Clostridium difficile infection: An emerging cause of death in the twenty-first century*

Viytta N. Abdullatif
Andrew Noymer†

Department of Population Health and Preventive Medicine
University of California
Irvine, CA 92697-3957

Abstract

Enterocolitis due to *Clostridium difficile* is major emerging cause of death in the US. Between 1999 and 2012, *C. diff.* deaths rose by a staggering almost-10-fold increase, to 7,739 from 793. This paper has three goals. First, we present a demographic description of *C. diff.* mortality in the US since 1999. Second, we test a hypothesis that the increase in *C. diff.* deaths is due to population aging. We find that the emergence of this cause of death follows a proportional hazard pattern, above age 40. Thus, population aging is not the only factor responsible for the increase in *C. diff.* deaths. This, combined with a contributory cause of death analysis, points towards healthcare-based strategies to combat *C. diff*. Third, we demonstrate a simple weighted least squares technique for estimating Gompertz models that gives parameter estimates that are closer to full maximum likelihood, compared to conventional approaches.

Clostridium difficile is a gram-positive spore-forming bacterium, and a clinically-significant enteric pathogen (Sunenshine and McDonald 2006). Clinical *C. diff.* infection is almost always associated with antibiotic use (Bartlett 2008*a*, Leffler and Lamont 2015). The Centers for Disease Control and Prevention (CDC) has identified *C. diff.* as one of three microoganisms at the highest

*Final author version. Published as *Biodemography and Social Biology* 62(2):198–207 (2016). http://www.tandfonline.com/doi/full/10.1080/19485565.2016.1172957

†To whom correspondence should be addressed: noymer@uci.edu

Table 1: Number of deaths: *C. diff.* infection (ICD10 A04.7), US, 1999–2012

	Number of deaths		
year	underlying	contributory	total
1999	793	752	1,545
2000	1,101	919	2,020
2001	1,332	988	2,320
2002	2,195	1,331	3,526
2003	2,776	3,601	6,377
2004	4,062	3,992	8,054
2005	5,332	3,204	8,536
2006	6,225	3,582	9,807
2007	6,372	3,507	9,879
2008	7,476	4,103	11,579
2009	7,251	4,089	11,340
2010	7,298	4,198	11,496
2011	8,085	4,616	12,701
2012	7,739	4,850	12,589

threat level, "urgent", along with carbapenem-resistant Enterobacteriaceae (CRE) and drug-resistant *Neisseria gonorrhoeae*, noting: "Although *C. difficile* is not currently significantly resistant to antibiotics used to treat it, it was included in the threat assessment because of its unique relationship with resistance issues, antibiotic use, and its high morbidity and mortality." (National Center for Emerging Zoonotic and Infectious Diseases 2013).

Table 1 shows the stark rise in deaths due to *Clostridium difficile* enterocolitis (ICD-10 A04.7) over a 14-year period. The United States began using ICD-10 for death codes in 1999, before which it is difficult to track *C. diff.* mortality. The reason is that in the ICD-9 coding scheme, *C. diff.* deaths were coded as 008.45 (5-digit code), whereas the mortality detail files report only 4-digit ICD-9 codes. For example, in 1998 there were 600 deaths in the US with cause of death ICD-9 008.4, "intestinal infections due to other spec-

ified bacteria", but we have no way of knowing how many were due to *C. diff.* specifically. Note, however, 600 deaths for 1998 is in-line with the level and trend seen in table 1. The ICD-9 coding scheme originated in 1975, but *C. diff.* was fully recognized as a human pathogen two years later (Bartlett et al. 1977), so it is not surprising that it was relegated to a 5-digit ICD code. Only with promotion to 4-digit ICD code status in ICD-10 can demographers specifically track levels and trends of *C. diff.* mortality.

Wysowski (2006) noted the increase in mortality due to *C. diff.*, 1999–2002. This work was expanded upon by Redelings et al. (2007), who described the increase in *C. diff.* mortality, 1999–2004, and included contributory as well as underlying deaths. Hall et al. (2012) updated the picture using data through 2007, and contrasted the relative lack of seasonality of *C. diff.* deaths compared to norovirus-associated mortality, which peaks in the winter. Using data up to 2009, Peery et al. (2012) noted "The toll of *C. difficile* infection is large and becoming more prominent". Writing in the official National Center for Health Statistics (NCHS) summary of mortality data, Murphy et al. (2012), similarly noted, "Enterocolitis due to *Clostridium difficile* ... has become a growing concern in recent years"; it is unusual for the NCHS to call attention to a specific cause of death in this way.

A goal of this study is to present a descriptive analysis of levels and changes in *C. diff.* mortality in the United States during the ICD-10 era for mortality coding (1999–2012). Because *C. diff.* death rates rise with age, the aging of the US population (Ortman et al. 2014) could potentially account for an increase in the number of *C. diff.* deaths even if age- and cause-specific

rates do not change. We know, *a priori*, that a nearly 10-fold increase in deaths in 14 years (cf. table 1) must have a causal component beyond population aging, but nonetheless it is desirable to describe the age-mortality profile for *C. diff.*, and whether it has changed over time.

We present underlying-cause and any-mention mortality rates for *C. diff.* infection, by age (5-year groups) and sex, calculated from data on every death in the US. We use cause-specific Gompertz models to test a proportional hazard hypothesis, and we introduce therein a weighted least squares technique for estimating Gompertz survival curves.

Clostridium difficile mortality

We used data from the multiple cause of death files of the National Center for Health Statistics (National Center for Health Statistics 2014), a database of all mortality in the United States. We extracted data on all deaths containing *C. diff.* infection (ICD10 A04.7) as an underlying or contributory cause, during the period 1999–2012. The start of the time span corresponds to the beginning of ICD-10 death coding in the United States, and the end of the span reflects the most recent available data. Using population data from the Human Mortality Database (2014), we calculated age-specific and age-adjusted death rates for *C. diff.* infection, separately by sex. The year-2000 United States standard population was used for the calculation of the age-adjusted rates. We also calculated all-mention death rates, i.e., where the numerators were all deaths involving *C. diff.* infection as an underlying or contributory cause (Wing and Manton 1981).

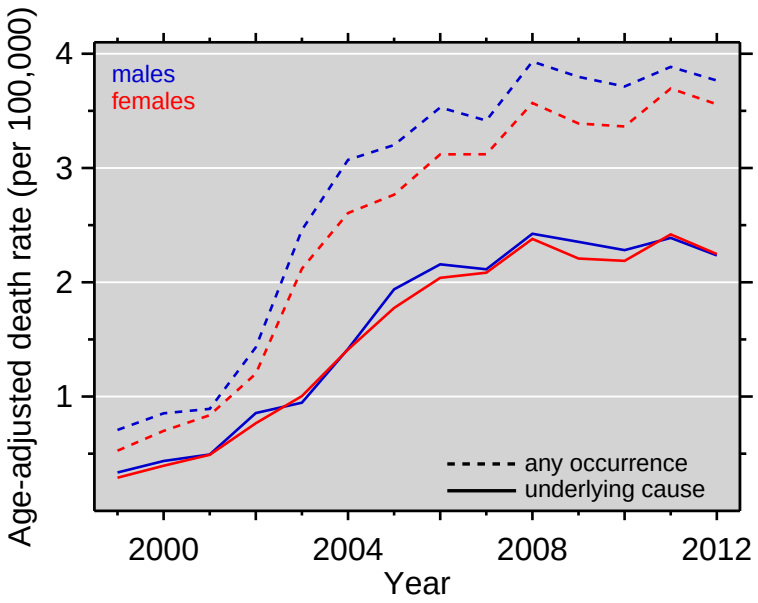

Figure 1: Age-adjusted death rate, *C. diff.* infection, 1999–2012.

Figure 1 shows the age-adjusted death rate for *C. diff.* infection, 1999–2012. This graph shows three important features of *C. diff.* mortality in the US. First, *C. diff.* age-adjusted death rates have risen about 7-fold since 1999, from less than 0.5 per 100,000 to just under 2.5 per 100,000. As table 1 shows, this corresponds to a 10-fold increase in the absolute number of deaths in which *C. diff.* is the underlying cause, and an 8-fold increase in all deaths involving *C. diff.*, in just a fourteen-year time span. The increase over time has been uneven, with a plateau from 2008–12. Second, there is no meaningful sex difference in the age-adjusted death rates for the underlying cause of death. Third, the all-mention age-adjusted death rates follow the underlying age-adjusted death rates over time in a roughly parallel fashion, but since 2003, a sex difference has emerged. Specifically, in

figure 1, males have higher all-mention *C. diff.* death rates than females after 2003; this is not seen meaningfully before 2003, nor in any year with the underlying-cause death rates. The reason(s) for the sex difference in all-mention rates after 2003 are not clear, but virtually all *C. diff.* mortality is healthcare-associated. Also, figure 1 shows that increases in *C. diff.* mortality since 1999 are not caused by a shift to underlying from contributory cause of death status.

Figure 2 illustrates the age-mortality profile of *C. diff.* infection (underlying cause) for two seven-year sub-periods, 1999–2005, and 2006–2012. The "lazy-J" pattern seen in the age-mortality profile is typical of many causes of death. What is noteworthy is that above age 40, the death rates for 2006–2012 parallel the 1999–2005 rates. In other words, the increase in *C. diff.* death rates over time, seen in figure 1, cannot be attributed to changes in a specific age group. In fact, above age 40, it is truly striking just how uniform by age is the increase in the *C. diff.* mortality rate; a proportional hazard phenomenon, in other words.

To test statistically the above assertion (i.e., that above age 40, the later period is a proportional hazard of the earlier period), we used Gompertz mortality models. We estimated the following model:

$$\log\big(M(x)\big) = \alpha + \beta x + \epsilon$$

where $M(\cdot)$ is the *C. diff.* cause-specific death rate, x is age, α and β are parameters to be estimated, and ϵ is an error term with presumed zero mean. Models were estimated separately by sex to sharpen the estimates, since

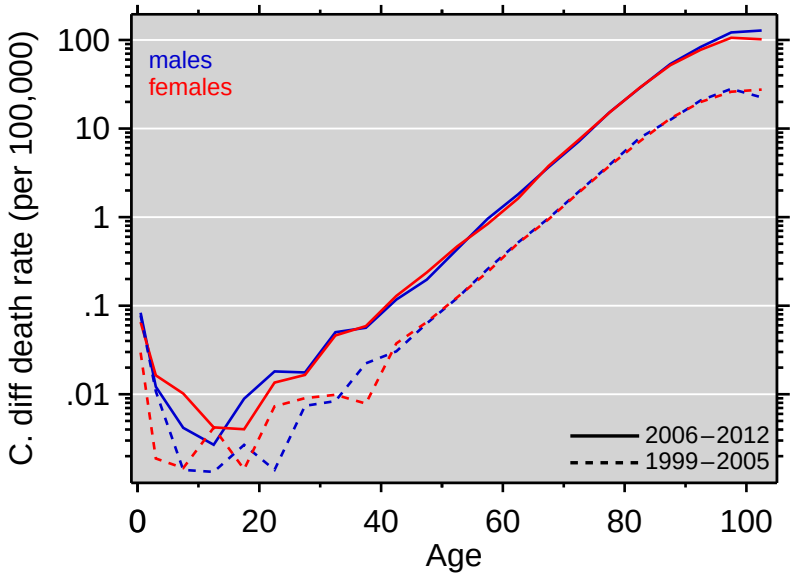

Figure 2: Age-mortality profile, *C. diff.* infection (underlying cause), 1999–2012.

male and female hazards often behave differently (Kohler and Kohler 2000), although in this case there is no qualitative distinction between the shapes or levels of the graph by sex, above age 40 (cf. figure 2). Five-year age groups, $40 \leq x \leq 99$ (i.e., 40–44, 45–49, ...), were used to smooth age heaping. In each age group, the mean age of *C. diff.* mortality in the interval was used as the x value in the regressions. For example, for females age 65–69 in the more recent time period, the mean age of death due to *C. diff.* was 67.153, so this value was used as x instead of the midpoint of the interval, 67.5. This "centering" of the x values is used to polish the estimates, although it has small effects. Such centering, combined with parametric estimation of mortality, can help avoid some of the methodological pitfalls outlined by Gelman and Auerbach (2016).

Table 2: Proportional hazard analysis of *C. diff.* mortality, US, 1999–2005 vs. 2006–2012.

	WLS					
	FEMALES			**MALES**		
	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ
age term (β)	0.124***	0.126***		0.129***	0.132***	
	(30.7)	(26.8)		(38.2)	(43.2)	
intercept (α)	−19.85***	−18.61***		−20.16***	−19.08***	
	(−59.8)	(−48.0)		(−75.6)	(−78.0)	
period term (γ)			1.240			1.080
			(1.80)			(2.34)
age×period (δ)			0.00157			0.00326
			(0.19)			(0.56)
N	12	12	24	12	12	24
R^2	0.9895	0.9863	0.9890	0.9932	0.9947	0.9952
residual d.f.	10	10	20	10	10	20
F	944.9	718.6	599.1	1459	1870	1392
RMSE	0.139	0.161	0.157	0.121	0.110	0.112

t statistics in parentheses

*** p<0.0001, ** p<0.001, * p<0.01

Table 2 gives the weighted least squares (WLS) regression results. The weights are the numbers of *C. diff.* deaths in each age×sex×period cell; this is numerator-weighted rate regression, in other words. The use of death counts as weights has two appealing properties. First, it effectively down-weights data at the extremes of the age range, where deaths are fewer. This has the desirable consequence of reducing the importance of subjective choices of age range (for instance, whether to use data beginning at age 40 or at age 45). Second, it moves the parameter estimates closer to those that would be achieved by maximum likelihood (ML) estimation — which can regarded as theoretically desirable (Brillinger 1986) — but without the computational expense of ML. To the best of our knowledge, this approach (using WLS expressly to approximate ML estimates) has not been

used before. Therefore, we provide, in the Appendix, a demonstration of our assertion that the WLS estimates are indeed better than — and not just different from — the OLS estimates (with the proviso that better in this context means closer to ML). Note also that, judging by the high R^2 values in table 2, the Gompertz specification seems to be very appropriate for *C. diff.* cause-specific mortality.

The test of the proportional hazard comes from an interactive model. Rather than run each period separately, we estimated the model using data from both periods, pooled:

$$\log\big(M(x)\big) = \alpha + \beta x + \gamma p + \delta(p \times x) + \varepsilon$$

where p is an indicator variable for the later period, γ and δ are additional parameters to be estimated, and ε is an error term like before. Thus, α is the intercept, β is the age term, γ is the period term, and δ is the age$\times$period intercept term. This model recapitulates the coefficients from the earlier models, but permits a statistical test for differences of slopes (i.e., $\delta \neq 0$). In model 3 of table 2, the α and β values are not shown because they coincide exactly with those of model 1. All analysis was conducted using Stata 13.1.

The proportional hazard assumption is upheld; we fail to reject the null hypothesis that $\delta = 0$ for both sexes. The differences in slopes across the two time periods are not only negligible in magnitude, but are not distinguishable statistically. As figure 2 demonstrates, and table 2 quantifies, changes in *C. diff.* mortality between 1999–2005 and 2006–2012 are about

the level but not about the shape (slope) of the Gompertzian pattern above age 40.

To complete our empirical analysis, table 3 summarizes the underlying causes of death when *C. diff.* infection was a contributory cause on the death certificate; the importance of multiple-cause analysis is noted by Désesquelles et al. (2014). Table 3 shows very clearly that *C. diff.* is involved in a wide variety of causes of death, especially those that occur in clinical settings. There is no single "typical" underlying cause of death when *C. diff.* is a contributory cause. The top 18 detailed (i.e., 4-digit) causes are listed in table 3, but these encompass only about half the deaths involving *C. diff.*, with the other half spread among 1,295 unique ICD codes. The clinical literature on *C. diff.*-associated morbidity clearly notes the association with healthcare (e.g. Sunenshine and McDonald 2006). Table 3 is fully consistent with this. For example, diseases of the circulatory system (ICD-10 Ixx.x) account for over 40% of deaths in which *C. diff.* is a contributory cause. While some mortality in this group may not be preceded by a hospital stay (for example, sudden death due to stroke or heart attack), admission to the emergency department and/or intensive care unit is common with these diseases (see ICD-10 "I" causes in table 3). Also, nearly 30% of deaths in which *C. diff.* is a contributory cause involve neoplasms (cancer, ICD-10 Cxx.x). Cancer is clearly a disease that involves interaction with healthcare prior to mortality. Moreover, immunosuppression is a risk factor for cancer (Vial and Descotes 2003) as well as infection. In short, table 3 shows that when *C. diff.* infec-

Table 3: Underlying cause of death when *C. diff.* infection (ICD10 A04.7) is a contributory cause of death, US, 1999–2012

Underlying Cause of Death	ICD10 code	Percent
Atherosclerotic heart disease	I25.1	7.77
Chronic obstructive pulmonary disease	J44.9	5.63
Septicaemia	A41.9	5.00
Acute myocardial infarction	I21.9	3.95
Lung cancer	C34.9	3.60
Dementia (unspecified)	F03	3.23
Urinary tract infection	N39.0	2.83
Stroke	I64	2.72
Alzheimer's disease	G30.9	2.37
Aspiration pneumonia	J69.0	2.27
Congestive heart failure	I50.0	2.03
Colon cancer	C18.9	1.40
Diabetes mellitus (without complications)	E14.9	1.26
Non-Hodgkin's lymphoma	C85.9	1.18
Angina pectoris	I25.0	1.06
COPD with ALRI	J44.0	1.03
Atrial fibrillation and flutter	I48	1.00
Colitis/enteritis/enterocolitis	K55.9	1.00
Other (1,295 unique ICD codes)		50.68
Of which:		
Certain infectious and parasitic diseases	A	3.2
Certain infectious and parasitic diseases	B	3.6
Neoplasms	C	23.1
Diseases of the blood and blood-forming organs and etc.	D	2.4
Endocrine, nutritional and metabolic diseases	E	5.9
Mental and behavioural disorders	F	1.5
Diseases of the nervous system	G	4.3
Diseases of the eye and adnexa, and the ear and mastoid process	H	<0.05
Diseases of the circulatory system	I	21.7
Diseases of the respiratory system	J	7.8
Diseases of the digestive system	K	10.4
Diseases of the skin and subcutaneous tissue	L	0.8
Diseases of the musculoskeletal system and connective tissue	M	2.3
Diseases of the genitourinary system	N	7.7
Pregnancy, childbirth and the puerperium	O	<0.05
Certain conditions originating in the perinatal period	P	0.2
Congenital malformations, deformations and chromosomal abnormalities	Q	0.4
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	R	0.3
External causes	V	0.9
	W	1.3
	X	1.6
	Y	0.7

tion a contributory cause of death, the underlying cause is likely to be any condition involving health care, especially circulatory conditions or cancer.

Conclusions

The increase in *C. diff.* mortality rates since 1999 in the US has occurred at all ages > 40. The pronounced proportional hazard pattern (figure 2) means that increases in the size of the elderly population are not the only driving factor behind the increase in *C. diff.* deaths seen in table 1. The emergence of virulent strains in the 2000s, such as NAP1/BI/02 (Blossom and McDonald 2007, Pepin 2008), is a better explanation for the mortality patterns documented herein. Moreover the shift, after about 2006, to vancomycin (from metronidazole) for treatment of *C. diff.* infection (Zar et al. 2007), can account for the plateau seen in figure 1. Vancomycin use has its own potential pitfalls, such as the emergence of vancomycin resistant enterococci (Cattoir and Leclercq 2013), but it seems to account for the stabilization of *C. diff.* mortality since 2008.

In summary, in addition to causing significant morbidity (Lessa et al. 2015), *Clostridium difficile* is an emerging cause of death in the United States. *C. diff.* mortality rates have increased profoundly in the last 14 years, evenly for all ages above 40, and for both sexes. Despite the overall increase, *C. diff.* mortality rates have plateaued in the most recent years. The emergence of *C. diff.* mortality is associated with new strains, and the plateau with the switch to vancomycin therapy (Bartlett 2008*b*). Promising experimental therapies include strain-specific probiotics (Gerding et al. 2015), as well as more di-

verse gut microbiome-related approaches (e.g., Rupnik 2015, Buffie et al. 2015).

One limitation of this study is that *C. diff.* mortality is probably the "tip of the iceberg" of infection, on which we do not have data. Moreover, the elderly, who are hospitalized disproportionately and for longer stays, may play a catalyzing role in the epidemiology of *C. diff.* in ways that are not captured by the mortality rates, which we have asserted are not rising in an age-specific way.

Especially because of the broad age range of increased mortality and the nonspecific nature of observed co-morbidities, public health agencies should closely monitor *C. diff.* mortality. The most important substantive lesson from this study is that the emergence of *C. diff.* as a cause of death in the United States is not due to population aging, but is healthcare-associated. We also demonstrate weighted least squares (WLS) regression as an approximation to full maximum likelihood for Gompertz parameter estimation, and we strongly recommend WLS over OLS whenever weights are available.

Works Cited

Bartlett, John G. 2008*a*. "Historical perspectives on studies of *Clostridium difficile* and *C. difficile* infection." *Clinical Infectious Diseases* 46(Suppl 1):S4–S11.

———. 2008*b*. "The case for vancomycin as the preferred drug for treatment of *Clostridium difficile* infection." *Clinical Infectious Diseases* 46(10):1489–1492.

Bartlett, John G., Andrew B. Onderdonk, Ronald L. Cisneros, and Dennis L. Kasper. 1977. "Clindamycin-associated colitis due to a toxin-producing

species of *Clostridium* in hamsters." *Journal of Infectious Diseases* 136(5):701–705.

Blossom, David B. and L. Clifford McDonald. 2007. "The challenges posed by reemerging *Clostridium difficile* infection." *Clinical Infectious Diseases* 45(2):222–227.

Brillinger, David R. 1986. "The natural variability of vital rates and associated statistics. With discussion." *Biometrics* 42(4):693–734.

Buffie, Charlie G., Vanni Bucci, Richard R. Stein, Peter T. McKenney, Lilan Ling, Asia Gobourne, Daniel No, Hui Liu, Melissa Kinnebrew, Agnes Viale, Eric Littmann, Marcel R. M. van den Brink, Robert R. Jenq, Ying Taur, Chris Sander, Justin R. Cross, Nora C. Toussaint, Joao B. Xavier, and Eric G. Pamer. 2015. "Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*." *Nature* 517(7533):205–208.

Cattoir, Vincent and Roland Leclercq. 2013. "Twenty-five years of shared life with vancomycin-resistant enterococci: Is it time to divorce?" *Journal of Antimicrobial Chemotherapy* 68(4):731–742.

Désesquelles, Aline, Elena Demuru, Viviana Egidi, Luisa Frova, France Meslé, Marilena Pappagallo, and Michele Antonio Salvatore. 2014. "Cause-specific mortality analysis: Is the underlying cause of death sufficient?" *Quetelet Journal* 2(1):119–135.

Gelman, Andrew and Jonathan Auerbach. 2016. "Age-aggregation bias in mortality trends." *Proceedings of the National Academy of Sciences of the United States of America* 113(7):E816–E817.

Gerding, Dale N., Thomas Meyer, Christine Lee, Stuart H. Cohen, Uma K. Murthy, Andre Poirier, Trevor C. Van Schooneveld, Darrell S. Pardi, Antonio Ramos, Michelle A. Barron, Hongzi Chen, and Stephen Villano. 2015. "Administration of spores of nontoxigenic *Clostridium difficile* strain M3 for prevention of recurrent *C difficile* infection: A randomized clinical trial." *Journal of the American Medical Association* 313(17):1719–1727.

Hall, Aron J., Aaron T. Curns, L. Clifford McDonald, Umesh D. Parashar, and Ben A. Lopman. 2012. "The roles of *Clostridium difficile* and norovirus among gastroenteritis-associated deaths in the United States, 1999–2007." *Clinical Infectious Diseases* 55(2):216–223.

Horiuchi, Shiro and Ansley J. Coale. 1982. "A simple equation for estimating the expectation of life at old ages." *Population Studies* 36(2):317–326.

Human Mortality Database. 2014. http://www.mortality.org/. Accessed 9 July.

Kohler, Hans-Peter and Iliana Kohler. 2000. "Frailty modelling for adult and old age mortality: The application of a modified DeMoivre hazard function to sex differentials in mortality." *Demographic Research* 3(8).

Leffler, Daniel A. and J. Thomas Lamont. 2015. "*Clostridium difficile* infection." *New England Journal of Medicine* 372(16):1539–1548.

Lessa, Fernanda C., Yi Mu, Wendy M. Bamberg, Zintars G. Beldavs, Ghinwa K. Dumyati, John R. Dunn, Monica M. Farley, Stacy M. Holzbauer, James I. Meek, Erin C. Phipps, Lucy E. Wilson, Lisa G. Winston, Jessica A. Cohen, Brandi M. Limbago, Scott K. Fridkin, Dale N. Gerding, and L. Clifford McDonald. 2015. "Burden of *Clostridium difficile* infection in the United States." *New England Journal of Medicine* 372(9):825–834.

Murphy, Sherry L., Jiaquan Xu, and Kenneth D. Kochanek. 2012. "Deaths: Preliminary Data for 2010." *National Vital Statistics Reports* 60(4). National Center for Health Statistics.

National Center for Emerging Zoonotic and Infectious Diseases. 2013. *Antibiotic resistance threats in the United States, 2013*. Centers for Disease Control and Prevention, Atlanta.

National Center for Health Statistics. 2014. *Mortality multiple cause-of-death data files,* http://www.cdc.gov/nchs/nvss/mortality_public_use_data.htm. National Center for Health Statistics. Accessed 14 April 2014.

Ortman, Jennifer M., Victoria A. Velkoff, and Howard Hogan. 2014. "An aging nation: The older population in the United States. Population estimates and projections." *Current Population Reports* P25-1140. U.S. Census Bureau.

Peery, Anne F., Evan S. Dellon, Jennifer Lund, Seth D. Crockett, Christopher E. McGowan, William J. Bulsiewicz, Lisa M. Gangarosa, Michelle T. Thiny, Karyn Stizenberg, Douglas R. Morgan, Yehuda Ringel, Hannah P. Kim, Marco Dacosta DiBonaventura, Charlotte F. Carroll, Jeffery K. Allen, Suzanne F. Cook, Robert S. Sandler, Michael D. Kappelman, and Nicholas J. Shaheen. 2012. "Burden of gastrointestinal disease in the United States: 2012 update." *Gastroenterology* 143(5):1179–1187.e3.

Pepin, Jacques. 2008. "Vancomycin for the treatment of *Clostridium difficile* infection: For whom is this expensive bullet really magic?" *Clinical Infectious Diseases* 46(10):1493–1498.

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and modeling population processes*. Blackwell, Oxford.

Redelings, Matthew D., Frank Sorvillo, and Laurene Mascola. 2007. "Increase in *Clostridium difficile*-related mortality rates, United States, 1999–2004." *Emerging Infectious Diseases* 13(9):1417–1419.

Rupnik, Maja. 2015. "Toward a true bacteriotherapy for *Clostridium difficile* infection." *New England Journal of Medicine* 372(16):1566–1568.

Sunenshine, Rebecca H. and L. Clifford McDonald. 2006. "*Clostridium difficile*-associated disease: New challenges from an established pathogen." *Cleveland Clinic Journal of Medicine* 73(2):187–197.

Vial, Thierry and Jacques Descotes. 2003. "Immunosuppressive drugs and cancer." *Toxicology* 185(3):229–240.

Wachter, Kenneth W. 2014. *Essential demographic methods*. Harvard University Press, Cambridge.

Wing, Steve and Kenneth G. Manton. 1981. "A multiple cause of death analysis of hypertension-related mortality in North Carolina, 1968–1977." *American Journal of Public Health* 71(8):823–830.

Wysowski, Diane K. 2006. "Increase in deaths related to enterocolitis due to *Clostridium difficile* in the United States, 1999–2002." *Public Health Reports* 121(4):361–362.

Zar, Fred A., Srinivasa R. Bakkanagari, K. M. L. S. T. Moorthi, and Melinda B. Davis. 2007. "A comparison of vancomycin and metronidazole for the treatment of *Clostridium difficile*-associated diarrhea, stratified by disease severity." *Clinical Infectious Diseases* 45(3):302–307.

Appendix

Here we present a regression tables of the same models as in table 2, except estimated using two alternate ways from the WLS presented therein. The first is maximum likelihood (ML). Specifically, we estimate the following Poisson regression models:

$$\log(\text{deaths}) = \alpha + \beta x + \log(\text{exposure})$$
$$\log(\text{deaths}) = \alpha + \beta x + \gamma p + \delta(p \times x) + \log(\text{exposure}),$$

as an alternative to the WLS models of logged mortality rates, as in the main body of the paper; these are presented in table A-1. In table A-2, we estimate Gompertz estimates by OLS in the typical way (see, e.g., Wachter 2014, p.69; Preston et al. 2001, p.193; Horiuchi and Coale 1982). These are the exact same models as in the main text of the paper, except OLS is used instead of the weighted approach.

For both males and females, the six coefficients in the ML estimation in table A-1 (α, β for each time period, and the interactive coefficients γ, δ) are closer to the WLS (table 2) than to the OLS coefficients of table A-2, as was asserted. The ML estimation treats each death as an observation, whereas the WLS treats each age-cell as a single data point (see "N cells", "N deaths" in table A-1). This necessarily results in much higher z-statistics in table A-1, compared to the t-statistics in table 2 or table A-2; the usual cautions apply, about null hypothesis significance testing with large sample sizes. The likelihood-based estimation of coefficients, using Poisson likelihood (see Brillinger 1986), where each death influences the estimates, is approximated by using each death as part of a weighting. The WLS approach is

much more efficient, computationally, so has the appeal of being easier, and closer to the ML estimates. Given the power of modern computers, the importance of this efficiency is perhaps debatable, but none of the substantive conclusions are changed. Based on this experience, our recommendation is to use ML or WLS, but never OLS (except in cases where the weights are unavailable).

Table A-1: Maximum likelihood version of table 2

	Maximum Likelihood					
	Females			**Males**		
	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ
age term (β)	0.124***	0.126***		0.129***	0.132***	
	(116)	(250)		(95.0)	(211)	
intercept (α)	−19.83***	−18.61***		−20.14***	−19.07***	
	(−225)	(−448)		(−188)	(−381)	
period term (γ)			1.222***			1.074***
			(12.5)			(9.08)
age×period (δ)			0.00176			0.00336
			(1.49)			(2.25)
N cells	12	12	24	12	12	24
N deaths	7,572	33,788	41,360	4,527	21,371	25,898
χ^2	19695	90687	125499	12158	59790	81752
df_m	1	1	3	1	1	3
LL	−111.4	−444.5	−555.9	−72.38	−166.9	−239.3

z statistics in parentheses

*** p<0.0001, ** p<0.001, * p<0.01

Table A-2: OLS version of table 2

	OLS					
	FEMALES			MALES		
	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ	(1) Early (1999–2005)	(2) Late (2006–12)	(3) Δ
age term (β)	0.127***	0.130***		0.130***	0.134***	
	(48.6)	(44.7)		(47.6)	(52.0)	
intercept (α)	−20.17***	−19.05***		−20.31***	−19.27***	
	(−107)	(−91.1)		(−104)	(−105)	
period term (γ)			1.125**			1.039**
			(4.01)			(3.87)
age×period (δ)			0.00294			0.00398
			(0.75)			(1.06)
N	12	12	24	12	12	24
R^2	0.9958	0.9950	0.9958	0.9956	0.9963	0.9963
residual d.f.	10	10	20	10	10	20
F	2364	1997	1572	2267	2701	1791
RMSE	0.155	0.172	0.164	0.161	0.152	0.157

t statistics in parentheses

*** p<0.0001, ** p<0.001, * p<0.01