding Site
	Ridascreen <i>Clostridium difficile</i> toxin A/B	Toxins A and B	Biopharm
	Toxin A/B II	Toxins A and B	Techlab
	Remel ProSpecT	Toxins A and B	Oxoid
Automated immunoassay	Vidas <i>C. difficile</i> Tox A/B	Toxins A and B	Vidas
Membrane assay	Remel Xpect	Toxins A and B	Oxoid
	Tox A/B Quik Chek	Toxins A and B	Techlab
	Premier Immunocard Stat toxin A and B	Toxins A and B	Meridian
Well-type EIA	C. diff Chek-60	GDH	Techlab
PCR	GeneOhm <i>C. difficile</i>	<i>tcdB</i> gene	BD

MATERIALS AND METHODS

Ethical approval. Ethical approval for this study was granted from the North Sheffield Research Ethics committee. Part of this study was completed on behalf of the Centre of Evaluations and Purchasing, part of the Purchasing and Supply Agency of the United Kingdom National Health Service (NHS) (18).

Sample selection. Ten fecal samples were selected daily for inclusion in the study, between April and September 2008, from those received by the routine enteric laboratory at the Microbiology Department of Leeds Teaching Hospitals NHS Trust. All samples included were diarrheal (adopting the shape of the container), <48 h old, submitted for CYT testing, had been stored at 2 to 5°C, and had sufficient fecal material to allow testing with all the assays (4). The 10 daily samples were randomly chosen from each daily cohort of samples (~20 to 25) that had sufficient volume of fecal matter to permit multiple testing by the biomedical scientist working in the enteric laboratory. Each sample was dispensed into a new container labeled with only a study number, with no link to patient details, before the repackaged samples were distributed to the investigators. To increase the statistical robustness of the study, the number of positive samples was increased by including samples from patients with a previous *C. difficile*-positive fecal sample ($n = 2$) and samples identified as toxin positive at another hospital ($n = 13$). Specimens were stored at 4°C for 1 week (in case retesting was required) and then frozen at -20°C.

Supply of kits and equipment. After identifying all commercially available *C. difficile* toxin detection kits in the United Kingdom, manufacturers were approached and asked to supply the kits and supplementary equipment needed for the study. No manufacturer approached declined to be included in the study (Table 1). The Vidas *C. difficile* toxin A and B assay was performed on the mini Vidas, provided by Biomérieux, and the Premier toxin A + B assay was performed on the DS2 instrument, provided by Launch Diagnostics. An automated washer (Wellwash; Labsystems) was used to complete the wash steps in the four manual toxin detection EIAs, Techlab Toxin A/B II, Remel ProSpecT, Ridascreen toxin A/B, and GA *Clostridium difficile* antigen, and the manual GDH EIA Techlab C. diff Chek-60 kit. The final optical density (OD) values for these five assays were read on a microplate reader (230s; Organon Teknika) at the wavelength specified for the assay. The PCR assay GeneOhm was performed on the Smartcycler (Cepheid, United Kingdom; supplied by BD at the time of the study). All assays were performed according to the manufacturers' operational instructions, although use of freeze-thawed samples for the PCR and GDH assays is not recommended.

Testing protocol. Every sample was tested using each toxin detection assay and by *C. difficile* culture. All *C. difficile* toxin detection assays were performed on the same day for each batch of 10 samples. A prestudy evaluation elucidated the optimum order in which to perform the assays. Selected samples were tested using the PCR ($n = 554$) and GDH ($n = 558$) assays, due to lack of sample volume in a minority of cases. The PCR and GDH assays were performed together at a later date on batched, previously frozen fecal samples (stored for <8 months at -20°C) that had never been defrosted before this occasion.

Cytotoxin assay. Fecal samples were diluted 1:5 in phosphate-buffered saline and centrifuged before 20 μ l of supernatant was added to duplicate Vero cell monolayers, one of which had been protected by the addition of 20 μ l *Clostridium sordelli* antitoxin (Prolab Diagnostics, United Kingdom). Vero cells were grown in 96-well flat-bottomed microtiter trays in 160 μ l of Dulbecco medium.

Samples were filtered before testing if the supernatant was cloudy. A positive result was recorded if cell rounding was seen, only in the unprotected cells, after 24 or 48 h of incubation in a 37°C CO₂ incubator.

Culture. All samples were cultured, following alcohol shock in 50:50 [vol/vol] absolute ethanol and water, on CCEYL (Braziers CCEYL agar base; Bioconnections, Wetherby, United Kingdom), supplemented with 8 mg/liter cefoxitin and 250 mg/liter cycloserine, 2% lysed horse blood (E & O Laboratories, United Kingdom), and 5 mg/liter lysozyme (Sigma, United Kingdom). Egg yolk was not added as a supplement. Plates were incubated in a MK3 anaerobic workstation (Don Whitley, Shipley, United Kingdom) and inspected for growth after 48 h. Gray-brown colonies growing on CCEYL with an irregular edge and a characteristic horse manure odor were identified as *C. difficile*. The Microgen *C. difficile* latex agglutination kit (Microgen Bioproducts Ltd., Camberley, United Kingdom) was used for confirmation of *C. difficile* identity where there was doubt about an isolate.

Spores were harvested into 10% glycerol broth from isolates subcultured onto fresh blood agar (FBA [E & O Laboratories, United Kingdom]), incubated in the anaerobic cabinet at 35°C for 7 days, and stored at -70°C. All isolates were typed by PCR ribotyping following the protocol from the *Clostridium difficile* Ribotyping Network for England (CDRNE) laboratory which is situated in the same department.

Cytotoxicigenic culture. *C. difficile* isolates that grew on CCEYL from CYT-negative samples were tested for the production of toxin. Isolates were inoculated into brain heart infusion broth (BHI) (Oxoid, Basingstoke, United Kingdom) and incubated for 48 h in an anaerobic cabinet. Culture supernatants were centrifuged before adding to duplicate Vero cell monolayers, one protected by *C. sordelli* antitoxin. A positive result was recorded if cell rounding was seen only in the unprotected cells after 24 or 48 h of incubation in a 37°C CO₂ incubator.

Result interpretation. Optical density readings were recorded for the manual EIAs and compared with the cutoff value for the assay to determine the result. Three assays used a fixed cutoff value; any sample with an OD above this cutoff was positive. Results were recorded as equivocal if the OD was equal to the cutoff value. For the remaining two EIAs, Ridascreen toxin A/B and GA *Clostridium difficile* antigen, a correction factor set by the manufacturer had to be added to the OD value for the negative control each time the assay was performed to calculate the cutoff. Any value higher than $1.1 \times$ cutoff was recorded as positive; below $0.9 \times$ cutoff was recorded as negative. Any value between 0.9 and $1.1 \times$ cutoff was recorded as equivocal. The lateral-flow assays were read visually by three evaluators to ensure an objective result. Results were recorded as positive (line or color change), negative (no line or color change), or equivocal (if unclear). When the operators could not agree, the majority result was recorded. The Vidas and DS2 instruments used algorithms set by the manufacturer to calculate the result and were recorded as positive, negative, or equivocal. The software on the Smartcycler (Cepheid, United Kingdom) recorded the results of the PCR assay as positive, negative, or unresolved.

Discordant results. Samples with equivocal results from the toxin detection kits were rediluted and retested; if a result was still equivocal, it was recorded as such. Specimens that gave discordant results when tested by the assays under evaluation were retested in duplicate on the same specimen, where sufficient specimen was available, to exclude the possibility of technical error. The BD GeneOhm *C. difficile* PCR assay was repeated when the result was unresolved, as

TABLE 2. Comparison of commercial *C. difficile* detection assays with results of cytotoxin testing

Assay	No. of samples with the following assay result:				Cytotoxin test result	No. of samples tested	Sensitivity (%) (95% CI) ^a	Specificity (%) (95% CI) ^a
	Positive	Negative	Equivocal	Invalid				
Premier toxin A+B	99	9	0	0	Positive	108	91.7 (84.7–96.1)	97.1 (95.1–98.4)
	14	474	0	0	Negative	488		
GA <i>Clostridium difficile</i> antigen	83	25	0	0	Positive	108	76.8 (67.7–84.4)	90.9 (88.0–93.3)
	44	444	0	0	Negative	488		
Ridascreen toxin A/B	72	36	0	0	Positive	108	66.7 (56.9–75.4)	95.1 (92.6–96.7)
	23	464	1	0	Negative	488		
Techlab toxin A/B II	98	10	0	0	Positive	108	90.7 (83.6–95.5)	95.7 (93.4–97.3)
	21	467	0	0	Negative	488		
Remel ProSpecT	97	11	0	0	Positive	108	89.8 (82.5–94.8)	92.6 (89.8–94.7)
	36	452	0	0	Negative	488		
Vidas <i>C. difficile</i> toxin A/B	97	2	9	0	Positive	108	89.8 (82.5–94.8)	96.7 (94.6–98.0)
	5	472	11	0	Negative	488		
Remel Xpect	84	17	7	0	Positive	108	77.8 (68.8–85.2)	98.8 (97.2–99.5)
	5	482	1	0	Negative	488		
Techlab Tox A/B Quik Chek	91	17	0	0	Positive	108	84.3 (76.0–90.6)	98.6 (96.9–99.4)
	5	481	1	1	Negative	488		
Premier Immunocard A + B	84	15	0	9	Positive	108	77.8 (68.8–85.2)	92.8 (90.1–94.9)
	4	453	0	31	Negative	488		
Techlab <i>C. diff</i> Chek-60	82	9	0	0	Positive	91	90.1 (81.6–95.1)	92.9 (90.1–95.0)
	33	434	0	0	Negative	467		
BD GeneOhm <i>C. difficile</i>	83	7	0	0	Positive	90	92.2 (84.1–96.6)	94.0 (91.3–95.9)
	22	436	0	6	Negative	464		

^a The sensitivity and specificity of the assays compared to the gold standard cytotoxin test results are shown. The sensitivity and specificity are given as percentages, and the 95% confidence interval (95% CI) values are shown in parentheses.

recommended by the manufacturer. None of the samples were repeated on the Techlab *C. diff* Chek-60 GDH kit (due to insufficient sample).

Statistical analysis. Sensitivity and specificity were calculated for each kit against both gold standard assays (CYT and CYTGC). The difference in both sensitivity and specificity between each pair of toxin detection assays was determined using McNemar's test for paired proportions, with exact binomial *P* values, due to the potentially small numbers of discordant samples. The sensitivity and specificity data were used to calculate the PPV and negative predictive value (NPV) for different prevalence rates of *C. difficile* toxin-positive fecal samples to reflect the prevalences seen in the community and hospital settings. No statistical comparisons were performed between the toxin detection assays and the PCR and GDH assays, as these measure different targets.

RESULTS

Six hundred fecal samples were included in the evaluation, of which 108 were positive by CYT and 125 were positive by CYTGC. All 600 samples were tested by the toxin detection kits, but due to insufficient sample volume, only 558 and 564 samples were tested using the BD GeneOhm PCR and Techlab *C. diff* Chek-60 GDH assays, respectively. Four samples were removed from the CYT gold standard analysis, as these samples were considered to be false-positive CYT results and may have skewed the data. Sensitivity and specificity data against the CYT gold standard are shown in Table 2, and sensitivity and specificity data against CYTGC are shown in Table 3. The CYT was also evaluated in comparison with CYTGC (Table 3).

Sensitivity and specificity of toxin detection assays. Statistical analysis showed that the Premier toxin A+B, Vidas *C. difficile* Tox A/B, Techlab toxin A/B II, and Remel ProSpecT assays were more sensitive ($P = >0.05$) than the other five

toxin detection assays (Tables 2 and 3; statistical data not shown). The Ridascreen toxin A/B assay was the least sensitive assay compared against either gold standard method, while the GA *Clostridium difficile* antigen assay was the least specific (Tables 2 and 3). The two lateral-flow assays demonstrated better specificity than any of the EIAs (Tables 2 and 3), although the differences were not significant (statistical data not shown). The cytotoxin assay had higher sensitivity and specificity values than any of the commercial toxin detection assays, except the Remel Xpect, when compared with CYTGC (Tables 2 and 3).

Sensitivity and specificity of the PCR assay. The BD GeneOhm PCR assay was more sensitive than any of the toxin detection assays in comparison with both gold standard methods (Tables 2 and 3). The assay did not perform as well for specificity, however, being in the middle of the range of specificity values seen with the toxin detection kits (Tables 2 and 3).

Sensitivity and specificity of the GDH detection assay. The sensitivity of the Techlab GDH assay was similar to that of the Techlab toxin detection assay (which had the highest sensitivity of the EIAs), when using the CYT gold standard, but it is less specific (Table 2). When assessed against CYTGC, the GDH assay is more sensitive than the Techlab toxin detection assay, but it remains less specific (Table 3).

Positive and negative predictive values. Table 4 shows the changing PPV values for each assay tested in both the community and the hospital setting (2% and 10% prevalence of *C. difficile*-positive fecal samples, respectively). When CYT is

TABLE 3. Comparison of commercial *C. difficile* detection assays with results of cytotoxigenic culture

Assay	No. of samples with the following assay result:				Cytotoxigenic culture result	No. of samples tested	Sensitivity (%) (95% CI) ^a	Specificity (%) (95% CI) ^a
	Positive	Negative	Equivocal	Invalid				
Cytotoxin test	108	17	0	0	Positive	125	86.4 (79.1–91.9)	99.2 (97.9–99.8)
	4	471	0	0	Negative	475		
Premier toxin A+B	101	24	0	0	Positive	125	80.8 (72.3–87.3)	97.5 (95.5–98.6)
	12	463	0	0	Negative	475		
GA <i>Clostridium difficile</i> antigen	86	39	0	0	Positive	125	68.8 (59.9–76.8)	91.4 (88.4–93.7)
	41	454	0	0	Negative	475		
Ridascreen toxin A/B	75	50	0	0	Positive	125	60.0 (50.9–68.7)	95.6 (93.2–97.2)
	20	434	1	0	Negative	475		
Techlab toxin A/B II	100	25	0	0	Positive	125	80.0 (71.9–86.6)	96.0 (93.7–97.5)
	19	456	0	0	Negative	475		
Remel ProSpecT	102	23	0	0	Positive	125	81.6 (73.7–87.9)	93.3 (90.6–95.4)
	32	443	0	0	Negative	475		
Vidas <i>C. difficile</i> Tox A/B	100	16	9	0	Positive	125	80.0 (71.9–86.6)	97.3 (95.2–98.5)
	2	462	11	0	Negative	475		
Remel Xpect	86	31	8	0	Positive	125	68.8 (59.9–76.8)	99.4 (98.2–99.9)
	3	472	0	0	Negative	475		
Techlab Tox A/B Quik Chek	93	32	0	0	Positive	125	74.4 (65.8–81.78)	98.9 (97.6–99.7)
	3	470	1	1	Negative	475		
Premier Immunocard A + B	86	29	0	10	Positive	125	68.8 (59.9–76.8)	93.0 (90.4–95.2)
	2	442	0	31	Negative	475		
Techlab <i>C. diff</i> Chek-60	92	13	0	0	Positive	105	87.6 (72.4–93.0)	94.3 (91.7–96.2)
	26	433	0	0	Negative	459		
BD GeneOhm <i>C. difficile</i>	92	11	0	1	Positive	104	88.5 (80.3–93.6)	95.4 (92.9–97.0)
	16	433	0	5	Negative	454		

^a The sensitivity and specificity of the assays compared to the gold standard cytotoxigenic culture results are shown. The sensitivity and specificity are given as percentages, and the 95% confidence interval (95% CI) values are shown in parentheses.

used as the gold standard, the two lateral-flow assays give the highest PPVs across the range of prevalence values. It should be noted, however, that at a prevalence of 2%, the highest PPV is still only 56.3%, rising to 87.5% at 10% prevalence. When CYTGC is used as the gold standard, the CYT and the Remel

Xpect lateral-flow assay give the highest PPVs, 92.0% and 92.4% for 10% prevalence, respectively.

The NPVs for all the assays were much higher than the PPVs, ranging from 99.3 to 99.8% at 2% prevalence to 96.3 to 99.1% at 10% prevalence, against the CYT gold standard (Table 4).

TABLE 4. Positive and negative predictive values of assays in comparison with both gold standard methods at both low (2%) and high (10%) prevalence of *C. difficile*

Assay	PPV and NPV values compared to cytotoxin testing at the indicated prevalence of <i>C. difficile</i>				PPV and NPV values compared to cytotoxigenic culture at the indicated prevalence of <i>C. difficile</i>			
	PPV		NPV		PPV		NPV	
	2%	10%	2%	10%	2%	10%	2%	10%
Cytotoxin test					67.7	92.0	99.7	98.5
Premier toxin A+B	39.5	78.0	99.8	99.1	39.5	78.0	99.6	97.9
GA <i>Clostridium difficile</i> antigen	14.8	48.6	99.5	97.3	14.0	47.0	99.3	96.3
Ridascreen toxin A/B	21.7	60.1	99.3	96.3	21.7	60.1	99.2	95.6
Techlab toxin A/B II	30.1	70.1	99.8	98.9	29.0	69.0	99.6	97.7
Remel ProSpecT	19.9	57.5	99.8	98.8	19.8	57.4	99.6	97.9
Vidas <i>C. difficile</i> Tox A/B	35.9	75.3	99.8	98.8	37.3	76.4	99.6	97.8
Remel Xpect	56.3	87.5	99.5	97.6	69.0	92.4	99.4	96.6
Techlab Tox A/B Quik Chek	54.6	86.8	99.7	98.3	59.1	88.7	99.5	97.2
Premier Immunocard A + B	18.1	54.7	99.5	97.4	16.8	52.4	99.3	96.4
Techlab <i>C. diff</i> Chek-60	20.6	58.6	99.8	98.8	24.0	63.1	99.7	98.6
BD GeneOhm <i>C. difficile</i>	23.8	63.0	99.8	99.1	28.1	68.0	99.7	98.7

Discordant and equivocal results for commercial toxin detection assays. All nine of the commercial toxin detection assays tested gave discordant results against both gold standards ($n = 23$ to 69 against CYT and $n = 29$ to 65 against CYTGC). In the majority of cases, these results were unaffected by repeat testing (mean of 68.6% and range of 48.9 to 93.3% against CYT; mean of 66.5% and range of 28.1 to 95.0% against CYTGC). There were 30 equivocal results generated by the commercial toxin detection assays, of which 19 (63%) remained equivocal on repeat testing.

Discordant results for BD GeneOhm PCR assay. Initial and repeat unresolved rates were 1.1% (6/554) and 0%, respectively. Five of the six samples that gave an “unresolved” result on first testing gave a negative result on repeat testing, with one sample giving a positive result. These results matched the CYTGC results for these samples. Of the 16 samples that gave a false-positive result against CYTGC, two could not be repeated due to lack of sample, five were repeated and gave a negative result, and 10 were repeated and gave a positive result. Two of the 10 repeat false-positive results were positive by CYT, but these results had been regarded as false CYT positives in the original analysis, as no other test, including culture, was positive for these samples.

Optical density values for commercial EIA well-based toxin detection assays. The OD values for each sample, in each of the commercial toxin detection assays, were recorded and initially plotted in scatter grams to show the distributions of false-positive and false-negative results around the respective assay cutoff values (data not shown). The results demonstrated that the samples that yielded false-positive and false-negative results varied markedly between assays. Thus, it was not a core group of samples that repeatedly yielded false-positive results. The data also show that the false-positive results were not solely associated with samples that yielded low OD results in some assays.

Ribotypes of isolates. There were 128 culture-positive samples, 125 of which were CYT positive. The most common ribotypes identified during this evaluation were 106 (26.6%), 027 (18.8%), and 002 (6.3%). The total numbers of each ribotype were 001 ($n = 4$), 002 ($n = 8$), 003 ($n = 1$), 005 ($n = 7$), 014 ($n = 1$), 014/20 ($n = 5$), 015 ($n = 7$), 018 ($n = 3$), 023 ($n = 3$), 027 ($n = 24$), 044 ($n = 1$), 050 ($n = 2$), 054 ($n = 2$), 070 ($n = 1$), 078 ($n = 4$), 084 ($n = 1$), 094 ($n = 1$), 106 ($n = 34$), 118 ($n = 4$), 140 ($n = 1$), and sporadic types ($n = 14$). There did not appear to be any differences between the assays at identifying positive results in different ribotypes, but the numbers were too small to calculate statistical significance (data not shown).

DISCUSSION

There have been many reports on the performance of *C. difficile* (toxin, GDH, or toxin gene) detection kits, but these invariably involve a very small number of comparator methods (11). Planche et al. recently carried out a systematic review of such studies and concluded that single assays had unacceptably low PPVs (11). Our evaluation, which we understand is the largest study of *C. difficile* detection methods performed up to this point, supports the findings of Planche et al. Thus, combining the results from multiple small studies provides a very similar outcome to our evaluation. NPVs are generally much

higher in both our evaluation and the recent systematic review, as would be expected when using routine diagnostic samples, the great majority of which will yield negative results for *C. difficile*, despite the presence of diarrhea. Importantly, the PPV of any test is affected by the prevalence of the disease in the population (i.e., by the test positivity rate). In most of the published comparison studies, the prevalence of toxigenic *C. difficile*-positive samples is higher than routinely seen in the clinical setting (9, 14, 15), leading to falsely inflated PPVs (11). Although our study had a prevalence of toxigenic *C. difficile*-positive samples of 20%, we used our data to calculate the PPV and NPV with changing levels of prevalence. In the hospital setting, with an expected prevalence of 10%, the mean PPV of a commercial toxin detection kit was 68.7% (range, 48.6 to 86.8%), which is comparable to the findings of Planche et al. (11). In the community setting, the prevalence is nearer 2% (19); as such, the mean PPV falls to 32.3% (range, 14.8 to 56.3%), making testing for *C. difficile* using current methods, especially single tests, extremely unreliable. The scatter grams of the OD values for the EIA toxin detection assays indicate that the false-positive and false-negative results were generally not seen in the same sample tested by different assays (data not shown). This implies that in general the incorrect results were due to the inaccuracy of the assay and not due to the sample or the evaluator. Although only one evaluator processed the samples on each day, two different evaluators worked on the project, and the number of discordant results was similar between them (data not shown).

As a result of increased awareness of *C. difficile* infection, including the epidemic spread of virulent *C. difficile* clones, additional testing has likely occurred. A recent United Kingdom survey showed that testing for *C. difficile* has increased by 39% in the microbiology laboratories of 18 hospitals; the median number of tests increased from 3,613 in 2005 and 2006 to 5,020 in 2007 and 2008 (8). Consequently, if the prevalence of true-positive samples tested decreases, this potentially increases the chance of obtaining a false-positive result using methods that have suboptimal accuracy. Furthermore, many laboratories have abandoned CYT testing in favor of the newer rapid toxin detection assays. As the reference laboratory for the CDRNE, we noted a high level of toxin detection assay-positive samples being sent for ribotyping that were culture negative for *C. difficile*. Conversely, local data show that we successfully isolate *C. difficile* from >95% of CYT-positive samples. We therefore became concerned about the apparent level of false-positive results for toxin detection kits, which stimulated our interest in conducting this study.

It should be noted that the clinical criteria used in definitions of *C. difficile* infection vary, notably, the frequency of diarrhea that is needed to satisfy a positive case diagnosis. Interpretation of clinical symptoms will clearly be made more difficult by inaccurate laboratory results, thus potentially affecting patient management. A false-positive result may lead to unnecessary treatment and isolation. The true cause of the patient's diarrhea may also not be further investigated if a diagnosis of *C. difficile* infection is made. In hospitals where *C. difficile* infection patients are isolated together within one ward, due to insufficient availability of single-room isolation facilities, a false-positive result could lead to a patient being at increased risk of cross-infection from patients who have true *C. difficile*

infections. Conversely, false-negative results may lead to cross-infection to other patients and overtreatment with empirical antibiotics. This has implications for patients not receiving appropriate treatment and for the hospital.

It is important to note that there is not a universally accepted method for the CYT. Sample selection and preprocessing, choice of cell line, and different interpretive end points may all affect test performance. Such issues, coupled with the long waiting time for a test result and the need to maintain a cell line, have contributed to the decreased availability of the CYT in diagnostic laboratories. In this evaluation, the CYT was the best performing toxin detection assay, compared with CYTGC, with a sensitivity and specificity of 86.4 and 99.2%, respectively. The reported performance of the CYT has, however, been variable. Peterson et al. reported poor sensitivity of the CYT (76.7%) using a gold standard of at least two positive tests plus the presence of diarrhea; such discordant analysis may affect the calculated sensitivity of the CYT (10). Stamper et al. also reported a low sensitivity of 67.2% for CYT (Wampole tox-B assay; Techlab) in comparison with CYTGC (15). Higher sensitivity and specificity for the CYT (98% and 99%, respectively) have also been reported (9). Interpretation of such published results is clearly difficult, however, given the difficulty of assessing the accuracy of a gold standard method. For example, comparison of results of a CYT with CYTGC may be misleading, as although these tests both measure toxin, the former samples the toxin present in vivo, while the latter involves in vitro production.

The BD GeneOhm PCR assay was more specific than the toxin detection kits, as would be expected for a molecular detection method. The sensitivity, specificity, PPV, and NPV found in this evaluation are comparable to those found by Stamper et al. when comparing the GeneOhm PCR assay to both CYT testing and CYTGC (15). The PCR assay lacks specificity, however, leading to low PPVs, comparable with the toxin detection assays. The PCR assay has potential as a negative screening assay, as it has the highest NPV (99.1%) of any assay in this evaluation (at 10% prevalence, versus CYT). Other PCR assays have been developed and are becoming commercially available. Sloan et al. used a PCR assay to detect the *tcdC* gene of *C. difficile* and found sensitivity and specificity values of 86% and 97%, respectively, which are comparable with those in our study (14). The reported PPV of 90% is much higher than we found for the BD assay, but their study had a prevalence of toxin-positive stool of 22%, approximately double that found in hospital clinical samples (14). Any PCR assay can detect only the presence of the toxin gene, however, not the presence of toxin, and the relevance of a positive result has yet to be elucidated. Peterson et al. evaluated a real-time PCR assay for *tcdB* detection in diarrheal samples and found good correlation with the clinical status of the patients with a PPV of 75.7% and NPV of 99.4% when the prevalence of toxin-positive samples was 13.1% (10). The clinical details of the patients in our study were unavailable, as the samples were placed into new containers with no link to patient details before testing was performed. It was therefore not within the scope of this evaluation to investigate the diagnosis of those patients in whom only the PCR assay was positive. Isolates in this evaluation were typed using PCR ribotyping. The ribotypes seen and their proportions are representative of the local epidemiology.

There did not appear to be any difference in the abilities of the assays to detect different ribotypes of *C. difficile*.

The GDH detection assay (Techlab *C. diff* Chek-60) also has potential as a negative screening tool with an NPV of 98.8% (at 10% prevalence, versus CYT). An assay PPV of 63.1% (at 10% prevalence, versus CYTGC) was comparable with those of the toxin detection assays and echoes the results of Zheng et al. (21). The GDH assay did not perform as well compared to the CYT. Clearly, the assays are detecting different targets; one detects an enzyme of *C. difficile* (GDH), and the other detects toxins produced by the organism. Although GDH may be useful in establishing the presence of *C. difficile* in a fecal sample, it does not indicate whether the organism has the potential to cause disease, as it cannot detect the presence of toxin. This assay can therefore be used only as part of a two-step testing algorithm. Fenner et al. showed that a rapid two-step algorithm using GDH followed by toxin testing could produce a result for 92% of samples in 4 h (3). However, GDH assays are reported to have variable sensitivities. In comparison with CYTGC, the Techlab GDH test had a sensitivity of 88% in our hands (using freeze-thawed samples). Using a different GDH assay (Triage [no longer available]), Sloan et al. reported a sensitivity of only 76% for *C. difficile* culture-positive (fresh or freeze-thawed) fecal samples (14).

Two- or three-step approaches to the diagnosis of *C. difficile* infection, e.g., rapid detection of toxin, bacterium, or toxin gene, with subsequent confirmation of the presence of toxin, will increase laboratory costs, but these might be offset by reduced total health care costs for *C. difficile* infection. A two-step approach will also increase the time for a final result, but this may be acceptable, particularly if interim results are made available. The algorithms tested so far include GDH testing as the first, negative screening step, due to its high NPV. In the algorithm used by Wren et al., the GDH assay performed well as the initial screen, with only four false-negative results (20). The authors hypothesize that these are due to colonization with *C. difficile*, rather than *C. difficile* infection, as the lactoferrin test (a marker of inflammatory diarrhea) gave negative results in these samples (20). There has been no consensus on the best assay for the second step, with studies using EIA toxin assays, cell culture CYTs, and toxin detection plus lactoferrin detection (3, 5, 12, 20). Algorithms including a toxin detection EIA as the second, confirmatory test, may not be optimal, as it relies on an assay which we have shown to have poor PPV. Published studies have compared one algorithm to the current methods of single assay diagnosis. There has been no evaluation of different combinations of assays, and further work is required to assess which assays would provide the optimal two-step laboratory diagnosis of *C. difficile* infection.

Planche et al. indicated that many of the published comparative studies made modifications to the manufacturer's instructions and used various methods for the cell CYT which makes comparison between studies difficult (11). We ensured that all assays in this evaluation were performed according to manufacturers' instructions and compared all assays to one method of CYT. We have conducted the first study to compare all the currently available (in the United Kingdom) *C. difficile* detection kits on the same sample set. Previous comparison studies have used smaller sample sets and high prevalence rates of *C.*

difficile-positive samples, not reflective of the clinical setting (9, 14, 15). All kits had low PPVs compared with CYT, assuming a positive sample prevalence of 10% (representing the hospital setting), which compromises the clinical utility of single tests for the laboratory diagnosis of *C. difficile* infection. As the prevalence of *C. difficile* infection decreases, this will exacerbate the issue of false-positive results in suboptimal assays. The optimum rapid single test was PCR for toxin B gene, as this had the highest negative predictive value. Diagnostic algorithms that optimize test combinations for the laboratory diagnosis of *C. difficile* infection need to be defined.

ACKNOWLEDGMENTS

We thank all of the suppliers involved for supplying kits and equipment for the evaluation. Part of this evaluation was funded by the Centre of Evaluations and Purchasing, Purchasing and Supply Agency, National Health Service (NHS).

We also thank Keith Perry at the Health Protection Agency's Microbiological Diagnostics Assessment Service, London, United Kingdom, for invaluable advice.

REFERENCES

- Bouza, E., T. Peláez, R. Alonso, P. Catalán, P. Muñoz, and M. Rodríguez Créixems. 2001. "Second look" cytotoxicity: an evaluation of culture plus cytotoxin assay of *Clostridium difficile* isolates in the laboratory diagnosis of CDAD. *J. Hosp. Infect.* **48**:233–237.
- Chang, T. W., J. G. Bartlett, S. L. Gorbach, and A. B. Onderdonk. 1978. Clindamycin-induced enterocolitis in hamsters as a model of pseudomembranous colitis in patients. *Infect. Immun.* **20**:526–529.
- Fenner, L., A. F. Widmer, G. Goy, S. Rudin, and R. Frei. 2008. Rapid and reliable diagnostic algorithm for detection of *Clostridium difficile*. *J. Clin. Microbiol.* **46**:328–330.
- Freeman, J., and M. H. Wilcox. 2003. Effects of storage conditions on viability of *Clostridium difficile* vegetative and spore cells and toxin activity in human faeces. *J. Clin. Pathol.* **56**:126–128.
- Gilligan, P. H. 2008. Is a two-step glutamate dehydrogenase antigen-cytotoxicity neutralization assay algorithm superior to the Premier toxin A and B enzyme immunoassay for laboratory detection of *Clostridium difficile*? *J. Clin. Microbiol.* **46**:1523–1525.
- Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. René, Y. Monczak, and A. Dascal. 2005. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. *N. Engl. J. Med.* **353**:2442–2449.
- McDonald, L. C., G. E. Killgore, A. Thompson, R. C. Owens, Jr., S. V. Kazakova, S. P. Sambol, S. Johnson, and D. N. Gerding. 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N. Engl. J. Med.* **353**:2433–2441.
- National Pathology Benchmarking Service. 2008. Microbiology benchmarking report 2007–2008. National Pathology Benchmarking Service, Keele University, Keele, United Kingdom.
- O'Connor, D., P. Hynes, M. Cormican, E. Collins, G. Corbett-Feeney, and M. Cassidy. 2001. Evaluation of methods for detection of toxins in specimens of feces submitted for diagnosis of *Clostridium difficile*-associated diarrhea. *J. Clin. Microbiol.* **39**:2846–2849.
- Peterson, L. R., R. U. Manson, S. M. Paule, D. M. Hacek, A. Robicsek, R. B. Thomson, Jr., and K. L. Kaul. 2007. Detection of toxigenic *Clostridium difficile* in stool samples by real-time polymerase chain reaction for the diagnosis of *C. difficile*-associated diarrhoea. *Clin. Infect. Dis.* **45**:1152–1160.
- Planche, T., A. Aghaizu, R. Holliman, P. Riley, J. Poloniecki, A. Breathnach, and S. Krishna. 2008. Diagnosis of *Clostridium difficile* infection by toxin detection kits: a systematic review. *Lancet Infect. Dis.* **8**:777–784.
- Reller, M. E., C. A. Lema, T. M. Perl, M. Cai, T. L. Ross, K. A. Speck, and K. C. Carroll. 2007. Yield of stool culture with isolate toxin testing versus a two-step algorithm including stool toxin testing for the detection of toxigenic *Clostridium difficile*. *J. Clin. Microbiol.* **45**:3601–3605.
- Rupnik, M., D. Gerding, and M. H. Wilcox. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat. Rev. Microbiol.* **7**:526–536.
- Sloan, L. M., B. J. Duresko, D. R. Gustafson, and J. E. Rosenblatt. 2008. Comparison of real-time PCR for detection of the *tcdC* gene with four toxin immunoassays and culture in diagnosis of *Clostridium difficile* infection. *J. Clin. Microbiol.* **46**:1996–2001.
- Stamper, P. D., R. Alcabasa, D. Aird, W. Babiker, J. Wehrin, I. Ikepama, and K. C. Carroll. 2009. Comparison of a commercial real-time PCR assay for *tcdB* detection to a cell culture cytotoxicity assay and toxigenic culture for direct detection of toxin-producing *Clostridium difficile* in clinical samples. *J. Clin. Microbiol.* **47**:373–378.
- Reference deleted.
- Reference deleted.
- Wilcox, M. H., and K. A. Eastwood. 2009. Evaluation report: *Clostridium difficile* toxin detection assays. Centre for Evidence-based Purchasing publication no. CEP08054. Centre for Evidence-based Purchasing, NHS Purchasing and Supplies Agency, National Health Service, London, United Kingdom. <http://www.pasa.nhs.uk/pasa/Doc.aspx?Path=%5bMN%5d%5bSP%5d/NHSprocurement/CEP/CEP08054.pdf>.
- Wilcox, M. H., L. Mooney, R. Bendall, C. D. Settle, and W. N. Fawley. 2008. A case-control study of community-associated *Clostridium difficile* infection. *J. Antimicrob. Chemother.* **62**:388–396.
- Wren, M. W. D., M. Sivapalan, R. Kinson, and N. P. Shetty. 2009. Laboratory diagnosis of *Clostridium difficile* infection. An evaluation of tests for faecal toxin, glutamate hydrogenase, lactoferrin and toxigenic culture in the diagnostic laboratory. *Br. J. Biomed. Sci.* **66**:1–5.
- Zheng, L., S. F. Keller, D. M. Lyerly, R. J. Carman, C. W. Genheimer, C. A. Gleaves, S. J. Kohlhepp, S. Young, S. Perez, and K. Ye. 2004. Multicenter evaluation of a new screening test that detects *Clostridium difficile* in fecal specimens. *J. Clin. Microbiol.* **42**:3837–3840.