Open Forum Infectious Diseases

BRIEF REPORT

Using Dog Scent Detection as a Point-of-Care Tool to Identify Toxigenic *Clostridium difficile* in Stool

Maureen T. Taylor,[1,2] Janine McCready,[1,2] George Broukhanski,[3,4] Sakshi Kirpalaney,[5] Haydon Lutz,[5] and Jeff Powis[1,2]

[1]Division of Infectious Diseases, Department of Medicine, Michael Garron Hospital, Toronto, Ontario, Canada; [2]Faculty of Family Medicine, McMaster University, Hamilton, Canada; [3]Public Health Ontario Laboratories, Toronto, Canada; [4]Department of Laboratory Medicine and Pathobiology, University of Toronto, Canada; [5]Royal College of Surgeons in Ireland School of Medicine, Dublin, Ireland

We evaluated the operating characteristics of 2 comparably trained dogs as a "point-of-care" diagnostic tool to detect toxin gene-positive *Clostridium difficile*. Although each dog could detect toxin gene-positive *C difficile* in stool specimens with sensitivities of 77.6 and 92.6 and specificities of 85.1 and 84.5, respectively, interrater reliability is only modest (Cohen's kappa 0.52), limiting widespread application.

Keywords. *C difficile* detection; canine olfactory capabilities; interrater reliability; scent dog detection.

Clostridium difficile infection (CDI) is a common nosocomial infection with presentations ranging from mild diarrhea to fulminant pseudomembranous colitis [1]. Over the last decade, there has been emergence of more severe disease associated with CDI outbreaks and increased morbidity and mortality [1, 2]. Early detection and diagnosis are crucial for the initiation of appropriate infection control measures and to improve patient outcomes.

The diagnosis of CDI is most commonly delayed due to challenges with the collection of stool samples or laboratory processing of samples. Consequently, the mean time from onset of symptoms to the start of treatment is approximately 2 days [3, 4]. This diagnostic delay can perpetuate transmission and impede patient flow due to unnecessary isolation and increased length of stay [5].

The goal of this study is to evaluate the operating characteristics of 2 comparably trained dogs as a "point-of-care" diagnostic tool to detect toxin gene-positive *C difficile*. Dogs have been successfully trained to detect the scent of various substances including drugs, plant and animal matter, and bed bugs and are increasingly being evaluated as diagnostic tools in medicine [6]. Based on current literature, only 2 dogs in separate countries have been trained to detect toxigenic *C difficile* [7–9]. Although these studies evaluated sensitivity and specificity, none of them have addressed potential variability of each dog's ability to detect toxin gene-positive *C difficile* because only a single dog was evaluated in each trial. Interrater reliability is a critical operating characteristic that is required to determine the generalizability of diagnostic tests and must be evaluated before dogs could be considered a valuable tool to detect CDI in patients or in the hospital environment.

METHODS

Sample Preparation

All samples were obtained from clinical stool specimens received from the provincial public health laboratory. Positive samples were identified as being positive for both glutamate dehydrogenase (GDH) enzyme immunoassay (EIA) positive using C. DIFF CHEK-60 test (TechLab, Blacksburg, VA) and illumigene *C difficile* deoxyribonucleic acid amplification assay (Meridian Bioscience, Cincinnati, OH). Negative controls consisted of equal proportions of GDH EIA-positive, gene amplification-negative and GDH EIA-negative, gene amplification-negative samples. *Clostridium difficile* strains isolated from toxin gene-positive samples were typed using capillary-based ribotyping [10]. Control samples were cultured using CHROMagar C. difficile fluorogenic culture medium (CHROMagar, Paris, France) to confirm they did not grow toxin-producing *C difficile*. Stool samples were applied to cellulose sponges inside scent detection vials, which have a fine mesh cap that allows for the odor to escape. The vials were then placed within visually identical metal scent boxes. Refrigerated stool specimens were received from the provincial public health laboratory throughout the training and validation study phases. Samples were refrigerated for up to 56 days and were never frozen to ensure stability of toxin levels [11]. Beyond that time, unused samples were discarded. The same methodology was used to prepare specimens for training as that for the validation study, but none of the specimens used in training were reused in the validation study.

Dog Training

Two rescue dogs underwent training in this study: a 3-year-old German Shepherd (Figure 1) and a 3-year-old Border Collie Pointer (Figure 2). A total of 3 professional dog instructors participated in training, including the dog owner, using a reward-based program in which the correct behavior was positively reinforced. The dogs were initially trained to detect the specific odor of toxin

Received 2 March 2018; editorial decision 16 July 2018; accepted 19 July 2018.

Correspondence: M. Taylor, BHSc(PA), Michael Garron Hospital, Toronto East Health Network, 825 Coxwell Avenue, Toronto, ON M4C 3E7, Canada (maureen.taylor@tehn.ca).

Open Forum Infectious Diseases®

DOI: 10.1093/ofid/ofy179

gene-positive *C difficile* strains in stool samples. Once this was achieved, they were introduced to negative samples for proofing. Finally, the dogs completed both positive and negative sample proofing sessions with one dog handler who was blinded to the location and number of positive and negative specimens. This final training phase occurred in a decommissioned ward with hospital beds and equipment but no patients or staff. The same decommissioned ward was used for the validation study with the same dog handler.

Validation Study

We formally tested diagnostic accuracy of each dog on 300 samples at an allocation ratio of approximately 30% positive to 70% negative samples. Each detection round consisted of 10 samples with a randomized number of positives (1–5). We conducted no more than 3 detection rounds per day to prevent dog fatigue. Prepared specimens were retained and refrigerated for up to 2 days before being replaced by fresh specimens. Scent boxes were placed randomly within rooms, and there was rerandomization of number of positive specimens and room assignment before each detection round. The dog trainer was unaware of the number of positive specimens in each round and the status of the sample in each room. The investigator was visually isolated from the trainer and the dog during the trial process. The trainer guided each dog independently along the ward and announced the dog's response as either positive (dog sits) or negative (dog did not sit) Supplementary Video. The dogs were allowed to "sniff" each sample as long as was required in order for them to make a determination. In most cases, this required less than 10 seconds. If the dog correctly identified a positive specimen, as announced by the dog trainer, the investigator acknowledged the correct response so that the dog could receive a food reward. There was no reward for an incorrect or correct negative response. Sensitivity, specificity, and interrater reliability were calculated. Probability of positive allocation of positive specimens was correlated to GDH EIA toxin levels and probability of correct negative allocation to GDH positivity. Interrater reliability was quantified using Cohen's kappa (κ). All statistical analyses were completed using R version 3.4.4.

RESULTS

A dendrogram of toxin gene-positive *C difficile* specimens used during the training and subsequent validation study is presented as Supplementary Appendix Figure 1. The most common ribotypes were North American pulsed-field gel electrophoresis type 1 (NAP 1) [12], NAP 4 [13] and NAP 11 [14] at 9.5%, 13.1% and 10.7% , respectively. The operating characteristics of each dog and interrater reliability are presented in Table 1. The interrater reliability was moderate with a Cohen's kappa of 0.52. Among positive samples, there was no association between GDH EIA levels or ribotype and probability of correct allocation by either dog. Among positive samples, there was no association between GDH EIA levels (Dog 1 $r = 0.19$, $P = .17$; Dog 2 $r = 0.04$, $P = .79$) or ribotype (Dog 1 $P = .62$; Dog 2 $P = .18$) and correct allocation by either dog. There was no association between the probability of correct identification of a negative sample and GDH positivity (Dog 1 $P = .30$; Dog 2 $P = .64$). None of the samples identified concordantly as false positive by both dogs grew a toxin gene-positive *C difficile*.

DISCUSSION

This study demonstrates that trained dogs can detect the presence of toxin gene-positive *C difficile* in stored stool samples with a sensitivity ranging from 77.6 to 92.6 and specificity of 84.4 to 85.1. In our institution, in year the study was completed the prevalence of stool specimens that were *C difficile* toxin gene positive was 13.7%. Using this information, the positive predictive value for Dog 1 would be 45.2% and 49.6% for Dog 2. The negative predictive value for Dog 1 was 96.1% and 98.7% for Dog 2. Our study was the first to simultaneously train 2 dogs and demonstrated only a moderate interrater reliability ($\kappa = 0.53$).

Our study demonstrated operating characteristics of dogs to detect toxin gene-positive *C difficile* similar to prior studies. A study in a large Dutch hospital during a *C difficile* outbreak showed a single male Beagle detected CDI in hospitalized patients with a sensitivity and specificity of 86% and 97%, respectively [8]. More recently, a Springer Spaniel in Canada was able to detect *C difficile* with a search capability sensitivity of 80% and a specificity of 92.9% when samples were hidden in the hospital environment [9]. However, none of these other studies assessed interrater reliability because only a single dog was trained. Our study demonstrates that individual dogs likely have variable ability to detect toxin gene-positive *C difficile* in stool specimens leading to our demonstrated moderate interrater reliability.

Table 1. Operating Characteristics of Two Dogs Used to Detect Toxigenic *Clostridium difficile* in 300 Stool Specimens[a]

	Specimen Distribution				
Dog	GDH^+ $Toxin^+$	GDH^+ $Toxin^-$	GDH^- $Toxin^-$	Sensitivity (95% CI)	Specificity (95% CI)
1	85	109	106	77.6 (67.3–86.0)	85.1 (79.6–89.6)
2	81	108	111	92.6 (84.6–97.2)	84.5 (79.0–89.0)

Abbreviations: CI, confidence interval; GDH, glutamate dehydrogenase.

[a]Dog 1 was a 3-year-old Border Collie Pointer, and dog 2 was a 3-year-old German Shepherd.

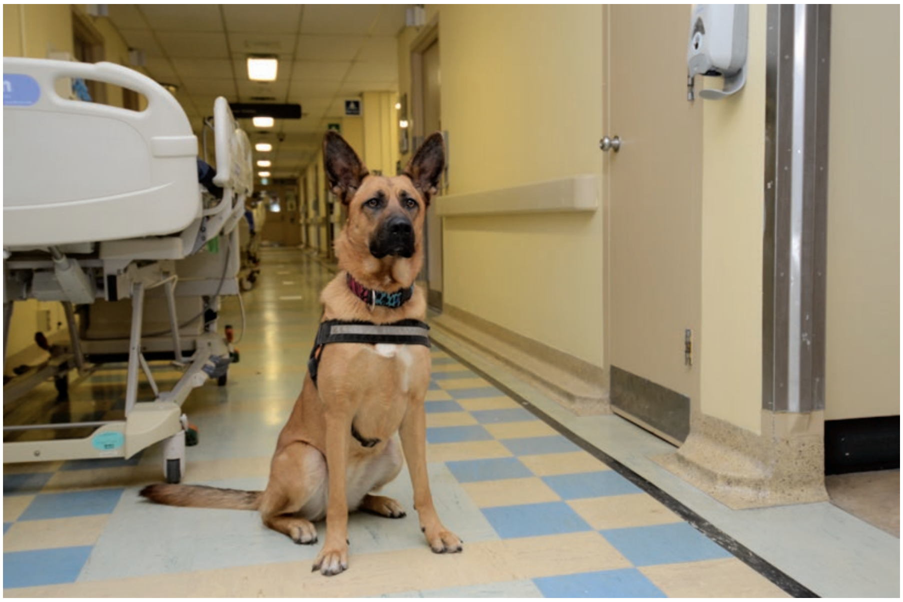

Figure 1. Piper, a 3-year-old German Shepherd, one of 2 dogs in our study.

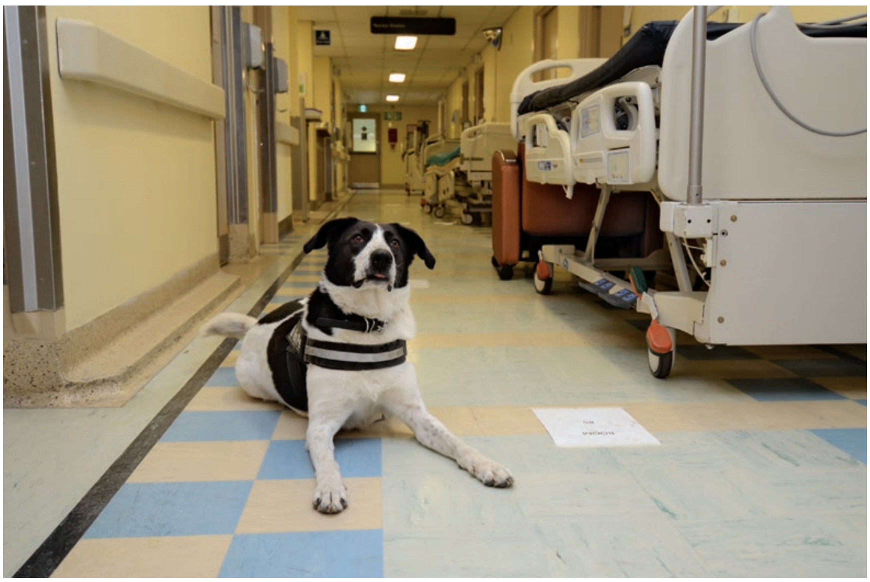

Figure 2. Chase, a 3-year-old Border Collie Pointer, one of 2 dogs in our study.

The inconsistency in each dog's ability to correctly allocate specimens is a major limitation to the widespread use of dogs to detect toxigenic *C difficile* in clinical settings. The variability in the operating characteristics of dogs as a diagnostic tool in medicine has been noted previously in studies to detect cancer; however, positive studies with more than 1 dog did not specifically evaluate interrater reliability [15]. The reason for variability in diagnostic accuracy is uncertain and may be due to either the individual dog's ability to learn a new task, distractibility of the specific animal, or the sensitivity of different breeds' olfactory systems [16, 17]. If each dog's ability to detect *C difficile* is unique, then every dog would need to be independently validated, in a fashion similar to our study, before using them for toxin gene-positive *C difficile* detection.

Our study has several limitations. The use of refrigerated rather than fresh stool limits the generalizability of our findings to an actual clinical scenario. In addition, the relatively small number of positive samples limits the precision with which we can measure sensitivity and specificity. Although our paper is the only one to have evaluated interrater reliability, we still included only 2 dogs in our study. There may have been unique characteristics of one of our dogs that led to our study's modest interrater reliability. Furthermore, the degree of unpredictability in animal behavior is itself an inherent drawback in this study. Despite being highly trained, dogs are vulnerable to distractions and other foreign stimuli in a unique social environment [16]. Our study was completed in a decommissioned hospital ward where the probability of distraction is lower than in a usual clinical setting, and hence our results likely represent an overestimation of sensitivity and specificity. For those wishing to pursue dog olfactory detection for *C difficile*, future studies should involve a greater number of (ideally fresh) stool specimens and a greater number of dogs. Finally, although we did attempt to blind both the dogs and the dog trainer to the status of each sample in the validation trials, the samples were reused several times over a 2-day period. Therefore, it is possible the dogs reacted to a unique odor in a sample that may have been unrelated to its toxin gene-positive *C difficile* status.

CONCLUSIONS

Our study confirms that dogs can detect toxin gene-positive *C difficile* in stool specimens with reasonable operating characteristics; however, more importantly, it demonstrates that interrater reliability is only modest. This finding limits the practical value of using dogs as a point-of-care CDI test. Dogs will never reliably achieve the accuracy of current highly sensitive molecular diagnostic tests for *C difficile*, and strategies that accelerate the testing process, such as more timely specimen collection or test turnaround time, would seem a more promising area for future research than canine detection.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Acknowledgments

We thank the following: Keith Wardrop, Crystal McIndoo, and Jessica Kortleve of K9 Services who rescued, trained, and handled the dogs in our study; Suzanne Gill and Grace Ho who assisted in search training on hospital wards and in the validation study; John Matelski who provided statistical analysis and Public Health Ontario Laboratories; and Dr. Vanessa Allen for providing stool specimens for training and study purposes.

Author contributions. J. M. and J. P. conceived the original idea for the study. J. M., J. P., G. B., S. K., and H. L. designed the study. G. B. obtained and verified microbiological specimens. M. T. T. completed the data collection. J. P. and M. T. T. undertook the data analysis, with statistical assistance from J. M. J. P., M. T. T., and J. M. interpreted the results. The initial draft of the manuscript was prepared by M. T. T. and J. P. and then circulated repeatedly to J. M. and G. B. for critical revisions, and all read and approved the final manuscript. J. P. is the guarantor.

Financial support. This work was funded by a research grant from the Michael Garron Hospital Foundation.

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References

1. Loo VG, Poirier L, Miller MA, et al. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. N Engl J Med **2005**; 353:2442–9.
2. Freeman J, Bauer MP, Baines SD, et al. The changing epidemiology of *Clostridium difficile* infections. Clin Microbiol Rev **2010**; 23:529–49.
3. Scheurer D. Diagnostic and treatment delays in recurrent *Clostridium difficile*-associated disease. J Hosp Med **2008**; 3:156–9.
4. Kundrapu S, Jury LA, Sitzlar B, et al. Easily modified factors contribute to delays in diagnosis of *Clostridium difficile* infection: a cohort study and intervention. J Clin Microbiol **2013**; 51: 2365–70.
5. Mitchell BG, Gardner A. Prolongation of length of stay and *Clostridium difficile* infection: a review of the methods used to examine length of stay due to healthcare associated infections. Antimicrob Resist Infect Control **2012**; 1:14.
6. Teodoro-Morrison T, Diamandis EP, Rifai N, et al. Animal olfactory detection of disease: promises and pitfalls. Clin Chem **2014**; 60:1473–9.
7. Bomers MK, van Agtmael MA, Luik H, et al. Using a dog's superior olfactory sensitivity to identify *Clostridium difficile* in stools and patients: proof of principle study. BMJ **2012**; 345:e7396.
8. Bomers MK, van Agtmael MA, Luik H, et al. A detection dog to identify patients with *Clostridium difficile* infection during a hospital outbreak. J Infect **2014**; 69:456–61.
9. Bryce E, Zurberg T, Zurberg M, et al. Identifying environmental reservoirs of *Clostridium difficile* with a scent detection dog: preliminary evaluation. J Hosp Infect **2017**; 97:140–5.
10. Indra A, Huhulescu S, Schneeweis M, et al. Characterization of *Clostridium difficile* isolates using capillary gel electrophoresis-based PCR ribotyping. J Med Microbiol **2008**; 57:1377–82.
11. Freeman J, Wilcox MH. The effects of storage conditions on viability of *Clostridium difficile* vegetative cells and spores and toxin activity in human faeces. J Clin Pathol **2003**; 56:126–8.
12. McDonald LC, Killgore GE, Thompson A, et al. An epidemic, toxin gene-variant strain of *Clostridium difficile*. N Engl J Med **2005**; 353:2433–41.
13. Jassem AN, Prystajecky N, Marra F, et al. Characterization of *Clostridium difficile* strains in British Columbia, Canada: a shift from NAP1 majority (2008) to novel strain types (2013) in one region. Can J Infect Dis Med Microbiol **2016**; 2016:8207418.
14. Kociolek LK, Gerding DN, Hecht DW, Ozer EA. Comparative genomics analysis of *Clostridium difficile* epidemic strain DH/NAP11/106. Microbes Infect **2018**; 20:245–53.
15. Lippi G, Cervellin G. Canine olfactory detection of cancer versus laboratory testing: myth or opportunity? Clin Chem Lab Med **2012**; 50:435–9.
16. Hackner K, Pleil J. Canine olfaction as an alternative to analytical instruments for disease diagnosis: understanding 'dog personality' to achieve reproducible results. J Breath Res **2017**; 11:012001.
17. Polgár Z, Kinnunen M, Újváry D, et al. A test of canine olfactory capacity: comparing various dog breeds and wolves in a natural detection task. PLoS One **2016**; 11:e0154087.