Description and validation of a new automated surveillance system for *Clostridium difficile* in Denmark

M. CHAINE[1]*, S. GUBBELS[1], M. VOLDSTEDLUND[1], B. KRISTENSEN[1],
J. NIELSEN[1], L. P. ANDERSEN[2], S. ELLERMANN-ERIKSEN[3], J. ENGBERG[4],
A. HOLM[5], B. OLESEN[6], H.C. SCHØNHEYDER[7,8], C. ØSTERGAARD[9],
S. ETHELBERG[1] AND K. MØLBAK[10]

[1] *Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut, Copenhagen, Denmark*
[2] *Department of Clinical Microbiology, Rigshospitalet, Copenhagen, Denmark*
[3] *Department of Clinical Microbiology, Aarhus University Hospital, Aarhus, Denmark*
[4] *Department of Clinical Microbiology, Slagelse Hospital, Slagelse, Denmark*
[5] *Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark*
[6] *Department of Clinical Microbiology, Herlev and Gentofte Hospital, Copenhagen, Denmark*
[7] *Department of Clinical Microbiology, Aalborg University Hospital, Aalborg, Denmark*
[8] *Department of Clinical Medicine, Aalborg University Hospital, Aalborg, Denmark*
[9] *Department of Clinical Microbiology, Hvidovre Hospital, Hvidovre, Denmark*
[10] *Division of Infectious Disease Preparedness, Statens Serum Institut, Copenhagen, Denmark*

SUMMARY

The surveillance of *Clostridium difficile* (CD) in Denmark consists of laboratory based data from Departments of Clinical Microbiology (DCMs) sent to the National Registry of Enteric Pathogens (NREP). We validated a new surveillance system for CD based on the Danish Microbiology Database (MiBa). MiBa automatically collects microbiological test results from all Danish DCMs. We built an algorithm to identify positive test results for CD recorded in MiBa. A CD case was defined as a person with a positive culture for CD or PCR detection of toxin A and/or B and/or binary toxin. We compared CD cases identified through the MiBa-based surveillance with those reported to NREP and locally in five DCMs representing different Danish regions. During 2010–2014, NREP reported 13 896 CD cases, and the MiBa-based surveillance 21 252 CD cases. There was a 99·9% concordance between the local datasets and the MiBa-based surveillance. Surveillance based on MiBa was superior to the current surveillance system, and the findings show that the number of CD cases in Denmark hitherto has been under-reported. There were only minor differences between local data and the MiBa-based surveillance, showing the completeness and validity of CD data in MiBa. This nationwide electronic system can greatly strengthen surveillance and research in various applications.

Key words: *Clostridium difficile*, gastrointestinal infections, hospital-acquired (nosocomial) infections, surveillance system.

* Author for correspondence: M. Chaine, Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut, Copenhagen, Denmark.
(Email: manc@ssi.dk)

INTRODUCTION

Clinical infection with *Clostridium difficile* (CD) usually occurs due to a disruption of the intestinal flora following antibiotic treatment, mostly with

broad-spectrum antibiotics [1–3]. CD is a Gram-positive bacterium, which is able to produce spores, two types of toxins, A and B, and, in some strains, a binary toxin [4]. Concerns have been raised specifically for the PCR ribotype 027 and other hypervirulent strains that may cause higher mortality, and may have a greater potential to cause hospital outbreaks [5–8].

In the USA, it was estimated in 2013 that 250 000 cases were annually attributable to a healthcare-associated CD infection [9]. An incidence of 7·4 per 10 000 patient-days of healthcare-associated CD infection was reported based on data from three states in the USA in 2010 [10]. In 2008, a European study aimed to assess the extent of CD infections (CDI) and tested patients with suspicion of CDI as well as patients developing diarrhoea 3 or more days after hospital admission. Three major Danish hospitals participated and found that 8% of tested patients had an infection with toxin-positive CD [11]. In the same year, a rise in number of patients with CDI was observed in the Capital Region of Denmark [12, 13], which was attributed to several hospital outbreaks. In 2009, the number of CDI started to rise in the neighbouring Region Zealand, probably due to the transfer of patients [12]. Results from studies using whole genomic sequencing techniques suggest that the community serves as a large reservoir of CD carriers, where transmission routes between patients within and outside hospitals are complex. Therefore, nationwide surveillance of CD, not only focusing on hospital-associated cases, is important [14].

Kola *et al.* reported that in 2011, 14 European countries had an ongoing nationwide surveillance system for CDI [15]. The surveillance systems were either case-based, laboratory-based or a combination of both. All systems required manual or active reporting in the data collection. Increasing evidence suggests that automatic electronic surveillance based on laboratory data is superior to conventional surveillance with respect to completeness and timeliness [16]. For the current Danish surveillance system, Departments of Clinical Microbiology (DCMs) actively report findings of gastrointestinal bacteria including CD to Statens Serum Institut (SSI). Data are entered in a database called the National Registry of Enteric Pathogens (NREP), which is used for surveillance of gastrointestinal pathogens in general [17]. Following the outbreaks in 2008, the DCMs were requested from 2009 to submit selected isolates for further typing to the national reference laboratory at SSI [13, 18].

In this study, we validated a new electronic surveillance system for CD based on the Danish Microbiology Database (MiBa) [19]. The aim of the new surveillance system (MiBa-based surveillance) is to improve national surveillance of infectious diseases in terms of completeness and timeliness. It is also envisioned to reduce the burden of reporting by machine-to-machine communication and replace manual procedures of reporting. The aim of this paper is to validate a computer algorithm identifying CD cases with toxigenic CD from MiBa, and compare the MiBa-based surveillance with the existing surveillance system for CD. Furthermore, we determined the completeness and quality of MiBa to local extracts from five DCMs.

METHODS

MiBa-based surveillance

MiBa is a real-time database that automatically collects microbiological test results from each Danish DCM at the time the electronic report is sent to the physician who requested the analysis. Reports are sent by a national standard transfer protocol. MiBa started collecting data from 1 January 2010 [19]. Patients are identified with their Danish civil registration number (CPR number), a unique number that each Danish resident gets upon birth or immigration [20]. For this study, data were extracted from MiBa from 1 January 2010 to 31 December 2014. The extract was made on 22 December 2016.

We defined a CD case as a person (i.e. a unique CPR number) with a positive CD culture or PCR detection of toxin A and/or B and/or binary toxin. Results specifying that the strain was non-toxigenic were excluded. Table 1 shows the steps in the computer algorithm used to identify positive test results from MiBa.

Data were imported from three related tables within the MiBa data model. The DCMs use different laboratory information systems with variable levels of compatibility with the MiBa data model. Although, the input data had differences in data structure and content, the majority of the information in the reports are structured and coded. A first step removed records where the strain was indicated to be non-toxigenic. Subsequently, results for positive CD cultures were retrieved. The next steps aimed at extracting the positive results for CD PCR, both from coded information, from pre-set text strings and free text. Finally, all data selected along each step were combined into one dataset containing all results positive for CD.

Table 1. *Summary of the algorithm used to identify CD results (excluding non-toxigenic strains) from MiBa*

Rule number	Description
1	Data import of the main table of MiBa (header), the table with identification of isolates (isolate) and the table with results (quantitative) based on tests performed for CD: records with the following investigation codes (code used for requisition): 10800 culture (pathogenic intestinal bacteria); 10801 culture (pathogenic intestinal bacteria and *Escherichia coli*); 10834 CD (culture); 12115 CD DNA/RNA; 15100 CD toxin A; 15111 CD toxins; 41150 microscopy and culture (parasites and pathogenic intestinal bacteria); 59001 diarrhoeal investigation; 59002 intestinal bacteria DNA/RNA.
	OR
	Records with the following analysis codes: 113 CD binary toxin; 114 CD PCR; 115 CD ribotype 027; 116 CD toxin; 117 CD toxin A + B; 118 CD toxin B.
	OR
	Records with the microorganisms code for CD: 5316 CD.
2	Records with specific free-text strings describing non-toxigenic strains and PCR negative results were excluded.
3	The sampling date was determined.
	If the sampling date was missing, the date of receipt in the Department of Clinical Microbiology was used.
	If the sampling date was more than 7 days before the date of receipt, then the date of receipt was used.
	If the sampling date was after the date of receipt, the date of receipt was used.
4	The time of sampling was determined.
	If the sampling time was not available, and the date of sampling and receipt were the same then the time of sampling was set to 4 h (parameter) before receipt.
	Otherwise, the time of sampling was set to 8:00.
5	All culture results indicating the microorganism code for CD were selected into a subset. Of these, the ones with a text that described a negative PCR were excluded.
6 and 7	Of the remaining records, pre-defined text strings for positive PCR results were identified from two fields and saved in subsets.
8, 9 and 10	Of the remaining records, free texts strings for positive PCR and culture results were identified from three fields and saved in subsets.
11	All subsets of data extracted from rules 5 through 10 were appended into one dataset containing all positives results for CD.
12	All results from the same date and time were combined into one record. The information of all records was maintained.

CD, *Clostridium difficile*; MiBa, Danish Microbiology Database.

National surveillance data for the evaluation of the surveillance system

To assess the performance of the MiBa-based surveillance on the national level, we compared it to the existing surveillance from NREP. Surveillance for gastrointestinal pathogens is laboratory based, and all DCMs report to NREP on a weekly basis. According to the ministerial order for notifiable diseases, findings of enteropathogenic bacteria need to be notified, including *Salmonella enterica*, *Campylobacter jejuni/coli*, *Yersinia enterocolitica*, *Shigella* spp., *Vibrio cholerae*, diarrhoeagenic *Escherichia coli*, as well as other bacteria that by the diagnostic laboratory are judged to cause gastrointestinal disease [21]. Because CD is not specifically mentioned in the list of agents, it is interpreted to fall under the wording 'other bacteria' in agreement with the local DCM. When a CD case is registered in NREP, it includes patients with a positive culture for CD or PCR detection of toxin A and/or B and/or binary toxin for CD. Patients are identified with their CPR number. NREP records a new CD case if it occurred 6 months or more after the first positive sample of the previous case. The study period of the national surveillance data was from 1 January 2010 to 31 December 2014.

Regional data for validation of data quality of the new electronic surveillance system

To validate the quality and completeness of the MiBa-based surveillance, we obtained data from five DCMs representative of different parts of Denmark: Capital Region of Denmark (Hvidovre, Herlev and Rigshospitalet), North Denmark Region and Region Zealand. The extracts included all results (from cultures and PCR) positive for CD, including information on

Table 2. *Demographic description of Clostridium difficile cases in the MiBa-based surveillance and NREP (based on the first positive sample) by sex, age and region of residence for the period 2010–2014*

	MiBa-based surveillance* N = 21 252	NREP[†] N = 13 896
Sex – # patients (%)		
F	11 749 (55)	7777 (56)
M	9503 (45)	6119 (44)
Age – median (interquartile range)		
All	73 (59–83)	74 (60–83)
F	75 (59–85)	75 (60–84)
M	72 (59–81)	72 (60–81)
Region of residence – # patients (%)[‡]		
Capital Region of Denmark	9026 (43)	3944 (28)
Central Denmark Region	2542 (12)	2181 (16)
North Denmark Region	2090 (10)	1945 (14)
Region Zealand	3493 (17)	2820 (20)
Region of Southern Denmark	3862 (18)	3006 (22)

F, females; M, males; MiBa, Danish Microbiology Database.

* MiBa-based surveillance: surveillance system based on the Danish Microbiology Database.

† NREP: National Registry of Enteric Pathogens.

‡ The geographic region of residence could not be retrieved for 139 patients from the MiBa-based surveillance.

toxin production/presence of toxin genes. If specified as non-toxigenic, results were excluded. For North Denmark Region and Region Zealand, the study period was 2010–2014 and for Capital Region of Denmark 2011–2014. The data linkage to the MiBa-based surveillance for the DCM North Denmark Region and Region Zealand was done on the sample identifier and the CPR number. For the DCM in the Capital Region of Denmark, the data linkage was only available on the CPR number.

Data analysis

Children <2 years old were excluded from all datasets as well as temporary CPR numbers.

The European surveillance protocol for CDI excludes children <2 years old with positive CD test results unless clinical evidence to prove otherwise [22]. Children <2 years old have been reported in several studies as asymptomatic carriers [23–26].

The number of CD cases, as identified through the MiBa-based surveillance and NREP, was described by age, sex and geographic region of residence, which were retrieved from the Danish Civil Registration System [20].

The overlap between systems was assessed and discrepant cases were further investigated in order to understand the reasons for the discrepancies.

Data management and analysis were carried out by use of the statistical software SAS/STAT™ 9·4 (SAS Institute Inc., Cary, North Carolina, USA).

RESULTS

There were 60 698 positives samples and 22 748 cases in MiBa. We excluded 163 cases with a non-valid CPR (503 samples) and 1333 children <2 years of age (3259 samples).

The MiBa-based surveillance system *vs* the current national surveillance NREP

We obtained 21 252 CD cases from the MiBa-based surveillance compared with 13 896 CD cases reported to NREP (Table 2).

The median age of patients in the MiBa-based surveillance was 73 years [interquartile range (IQR) 59–83 years], similar to the age distribution of cases in NREP (74 years, IQR 60–83 years). The proportion of men and women was similar in both surveillance systems (Table 2).

Table 3 shows the number of concordant and discordant CD cases between the current national surveillance NREP and the MiBa-based surveillance system.

The comparison between NREP and MiBa-based surveillance revealed an under-reporting to the NREP,

Table 3. *Comparison of Clostridium difficile cases identified in the surveillance based on the MiBa, local data and prevailing surveillance databases for the period 2010–2014*

Validation data	Period	MiBa-based surveillance +/validation data− (# patients)	MiBa-based surveillance −/validation data+ (# patients)	MiBa-based surveillance +/validation data+ (# patients)	Total study population
NREP	2010	2071	370	1670	4111
	2011	2505	173	2284	4962
	2012	2451	77	2443	4971
	2013	1302	30	3694	5026
	2014	1107	15	3655	4777
	All years	7968	612	13 284	21 864
Three DCMs in Capital Region of Denmark*	2011–2014	31	8	7215	7254
DCM North Denmark Region†	2010–2014	15	3	5454	5472
DCM Region Zealand†	2010–2014	18	11	6388	6417

DCM, Department of Clinical Microbiology; MiBa-based surveillance, surveillance system based on the Danish Microbiology Database; NREP, National Registry of Enteric Pathogens.

* DCM Hvidovre, DCM Herlev and DCM Rigshospitalet.

† Comparisons based on sample identifier instead of unique patient (CPR number).

especially in the period 2010–2012 (Table 3 and Fig. 1). This under-reporting was most significant for the Capital Region of Denmark and may have been due to cumbersome manual extraction. When it was clarified in 2013 that also PCR-based detection of CD toxins should be reported, completeness was improved. Figure 1 shows that the national and regional trends with NREP were consistently lower than the MiBa-based surveillance but with similar trends. Completeness increased over time and the highest number of discrepant CD cases found only in NREP was from 2010.

Validation of the data quality in the MiBa-based surveillance system

Table 3 shows the number of concordant and discordant CD cases and positives samples identified in the MiBa-based surveillance and the regional data. The discrepancies are described in detail below.

The MiBa-based surveillance *vs* three DCMs from Capital Region of Denmark

Of the eight cases found in the database from the Capital Region of Denmark, but not identified by the MiBa-based surveillance, four patients were not recorded in MiBa at all, and four were PCR positive

in the local extract but PCR negative in MiBa. For the 31 patients found only in the MiBa-based surveillance, three were reported in late December 2014 in MiBa and in early January 2015 in the DCM's extract. This discrepancy was found because the three samples were recorded in the Capital Region of Denmark database 1 day after the cut-off date of our comparison. For the remaining 28, no clear explanation could be found for the discrepancy.

The MiBa-based surveillance *vs* DCM from North Denmark Region

Among the three samples found only in the data from DCM North Denmark Region, all were reported in MiBa as a negative PCR for CD. Of the 15 samples found only in the MiBa-based surveillance, six samples were recorded on 31 December 2014 in MiBa and recorded on 1 January 2015 in the DCM's extract. For seven samples, the cultures were performed on a biopsy, and were therefore not included into the local extract. For the last two samples, there was no clear reason.

The MiBa-based surveillance *vs* DCM from Region Zealand

The 11 samples recorded only in DCM Region Zealand included a borderline test result of the PCR
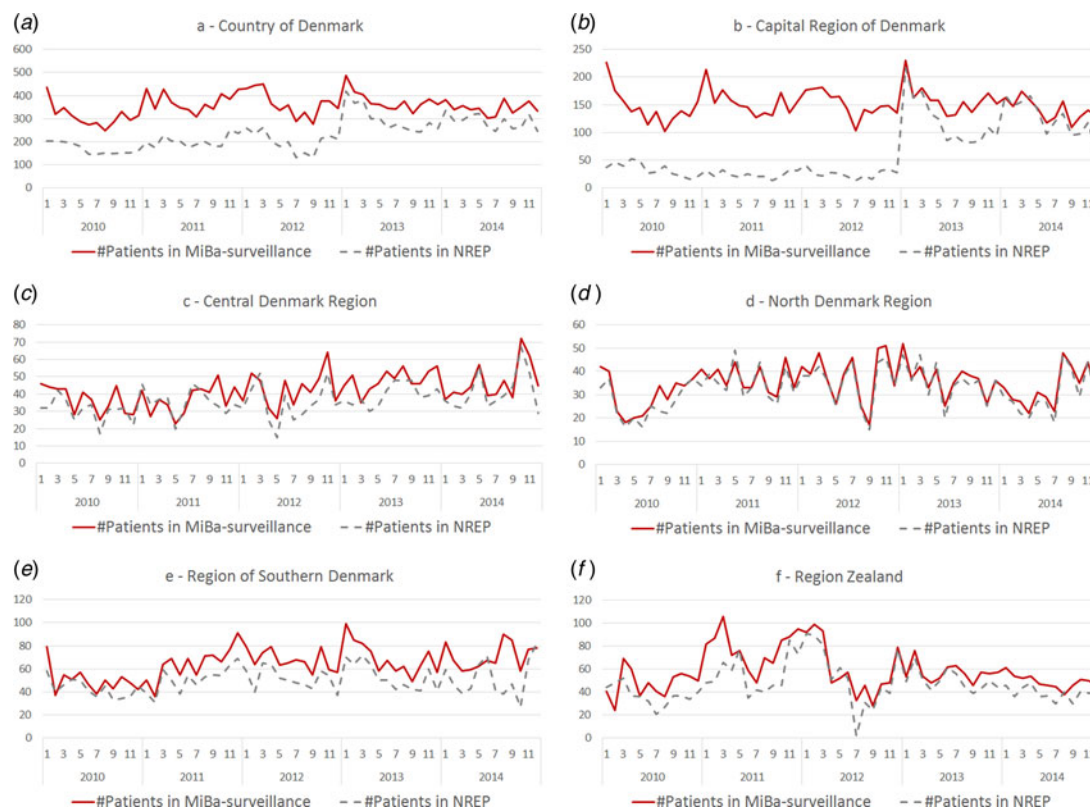
Fig. 1. Number of patients with a *Clostridium difficile* infection in the surveillance based on the Danish Microbiology Database (MiBa-based surveillance) and the National Registry of Enteric Pathogens (NREP) by month of first positive sample for Denmark as a whole and by region for the period 2010–2014.

reaction, which was manually corrected before the report was sent to MiBa, while the original test result was kept in the local database. Therefore, the result in MiBa was the correct one. The 18 samples found only in the MiBa-based surveillance were recorded as a positive PCR result in MiBa, but negative in the local extract. In these cases, the final report had been manually corrected from negative to positive, based on a manual interpretation of a weak but significant PCR signal. The original test result was kept in the local database and the result in MiBa was the correct one.

DISCUSSION

In this study, we validated a new fully automated surveillance system for CD cases based on laboratory reports in MiBa.

When compared with the current surveillance through the NREP, we found that the MiBa-based surveillance recorded one-third more patients with a positive test for CD than did NREP. This suggests that the number of CD cases in Denmark has been underreported, possibly because of cumbersome manual reporting procedure. When looking in detail at the few patients that were identified by NREP, but not by the MiBa-based surveillance, most of the test results were from 2010, suggesting that start-up problems with electronic transfers of reports to MiBa have been solved in later years. The trends observed on regional and national level, for the MiBa-based surveillance and NREP suggested that reporting to MiBa is more complete than to NREP.

By comparing the cases identified in MiBa with extracts from five of the 13 DCMs, we were able to assess the quality and completeness of the MiBa-based surveillance system. Data from the MiBa-based surveillance almost completely matched the positive tests for CD in these five selected regional datasets. Finally, most discrepancies in the comparison analysis could be explained, which provides reassurance for the use of MiBa as a future data source for laboratory-based surveillance of CD cases.

There are two aspects of the algorithm that could theoretically lead to discrepancies, although this was not seen in the presented validation studies. In the

first place, rule 2 of the algorithm deletes records specifying non-toxigenic strains. However, if a DCM does not report this information to MiBa, the algorithm cannot distinguish non-toxigenic strains and will keep them in the dataset – and thereby overestimate the number of toxigenic cases in MiBa. In order to solve this issue, there is a need for DCMs to report laboratory results in a more detailed manner to MiBa. In the second place, rule 2 can only remove cases where a positive culture was followed by a negative PCR for toxins, if these two results were reported in one record. A more advanced algorithm would need to be developed to handle situations in which these two results are reported in separate records with different dates. Finally, rules 6 and 7 depend on free text. This makes the system sensitive to changes in the content applied at the local DCM. It could lead to a loss of records if a DCM introduces new texts and requires regular evaluation of text strings and updates of the algorithm.

However, compared with earlier reported electronic surveillance systems for CDI [16, 27], data in MiBa are complete at a national level. Most information in the reports sent to MiBa is compliant to a national standard transfer protocol, rendering data structured and coded, making the data delivery relatively robust over time. At four CDC Prevention Epicenters hospitals, an automated surveillance system for CDI was developed [16], which was found more reliable to use, compared with manual reporting. However, different algorithms were used for each hospital to cope with local differences in available data, potentially resulting in different disease rates. Another automated surveillance was developed in the USA [27]. This system is also based on a selected cluster of hospitals, and only CDI related to an admission is captured. The majority of data sources were based on free text, and data were captured by text mining. In neither system were the data population based, and it was not possible to trace individuals across hospitals systems. However, in both systems, cases were classified according to hospital admissions. Since 2015, the presented MiBa-based surveillance has been integrated into the hospital-acquired infection database (HAIBA), where individual cases are linked to patient administrative data from hospital systems, to monitor hospital-onset hospital-acquired CDI and community-onset hospital-acquired CDI [28]. To our knowledge, few countries have developed nationwide population-based electronic surveillance systems, including individual person identification, making it possible to trace individuals across institutions and link data across registries. The Infection Intelligence Platform being developed in Scotland shares many of the same features as the Danish MiBa-based surveillance and population-based studies at national level on CD are taking place [29, 30].

It is a limitation of the current NREP surveillance system and the MiBa-based surveillance system that they do not provide information on subtyping. An additional surveillance is run at the national reference laboratory at SSI with specific focus on the surveillance of CD subtypes, based on selected CD isolates referred from the DCMs to SSI for subtyping. In the near future, MiBa will be able to handle information on subtypes in a structured and coded form. Integrating this information in the MiBa-based surveillance would mean a great improvement in the usefulness of the surveillance data.

In the present study, we did not assess the improvements in timeliness or the reduction in burden of reporting by the laboratories. However, we are certain that the MiBa-based surveillance is by far superior to the current NREP surveillance, in both respects, although automated systems do have important pitfalls that deserve attention. The landscape of modern healthcare is highly dynamic, and organisations, laboratory techniques and IT systems change constantly. One must not forget the centrally and locally based resources that are required to maintain the MiBa-based surveillance and to secure data quality.

In conclusion, despite different local laboratory information systems, we show that automated laboratory-based national surveillance for CD is both feasible and complete. The benefit of such a system is that it does not require active reporting. The existing surveillance systems underestimated the incidence of CDI. Based on the new and validated MiBa-based surveillance system, we found that CD, with about 4500 cases per year, may be the most common cause of bacterial gastroenteritis in Denmark, probably exceeding the annual number of confirmed *Campylobacter* infections that ranged between 3730 and 4350 in the period from 2010 to 2015. We also concluded that the MiBa-based surveillance system for CD can effectively replace the current surveillance system.

This nationwide MiBa-based surveillance system for CDI can greatly strengthen surveillance and research in various applications. For example, when linked to patient administrative data, the relation to the healthcare system can be assessed as it is shown in HAIBA.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None declared.

ETHICAL CONSIDERATIONS

This study was approved by the Danish Data Protection Authority as part of the development of the Danish Hospital-Acquired Infections Database (registration number 2015-54-0942).

REFERENCES

1. **Kelly CP, et al.** *Clostridium difficile* colitis. *New England Journal of Medicine* 1994; **330**: 257–262.

2. **Olson MM, et al.** Ten years of prospective *Clostridium difficile*-associated disease surveillance and treatment at the Minneapolis VA Medical Center, 1982–1991. *Infection Control and Hospital Epidemiology* 1994; **15**: 371–381.

3. **Martinez FJ, et al.** *Clostridium difficile* outbreaks: prevention and treatment strategies. *Risk Management and Healthcare Policy* 2012; **5**: 55–64.

4. **Akerlund T, et al.** Correlation of disease severity with fecal toxin levels in patients with *Clostridium difficile*-associated diarrhea and distribution of PCR ribotypes and toxin yields *in vitro* of corresponding isolates. *Journal of Clinical Microbiology* 2006; **44**: 353–358.

5. **Warny M, et al.** Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *The Lancet* 2005; **366**: 1079–1084.

6. **Labbe A-C, et al.** *Clostridium difficile* infections in a Canadian tertiary care hospital before and during a regional epidemic associated with the BI/NAP1/027 strain. *Antimicrobial Agents and Chemotherapy* 2008; **52**: 3180–3187.

7. **Bacci S, et al.** Binary toxin and death after *Clostridium difficile* infection. *Emerging Infectious Diseases* 2011; **17**: 976–982.

8. **Walker AS, et al.** Relationship between bacterial strain type, host biomarkers, and mortality in *Clostridium difficile* infection. *Clinical Infectious Diseases* 2013; **56**: 1589–1600.

9. **Center of Diseases Control and Prevention (CDC).** Antibiotic resistance threats in the United States, 2013 (http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf). Accessed 19 January 2017.

10. **Centers for Disease Control and Prevention (CDC).** Vital signs: preventing *Clostridium difficile* infections. *MMWR Morbidity and Mortality Weekly Report* 2012; **61**: 157–162.

11. **Bauer MP, et al.** *Clostridium difficile* infection in Europe: a hospital-based survey. *The Lancet* 2010; **377**: 63–73.

12. **St-Martin G, et al.** *Clostridium difficile*. Statens Serum Institut EPI-NEWS; 25 March: No. 13-2009.

13. **Olsen KEP.** *Clostridium difficile* 2009–2011. Statens Serum Institut EPI-NEWS; 22 February: No 7-8-2012.

14. **Eyre DW, et al.** Diverse sources of *C. difficile* infection identified on whole-genome sequencing. *The New England Journal of Medicine* 2013; **369**.

15. **Kola A, et al.** Survey of *Clostridium difficile* infection surveillance systems in Europe, 2011. *Eurosurveillance* 2016; **21**.

16. **Dubberke ER, et al.** Implementing automated surveillance for tracking *Clostridium difficile* infection at multiple healthcare facilities. *Infection Control and Hospital Epidemiology* 2012; **33**: 305–308.

17. **Bager P, et al.** Cesarean delivery and risk of intestinal bacterial infection. *The Journal of Infectious Diseases* 2010; **201**: 898–902.

18. **Tema om *Clostridium difficile*.** Statens Serum Institut (http://www.ssi.dk/Aktuelt/Temaer/Generelle20temaer/Clostridium20difficile.aspx). Accessed 17 July 2016.

19. **Voldstedlund M, et al.** The Danish Microbiology Database (MiBa) 2010 to 2013. *Eurosurveillance* 2014; **19**.

20. **Pedersen CB, et al.** The Danish Civil Registration System. A cohort of eight million persons. *Danish Medical Bulletin* 2006; **53**: 441–449.

21. **Ministeriet for Sundhed og Forebyggelse.** Bekendtgørelse om lægers anmeldelse af smitsomme sygdomme m.v. BEK nr 277; 14 April (https://www.retsinformation.dk/Forms/R0710.aspx?id=21406). Accessed 3 June 2016.

22. **European Centre for Disease Prevention and Control.** European Surveillance of *Clostridium difficile* infections. Surveillance protocol version 22, 2015.

23. **Larson HE, et al.** Epidemiology of *Clostridium difficile* in infants. *Journal of Infectious Diseases* 1982; **146**: 727–733.

24. **Hall IC, O'Toole E.** Intestinal flora in new-born infants: with a description of a new pathogenic anaerobe,

bacillus difficilis. *American Journal of Diseases of Children* 1935; **49**: 390–402.

25. **Collignon A, et al.** Heterogeneity of *Clostridium difficile* isolates from infants. *European Journal of Pediatrics* 1993; **152**: 319–322.

26. **Shim JO.** *Clostridium difficile* in children: to treat or not to treat? *Pediatric Gastroenterology, Hepatology & Nutrition* 2014; **17**: 80–84.

27. **Benoit SR, et al.** Automated surveillance of *Clostridium difficile* infections using BioSense. *Infection Control and Hospital Epidemiology* 2011; **32**: 26–33.

28. **Gubbels S, et al.** HAIBA - Hospital-Acquired Infection database. Statens Serum Institut EPI-NEWS; 25 February: No 9-2015.

29. **Information Services Division Scotland.** Infection Intelligence Platform (http://www.isdscotland.org/Health-Topics/Health-and-Social-Community-Care/Infection-Intelligence-Platform/About-IIP/). Accessed 3 May 2017

30. **Banks A-L, et al.** National study of outcome after *Clostridium difficile* infection (CDI) in Scotland. Information services division Scotland.