Contribution to *Clostridium difficile* transmission of symptomatic patients with toxigenic strains who are fecal toxin negative

*Damian PC Mawer[1], *David W Eyre[2,3], David Griffiths[2,3], Warren N Fawley[1,4], Jessica SH Martin[5], T Phuong Quan[2,3], Timothy EA Peto[2,3], Derrick W Crook[2,3,6], A Sarah Walker[2,3], Mark H Wilcox[1,5]

[1]Department of Microbiology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[2]Nuffield Department of Medicine, University of Oxford, Oxford, UK
[3]NIHR Oxford Biomedical Research Centre, University of Oxford, UK
[4]Leeds Regional Microbiology Laboratory, Public Health England, Leeds, UK
[5]Leeds Institute of Biomedical and Clinical Sciences, University of Leeds, Leeds, UK
[6]Public Health England, Colindale, UK

*Drs Mawer and Eyre contributed equally to the manuscript.

Key words

Clostridium difficile; infection; fecal toxin; transmission

Running title

C. difficile fecal toxin status & transmission

Corresponding Authors

Dr Damian Mawer
Department of Microbiology, Leeds Teaching Hospitals NHS Trust
Old Medical School, Thorseby Place
Leeds, West Yorkshire
LS1 3EX, United Kingdom
Telephone: +44 113 3928663; Fax: +44 113 3922696; E-mail: damian.mawer@nhs.net

Dr David Eyre
Nuffield Department of Clinical Medicine, University of Oxford
John Radcliffe Hospital
Headley Way
Oxford, Oxfordshire
OX3 9DU, United Kingdom
Telephone: +44 1865 220855; E-mail: david.eyre@ndm.ox.ac.uk

Article summary point

Using whole-genome sequencing, in two UK hospitals, patients with diarrhea, toxigenic *Clostridium difficile*, but a negative fecal toxin result, were potential sources for 3% of infections; toxin-positive cases were potential sources for 10%, and another 6% were linked to both groups.

Abstract

Background

The role of symptomatic patients who are toxigenic strain-positive (TS+) but fecal toxin-negative (FT-) in transmission of *Clostridium difficile* is currently unknown.

Methods

We investigated the contribution of symptomatic TS+/FT- and TS+/FT+ patients in *C. difficile* transmission in two UK regions. From two-step testing, all glutamate dehydrogenase (GDH)-positive specimens, regardless of fecal toxin result, from Oxford (April2012-April2013) and Leeds (July2012-April2013) microbiology laboratories underwent culture and whole-genome sequencing (WGS), using WGS to identify toxigenic strains. Plausible sources for each TS+/FT+ case, including TS+/FT- and TS+/FT+ patients, were determined using WGS, with and without hospital admission data.

Results

1447/12772(11%) fecal samples were GDH-positive, 866/1447(60%) contained toxigenic *C. difficile* and fecal toxin was detected in 511/866(59%), representing 235 Leeds and 191 Oxford TS+/FT+ cases. TS+/FT+ cases were three times more likely to be plausibly acquired from a previous TS+/FT+ case than a TS+/FT- patient. 51(19%) of 265 TS+/FT+ cases diagnosed >3 months into the study were genetically-related (≤2 single nucleotide polymorphisms) to ≥1 previous TS+/FT+ case or TS+/FT- patient: 27(10%) to only TS+/FT+ cases, 9(3%) to only TS+/FT- patients, and 15(6%) to both. Only 10/265(4%) were genetically-related to a previous TS+/FT+ or TS+/FT- patient and shared the same ward simultaneously or within 28 days.

Conclusions

Symptomatic TS+/FT- patients were a source of *C. difficile* transmission, although they accounted for less onward transmission than TS+/FT+ cases. Although transmission from symptomatic patients with either fecal toxin status accounted for a low overall proportion of new cases, both groups should be infection control targets.

Background

Clostridium difficile infection (CDI) remains a significant concern for patients and healthcare providers, despite recent falls in incidence in some settings, including the UK.[1] Three UK studies using whole-genome sequencing (WGS) have shown, in endemic settings with routine infection control policies, only a minority of cases are acquired from other, known, cases: 35% of cases in Oxford[2] and Leeds[3], and 37% of ribotype-027 cases in Liverpool[4] were genetically-linked to a previous case. Only a subset of these cases also shared time on the same hospital ward. Studies using other genotyping techniques have found similar results.[5-7] Such findings question the sources of *C. difficile* responsible for most CDIs.

While hospitalized asymptomatically colonized patients are a potential source,[7-9] another group of patients with enhanced potential to transmit *C. difficile* are symptomatic patients who are toxigenic-strain positive (TS+), but fecal toxin negative (FT-). These patients are identified by two-step algorithms for CDI diagnosis.[10] An initial screen (e.g. glutamate dehydrogenase (GDH) enzyme immunoassay (EIA), or toxin gene nucleic acid amplification test [NAAT]) detects the presence of *C. difficile*; the second confirmatory step detects fecal toxin using either EIA or a cell cytotoxin assay (CCT). In the UK TS+/FT- patients are usually regarded as being colonized with *C. difficile* but not infected, based on a large multi-center prospective study showing only patients with detectable fecal toxin had adverse outcomes.[11] However, outside the UK, such patients, typically identified with NAATs, are often,[12] but not universally,[13] regarded as having CDI, and NAAT testing has been recommended in some guidelines.[14] Resolving the disease state of TS+/FT- patients is not a focus of this study; instead we investigated their contribution to onward transmission of *C. difficile*.

We undertook WGS of consecutive *C. difficile* GDH-positive fecal samples, irrespective of the subsequent fecal toxin assay result, in two UK centers, over 9-12 months. WGS, combined with hospital admission and ward movement data, were used to assess the contribution of *C. difficile* TS+/FT- and TS+/FT+ patients to onward transmission.

Methods

Samples and setting

Consecutive hospital and community samples submitted for *C. difficile* diagnostic testing were obtained from the microbiology laboratories of 2 UK teaching hospitals following the introduction of two-step testing: Leeds Teaching Hospitals, serving Leeds (population 750,000, 07-July-2012 to 06-April-2013), and Oxford University Hospitals, serving Oxfordshire (population 600,000, 01-April-2012 to 31-March-2013). In Leeds and Oxford, repeat samples from the same patient ≤14 and ≤28 days, respectively, following a toxin-positive sample were not routinely processed. Patient admissions and hospital ward movements were obtained from hospital administration systems. Inclusion of community samples allowed cases diagnosed in the community, but potentially acquired in hospital, to be identified.

In Leeds, any patient with ≥1 episode of unexplained diarrhea was isolated and a fecal sample sent for *C. difficile* testing. TS+/FT+ cases were isolated for the duration of hospital admission. Ward staff could isolate TS+/FT- patients if they were considered a transmission risk. In Oxford, patients with unexplained diarrhea (≥3 unformed stools in 24 hours) were isolated and treated empirically with oral vancomycin. TS+/FT+ cases remained isolated until 48 hours following resolution of diarrhea. Treatment and isolation were discontinued in TS+/FT- patients unless clinical suspicion of CDI remained high.

Diagnostic testing and WGS

Leeds samples were tested with GDH EIA, C. diff Chek (Techlab, Blacksburg, VA), and when GDH-positive an in-house cell cytotoxicity assay, and Oxford samples with Premier *C. difficile* GDH and GDH-positive samples with Premier Toxins A&B EIA (Meridian Bioscience, Cincinnati, OH). At both centers, GDH-positive samples were cultured as described previously[15] and whole-genome sequenced using Illumina technology. In Leeds, isolates were confirmed as *C. difficile* with MALDI-TOF mass-spectrometry; in Oxford WGS was used. Sequences were mapped to the 630 reference genome[16], and assembled *de novo*[17] (see Supplementary Methods for details). Multi-locus sequence types, STs,[15] were determined *in silico*.

Toxigenic strains were identified using BLAST searches of *de novo* assemblies (≥1000 nucleotide identities with toxin A or B genes). Non-toxigenic strains were excluded (n=249, most common STs ST15(n=66,27%), ST26(n=66,27%), ST7(n=51,20%), and ST3(n=11,4%); the remainder were recognized non-toxigenic STs).

Definitions

Patients with toxigenic *C. difficile* were classified according to fecal toxin result: as TS+/FT+ and TS+/FT-. In patients diagnosed with more than one *C. difficile* strain, as defined by WGS (see below), each was considered separately. Some patients had several samples with the same strain, and could be consistently fecal toxin-negative, consistently toxin-positive, or have both fecal toxin-negative and toxin-positive samples. Each TS+/FT+ CDI's origin was determined using standard surveillance definitions.[18] Cases were defined as healthcare-associated if sampled >48 hours after admission or discharged within ≤4 weeks, as indeterminate if discharged 4-12 weeks previously, and as community-associated if discharged >12 weeks prior to sampling, or without any hospital admission.

Analysis

Single nucleotide polymorphisms (SNPs) between sequences were determined from maximum likelihood phylogenies constructed with phyML[19] after correction for recombination with ClonalFrameML.[20] Sequences related to a previous sequence within ≤2 SNPs were considered consistent with plausible direct transmission; ≤2 SNPs is expected between transmitted strains obtained ≤123 days apart.[2] Results for sequences related to previous sequences within varying thresholds (0-10 SNPs) were generated as a sensitivity analysis. In patients with multiple samples, sequences >10 SNPs different to a previous sequence from the same patient were considered to represent acquisition of a new strain; 10 SNPs is considerably more variation than would be expected from within-host diversity and mutation over the one year study period.[2]

Where the only possible genetically-related sources of a TS+/FT+ case were TS+/FT- patients, the origin was attributed to TS+/FT- patients; similarly, if all possible genetically-

related sources were TS+/FT+ cases, the origin was attributed to TS+/FT+ cases. Where a TS+/FT+ case was genetically-linked to either a TS+ patient with both fecal toxin-positive and toxin-negative samples, or several patients including ≥1 TS+/FT+ case and ≥1 TS+/FT- patient, the origin was denoted as either a TS+/FT+ case or TS+/FT- patient.

Patients with toxigenic *C. difficile* who shared time on the same ward following the diagnosis of the first patient and before the diagnosis of the second were considered to have had ward contact. Patients admitted to the same ward, but up to 28 days apart, were considered related by possible ward contamination if the first patient was diagnosed before their ward discharge, and the second patient following their admission to the same ward.[5] Patients who shared time in the same hospital, but had no ward or ward contamination contact, were considered to have hospital contact. A sensitivity analysis assumed ward contamination persisted for 365 days.

Logistic regression was used to test for associations between ST and the proportion of TS+/FT+ cases genetically-related to a previous TS+/FT+ case or TS+/FT- patient, for the 9 most common STs (all with ≥10 cases).

Ethics

The study was approved by the Berkshire Research Ethics Committee (10/H0505/83) and the Health Research Authority (8-05(e)2010).

Results

8068 hospital and community samples were submitted for *C. difficile* testing in Leeds, and 4704 samples in Oxford. 771(10%) and 637(14%) samples were GDH-positive respectively, and, of these, 488(63%) and 372(58%) contained toxigenic *C. difficile* by WGS (Figure 1). Leeds samples were obtained from 367 patients (220 female,60%), median (interquartile range, IQR) 72(52-82) years old, representing 382 genetically distinct infections/colonizations, and Oxfordshire samples from 297 patients (167 female,56%), 78(62-86) years old, 302 genetically distinct infections/colonizations.

In both laboratories, 59% of samples containing toxigenic *C. difficile* had fecal toxin detected despite using different assays, EIA in Oxford (218/372) and CCT in Leeds (289/488). These samples represented 235 distinct TS+/FT+ cases in Leeds, with 3.7 healthcare-associated/indeterminate cases per 10000 bed-days and 7.9 community-associated cases per 100000 person-years, and 191 distinct TS+/FT+ cases in Oxfordshire, 3.2/10000 bed-days and 7.0/100000 person-years, respectively (Figure 1).

There was considerable genetic diversity amongst the *C. difficile* causing the 426 TS+/FT+ cases, with 52 different STs identified. The 10 most frequently isolated STs (common ribotype equivalents) accounted for 285(67%) of cases, and were (in rank order) ST2(014/020), ST8(002), ST6(005), ST11(078), ST10(015), ST5(023/069), ST44(015), ST3(001/072), ST14(014), ST17(018). The epidemic ST1(027/NAP1) strain was only found in three (Leeds) cases.

Genetic relationships between infections/colonizations

Samples were compared with all prior samples from the same center over the study periods, but potential sources were sought only for new TS+/FT+ infections from 3 months into the study at each center (Leeds n=142, Oxfordshire n=123), to ensure sufficient time for their possible sources to have been sampled. Using a threshold of ≤2 SNPs to determine genetic relatedness, overall 51/265(19.2%, 95%CI, 14.7-24.5%) TS+/FT+ cases were genetically-related to ≥1 sequenced previous TS+/FT+ case or TS+/FT- patient (Table 1). 9/265(3.4%, 1.6-6.3%) of TS+/FT+ cases were genetically linked only to TS+/FT- patients and not to previous TS+/FT+ cases. In contrast, 27/265(10.2%, 6.8-14.5%) TS+/FT+ cases were genetically linked to other TS+/FT+ cases, and 15/265(5.7%, 3.2-9.2%) to both TS+/FT+ cases and TS+/FT- patients. There was no evidence of a difference in sources between Leeds and Oxford (Table 1; exact p=0.27).

Considering the source of *C. difficile* for all patients, TS+/FT- patients as well as TS+/FT+ cases, results were similar (Table S1; exact p=0.85 comparing all patients vs. TS+/FT+ cases alone): 75/433(17%) patients could be linked to a previously sequenced TS+/FT+ case or

TS+/FT- patient, 16(4%) to only TS+/FT- patients, 36(8%) to only previous TS+/FT+ cases and 23(5%) to both.

There were 13 ST44 infections, none of which were genetically-related to a prior TS+/FT+ case, the remaining 8 most common STs were compared with all other STs as the reference group. Within the limits of the relatively small numbers of TS+/FT+ cases within each ST, there was no evidence that CDI caused by any of these STs were more or less likely, to be genetically-related to a previous TS+/FT+ case or TS+/FT- patient ($p \geq 0.18$; Table 2), or that CDI source was associated with patient age, sex or healthcare/community-associated disease (Table 3).

Over the whole study period at both centers, considering all 684 TS+/FT+ cases and TS+/FT- patients, 535 were not related to any other TS+/FT+ case or TS+/FT- patient within ≤2 SNPs. The remaining 149 TS+/FT+ cases and TS+/FT- patients were clustered: sequences included in a cluster were related to ≥1 other sequence within ≤2 SNPs in the cluster, but not necessarily to all of them. Most clusters contained 2 or 3 patients; 14(9%) patients were in clusters consisting of exclusively TS+/FT- patients, 45(30%) were in exclusively TS+/FT+ clusters, and 90(60%) were in clusters with both TS+/FT- patients and TS+/FT+ cases (Figure 3).

Epidemiological relationships between genetically-related infections/colonizations

Only a subset of TS+/FT+ cases and plausible TS+/FT+ or TS+/FT- sources related within ≤2 SNPs shared a hospital-based epidemiological link. Considering all 265 TS+/FT+ cases from both Leeds and Oxfordshire from 3 months into the study, 27(10%) were genetically-related to only previous TS+/FT+ cases. However, only 6(2%) were genetically-related and shared time on the same ward. A further 4(2%) were genetically-related and were inpatients on the same ward at different times within 28 days. 8(3%) were not admitted to the same ward within 28 days, but were admitted to the same hospital at the same time (Table 1).

Another 9(3%) TS+/FT+ cases were genetically-related to only previously TS+/FT- patients: 5(2%) sharing time on a ward, 1(0.4%) the same ward at different times within 28 days, and

1(0.4%) time in the same hospital as above. There was a trend towards potential TS+/FT- sources being more likely to share time on the same ward as the subsequent TS+/FT+ case, compared with potential TS+/FT+ sources (5/9 vs. 6/27, exact p=0.10). An additional 15(6%) TS+/FT+ cases were genetically-related to both a TS+/FT- patient and a TS+/FT+ case, but 14 had no hospital-based links with the genetically-related sources, suggesting these patients may share a common indirect source rather than direct hospital-based contact. No additional epidemiological links between genetically-related TS+/FT+ cases and TS+/FT+ cases or TS+/FT- patients were identified if ward contamination could persist for up to 365 days.

To test the robustness of our observations to the SNP threshold used to define plausible direct transmission, the number of TS+/FT+ cases genetically-related to a previous TS+/FT+ case or TS+/FT- patient within varying SNP thresholds from 0 to 10, and any associated hospital-based epidemiological links, were determined (Figure 3). As expected, as the number of SNPs used to define plausible direct transmission increased, the percentage of TS+/FT+ cases genetically-related to a previous TS+/FT+ or TS+/FT- patient increased. However, the number of TS+/FT+ cases genetically-related and with plausible epidemiological contact, i.e. sharing hospital wards, remained relatively constant from 2 SNPs onwards, supporting the 2 SNP threshold used for the main analysis.

Discussion

We used WGS and ward admission data to investigate the proportion of CDI cases potentially acquired from symptomatic patients with toxigenic *C. difficile*, but with no detectable fecal toxin. TS+/FT+ CDI cases were three times more likely to be genetically-related to a previous TS+/FT+ case (27/265) than a TS+/FT- patient (9/265). Considering the subset of potential sources that also shared time on the same ward, or were admitted to the same ward within 28 days, i.e. the most probable of the genetically-plausible transmission events, CDI cases were 1.7 times more likely to be related to a previous TS+/FT+ case compared with a TS+/FT- patient (10/265 vs. 6/265). However, ~~As~~ there were also 1.7 times as many TS+/FT+ as TS+/FT- patients in the study, making linkage to a previous TS+/FT+ case 1.7 times more likely than to a TS+/FT- patient based on the relative

frequencies of TS+/FT+ and TS+/FT- patients alone. Therefore, the rate of transmission, on a per patient basis, from each TS+/FT+ or TS+/FT- patient is likely to be very similar. By contrast, asymptomatically colonized patients are likely less infectious. Using national databases and a transmission model, individual hospitalized CDI cases have been estimated to transmit *C. difficile* at a rate 15 (95%CI 7.2-32) times that of hospitalized asymptomatically colonized patients.[21] However, as asymptomatic carriage is more common than CDI (e.g. 8-fold in hospitals[9]), colonized patients, as a group, could still account for a substantial amount of transmission. In a Canadian study, isolation of all asymptomatically colonized patients reduced CDI incidence by 62% compared with historic controls.[9]

The overall number of our TS+/FT+ CDI cases potentially attributable to the combination of TS+/FT+ cases and TS+/FT- patients with diarrhea was low: 19% of TS+/FT+ CDI cases were genetically-related to a previous TS+/FT+ or TS+/FT- patient, only 6% also shared a hospital ward at the same time or within 28 days, and only 10% had any form of hospital contact. This supports previous WGS-based studies, at both our hospitals[2,3] and others[4], that found that only a minority of CDIs are acquired from other cases in endemic settings. The proportion in the present study is lower than the 35-37% identified previously. The most likely explanation is the very small number of infections with the epidemic ST1(027/NAP1) strain, reflecting falling UK incidence[22,23], and the burden of transmissions attributable to ST1 in previous studies.[3]

Our study has several limitations. Only patients with diarrhea were sampled, and at the discretion of individual practitioners. However, the ratio of toxin-positive stools sequenced to samples tested was 3.6%(289/8068) in Leeds, and 4.6%(218/4704) in Oxford, suggesting rates of testing were high, including compared with the UK average from 2008 of 6.45%, when testing was principally based on toxin detection.[24] Of those tested, some patients with *C. difficile* will have been missed by the GDH assay (sensitivity 92.3-97.1%[11,25]). In addition, 2.6% of isolates failed WGS and were excluded. We therefore may have missed some links between TS+/FT- or TS+/FT+ patients and TS+/FT+ CDI cases, modestly underestimating the frequency with which this occurs.[5] However, if cases were missed at

random, we believe the relative amount of transmission attributable to TS+/FT+ cases and TS+/FT- patients has been robustly estimated. We did not gather data on factors that might influence a TS+/FT- patient's potential to transmit *C. difficile*, including duration and severity of diarrhea, antibiotic exposure, or the timing and duration of isolation. In addition, systematic serial sampling was not undertaken to allow an assessment of the duration of detectable *C. difficile*. Our study was performed in a setting where the majority of CDI arises from a diverse range of endemic strains; findings may vary in higher incidence settings, including where the epidemic ST1(027/NAP1) strain dominates.

Despite these limitations, we demonstrate that patients with toxigenic *C. difficile* without detected fecal toxin account for a quarter or more of potential within hospital transmission events from symptomatic patients. More intensive infection control interventions around such cases, including routine isolation, should be considered to mitigate transmission risk. Compared with asymptomatically colonized patients, TS+/FT- patients represent a good initial target for expanding infection control efforts, as they are less numerous, and, as discussed above, appear more infectious[21] on a per patient basis. However, ultimately if the findings of [9] can be replicated, isolation of asymptomatically colonized patients, who are each less infectious, but more numerous, may result in greater reductions in transmission. Substantially greater resource requirements limit the later approach. Some GDH-positive fecal toxin-negative patients may carry non-toxigenic *C. difficile* and not pose an infection control risk. Patients with toxigenic *C. difficile* could be identified by screening with a toxin gene NAAT, or using a three-step strategy (GDH-positive, fecal toxin-negative samples tested with a toxin gene NAAT).

The results of this and previous studies in both Oxford and Leeds suggest CDI cases, and also symptomatic patients with toxigenic *C. difficile* with a negative fecal toxin result, are not sources for the majority of CDI. Major unanswered questions remain, including what proportion of CDI cases can be explained by healthcare-associated and community contact with asymptomatically colonized people, and the extent to which other possible sources including food[26,27] and the environment[28] contribute to CDI. In addition to reducing the risk of CDI through antimicrobial stewardship,[23] understanding the relative

importance of each of these reservoirs across a range of settings is required to develop rational control polices and reduce the incidence of CDI. Meanwhile, efforts to reduce hospital transmission from symptomatic patients with toxigenic *C. difficile* with a negative fecal toxin result should be implemented.

Funding

This study was supported by the UK Clinical Research Collaboration (Wellcome Trust [grant 087646/Z/08/Z]; Medical Research Council [grant G0800778]; and the National Institute for Health Research); and the NIHR Oxford Biomedical Research Centre. DWC and TEAP are NIHR senior investigators. DWE is a NIHR clinical lecturer.

Acknowledgements

The authors thank Claire Berry and Faye Pinker for their help collecting hospital admissions data in Leeds.

Potential conflicts of interest

M.H.W has received consulting fees from Actelion, Astellas, MedImmune, Merck, Pfizer, Sanofi-Pasteur, Seres, Summit, and Synthetic Biologics; lecture fees from Alere, Astellas, Merck & Pfizer; and grant support from Actelion, Astellas, bioMerieux, Da Volterra, Merck and Summit.

D.P.C.M, D.W.E, D.G, W.N.F, J.S.H.M, T.P.Q, T.E.A.P, D.W.C and A.S.W all have no conflicts of interest to declare.

Data deposition

Sequences generated during this study can be found on the NCBI short read archive under BioProject PRJNA327723.

References

1. Public Health England. Annual Epidemiological Commentary: Mandatory MRSA, MSSA and E. coli bacteraemia and C. difficile infection data, 2014/15. 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/442952/Annual_Epidemiological_Commentary_FY_2014_2015.pdf. Accessed 21 June 2016.

2. Eyre DW, Cule ML, Wilson DJ, et al. Diverse Sources of C. difficileInfection Identified on Whole-Genome Sequencing. N Engl J Med **2013**; 369:1195–1205.

3. Martin J, Eyre DW, Fawley W, et al. 2016. *C. difficile* ribotypes exhibit variable patient-to- patient transmission rates, as determined by whole genome sequencing, suggesting differing reservoirs and modes of acquisition [abstract 0557]. 26th European Congress of Clinical Microbiology and Infectious Disease (ECCMID), 9-12 April, Amsterdam, Holland.

4. Kumar N, Miyajima F, He M, et al. Genome-based infection tracking reveals dynamics of Clostridium difficile transmission and disease recurrence. Clin Infect Dis **2015**; 62:1031–752.

5. Walker AS, Eyre DW, Wyllie DH, et al. Characterisation of Clostridium difficile hospital ward-based transmission using extensive epidemiological data and molecular typing. PLoS Med **2012**; 9:e1001172:1–12.

6. Norén T, Akerlund T, Bäck E, et al. Molecular epidemiology of hospital-associated and community-acquired Clostridium difficile infection in a Swedish county. J Clin Microbiol **2004**; 42:3635–3643.

7. Curry SR, Muto CA, Schlackman JL, et al. Use of multilocus variable number of tandem repeats analysis genotyping to determine the role of asymptomatic carriers in Clostridium difficile transmission. Clin. Infect. Dis. **2013**; 57:1094–1102.

8. Eyre DW, Griffiths D, Vaughan A, et al. Asymptomatic Clostridium difficile colonisation and onward transmission. PLOS One **2013**; 8:e78445.

9. Longtin Y, Paquet-Bolduc B, Gilca R, et al. Effect of Detecting and Isolating Clostridium difficile Carriers at Hospital Admission on the Incidence of C difficile Infections: A Quasi-Experimental Controlled Study. JAMA Intern Med **2016**;

10. Wilcox MH, Planche T. Summary of research underpinning the Department of Health's new *Clostridium difficile* testing guidance. 2011. Available at: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317132979562. Accessed 25 July 2012.

11. Planche TD, Davies KA, Coen PG, et al. Differences in outcome according to Clostridium difficile testing method: a prospective multicentre diagnostic validation study of C difficile infection. Lancet Infect Dis **2013**; 13:936–945.

12. Agency for Healthcare Research and Quality. Early Diagnosis, Prevention, and Treatment of *Clostridium difficile*: Update. 2016. Available at: https://www.effectivehealthcare.ahrq.gov/ehc/products/604/2208/c-difficile-update-report-160502.pdf. Accessed 21 October 2016.

13. Polage CR, Gyorke CE, Kennedy MA, et al. Overdiagnosis of Clostridium difficile Infection in the Molecular Test Era. JAMA Intern Med **2015**; 175:1792–1801.

14. Surawicz CM, Brandt LJ, Binion DG, et al. Guidelines for Diagnosis, Treatment, and Prevention of Clostridium difficile Infections. Am J Gastroenterol **2013**; 108:478–498.

15. Griffiths D, Fawley W, Kachrimanidou M, et al. Multilocus sequence typing of Clostridium difficile. J Clin Microbiol **2010**; 48:770–778.

16. Sebaihia M, Wren BW, Mullany P, et al. The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. Nat Genet **2006**; 38:779–786.

17. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **2008**; 18:821–829.

18. McDonald LC, Coignard B, Dubberke E, et al. Recommendations for surveillance of Clostridium difficile-associated disease. Infect Control Hosp Epidemiol **2007**; 28:140–145.

19. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol **2010**; 59:307–321.

20. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLoS Comput. Biol. **2015**; 11:e1004041.

21. Durham DP, Olsen MA, Dubberke ER, Galvani AP, Townsend JP. Quantifying Transmission of Clostridium difficile within and outside Healthcare Settings. Emerging Infect Dis **2016**; 22:608–616.

22. Public Health England. Clostridium difficile Ribotyping Network (CDRN) for England and Northern Ireland: Biennial Report (2013-2015). 2016. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/491253/CDRN_2013-15_Report.pdf. Accessed 21 June 2016.

23. Dingle KE. Clinical and epidemiological characteristics of new epidemic Clostridium difficile strains. 2015. 25th European Congress of Clinical Microbiology and Infectious Disease (ECCMID), 25-28 April, Copenhagen, Denmark.

24. Goldenberg SD, French GL. Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. J Hosp Infect **2011**; 79:4–7.

25. Davies KA, Planche T, Crook D, Monahan I, Pope C, Wilcox MH. 2013. Comparison of

Premier *C. difficile* GDH enzyme immunoassay and Illumigene loop mediated isothermal amplification assay with two reference methods for the laboratory detection of *C. difficile* [abstract 1876]. 27-30 April, Berlin, Germany.

26. Hensgens MPM, Keessen EC, Squire MM, et al. Clostridium difficile infection in the community: a zoonotic disease? Clin Microbiol Infect **2012**; 18:635–645.

27. Gould LH, Limbago B. Clostridium difficile in food and domestic animals: a new foodborne pathogen? Clin Infect Dis **2010**; 51:577–582.

28. Saif al N, Brazier JS. The distribution of Clostridium difficile in the environment of South Wales. J Med Microbiol **1996**; 45:133–137.

477 Tables

478

Possible transmission source	Genetically linked ≤2 SNPs (% n)	Epidemiological links between genetically linked cases, No. (% n, % genetically linked)		
		Shared time on same ward	Shared same ward within 28 days	Shared time in same hospital only
Leeds (CDI cases, n = 142)				
Prior TS+/FT+ cases	17 (12)	2 (1, 12)	2 (1, 12)	7 (5, 41)
TS+/FT- patients	5 (4)	2 (1, 40)	1 (1, 20)	1 (1, 20)
Both	11 (8)	0 (0, 0)	0 (0, 0)	1 (1, 9)
Total	33 (23)	4 (3, 12)	3 (2, 9)	9 (6, 27)
Oxford (CDI cases, n = 123)				
Prior TS+/FT+ cases	10 (8)	4 (3, 40)	2 (2, 20)	1 (1, 10)
TS+/FT- patients	4 (3)	3 (2, 75)	0 (0, 0)	0 (0, 0)
Both	4 (3)	0 (0, 0)	0 (0, 0)	0 (0, 0)
Total	18 (15)	7 (6, 39)	2 (2, 11)	1 (1, 6)
Combined (CDI cases, n = 265)				
Prior TS+/FT+ cases	27 (10)	6 (2, 22)	4 (2,15)	8 (3, 30)
TS+/FT- patients	9 (3)	5 (2, 56)	1 (1, 11)	1 (1, 11)
Both	15 (6)	0 (0, 0)	0 (0, 0)	1 (1, 7)
Total	51 (19)	11 (4, 22)	5 (2, 10)	10 (4, 20)

479

480 **Table 1. Proportion of toxigenic strain-positive, fecal toxin-positive (TS+/FT+) CDI cases**
481 **genetically (≤2 SNPs) and epidemiologically related to prior TS+/FT+ cases and TS+/FT-**
482 **patients.**

483

		Genetically-related to prior TS+/FT+ case or TS+/FT- patient		
ST	**n**	**Total (% n)**	**Odds ratio (95% CI)**	***P* value**
All other STs	114	24 (21)	1	-
2	33	9 (27)	1.41 (0.58, 3.42)	0.45
5	17	4 (24)	1.15 (0.34, 3.86)	0.82
6	23	2 (9)	0.36 (0.08, 1.63)	0.18
8	21	3 (14)	0.63 (0.17, 2.30)	0.48
10	14	1 (7)	0.29 (0.04, 2.32)	0.24
11	20	5 (25)	1.25 (0.41, 3.78)	0.69
14	10	3 (30)	1.61 (0.39, 6.69)	0.51
44	13	0 (0)	-	-

Table 2. Association between ST and proportion of CDI cases genetically-related to prior TS+/FT+ cases and TS+/FT- patients. Each ST in the table was compared to all other STs (the reference group) by logistic regression.

	No genetically linked source	TS+/FT- source	TS+/FT+ source	Both	p value	p value, any source vs no genetically-linked source
Classification (row %)					0.99	0.83
Community-associated	53 (83%)	1 (2%)	7 (11%)	3 (5%)		
Indeterminate	22 (85%)	1 (4%)	2 (8%)	1 (3%)		
Healthcare-associated	139 (79%)	7 (4%)	18 (10%)	11 (6%)		
Age					0.76	0.59
Median	75	82	79	78		
IQR	54 - 83	69 - 86	24 - 85	58 - 84		
Sex (row %)					0.5	0.35
Female	115 (79%)	4 (3%)	17 (12%)	10 (7%)		
Male	98 (84%)	5 (4%)	9 (8%)	5 (4%)		

Table 3. Patient demographics according to CDI source (n=265). Age and sex were not recorded for 2 patients. Exact p values are shown for classification and sex; p values for age were calculated with the Kruskal-Wallis rank test.

Figure legends

Figure 1. Samples and patient demographics for Leeds (panel A) and Oxfordshire (panel B). Each percentage uses the row above as denominator. Distinct infection is one >10 SNPs distinct to any previous infection in the same patient. HA, healthcare-associated. CA, community-associated. MALDI-TOF MS, matrix assisted laser desorption time of flight mass spectrometry. Age and sex were not recorded for 3 Oxfordshire patients.

Figure 2. Numbers of patients in clusters related within ≤2 SNPs. Clusters consisting exclusively of toxigenic strain-positive, fecal toxin-negative (TS+/FT-) patients are shown in blue, clusters consisting exclusively of TS+/FT+ cases in red, and clusters with both TS+/FT- patients and TS+/FT+ cases in orange.

Figure 3. Proportion of Leeds and Oxfordshire CDI cases genetically-related to a previous toxigenic strain-positive, fecal toxin-positive (TS+/FT+) case or TS+/FT- patient within varying SNP thresholds. Bars are shaded according to the fecal toxin status of the genetically-related potential sources of infection.