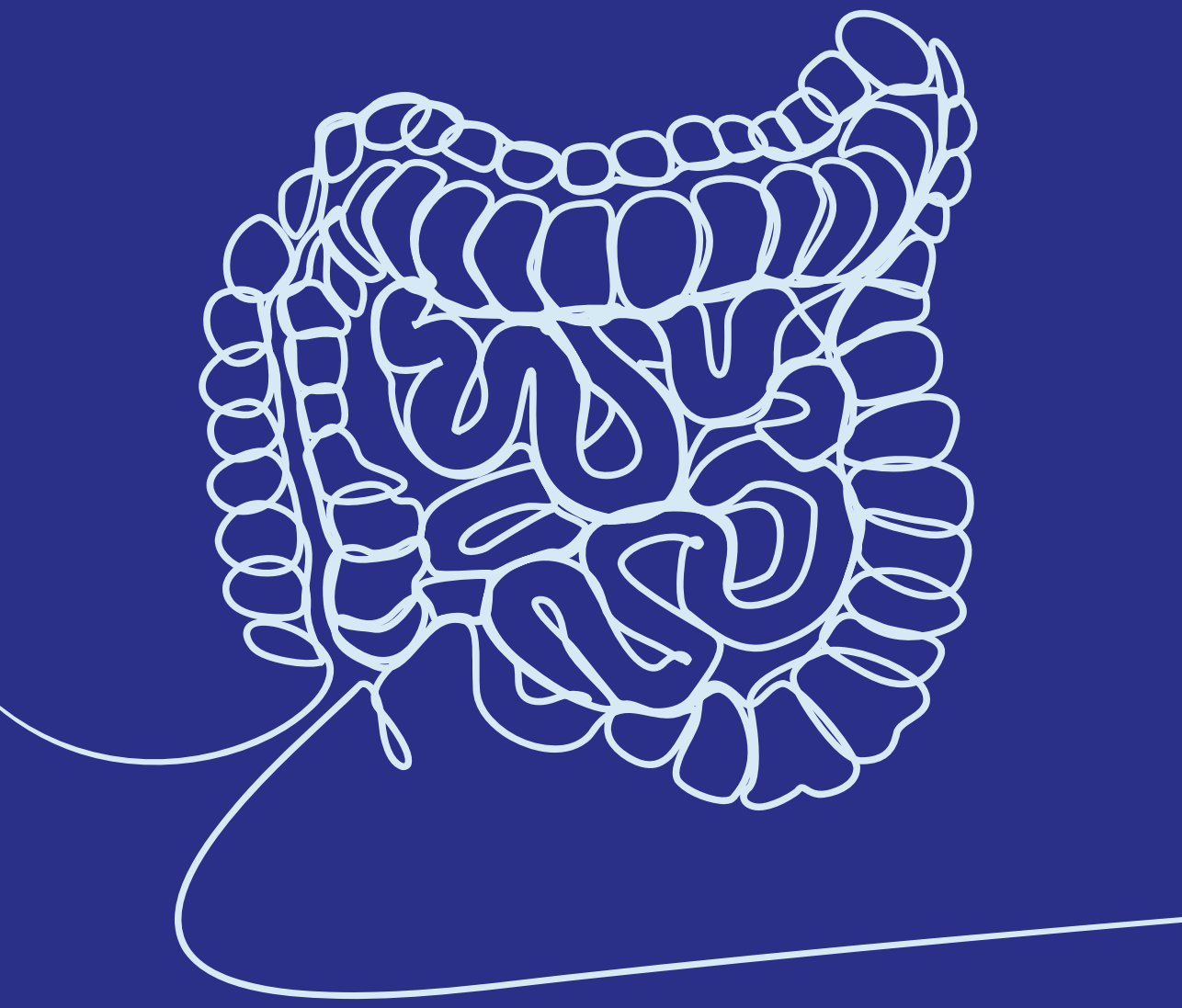
Towards understanding Clostridioides difficile colonization

Crobach, M.J.T.

CHAPTER 6

Screening for *Clostridioides difficile* colonization at admission to the hospital: a multi-centre study



Clinical Microbiology and Infection, 2023

Monique J.T. Crobach[1]
Bastian V.H. Hornung[1]
Cees Verduin[2]
Margreet C. Vos[3]
Joost Hopman[4]
Nitin Kumar[5]
Celine Harmanus[1]
Ingrid Sanders[1]
Elisabeth M. Terveer[1]
Mark D. Stares[5]
Trevor D. Lawley[5]
Ed J. Kuijper[1,6]

1 Department of Medical Microbiology, Leiden University Medical Center, PO Box 9600, 2300RC, Leiden, The Netherlands

2 former: Department of Medical Microbiology, Amphia Hospital Breda, The Netherlands

3 Department of Medical Microbiology and Infectious Diseases, Erasmus University Medical Centre, Rotterdam, The Netherlands

4 Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, The Netherlands

5 Host-Microbiota Interactions Laboratory, Wellcome Sanger Institute, Hinxton, United Kingdom

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

Clin Microbiol Infect, 2023 Jul;29(7):891-896

Abstract

Objective. This study aimed to assess the value of *C. difficile* colonization (CDC) screening at hospital admission in an endemic setting.

Methods. A multi-centre study was performed in 4 hospitals located across the Netherlands. Newly admitted patients were screened for CDC. The risk to develop *C. difficile* infection (CDI) during admission and one-year follow-up was assessed for colonized and non-colonized patients. *C. difficile* isolates from colonized patients were compared with isolates from incident CDI cases using core genome multi locus sequence typing (cgMLST) to determine if onwards transmission had occurred.

Results. CDC was present in 108/2211 admissions (4.9%), while colonization with a toxigenic strain (tCDC) was present in 68/2211 (3.1%) of admissions. Among these 108 colonized patients, diverse PCR ribotypes were found and no 'hypervirulent' RT027 was detected ((95% CI, 0- 0.028). None of the colonized patients developed CDI during admission (0/49, 95% CI 0-0.073) or one-year follow-up (0/38, 95% CI 0-0.93). Core genome MLST identified 6 clusters with genetically related isolates from tCDC and CDI patients, but in these clusters only one possible transmission event from a tCDC to a CDI patient was identified by epidemiological data.

Conclusion. In this endemic setting with a low prevalence of 'hypervirulent' strains screening on CDC at admission did not detect any CDC patient who progressed to symptomatic CDI and only one possible transmission event from a colonized patient to a CDI patient. Thus, screening on CDC at admission is not useful in this setting.

Introduction

Clostridioides difficile infection (CDI) remains an important source of healthcare and antibiotic associated diarrhoea. However, not every individual will develop symptomatic CDI after contact with *C. difficile* spores: patients with asymptomatic *C. difficile* colonization (CDC) outnumber symptomatic CDI patients (1). CDC patients do not exhibit symptoms, but might progress to symptomatic CDI upon disturbance of their microbiota. Also, they do shed *C. difficile* spores in their environment thereby acting as a reservoir and potential source for *C. difficile* (2, 3). Although infection control measures focus currently on symptomatic cases only (4), literature has shown that isolation of CDC patients may help in preventing nosocomial transmission (5). Notably, most studies on the importance of CDC patients are conducted in settings with high CDI incidence rates and/or a high proportion of hypervirulent ribotypes (6, 7). The contribution of CDC patients to the epidemiology of CDI is less well known in other settings. In this study, we investigate the value of a CDC screening program on hospital admission in an endemic setting. Factors that determine the need for such a screening program including the prevalence of colonization, the risk of colonized patients to progress to CDI and the chance of onwards transmission from CDC to CDI patients were taken into account.

Methods

Study design and patients

The study was performed in 4 acute care hospitals (3 university-affiliated, 1 general) located across the Netherlands. In one of these hospitals the Dutch reference laboratory for *C. difficile* is housed and all hospitals participate in national sentinel CDI surveillance. In each of the 4 hospitals, patients were enrolled during a 6 to 8 month period between January 2015 and December 2016. Adult patients admitted to predefined wards (medical and surgical) were eligible. Patients with CDI at admission or CDI diagnosed within the first 72hrs of admission were excluded. Additional exclusion criteria are listed in S1. Patients could be enrolled more than once if readmitted during the study period. Consenting subjects had stool samples (and in 1 hospital partly rectal swabs) collected within 72hrs of admission. If patients were discharged before spending 72hrs in the hospital, stool samples could be collected at home and returned to the hospital, no time limit was imposed on collection of these samples. Patients with a positive *C. difficile* culture but no diagnosis of CDI were considered *C. difficile* colonized (CDC). The subset of CDC patients with

a toxigenic strain in their stool cultures were considered toxigenic *C. difficile* colonized (tCDC). CDC patients were included as cases in the case control study after obtaining written informed consent. For each case, 3 controls were selected from the cohort that tested negative for *C. difficile* in their stool samples obtained at admission. These controls were the 3 consecutive patients who submitted a study stool sample to the laboratory and agreed to participate in the case control study.

Toxigenic *C. difficile* isolates from tCDC and CDI patients were compared to determine if transmission from tCDC patients to CDI patients had occurred. CDI cases were all hospitalized patients diagnosed with CDI during the study period and 3 months thereafter in each of the participating hospitals. Isolates of these CDI cases were collected and sent to the Dutch CDI reference laboratory as part of the national sentinel CDI surveillance (8). All CDI cases had to comply with definitions valid in the surveillance protocol (see S2). Test methods for diagnosing CDI in the 4 hospitals are described in S3. Samples from recurrent (>2 but <8 weeks after initial episode) or new (≥8 weeks after the initial episode) CDI episodes were once more included.

Microbiological analysis

Stool culture for the presence of *C. difficile* was performed on a daily basis; during weekends or holidays samples were stored at 4^{0}C until the following working day. Culture methods are described in S4. All identified isolates from (enrichment) culture were ribotyped by resolution capillary gel-based electrophoresis PCR-ribotyping, using the Dutch national reference laboratory library (9). In addition, a multiplex PCR to detect toxin genes *tcdA*, *tcdB*, *cdtA* and *cdtB* was performed on cultured isolates (10). Strains positive for *tcdA*, *tcdB* or *cdtA/cdtB* were defined as toxigenic strains, all other strains were defined as non-toxigenic strains.

Data collection

Patient information was collected at baseline via a patient's questionnaire and medical electronical records. For each patient, the Charlson's Comorbidity Index was calculated (11). Follow-up by patient's questionnaires was scheduled at 30 days and 1 year after enrolment to determine how many patients developed CDI.

Transmission analysis using core genome MLST

Methods used for reculturing, sequencing and construction of the core genome MLST are described in S5. In short, genomes were assembled as previously described (12), annotated with Prokka (13), and alleles for the cgMLST were predicted with a method compatible with SeqSphere (14). Library preparation is described in S6. Based on previous publications (15), ≤ 2 different alleles in the cgMLST were considered to be the same strain if the time frame of sampling was less than 124 days, and ≤ 3 different alleles if it was less than 1 year. Ward movement data of CDI and colonized patients were investigated if their isolates were genetically related. Criteria for epidemiologic linkage are described in S7.

Statistical analysis

Characteristics of CDC patients and tCDC patients were compared to their respective controls. All analyses were performed using STATA SE statistical software version 15.1 (Statacorp, Texas, USA). A p-value of <0.05 was considered statistically significant.

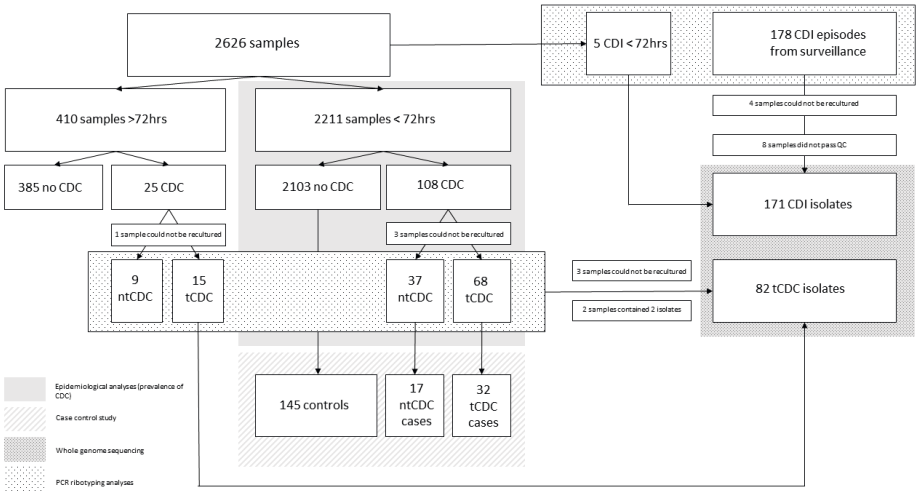


Figure 1. Flowchart of included samples.
CDI; *Clostridoides difficile* infection, CDC: *Clostridoides difficile* colonization, ntCDC: non-toxigenic *Clostridioides difficile* colonization, tCDC: toxigenic *Clostridoides difficile* colonizatiion

Ethical and Methodological Considerations

The study protocol was submitted to the Institutional Review Ethics Board that deemed that this research is not subject to the Medical Research Involving Human Subjects Act. They had no objection to the conduction of the research or collection of the stool samples on admission under verbal informed consent. *C. difficile* culture results were not disclosed to patients or treating physicians. Patients selected as cases and controls provided written informed consent. Stool samples from CDI patients were collected as part of routine care and PCR ribotyped for surveillance purposes. No additional consent was required for whole genome sequencing of samples.

Data availability

All genomic data has been uploaded to the European Nucleotide Archive under study number PRJEB25045.

Results

Included samples and prevalence of *C. difficile* colonization

In total 2626 samples were screened for CDC, ranging from 500 to 1011 samples per hospital (Table 1). 415 samples were excluded from epidemiological analyses (Figure 1). From the remaining 2211 samples, 1736 were stool samples, 467 were rectal swabs, and for 8 information about sampling method was lost. *C. difficile* was found in 108 samples, thus the prevalence of CDC at admission to the hospital was 4.9% (108/2211). Toxigenic strains were found in 68/108 samples. The prevalence of tCDC was therefore 3.1% (68/2211).

Table 1. Included samples and prevalence of CDC and tCDC per hospital

Hospital	Enrolment period	N included samples	N CDC	N tCDC	Prevalence CDC (%)	Prevalence tCDC (%)	N patients enrolled in case control study
LUMC	Jan 2015-Jul 2015	453	19	10	4,19	2,21	44
Erasmus	Sept 2015-Apr 2016	581	36	24	6,20	4,13	50
Amphia	Oct 2015-Mar 2016	786	33	20	4,20	2,54	72
Radboud	April 2016-Nov 2016	391	20	14	5,12	3,58	28
Total		2211	108	68	4,88	3,08	194

Apr, April; CDC: *Clostridioides difficile* colonization; Jan, January; Jul, July; LUMC, Leiden University Medical Center; Mar, March; Nov, November; Oct, October; Sep, September; tCDC: toxigenic *Clostridioides difficile* colonization.

Patient characteristics

In total 194 patients were enrolled in the case-control study: 32 tCDC patients , 17 patients colonized by non-toxigenic strains (ntCDC) and 145 controls (Figure 1). Results from univariate analysis are shown in Table 2.

CDI during follow-up

None of 49 colonized (95% CI 0-0.073) or 145 control patients (95% CI 0-0.025) developed CDI during admission or within the month after enrollment. Questionnaires at one year follow-up were returned by 152 (85% of alive) patients (38 CDC patients and 114 controls). None of these patients reported to have developed CDI during follow-up (0/38 CDC, 95% CI 0-0.093 and 0/114 controls, 95% CI 0-0.032). Chart review of deceased patients showed than one control patient developed CDI 2 months after negative admission screen.

PCR ribotyping and sequence typing

Forty-four different (known) PCR ribotypes were identified among 129 colonized patients. Colonization with the 'hypervirulent' RT027 was not identified, 4 patients were colonized with the 'hypervirulent' RT078 (all from different hospitals). During the study period and 3 months thereafter, 183 CDI episodes were identified and these samples were included for comparison with tCDC isolates. RT027 was also not found among CDI patients.

In total 253 strains were available for WGS analysis (82 isolates from tCDC patients and 171 isolates from CDI patients (Figure 1). Sequence types were assigned to all isolates (Figure 2). ST11 (RT078, RT826 and related ribotypes) was more frequently found among CDI patients than among tCDC patients (19.9% vs 4.9%, $p<0.01$) (Figure 2).

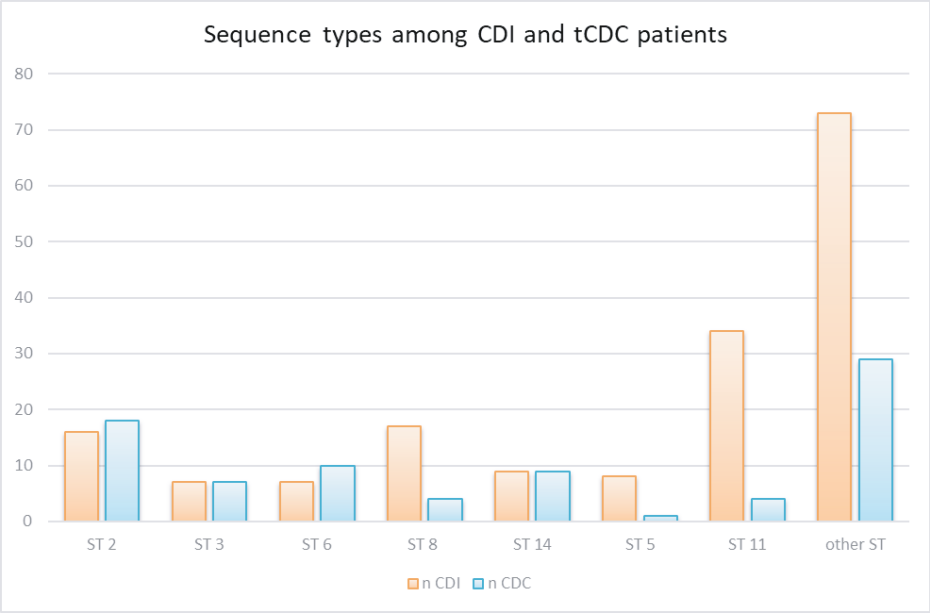


Figure 2. Sequence types among patients with *Clostridioides difficile* infection and toxigenic *Clostridioides difficile* colonization.
tCDC, toxigenic *Clostridoides difficile* colonization; CDI, *Clostridoides difficile* infection; ST, sequence type

Core genome MLST

Given the aforementioned cut-offs, in total 24 clusters could be identified (Figure S1). Six of these clusters (C1-C6) contained isolates from CDI and tCDC patients. Ward movement data for these clusters were investigated.

In cluster C1 (2 patients with RT020/220 from the same hospital), patients shared a ward 11 days before the first patient was found to be colonized and 37 days before the other patient was diagnosed with CDI (direction of transmission indeterminate). In cluster C2 (2 CDI and 2 CDC patients with RT265 from 2 different hospitals) a possible epidemiological link could be established between 2 of the patients: these patients were admitted to the same ward at the time of the first CDI positive sample and the second patient was found to be *C. difficile* colonized 42 days later at readmission (directional transmission from CDI to CDC).

In the other four clusters (C3-C6), no epidemiological link could be found (see legend Figure S1).

Thirteen clusters contained isolates from CDI episodes only (Figure S1). The largest of these clusters was earlier determined to be an outbreak of RT826 (7 samples RT826, one sample RT127) at a single ward in one of the hospitals (16).

Five pairs of genetically related isolates were detected in colonized patients (Figure S1). Four of these pairs were identified from the same hospital >20 days apart, but one of these pairs contained isolates from the same ward sampled only 2 days apart.

Discussion

In this multi-centre study, we screened 2211 patients on hospital admission and found that *C. difficile* colonization was present in 4.9% of admissions; colonization with toxigenic *C. difficile* strains was present in 3.1% of admissions, comparable to previously reported numbers (5, 17, 18). Identified strains among colonized patients were genetically diverse, indicating various reservoirs. Sixty-three percent of colonized patients were colonized by toxigenic strains, but in contrast to other studies (19, 20) not a single colonized patient developed CDI in the year after study enrolment. The reason why we could not confirm a high risk to develop CDI may either be that the number of colonized patients in our study was too low, and/or that there was truly no increased risk due to the local situation such as low numbers/absence of virulent strains circulating among colonized patients and low antimicrobial use (21).

We identified only one possible onwards transmission event from a colonized patient: a tCDC and CDI patient shared a ward before they tested positive for *C. difficile*. Our data are in contrast with published reports (6, 7), which could be explained by the low incidence setting in which our study was performed. During the study period, CDI incidence ranged from 1.87 to 4.59 CDI cases per 10,000 admission days among the hospitals (8, 22). Only one outbreak due to RT826 was detected (16). The hypervirulent RT027 was not detected in CDI nor in CDC patients. As higher transmission has been shown for certain lineages (23) the absence of these lineages may explain why transmission was infrequent in our study and also why no other large clusters between CDI patients were detected by cgMLST. Moreover, other local characteristics may play a role, like antimicrobial pressure and infection control policies.

Of note, we also detected a few genetically related pairs of isolates in colonized patients, suggesting a common source or transmission before admission, although the detection of genetical identical isolates on the same ward only 2 days apart raises the suspicion of transmission (either patient to patient or from the hospital environment) during admission in that particular case.

Our study had numerous strengths. We captured all CDI cases as all 4 hospitals participate in continuous sentinel CDI surveillance. Moreover, CDI diagnosis was not only based on laboratory tests, instead all cases underwent chart review by local infection control personnel and had clinical symptoms compatible with CDI. We included all CDI cases that occurred in the hospitals instead of CDI cases diagnosed on study wards only, as transmission may possibly extend beyond wards (24).

However, our study also has some limitations. First of all, we may have missed a substantial amount of *C. difficile* introductions into the hospitals due to study design (screening was performed on only a few specific wards per hospital) and difficulties in study execution (stool samples were only received from half of 5200 consenting subjects). During the study period, the total number of admissions in the four hospitals were 13987, 19424, 21220 and 25510, respectively, indicating that screening for colonization was not performed in the vast majority of these admissions. On the other hand, to account for *C. difficile* transmission extending beyond wards, all incident CDI cases from each entire hospital were included in cgMLST. Thereby, we are underestimating the contribution of colonized patients to CDI overall, as a source could possibly not be identified if a CDI case occurred on a ward where screening for colonization was not performed.

Furthermore, patients were only sampled once during the study. Consequently, we do not know how many patients were (a) transiently colonized, (b) persistent carriers, or (c) acquired colonization during admission, although this may affect both the risk for CDI progression and *C. difficile* transmission pressure. Moreover, we did only include hospital-onset CDI cases, thereby ignoring that transmission may not (directly) lead to symptomatic CDI. Patients that acquired *C. difficile* from a colonized patient during admission, but developed CDI only after discharge, have not been captured in our study.

Table 2. Univariate analysis of potential risk factors for CDC and tCDC.

	49 CDC patients	145 controls	CDC vs control (OR, 95% CI)	32 tCDC patients	95 controls	tCDC vs control (OR, 95% CI)
Male sex	23 (46.9%)	78 (53.8%)		16 (50.0)	56 (59.0%)	
Median age (IQR)	59 (47.5-67.5)	61 (52-68)		57 (48-71)	63 (52-70)	
Born in the Netherlands	45 (91.8%)	133 (91.7%)		30 (93.8)	88 (92.6%)	
Comorbidity						
Median Charlson Comorbidity Score (IQR)	3 (2-4)	3 (1-5)		3 (2-5)	3 (1-5)	
Solid organ transplant	12 (24.5%)	15 (10.3%)	2.8 (1.2-6.5)	10 (31.3)	11 (11.6%)	3.5 (1.3-9.2)
IBD	8 (16.3%)	7 (4.8%)	3.8 (1.3-11.2)	7 (21.9)	4 (4.2%)	6.4 (1.7-23.5)
Non-metastatic solid malignancy	9 (18.4%)	25 (17.2%)	1.1 (0.5-2.5)	5 (15.6)	18 (19.0%)	0.8 (0.3-2.3)
Metastatic solid malignancy	3 (6.1%)	16 (11.0%)	0.5 (0.1-1.9)	0 (0)	10 (10.5%)	-
Chronic kidney disease	15 (30.6%)	24 (16.6%)	2.2 (1.1-4.7)	11 (34.4)	17 (17.9%)	2.4 (0.98-5.9)
DM uncomplicated	7 (14.3%)	21 (14.5%)	1.0 (0.4-2.5)	5 (15.6)	11 (11.6%)	1.4 (0.5-4.4)
DM end-organ damage	3 (6.1%)	9 (6.2%)	1.0 (0.3-3.8)	2 (6.3)	7 (7.4%)	0.8 (0.2-4.3)
Myocardial infarction	5 (10.2%)	16 (11.0%)	0.9 (0.3-2.6)	3 (9.4)	11 (11.6%)	0.8 (0.2-3.0)
Peptic ulcer disease	3 (6.1%)	11 (7.6%)	0.8 (0.2-3.0)	2 (6.3)	9 (9.5%)	0.6 (0.1-3.1)
COPD	12 (24.5%)	19 (13.1%)	2.2 (0.95-4.8)	8 (25.0)	11 (11.6%)	2.5 (0.9-7.0)
Mild liver disease	4 (8.2%)	10 (6.9%)	1.2 (0.4-4.0)	4 (12.5)	7 (7.4%)	1.8 (0.5-6.6)
Severe liver disease	2 (4.1%)	5 (3.5%)	1.2 (0.2-6.3)	1 (3.3)	4 (4.2%)	0.7 (0.1-6.8)
HIV	0 (0%)	1 (0.7%)	-	0 (0)	1 (1.1%)	-
BMT or SCT	0 (0%)	2 (1.4%)	-	0 (0)	2 (2.1%)	-
Psychiatric disorder	6 (12.2%)	17 (11.7%)	1.1 (0.4-2.8)	4 (12.5)	10 (10.5%)	1.2 (0.4-4.2)
Previous diarrhea and CDI						
Diarrhea in previous 3 months	26 (53.1%)	59 (40.7%)	1.6 (0.9-3.2)	16 (50.0)	36 (37.9%)	1.6 (0.7-3.7)
Previous CDI	3 (6.1%)	1 (0.7%)	9.3 (0.9-91.9)	2 (6.3)	0 (0%)	-
Household member with previous CDI	0 (0%)	1 (0.7%)	1.0 (1.0-1.0)	0 (0)	1 (1.1%)	1.0 (1.0-1.0)

Table 2. *Continued.*

	49 CDC patients	145 controls	CDC vs control (OR, 95% CI)	32 tCDC patients	95 controls	tCDC vs control (OR, 95% CI)
Healthcare contact						
Previous hospital admission (last 12 months)	36 (73.5%)	72 (49.7%)	2.8 (1.4-5.7)	24 (75.0)	47 (49.5%)	3.1 (1.3-7.5)
Working in healthcare system	3 (6.1%)	17 (11.7%)	0.5 (0.1-1.8)	1 (3.1)	9 (9.5%)	0.3 (0.04-2.5)
Previous medication use (last 3 months)						
Antibiotics	34 (69.4%)	79 (54.5%)	1.9 (0.9-3.8)	23 (71.9)	51 (53.7%)	2.2 (0.9-5.3)
PPI or antacids	36 (73.5%)	87 (60.0%)	1.8 (0.9-3.8)	25 (78.1)	56 (59.0%)	2.5 (0.98-6.3)
Anti-cancer chemotherapy	3 (6.1%)	11 (7.6%)	0.8 (0.2-3.0)	2 (6.3)	7 (7.4%)	0.8 (0.2-4.3)
Immunosuppressants	25 (51.0%)	53 (36.6%)	1.8 (0.9-3.5)	19 (59.4)	34 (35.8%)	2.6 (1.2-6.0)
Animal contact						
Pet dog	22 (44.9%)	34 (23.5%)	2.7 (1.3-5.3)	17 (53.1)	21 (22.1%)	4.0 (1.7-9.3)
Pet cat	3 (6.1%)	18 (12.4%)	0.5 (0.1-1.7)	2 (6.3)	10 (10.5%)	0.6 (0.1-2.9)
Contact with livestock	3 (6.1%)	17 (11.7%)	0.5 (0.1-1.8)	3 (9.4)	11 (11.6%)	0.8 (0.2-3.0)
Children in household attending daycare	1 (2.0%)	5 (3.5%)	0.6 (0.1-5.1)	0 (0)	4 (4.2%)	-

BMT, bone marrow transplantation; CDC, *Clostridioides difficile* colonization; CDI, *Clostridioides difficile* infection; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; HIV, human immunodeficiency virus; IBD, inflammatory bowel disease; IQR, interquartile range; OR, Odds ratio; PPI, proton pump inhitor; SCT, stem cell transplantation; tCDC, toxigenic *C. difficile* colonization

Environmental swabs were not taken during our study though colonized patients may contaminate the hospital environment with spores that can persist for long times. A direct transmission link can be missing when *C. difficile* acquisition occurs at a later moment from this contaminated hospital environment.

Criteria to determine epidemiologic linkage were quite strict and did not take into account transmission beyond wards. In our study, data about patients' movements to other hospital areas (like the radiology department) were not available.

Another limitation includes the applied criteria to consider isolates to be the same strain, as these were originally based on single nucleotide variant analysis instead of cgMLST. As it is not known if the discriminatory power of both approaches is similar, we checked all comparisons with 3 or less allele differences in cgMLST. In all besides one comparison, one allele was equal to 1 single nucleotide polymorphism (SNP). In the last comparison, where only one allele difference was predicted, this allele had 2 SNPs. We therefore think that the criteria are still applicable to our study.

Implementing screening was difficult and burdensome, while those tCDC patients that were detected did not have a high risk of progressing to CDI themselves and were not identified as an important direct source for incident hospitalized CDI cases. However, tCDC patients may still contribute to *C. difficile* transmission by transmitting *C. difficile* to other patients who remain asymptomatically colonized instead of (directly) progressing to CDI, or by contaminating the hospital environment. The hospital environment can however be contaminated in many other ways, for example by CDI patients whose isolation precautions are lifted after resolution of symptoms but who are still shedding spores. Therefore, we think that we should focus on decreasing CDI susceptibility (e.g. by antimicrobial stewardship programs) and complying with general infection prevention measures to prevent spread from *C. difficile* and other nosocomial pathogens. Sentinel surveillance to monitor CDI incidence rates and circulating ribotypes and the use of molecular typing in case of suspected transmission is of value to detect clusters and outbreaks (25). A very typical example of this approach was the finding of the RT826 cluster that was already detected via sentinel surveillance (16), and turned out to be the only transmission between multiple CDI patients detected by cgMLST in this study.

Funding

This work was supported by the Netherlands Organization for Health Research and Development, ZonMW (grant 50-52200-98-035). This work was supported by Wellcome Sanger core funding (WT098051).

Conflicts of interest

All authors report no conflicts of interest.

Authors' contributions

MC and EK designed the study. MC, ET, CV, MV and JH were local investigators at the four study hospitals. CH and IS performed the culturing of the isolates and DNA extraction. MS performed the DNA sequencing, genome assembly and annotation. MS, NK and BH performed the bioinformatics analysis. MC performed the epidemiological analyses. MC and BH wrote the manuscript with the help and feedback of all other authors.

Acknowledgements

We thank the research nurses from participating hospitals for their support in obtaining patient materials and data. We thank the Wellcome Sanger Institute Pathogen Informatics team for informatics support.

Supplementary material

S1. Additional exclusion criteria

Patients were excluded if:

- admitted from other wards within the same hospital
- admitted for palliative care
- hemodynamically unstable
- residency outside the Netherlands
- unable to speak or read Dutch or English
- unable to participate in the verbal informed consent process on their own behalf or represented by a surrogate

S2. Definition of CDI according to the Dutch sentinel surveillance program

Patients had to have clinical suspicion of CDI (which included either diarrhea -defined as at least 3 loose stools per 24hrs for 2 days- or toxic megacolon, and the absence of an alternative explanation for diarrhea) in combination with a positive *C. difficile* test or the presence of pseudomembranous colitis by endoscopy/histopathology.

S3. Test methods used for diagnosing CDI in the 4 hospitals

Toxin A/B EIA (1 hospital), a Nucleic Acid Amplification Assay (1 hospital), a Nucleic Acid Amplfication Assay in combination with a Toxin A/B EIA (1 hospital) or a GDH EIA in combination with a Toxin A/B EIA (1 hospital).

S4. *C. difficile* culture methods

Stool samples and rectal swabs were plated directly on CLO plates (selective *C. difficile* medium containing cefoxitin, amphotericin B and cycloserine, BioMérieux, The Netherlands) and after ethanol shock on CNA (Columbia blood-agar containing colistin and nalidixic acid, BioMérieux , The Netherlands) and CLO plates. After 48 and 96hrs, plates were read. Presumptive colonies were identified by Matrix-Assisted Laser Desorption Ionization–Time Of Flight (MALDI-TOF) Mass Spectrometry and sent to the national reference laboratory for GluD PCR to confirm *C. difficile* presence. For rectal swabs, an additional enrichment

culture was performed. Swabs were inoculated in *C. difficile* enrichment modified broth (*Clostridium difficile* enrichment broth with 0.1% sodium taurocholate moxalactam (32mg/L), norfloxacin (12mg/L) and cysteine hydrochloride (500mg/L), Mediaproducts BV, Groningen, the Netherlands) for 5 days, and then subcultured on CLO plates.

S5. Methods for DNA preparation, sequencing and core genome MLST

C. difficile strains were anaerobically (re)cultured for 48hrs on TSS plates (Trypcase Soy agar + 5% sheep blood, Biomérieux). A single colony was picked, suspended in 9ml BHI (Brain Heart Infusion broth, Media Products Groningen) and incubated anaerobically overnight at 37°C with gentle shaking at 100rpm. The culture was centrifuged afterwards at 4000rpm for 10 minutes. The pellet was then resuspended in 800µl PBS (Phosphate-buffered saline Media Products Groningen), and 24µl of lysozyme (50 g/l) was added to a final concentration of 1,3g/l. The solution was incubated for 20 minutes at 37°C. Afterwards 15µl of proteinase K (20g/l, Roche) were added to a final concentration of 1,3g/l, and the solution was again incubated for 20 minutes at 37°C. DNA was extracted with the QiaSymphony and the program Complex800_OBL_V4_DSP. The extracted DNA was used for whole genome sequencing and ribotyping. Ribotyping was performed according to (1), with the following changes: PCR was performed on a BIO-RAD MyCyclerTM and analyzed on the Applied Biosystems 3500xL Genetic Analyzer. All samples were sequenced on an Illumina HiSeq X10 machine in paired-end mode, with a read length of 151.

Genomes were assembled using Velvet v1.2.10 (2), SSPACE v2.0 (3) and GapFiller v1.1 (4) as described previously (5), and annotated with Prokka, version 1.5 (6) with options -M n -V b. MLST types were determined with MLSTcheck (7) version 2.1.1706216. All reads were mapped to their respective genomes with bowtie2 v. 2.3.4.1 (8) and further converted to sorted bam files with samtools v. 1.6 (9). Optical duplicates were marked with PicardTools v. 1.124, and SNPs were called with HaplotypeCaller from the GATK package, v.4. 1.2 (10), with ploidy set to 2, to detect possibly mixtures of two different strains. Results were visualized with af-plot v 0.2.1, https://doi.org/10.5281/zenodo.3238297. Additionally all samples were profiled with checkm v1.0.13 (11). The typing scheme, sequences of alleles, and other relevant information for *C. difficile* core genome MLST (cgMLST) were downloaded from cgmlst.org. A blastn search (with standard parameters; Blast v 2.11) with all predicted *C. difficile* genes was performed against the cgMLST database. A gene was assigned to an allel if it was 100% identical over at least 95% of its length. A gene was tentatively assigned to an allel if it was at least 99% identical over 95% of its length. Assignment of cluster types

was performed via the hamming distance as implemented in Numpy/SciPy (12), based on the predicted allels. A minimum spanning tree was constructed via the networkX library (13), and was visualized in Cytoscape 3 (14). A limited amount of samples was also analysed in SeqSphere (15), for validation purposes. All genomic data has been uploaded to the European Nucleotide Archive under study number PRJEB25045.

We applied a threshold of <= 3 alleles in cgMLST as recent data show that this threshold better identifies outbreaks than the current threshold of 6 alleles, especially in ribotypes with lower mean intra-ribotype allele differences (Baktash, submitted).

S6. Library preparation: IHTP WGS NEB UltraII library process – DUAL/ QUAD processing using Sanger 168 tags and PE1 tag2 sets 1-4

- Samples quantified with Biotium Accuclear Ultra high sensitivity dsDNA Quantitative kit using Mosquito LV liquid platform, Bravo WS and BMG FLUOstar Omega plate reader and cherrypicked to 200ng / 120ul using Tecan liquid handling platform.
- Cherrypicked plates sheared to 450bp using a Covaris LE220 instrument.
- Post sheared samples purified using Agencourt AMPure XP SPRI beads on Agilent Bravo WS.
- Library construction (ER, A-tailing and ligation) using 'NEB Ultra II custom kit' on an Agilent Bravo WS automation system.
- PCR set-up using KapaHiFi Hot start mix and Sanger 168 tags (i7) and PE1.D1-PE1.D2 (Dual) or PE1.D1-PE1.D4 (Quad) tag2 (i5) tags on Agilent Bravo WS automation system.
- PCR cycles, 6 standard cycles,
- Post PCR plate purified using Agencourt AMPure XP SPRI beads on Beckman NX96 liquid handling platform.
- Libraries quantified with Biotium Accuclear Ultra high sensitivity dsDNA Quantitative kit using Mosquito LV liquid handling platform, Bravo WS and BMG FLUOstar Omega plate reader.
- 2 x 96 (Dual) or 4 x 96 (Quad) libraries pooled in equimolar amounts on a Beckman BioMek NX-8 liquid handling platform.
- Library pools normalised depending on sequencing platform. For NovaSeq, pools are normalised to 1.2nM (XP) or 2.25nM (Standard S4) and loaded on requested Illumina sequencing platform. For HiSeqX platforms pools normalised to 2.8nM.

S7. Definitions for epidemiological linkage

Epidemiologically linkage was plausible if (a) the donor-recipient pair shared a ward after the donor tested positive and before the recipient tested positive or (b) shared a ward before either tested positive or (c) if the recipient was admitted to a ward that had been occupied by the donor patient before. A maximum infectious period of 8 weeks and incubation period of 12 weeks were allowed. Both CDI and colonized patient were assumed to contaminate the ward for 26 weeks after testing positive (16).

Supplementary references

1. Fawley WN, Knetsch CW, MacCannell DR, Harmanus C, Du T, Mulvey MR, et al. Development and validation of an internationally-standardized, high-resolution capillary gel-based electrophoresis PCR-ribotyping protocol for Clostridium difficile. PLoS One. 2015;10(2):e0118150.
2. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18(5):821-9.
3. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics. 2011;27(4):578-9.
4. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. Genome Biol. 2012;13(6):R56.
5. Page. 2016 [Available from: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000083.
6. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.
7. Page AJ, Taylor J, Keane JA. Multilocus sequence typing by blast from de novo assemblies against PubMLST. The Journal of Open Source Software. 2016;1(8).
8. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.
9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
10. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491-8.
11. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043-55.
12. Bletz S, Janezic S, Harmsen D, Rupnik M, Mellmann A. Defining and Evaluating a Core Genome Multilocus Sequence Typing Scheme for Genome-Wide Typing of Clostridium difficile. J Clin Microbiol. 2018;56(6).

13. Hagberg AA, Schult DA, Swart PJ, editors. Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of the 7th Python in Science Conference; 2008; Pasadena, CA USA.
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.
15. Jünemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, et al. Updating benchtop sequencing performance comparison. Nat Biotechnol. 2013;31(4):294-6.
16. Walker AS, Eyre DW, Wyllie DH, Dingle KE, Harding RM, O'Connor L, et al. Characterisation of Clostridium difficile hospital ward-based transmission using extensive epidemiological data and molecular typing. PLoS Med. 2012;9(2):e1001172.

Description belonging to FIgure S1:
Colonized patients are shown as circles, CDI patients are shown as diamonds.
Recurrent episodes with identical isolates in one patient are shown as larger diamonds (the number of retrieved isolates in the patient is given, e.g. 2 x 001 means that the patient had one initial episode and one recurrence with an identical isolate of RT001).
Colonization with identical isolates in one patient is shown as a larger circle.
Both colonization and infection with identical isolates in one patient is shown as a diamond/circle combination.
'?' means that the ribotype was not recognized in our reference database.
Clusters with isolates from both colonized and infected patients are numbered (C1-C6).
In cluster C3 (2 patients with RT014 from the same hospital), CDC was detected almost 5 months after CDI.
In cluster C4 (2 patients with RT002 from 2 different hospitals) CDC was detected 7 months after CDI in the other hospital.
In cluster C5 (2 patients with RT001 from the same hospital) community-onset CDI was diagnosed more than 5 months after detection of the colonized patient and the CDI patient had not been admitted to the hospital in the 12 weeks preceding CDI diagnosis.
Cluster C6 consisted of 5 patients with RT001 from 3 different hospitals. In the first hospital CDI was diagnosed 4 months after CDC; patients had not been admitted to the same ward in the 12 weeks preceding CDI diagnosis. The other CDI patient in this cluster derived from a different hospital and CDC patients from the third hospital were sampled 6 months apart.
CDI: *Clostridoides difficile* infection, tCDC: toxigenic *Clostridioides difficile* colonization

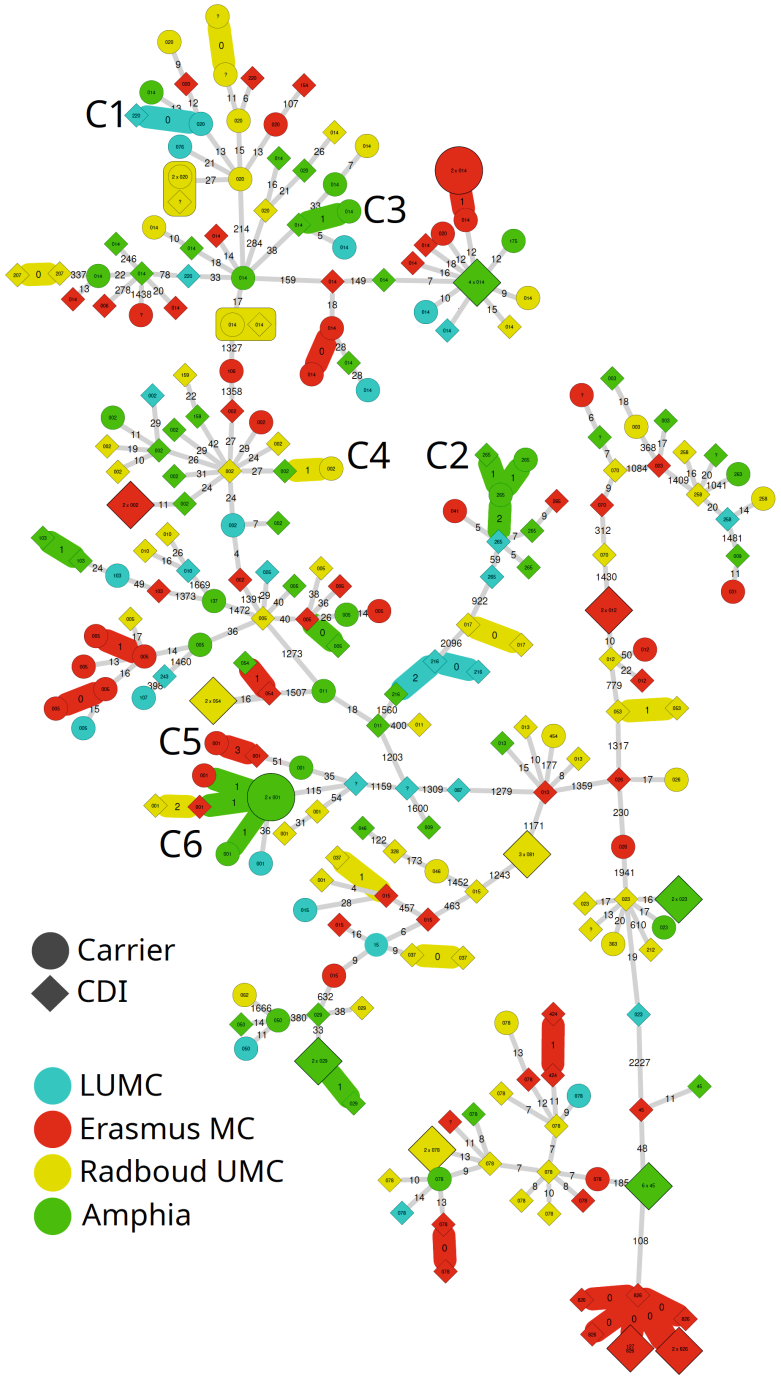


Figure S1. Core genome MLST of isolates from tCDC and CDI patients.
Description on left page.

References

1. Crobach MJT, Vernon JJ, Loo VG, Kong LY, Pechine S, Wilcox MH, et al. Understanding Clostridium difficile Colonization. Clinical microbiology reviews. 2018;31(2).
2. Riggs MM, Sethi AK, Zabarsky TF, Eckstein EC, Jump RL, Donskey CJ. Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic Clostridium difficile strains among long-term care facility residents. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2007;45(8):992-8.
3. Gilboa M, Houri-Levi E, Cohen C, Tal I, Rubin C, Feld-Simon O, et al. Environmental shedding of toxigenic Clostridioides difficile by asymptomatic carriers: A prospective observational study. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2020;26(8):1052-7.
4. Tschudin-Sutter S, Kuijper EJ, Durovic A, Vehreschild M, Barbut F, Eckert C, et al. Guidance document for prevention of Clostridium difficile infection in acute healthcare settings. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2018;24(10):1051-4.
5. Longtin Y, Paquet-Bolduc B, Gilca R, Garenc C, Fortin E, Longtin J, 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 internal medicine. 2016;176(6):796-804.
6. Kong LY, Eyre DW, Corbeil J, Raymond F, Walker AS, Wilcox MH, et al. Clostridium difficile: Investigating Transmission Patterns Between Infected and Colonized Patients Using Whole Genome Sequencing. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2019;68(2):204-9.
7. Curry SR, Muto CA, Schlackman JL, Pasculle AW, Shutt KA, Marsh JW, et al. Use of multilocus variable number of tandem repeats analysis genotyping to determine the role of asymptomatic carriers in Clostridium difficile transmission. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;57(8):1094-102.
8. Crobach MJ DvS, Terveer EM, Harmanus C, Sanders IMJG, Kuijper EJ, Notermans DW, Greeff de SC, Albas J, Dissel v JT. Eleventh Annual Report of the National Reference Laboratory for *Clostridium difficile* and results of the sentinel surveillance. 2017.
9. Fawley WN, Knetsch CW, MacCannell DR, Harmanus C, Du T, Mulvey MR, et al. Development and validation of an internationally-standardized, high-resolution capillary gel-based electrophoresis PCR-ribotyping protocol for Clostridium difficile. PLoS One. 2015;10(2):e0118150.
10. Persson S, Torpdahl M, Olsen KE. New multiplex PCR method for the detection of Clostridium difficile toxin A (tcdA) and toxin B (tcdB) and the binary toxin (cdtA/cdtB) genes applied to a Danish strain collection. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2008;14(11):1057-64.
11. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol. 2011;173(6):676-82.

12. Page. 2016 [Available from: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000083.
13. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.
14. Jünemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, et al. Updating benchtop sequencing performance comparison. Nat Biotechnol. 2013;31(4):294-6.
15. Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, et al. Diverse sources of C. difficile infection identified on whole-genome sequencing. N Engl J Med. 2013;369(13):1195-205.
16. Crobach MJT, Voor In 't Holt AF, Knetsch CW, van Dorp SM, Bras W, Harmanus C, et al. An outbreak of Clostridium difficile infections due to new PCR ribotype 826: epidemiologic and microbiologic analyses. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2018;24(3):309.e1-.e4.
17. Loo VG, Bourgault AM, Poirier L, Lamothe F, Michaud S, Turgeon N, et al. Host and pathogen factors for Clostridium difficile infection and colonization. N Engl J Med. 2011;365(18):1693-703.
18. Tschudin-Sutter S, Carroll KC, Tamma PD, Sudekum ML, Frei R, Widmer AF, et al. Impact of Toxigenic Clostridium difficile Colonization on the Risk of Subsequent C. difficile Infection in Intensive Care Unit Patients. Infect Control Hosp Epidemiol. 2015;36(11):1324-9.
19. Zacharioudakis IM, Zervou FN, Pliakos EE, Ziakas PD, Mylonakis E. Colonization with toxinogenic C. difficile upon hospital admission, and risk of infection: a systematic review and meta-analysis. Am J Gastroenterol. 2015;110(3):381-90; quiz 91.
20. Poirier D, Gervais P, Fuchs M, Roussy JF, Paquet-Bolduc B, Trottier S, et al. Predictors of Clostridioides difficile Infection Among Asymptomatic, Colonized Patients: A Retrospective Cohort Study. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2020;70(10):2103-210.
21. Within Europe, antibiotics use lowest in the Netherlands, data from Centraal Bureau voor Statistiek 2016 [Available from: https://www.cbs.nl/en-gb/news/2016/06/within-europe-antibiotics-use-lowest-in-the-netherlands.
22. Crobach MJ DvS, Harmanus C, Sanders IMJG, Kuijper EJ, Notermans DW, Greeff de SC, Albas J, Dissel v JT. Tenth Annual Report of the National Reference Laboratory for Clostridium difficile and results of the sentinel surveillance. 2016.
23. Martin JSH, Eyre DW, Fawley WN, Griffiths D, Davies K, Mawer DPC, et al. Patient and strain characteristics associated with Clostridium difficile transmission and adverse outcomes. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018.
24. Murray SG, Yim JWL, Croci R, Rajkomar A, Schmajuk G, Khanna R, et al. Using Spatial and Temporal Mapping to Identify Nosocomial Disease Transmission of Clostridium difficile. JAMA internal medicine. 2017;177(12):1863-5.
25. Krutova M, Wilcox MH, Kuijper EJ. A two-step approach for the investigation of a Clostridium difficile outbreak by molecular methods. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2019;25(11):1300-1.