

## *Clostridioides difficile* Prevalence in the United States: National Inpatient Sample, 2016 to 2018

Molly R. Petersen,<sup>1</sup> Sara E. Cosgrove,<sup>2</sup> Eili Y. Klein,<sup>3</sup> Xianming Zhu,<sup>1</sup> Thomas C. Quinn,<sup>2,4</sup> Eshan U. Patel,<sup>1,5</sup> M. Kate Grabowski,<sup>1,5</sup> and Aaron A. R. Tobian<sup>1,2,5</sup>

<sup>1</sup>Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA, <sup>2</sup>Department of Medicine, Division of Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA, <sup>3</sup>Department of Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA, <sup>4</sup>Laboratory of Immunoregulation, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA, <sup>5</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

Data from the National Inpatient Sample indicate that *Clostridioides difficile* prevalence decreased from 10.1 (95% confidence interval [CI] = 9.9–10.3) to 8.6 (95% CI = 8.5–8.8) per 1000 hospital discharges between 2016 and 2018, after accounting for age, sex, and race. There was heterogeneity in the prevalence and decrease in prevalence by geographic region in the United States.

**Keywords.** antimicrobial stewardship; *Clostridioides difficile*; National Inpatient Sample; NIS.

*Clostridioides difficile* infections pose a significant threat to public health and place a large burden on the economy. Up to 40% of people who acquire a *C difficile* infection in the community will need to be hospitalized, and the estimated mortality for *C difficile* infection is 5% [1]. Cases increased between 2000 and 2009; however, in the past decade, there have been many initiatives and policy-level interventions over the past decade to reduce *C difficile* [2].

Data from 2009 to 2017 across 10 states suggested that *C difficile* rates have been decreasing, and they vary by geographic location [3, 4]. However, recent nationally representative data are not available nor are *C difficile* rates by US census division or hospital-level characteristics. Understanding geographical variation in prevalence, as well as variation by hospital characteristics, can help identify the extent of the problem in different settings, potentially informing existing interventions.

This study assesses the prevalence of *C difficile* among hospital discharges in the United States using the 2016–2018 National Inpatient Sample (NIS), the largest publicly available inpatient healthcare dataset for the United States, and evaluates prevalence and trends by hospital-level characteristics such as geographic location.

### METHODS

#### Data Source

The NIS is a database developed by the Agency for Healthcare Research and Quality (AHRQ) and for the Healthcare Cost and Utilization Project (HCUP), and it uses a stratified probability sample of approximately 20% of all hospital discharges in participating hospitals, which cover 97% of the US population. Hospitals are stratified by teaching status, urban/rural location, bed size, ownership/control, and the 9 US census divisions. The years 2016–2018 were included in this study, because those were the most recent years available where the entire year uses *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM) codes. Each observation represents a hospital discharge; therefore, a single person may represent more than 1 observation of the analysis if they were discharged from a participating hospital more than once in a calendar year [5].

Occurrence of *C difficile* was identified using up to the first 30 ICD-10-CM diagnosis codes provided for each observation. Both the codes for “Enterocolitis due to *Clostridium difficile*, recurrent” (A4.71) and “Enterocolitis due to *Clostridium difficile*, not specified as recurrent” (A4.72) were used to identify *C difficile* infections. States partition into census divisions are listed in Supplemental Table 1. The NIS does not include discharges from Alabama or Idaho. In addition, New Hampshire did not submit their data in time to be included in the databases [5].

#### Statistical Analysis

Data were analyzed using Stata/MP, version 15 (Statacorp, College Station, TX), using “svy” commands with the weights as provided by HCUP. The overall crude and adjusted *C difficile* prevalence in 2016–2018 were examined, as well as the crude and adjusted change in prevalence between 2016 and 2018. When looking at change over time, both the relative chance (prevalence difference) and the relative change (prevalence ratio) were calculated. In addition, subgroup analyses were performed evaluating hospital teaching status/urban-rural location, bed size, ownership/control, and the 9 US census divisions. Adjustment was done for age, sex, and race to account for different distribution of these variables by hospital-level

Received 25 June 2021; editorial decision 23 July 2021; accepted 27 July 2021.

Correspondence: Aaron A. R. Tobian, MD, PhD, Department of Pathology, School of Medicine, Johns Hopkins University, 600 N. Wolfe St., Carnegie Room 437, Baltimore, MD, USA 21287 (atobian1@jhmi.edu).

#### Open Forum Infectious Diseases® 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com <https://doi.org/10.1093/ofid/ofab409>

characteristic. Change in prevalence was calculated using average marginal effects, which is the average change in predicted probability of *C difficile* for each value of the hospital-level characteristic over time, treating all observations as if they had the particular values of hospital and year of interest but keeping their observed value of age, sex, and race. This was done by first performing a logistic regression with terms for year, the hospital-level characteristic of interest, the interaction between year, and the hospital-level characteristic. The probability of *C difficile* was then calculated for any given year/census division combination.

### Patient Consent Statement

The NIS is a deidentified, publicly available data set; this study was deemed exempt from review from the Johns Hopkins Institutional Review Board. This analysis was conducted in accordance with the HCUP data use agreement guidelines.

## RESULTS

The distribution of characteristics by year are listed in [Supplemental Table 2](#). Race had the highest levels of missingness

of approximately 5%. The overall crude prevalence of *C difficile* decreased from 10.1 per 1000 hospital discharges (95% confidence interval [CI], 9.9–10.3) to 9.3 per 1000 hospital discharges (95% CI, 9.1–9.4) in 2017 and 8.6 per 1000 hospital discharges (95% CI, 8.5–8.8) in 2018, leading to an adjusted prevalence difference (aPD) of –1.6 (95% CI, –1.9 to –1.4) per 1000 over the study period.

Prevalence by US census division in 2016 and 2018, as well as both the relative and absolute change from 2016 to 2018 by hospital-level characteristics, are shown in [Table 1](#). In 2018, the prevalence was highest in the East North Central region with 9.9 cases per 1000 discharges (95% CI, 9.5–10.4) and lowest in the West South Central with 7.4 cases per 1000 discharges (95% CI, 6.9–7.8). The largest changes in *C difficile*-adjusted prevalence were seen in the Mountain region (aPD = –3.1 [95% CI, –4.2 to –1.9] per 1000 discharges) and Pacific region (aPD = –3.0 [95% CI, –3.6 to –2.3] per 1000 discharges). Despite having the largest decrease in prevalence, the Mountain region had the second highest prevalence of all census divisions in 2018.

Rural hospitals had a smaller change in *C difficile* prevalence compared with both urban teaching and urban nonteaching

**Table 1. Prevalence of *Clostridioides difficile* per 1000 Hospital Discharges in 2016 and 2018 by Hospital-Level Characteristics**

Characteristic	Prevalence per 1000 Hospital Discharges		2018 vs 2016		2018 vs 2016	
	2016 (95% CI)	2018 (95% CI)	PD (95% CI)	aPD <sup>a</sup> (95% CI)	PR (95% CI)	aPR <sup>a</sup> (95% CI)
Overall	10.1 (9.9–10.3)	8.6 (8.5–8.8)	<b>–1.5 (–1.7 to –1.2)</b>	<b>–1.6 (–1.9 to –1.4)</b>	<b>0.85 (0.83–0.88)</b>	<b>0.84 (0.82–0.86)</b>
Census Division						
New England	11.5 (10.6–12.4)	9.7 (8.9–10.4)	<b>–1.9 (–3.0 to –0.8)</b>	<b>–1.9 (–2.9 to –0.9)</b>	<b>0.84 (0.75–0.93)</b>	<b>0.82 (0.74–0.91)</b>
Middle Atlantic	9.8 (9.3–10.3)	8.5 (8.0–9.0)	<b>–1.3 (–2.0 to –0.5)</b>	<b>–1.4 (–2.0 to –0.7)</b>	<b>0.87 (0.80–0.94)</b>	<b>0.86 (0.80–0.92)</b>
East North Central	10.8 (10.3–11.2)	9.9 (9.5–10.4)	<b>–0.9 (–1.5 to –0.2)</b>	<b>–0.9 (–1.6 to –0.3)</b>	<b>0.92 (0.86–0.98)</b>	<b>0.91 (0.85–0.97)</b>
West North Central	10.1 (9.4–10.8)	9.4 (8.8–10.0)	–0.7 (–1.6 to 0.2)	–0.8 (–1.7 to 0.1)	0.93 (0.85–1.02)	0.92 (0.83–1.01)
South Atlantic	9.9 (9.5–10.3)	8.6 (8.2–8.9)	<b>–1.4 (–1.9 to –0.8)</b>	<b>–1.4 (–1.9 to –1.0)</b>	<b>0.86 (0.81–0.91)</b>	<b>0.85 (0.81–0.90)</b>
East South Central	10.2 (9.4–11.0)	9.0 (8.2–9.8)	<b>–1.2 (–2.4 to –0.1)</b>	<b>–1.2 (–2.2 to –0.3)</b>	<b>0.88 (0.77–0.98)</b>	<b>0.88 (0.78–0.97)</b>
West South Central	8.8 (8.2–9.3)	7.4 (6.9–7.8)	<b>–1.4 (–2.1 to –0.7)</b>	<b>–1.6 (–2.3 to –0.9)</b>	<b>0.84 (0.77–0.92)</b>	<b>0.83 (0.76–0.90)</b>
Mountain	11.3 (10.5–12.1)	8.9 (8.2–9.6)	<b>–2.4 (–3.5 to –1.3)</b>	<b>–3.1 (–4.2 to –1.9)</b>	<b>0.79 (0.70–0.87)</b>	<b>0.75 (0.67–0.83)</b>
Pacific	10.0 (9.6–10.5)	7.5 (7.1–7.9)	<b>–2.5 (–3.2 to –1.9)</b>	<b>–3.0 (–3.6 to –2.3)</b>	<b>0.75 (0.69–0.80)</b>	<b>0.73 (0.68–0.78)</b>
Beds Size						
Small	9.4 (9.0–9.7)	8.2 (7.9–8.5)	<b>–1.2 (–1.6 to –0.7)</b>	<b>–1.3 (–1.7 to –0.9)</b>	<b>0.88 (0.83–0.92)</b>	<b>0.86 (0.82–0.90)</b>
Medium	9.8 (9.5–10.1)	8.4 (8.2–8.7)	<b>–1.4 (–1.8 to –0.9)</b>	<b>–1.6 (–2.0 to –1.2)</b>	<b>0.86 (0.82–0.90)</b>	<b>0.84 (0.81–0.88)</b>
Large	10.5 (10.2–10.8)	8.9 (8.6–9.2)	<b>–1.6 (–2.0 to –1.2)</b>	<b>–1.7 (–2.1 to –1.3)</b>	<b>0.85 (0.81–0.88)</b>	<b>0.83 (0.81–0.86)</b>
Location/Teaching Status						
Rural	9.0 (8.6–9.3)	8.7 (8.4–9.1)	–0.2 (–0.8 to 0.3)	–0.3 (–0.8 to 0.1)	0.97 (0.92–1.03)	0.96 (0.91–1.02)
Urban Nonteaching	10.3 (10.0–10.6)	8.5 (8.2–8.8)	<b>–1.8 (–2.3 to –1.3)</b>	<b>–1.9 (–2.3 to –1.5)</b>	<b>0.83 (0.78–0.87)</b>	<b>0.81 (0.77–0.85)</b>
Urban Teaching	10.2 (9.9–10.4)	8.7 (8.4–8.9)	<b>–1.5 (–1.9 to –1.2)</b>	<b>–1.8 (–2.1 to –1.5)</b>	<b>0.85 (0.82–0.88)</b>	<b>0.81 (0.81–0.86)</b>
Hospital Control						
Government	9.6 (9.0–10.1)	8.7 (8.1–9.2)	<b>–0.9 (–1.8 to –0.1)</b>	<b>–1.2 (–2.0 to –0.4)</b>	<b>0.90 (0.82–0.99)</b>	<b>0.88 (0.81–0.96)</b>
Private, Nonprofit	10.5 (10.3–10.7)	9.0 (8.8–9.2)	<b>–1.5 (–1.8 to –1.2)</b>	<b>–1.6 (–1.9 to –1.4)</b>	<b>0.86 (0.83–0.88)</b>	<b>0.85 (0.82–0.87)</b>
Private, for-Profit	8.6 (8.2–9.0)	6.8 (6.5–7.2)	<b>–1.8 (–2.3 to –1.2)</b>	<b>–1.8 (–2.3 to –1.3)</b>	<b>0.80 (0.74–0.85)</b>	<b>0.79 (0.74–0.84)</b>

Abbreviations: aPD, adjusted prevalence difference; aPR, adjusted prevalence ratio; CI, confidence interval; PD, prevalence difference; PR, prevalence ratio.

NOTE: All data are weighted using survey weights provided by the Healthcare Cost and Utilization Project (HCUP). Bold typeface indicates statistical significance. *Clostridioides difficile* defined using *International Classification of Diseases, Tenth Revision, Clinical Modification* Code A4.71 (Enterocolitis due to *Clostridium difficile*, recurrent) or A4.72 (Enterocolitis due to *Clostridium difficile*, not specified as recurrent) during hospitalization. Alabama (East South Central) and Idaho (Mountain) do not participate in the National Inpatient Sample.

<sup>a</sup>Estimates were adjusted and standardized for age, sex, and race.

hospitals (rural: aPD = -0.3 [95% CI, -0.8 to 0.1] vs urban teaching: aPD = -1.8 [95% CI -2.1 to -1.5]); however, rural hospitals started off with a lower level in 2016, and the prevalence in 2018 was similar in urban and rural hospitals. Private for-profit hospitals had a lower prevalence of *C difficile* in 2018 compared with private, nonprofit, and government run hospitals.

## DISCUSSION

In this study, we found an overall decrease in *C difficile* prevalence among US hospital discharges from 2016 to 2018, although the extent of the decrease was variable by geographic location and hospital location/teaching status. One potential reason for the heterogeneity in the decrease of *C difficile* prevalence may be differential implementation of antimicrobial stewardship programs and infection control measures across states or hospital types. Furthermore, The Centers for Medicare and Medicaid Services Pay for Performance incentives to reduce *C difficile* rates may provide varying levels of motivation for hospitals depending on hospital-level factors such as ownership.

This study is consistent with other studies that have shown a decrease in *C difficile* in recent years, although research has suggested the decrease has been mostly in hospital-onset *C difficile*, whereas community-onset has remained the same [4, 6]. Data from NIS cannot distinguish between community-onset versus hospital-onset, so we were unable to determine whether declines observed were due to reductions in community- or hospital-onset infections. The insight from this study does provide more finite detail on the magnitude of changes over recent years, and it demonstrates the burden of *C difficile* in the hospital setting. More research on the effectiveness of interventions by geographic location and other hospital-level characteristics may give insight to the difference in trends found in this study, and it could provide information for how to better tailor interventions and allocate resources.

Although prevention of *C difficile* has generally focused on interventions targeting symptomatic patients in healthcare settings, such as isolation and environmental cleaning, increasing data suggest that symptomatic healthcare cases are not the only source for transmission [7]. Thus, interventions focused on symptomatic patients, although important and effective, will need augmentation from additional measures [8]. The variation in the proportion of cases that are community acquired, as well as differences in the likelihood that patients are symptomatic, suggest that interventions in the community and long-term care settings are both likely needed to aid transmission reduction [7, 9, 10]. In addition, improved understanding of the role of colonized people in transmission as well as how colonization contributes to individual risk of disease is important to decreasing *C difficile* [9].

Although data from this study are from a nationally representative sample and adjusted by age, sex, and race, there remain some limitations of the data. As with all NIS studies, data are based on ICD-10-CM billing codes and may not reflect the actual prevalence of infection. In addition, the study focused on prevalence because it is not possible to distinguish new versus recurrent *C difficile* infection. It has also been demonstrated that ICD-10 codes likely underestimate *C difficile* infection [11]. Method of diagnosis for each hospitalization is unknown. Different methods of diagnosis have varying level of sensitivity and specificity, as well as differing abilities to discriminate between *C difficile* colonization and *C difficile* infection [4, 12]. Nucleic acid amplification testing is a popular diagnostic test; however, standalone use has been demonstrated to overestimate *C difficile* infection rates by misclassifying those who are colonized as having infection [12]. The data only cover a short time period, making it difficult to determine whether there is truly a sustained decrease in prevalence over time in recent years.

## CONCLUSIONS

Overall, these promising data suggest that efforts to curb *C difficile* have been effective, and that despite the variability in prevalence and change by location, all geographic divisions in the United States saw a decrease over the study period. These updated prevalence estimates can help state health departments who may wish to prioritize *C difficile* reduction interventions.

## Supplementary Data

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

## Acknowledgments

We are grateful to the National Inpatient Sample study staff and participants, without whom this study would not have been possible.

**Disclaimer.** The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

**Financial support.** This work was funded in part by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases (NIAID) and extramural support from NIAID (R01AI120938 and R01AI128779 [to A. A. R. T.] and T32AI102623 [to E. U. P.]). E. Y. K. was supported by Centers for Disease Control and Prevention MInD-Healthcare Program (Grant Number U01CK000589).

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

## References

1. Leffler DA, Lamont JT. *Clostridium difficile* infection. *N Engl J Med* 2015; 372:1539-48.
2. Drozd EM, Inocencio TJ, Braithwaite S, et al. Mortality, hospital costs, payments, and readmissions associated with *Clostridium difficile* infection among medicare beneficiaries. *Infect Dis Clin Pract (Baltim Md)* 2015; 23:318-23.
3. Lessa FC, Mu Y, Bamberg WM, et al. Burden of *Clostridium difficile* infection in the United States. *N Engl J Med* 2015; 372:825-34.

4. Guh AY, Mu Y, Winston LG, et al.; Emerging Infections Program Clostridioides difficile Infection Working Group. Trends in U.S. burden of *Clostridioides difficile* infection and outcomes. *N Engl J Med* **2020**; 382:1320–30.
5. HCUP National Inpatient Sample (NIS). Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality, **2012**.
6. Solanki D, Kichloo A, and El-Amir Z, et al. *Clostridium difficile* infection hospitalizations in the United States: insights from the 2017 National Inpatient Sample. *Gastroenterology Res* **2021**; 14:87–95.
7. Eyre DW, Cule ML, Wilson DJ, et al. Diverse sources of *C. difficile* infection identified on whole-genome sequencing. *N Engl J Med* **2013**; 369:1195–205.
8. Garcia Reeves AB, Lewis JW, Trogon JG, et al. Association between statewide adoption of the CDC's core elements of hospital antimicrobial stewardship programs and rates of methicillin-resistant *Staphylococcus aureus* bacteremia and *Clostridioides difficile* infection in the United States. *Infect Control Hosp Epidemiol* **2020**; 41:430–7.
9. Crobach MJT, Vernon JJ, Loo VG, et al. Understanding *Clostridium difficile* colonization. *Clin Microbiol Rev* **2018**; 31:e00021-17.
10. Lessa FC, Mu Y, Winston LG, et al. Determinants of *Clostridium difficile* infection incidence across diverse United States geographic locations. *Open Forum Infect Dis* **2014**; 1:ofu048.
11. Jones G, Taright N, Boelle PY, et al. Accuracy of ICD-10 codes for surveillance of *Clostridium difficile* infections, France. *Emerg Infect Dis* **2012**; 18:979–81.
12. Boly FJ, Reske KA, Kwon JH. The role of diagnostic stewardship in *Clostridioides difficile* testing: challenges and opportunities. *Curr Infect Dis Rep* **2020**; 22:7.