Defining the Illicit Fentanyl Drug Supply: A Replication and Extension of Zoorob (2019)

Ali Crump

5/11/2020

Abstract

Zoorob (2019) shows that geography and fentanyl exposure explain much of the variation in increased overdose mortality rates between 2011 and 2017. I successfully replicated much of Zoorob's results, but I found discrepancies in the fentanyl exposure coefficients and the total death estimates for each model. My replication finds that the total death estimates are approximately 13% and 16% larger for each model and that the regression coefficients on fentanyl are slightly larger than those published. In addition to replicating Zoorob's work, this paper provides a Rubin causal model table to better understand the model framework, focusing on explaining the ordinary least-squares model. Next, the bulk of the extension investigates alternative definitions of fentanyl exposure while keeping all of Zoorob's other modeling choices the same. The original paper's definition of fentanyl exposure explains more variation in age-adjusted mortality rates than those proposed in the extension. This finding is important because many different methods of defining fentanyl exposure exist, however, the proposed alternative definitions in this extension do not appear to improve Zoorob's model.

Introduction

|z Fentanyl, a synthetic opioid 100 times more potent than morphine, is responsible for as much as two-thirds of the 46,802 overdose deaths in the United States in 2018 ((*What Is Fentanyl?*, n.d.)). However, the opioid epidemic does not affect the country equally, as eastern and northeastern states experience more overdose deaths and drug seizures than states west of the Mississippi River (Zoorob (2019)). Zoorob uses NFLIS forensic drug seizure data and age-adjusted mortality rate data[1] from state-years between 2011 and 2017 to estimate fentanyl overdose mortality rates in two models: an ordinary least-squares model and a two-stage least-squares model. Zoorob finds that longitude is highly predictive of fentanyl exposure and uses longitude as the instrumental variable in his second model. Zoorob estimates the causal effect of fentanyl exposure on mortality rates to be about 4.5 and 5.4 for each model respectively, interpreted as the difference in the effect of fentanyl exposure between two state-years that differ from each other by one standard deviation. Zoorob also provides estimates of mortality deaths for each year, which are broadly consistent with official mortality statistics.

Each of the models predicts overdose mortality as a function of a variable he calls fentanyl exposure. Zoorob defines fentanyl exposure by taking into account the state, year, an error term, and the natural logarithm of the number of seizure tests containing fentanyl (plus 1 to avoid an undefined logarithm of zero). Zoorob also divides by a state's population to generate a measure that allows comparisons across different states:

[1]Zoorob obtained the data used for his analysis through a Freedom of Information Act request. The data consist of state test results for drug seizures between 2011 and 2016, which he filters for test results containing fentanyl. Zoorob also uses age-adjusted mortality data from the National Center for Health Statistics. All the data used contain state and year information, and he uses state-annual populations to calculate mortality rates relative to a state's population in a particular year. The data and code that Zoorob used in his paper are available on the Harvard Dataverse

$$Fentanyl_{ij} = \log\left(\frac{S_{ij}}{P_{ij}} + 1\right)$$

Model 1 below is an ordinary least squares equation where α_i is state i and η_j is year j The standard errors are two-way clustered by state and year and include population weights (Zoorob (2019)).

$$Overdose_{ij} = \alpha_i + \eta_j + \beta_1 Fentanyl_{ij} + \epsilon_{ij}$$

The second model uses a two-stage least squares regression with longitude as the instrument variable:

$$\widehat{Fentanyl}_{ij} = \alpha_i + \eta_j + \beta_1(Longitude_i \cdot Year_j) + \epsilon_{ij} Overdose_{ij} = \alpha_i + \eta_j + \beta_2 \widehat{Fentanyl}_{ij} + \epsilon_{ij}$$

Findings in the paper show that much of the variation in the increased overdose mortality rates is explained by fentanyl exposure, getting stronger over time. Zoorob's analysis of police fentanyl seizures sheds light on the illegal drug supply and drug environments in the United States, showing that illicitly manufactured fentanyl contributed to the worsening overdose crisis. Zoorob determines that the drug supply is shaping the epicenter of the overdose crisis, for the epicenter shifted towards the eastern U.S., and states east of the Mississippi River tend to have greater fentanyl exposure and sharper increases in overdose deaths than states west of the Mississippi River.

To conduct my replication, I used R^2(R Core Team (2013)). I successfully replicated all of Zoorob's figures and one of the fentanyl exposure regression tables (Table 1); however, I was unable to replicate Table 2. The replicated total death estimates are approximately 13% and 16% larger for each model and the regression coefficients on fentanyl exposure are slightly larger than those published. The extension of the project proceeds with the model and regressions used in the corrected table.

The goals of this extension are to better understand Zoorob's models and to see if alternate definitions of fentanyl exposure could improve the models. The 'Rubin Table' frames the paper, which uses topics not covered in the scope of this course, in a context any reader familiar with statistical regressions will understand. The reasoning behind the second facet of the extension is that not just one definition of fentanyl exposure exists; one could apply seizure data to quantify a state's drug environment using other methods. Using three alternative definitions that seem reasonable, I conduct Zoorob's same analyses. He incorporates population weights in his regression analyses, so I examine a definition that does not use the state populations. I also test other plausible definitions that perform log transformations before dividing by state populations.

The back-of-the-envelope calculations from two of the alternative definitions are wildly unrealistic, estimating nearly a quarter of a million deaths up to as many as half a million deaths due to fentanyl between 2011 and 2017. Fentanyl only represents a portion of synthetic opioid deaths; even if fentanyl accounted for every single overdose death, official mortality statistics report 360,242 total overdose deaths, making these larger estimates impossible ((*Data Brief 33536. Drug Overdose Deaths in the United States, 1999-2018*, n.d.)). In contrast, the other definition used in the extension estimates total deaths approximately 80% smaller and 53% than Zoorob's models, respectively. Since official fentanyl overdose death numbers remain unknown, these estimates are not impossible, however, they are likely too extreme. Consistent with these extreme estimates, comparing R^2 values for each of the models strongly suggests that Zoorob's method for measuring fentanyl exposure explains more variation in mortality rates than the alternate definitions of fentanyl exposure and that his method is more feasible.

Literature Review

The number of drug overdose deaths in the United States has rapidly increased since 2014. However, the opioid epidemic did not affect all regions of the U.S. equally; according to the CDC, almost all states west of the Mississippi River did not see an increase while those to the east did. Similar work to Zoorob's by Rigg et al. (2018) investigates the role of geography in opioid-related mortalities. Rigg et al. find that the opioid

[2]More information on this project can be found on my Github repository., Github repository

epidemic in the United States is not disproportionately rural and that the epidemic varies greatly across different rural areas (Khary K. Rigg (2018)). Also, research by Barocas et al. on the effect of opioid use among patients with endocarditis also cites Zoorob's work. Barocas et al. find a decreased risk of overdose associated with the West and South compared to the Northeast, consistent with Zoorob's conclusions (Joshua A. Barocas (2020)). Other authors' claims are consistent with Zoorobs; for example, research by Pearce et. al (2020) and Gladden et al. (2019) cite Zoorob's claim that the drug supply is contributing to the overdose crisis in their work (Lindsay A. Pearce and Nosyk (2020))(R. Matt Gladden and Kariisa (2019)). However, other work such as that of Dasgupta et. al (2018) argues social and economic factors play a role in one's susceptibility to opioid addiction and overdose, whereas Zoorob claims that the geographical patterns point to drug supply also playing a primary role in the epidemic. Surprisingly, drug overdose deaths decreased in 2018 by 4.6% from 2017 in the United States, however, fentanyl deaths continued to rise ((*Drug Overdose Deaths*, n.d.))(Abby Goodnough (2019)). The amount of unknowns and variation in the opioid epidemic coupled with this unexpected change in 2018 shows that more testing, data, and further work are required to better understand the overdose crisis and the impact of synthetic opioids like fentanyl.

Replication

The replication was partially successful. I successfully replicated Figures 1, 2, 3, 4, and Table 1 from the paper as well as figures from the appendix. As an example, Figure 1 is shown in the appendix. However, I was unable to replicate the results from Table 2, which include the parameters of the statistical models for overdose and fentanyl exposure for both the least-squares model (Model 1) and the second stage of a two-stage least squares model (Model 2). The results for table 2 from the original paper and my replication are shown below:

A: Statistical Models of Fentanyl & Overdose

Dependent variable: Age-Adjusted Mortality Rate		
	(Model 1, OLS)	(Model 2, 2SLS)
Fentanyl Exposure	4.495***	5.426***
	(0.687)	(0.717)
State Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
Population Weights	Y	Y
Observations	357	357
Adjusted R^2	0.915	0.909
First-Stage *F*		31.6

B: Total Estimated Deaths Attributable to Fentanyl by Model.

	Model 1 Deaths	Model 2 Deaths
2011	2,295	2,705
2012	2,365	2,788
2013	3,312	3,904
2014	8,870	10,458
2015	15,446	18,211
2016	23,188	27,339
2017	30,398	35,841

*p < 0.1.
**p < 0.05.
***p < 0.01.

Figure 1: Zoorob (2019) Table 2

Zoorob's Model 1 estimates 85,874 total overdose deaths, while my replication estimates 96,969, approximately 13% higher. Likewise, for Model 2, his estimates total to 101,246, whereas mine total to 117,079, nearly 16% higher. The author confirmed slightly different models likely provided the published estimates than the ones included in the replication materials. In addition to the higher death estimates, my estimates

for the coefficient on fentanyl exposure are slightly higher than Zoorob's for both models, which might explain why my estimated deaths are also higher. My estimates are shown below.

Table 1

	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
fent_r	4.508***	
	(0.635)	
'fent_r(fit)'		5.443***
		(0.653)
Observations	357	357
R^2	0.928	0.923
Adjusted R^2	0.914	0.908
Residual Std. Error (df = 299)	5,372.861	5,545.678
Note:	*p<0.1; **p<0.05; ***p<0.01	

% latex table generated in R 3.6.1 by xtable 1.8-4 package % Mon May 11 05:52:46 2020

	Model 1 Deaths	Model 2 Deaths
2011	2580	3115
2012	2659	3210
2013	3723	4495
2014	9973	12041
2015	17367	20969
2016	26491	31985
2017	34176	41263

Extension

For my extension, I started by understanding the model using the framework of potential outcomes. Unlike more straightforward Rubin causal models with binary treatments, fentanyl exposure can theoretically take on any positive value. In the Rubin Causal Table below, the units are state-years for every state in the U.S. and for every year from 2011 to 2017. Fentanyl exposure represents the state's fentanyl exposure measured using seizure and population data. The table arbitrarily looks at the interval of fentanyl exposure from 2.0 to 2.6 for simplicity, but theoretically fentanyl exposure ranges from 0 to infinity. The values in the table for a particular state, year, and fentanyl exposure represent the overdose mortality rate in that state and year. Every state-year can only experience one level of fentanyl exposure, the fundamental problem of causal inference. Therefore, for every other fentanyl exposure level the state did not experience, the mortality rate is represented by a question mark; a state-year cannot experience more than one level of fentanyl exposure, so those mortality rates are unobservable unknown values. One works around this problem by using other states' mortality rates to fill in the question marks.

Rubin Causal Table

Estimating Overdose Mortality Rate Using Fentanyl Exposure

		Fentanyl Exposure							
Unit	Year	...2.0	2.1	2.2	2.3	2.4	2.5	2.6...	Causal Effect
Alabama	2011	?	?	?	4.7	?	?	?	?
Alaska	2011	?	?	?	?	?	?	2.9	?
Arizona	2011	1.3	?	?	?	?	?	?	?
Arkansas	2011	?	?	3.1	?	?	?	?	?
...	...	...	...	...	...	...	...	...	...
Alabama	2012	?	?	?	?	?	5.0	?	?
...	...	...	...	...	...	...	...	...	...
Wyoming	2017	?	2.0	?	?	?	?	?	?

For example, suppose Alabama had a fentanyl exposure of 2.4 and a mortality rate of 4.7 deaths per 100,000 people in 2011. We cannot observe what the mortality rate would have been if fentanyl exposure had been 2.0, 2.1, 2.2, etc. Instead, we use mortality rates from other states with other fentanyl exposures to "fill in the blanks." Controlling for state populations and state-fixed effects, we might predict Alabama's 2.3 fentanyl exposure mortality rate in 2011 using Arkansas' 3.1 mortality rate. Continuing this process for every "question mark," we can compute the causal effect of a 1 unit change in fentanyl exposure. This visualization helps make these complex models more accessible.

In the second part of the extension, I explore how Zoorob uses the NFLIS state seizure data to define a measure of fentanyl exposure, $Fentanyl_{ij} = \log\left(\frac{S_{ij}}{P_{ij}} + 1\right)$. In essence, fentanyl exposure is a measure of a state's fentanyl drug supply in a particular year, controlling for population sizes and other fixed effects. He uses this measure of fentanyl exposure in his regression for both Models 1 and Model 2. Zoorob explains that he uses this measure to compare fentanyl exposure across states of highly variable populations, and that the plus 1 avoids and undefined logarithm of zero. However, the decision to transform the variables in this manner remains partially unclear. For my extension, I define fentanyl exposure several ways:

1. $Fentanyl_{ij} = \log\left(S_{ij} + 1\right)$
2. $Fentanyl_{ij} = \frac{S_{ij}}{\log\left(P_{ij}+1\right)} + 1$
3. $Fentanyl_{ij} = \frac{\log\left(S_{ij}+1\right)}{\log\left(P_{ij}\right)}$

and run his same analyses. Zoorob's regression includes population weights and information already, so the first equation considers the choice of using P_{ij} to define fentanyl exposure. The second definition only takes the log of population, since populations are large compared to the number of fentanyl drug seizures. The third definition is similar to Zoorob's but instead takes the log separately for both values, S_{ij} and P_{ij}. The goal of this component of the extension is to see how different assumptions in defining fentanyl exposure affect the model results and to see if the original conclusions are robust to reasonable changes in the model.

Tables

% latex table generated in R 3.6.1 by xtable 1.8-4 package % Mon May 11 05:52:47 2020

Conclusion

My ordinary least squares estimated fentanyl deaths from extensions 1 and 3, totaling to 234,037 and 238,812 respectively, are much larger than Zoorob's corrected estimates, totaling to 96,969. Zoorob's corrected 2016 estimate, 26,491 deaths, is also more consistent with the official mortality statistic of 19,400 than either of the extension estimates: 46,641, 5,732, and 47,610 respectively. Estimates from extensions 1 and

	Dependent variable:				
	ext1				
	(1)	(2)	(3)	(4)	(5)
longitude	0.017**	0.047***	0.053***	0.066***	0.077***
	(0.009)	(0.009)	(0.011)	(0.011)	(0.011)
latitude	−0.015	−0.001	−0.049	0.003	−0.009
	(0.027)	(0.030)	(0.036)	(0.035)	(0.035)
MORT_2013	−0.028	0.048	0.029	0.022	0.008
	(0.030)	(0.033)	(0.040)	(0.038)	(0.039)
Constant	5.131***	7.059***	10.730***	10.775***	13.076***
	(1.446)	(1.587)	(1.913)	(1.849)	(1.883)
Observations	51	51	51	51	51
R^2	0.097	0.377	0.359	0.441	0.508
Adjusted R^2	0.040	0.338	0.319	0.405	0.477
Residual Std. Error (df = 47)	1.164	1.277	1.539	1.487	1.515
F Statistic (df = 3; 47)	1.690	9.496***	8.791***	12.348***	16.171***
Note:					*p<0.1; **p<0.05; ***p<0.01

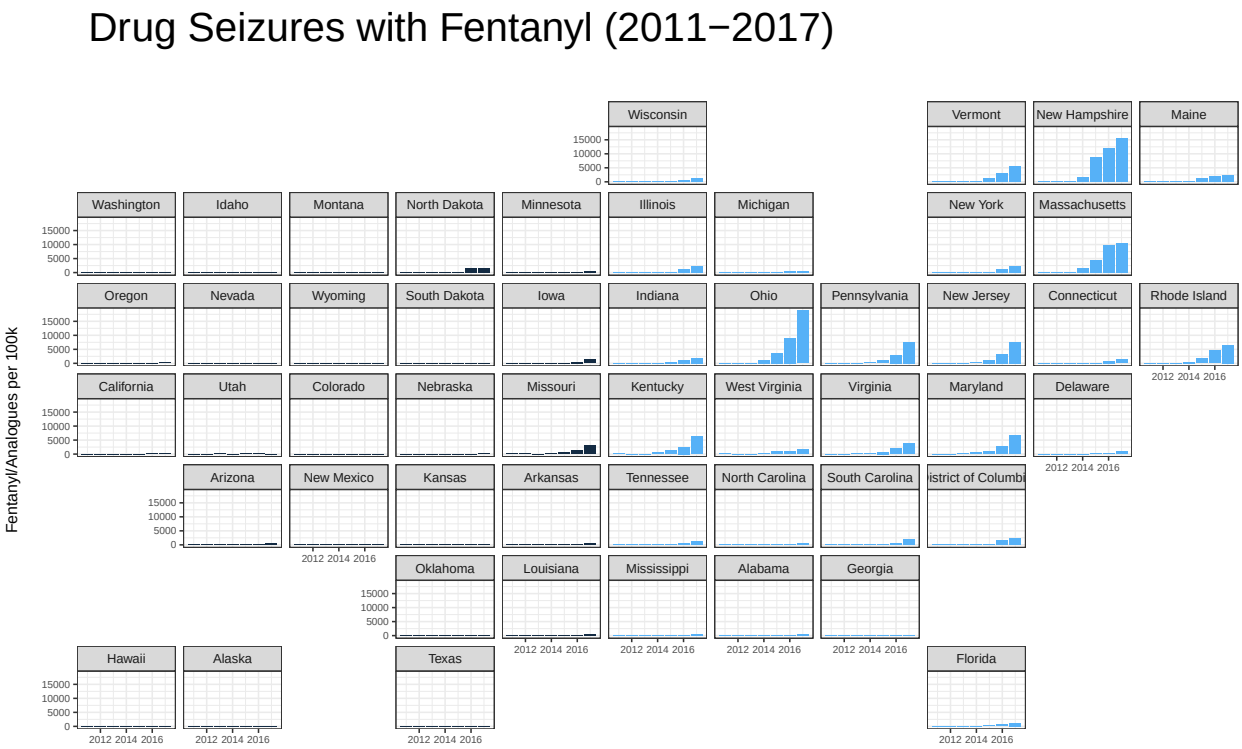
	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
ext1	2.380**	
	(0.709)	
‘ext1(fit)‘		5.549**
		(1.523)
Observations	357	357
R^2	0.866	0.781
Adjusted R^2	0.840	0.739
Residual Std. Error (df = 299)	7,319.545	9,357.110
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Model 1 Deaths	Model 2 Deaths
2011	19747	46039
2012	20795	48482
2013	23774	55429
2014	31860	74280
2015	39190	91371
2016	46641	108743
2017	52030	121305

3 appear to be too high, and estimates from extension 2 appear to be too low to be consistent with official mortality statistics.

Moreover, when comparing R^2 values for each model, Zoorob's method of defining fentanyl exposure appears to explain more of the variation in mortality rate than any of the extensions. Zoorob's R^2 for the OLS model is 0.915, whereas those of the extensions are 0.840, 0.883, and 0.846.

Appendix



Extension 2

Extension 3

	Dependent variable:				
	ext2				
	(1)	(2)	(3)	(4)	(5)
longitude	0.024*	0.241**	0.788**	1.928**	3.422**
	(0.012)	(0.107)	(0.335)	(0.810)	(1.569)
latitude	−0.014	0.149	0.602	1.586	2.200
	(0.039)	(0.340)	(1.064)	(2.569)	(4.977)
MORT_2013	−0.012	0.397	1.119	2.450	5.487
	(0.043)	(0.373)	(1.168)	(2.822)	(5.466)
Constant	4.270**	17.170	53.121	130.373	244.325
	(2.054)	(18.032)	(56.439)	(136.332)	(264.082)
Observations	51	51	51	51	51
R^2	0.079	0.125	0.130	0.130	0.118
Adjusted R^2	0.020	0.069	0.075	0.074	0.061
Residual Std. Error (df = 47)	1.653	14.509	45.411	109.693	212.481
F Statistic (df = 3; 47)	1.342	2.243*	2.342*	2.338*	2.091
Note:					*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
ext2	0.022***	
	(0.004)	
‘ext2(fit)‘		0.060***
		(0.014)
Observations	357	357
R^2	0.902	0.657
Adjusted R^2	0.883	0.592
Residual Std. Error (df = 299)	6,268.497	11,704.250
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Model 1 Deaths	Model 2 Deaths
2011	85	230
2012	86	232
2013	138	374
2014	717	1943
2015	2159	5850
2016	5732	15528
2017	11482	31104

	Dependent variable:				
	ext3				
	(1)	(2)	(3)	(4)	(5)
longitude	0.001*	0.003***	0.003***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
latitude	−0.0004	0.001	−0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
MORT_2013	−0.002	0.004*	0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.300***	0.413***	0.657***	0.640***	0.791***
	(0.088)	(0.092)	(0.116)	(0.105)	(0.102)
Observations	51	51	51	51	51
R^2	0.090	0.433	0.383	0.511	0.600
Adjusted R^2	0.032	0.397	0.344	0.480	0.574
Residual Std. Error (df = 47)	0.071	0.074	0.093	0.084	0.082
F Statistic (df = 3; 47)	1.558	11.982***	9.743***	16.381***	23.501***
Note:					*p<0.1; **p<0.05; ***p<0.01

References

Abby Goodnough, Margot Sanger-Katz, Josh Katz. 2019. *Drug Overdose Deaths Drop in U.s. For First Time Since 1990.* The New York Times. https://www.nytimes.com/interactive/2019/07/17/upshot/drug-overdose-deaths-fall.html.

Data Brief 33536. Drug Overdose Deaths in the United States, 1999-2018. n.d. Centers for Disease Control; Prevention. https://www.cdc.gov/nchs/data/databriefs/db356_tables-508.pdf#page=1.

Drug Overdose Deaths. n.d. Centers for Disease Control; Prevention. https://www.cdc.gov/drugoverdose/data/statedeaths.html.

Joshua A. Barocas, Jianing Wang, Jake R. Morgan. 2020. *Outcomes Associated with Medications for Opiod Use Disorder Among Persons Hospitalized for Infective Endocarditis. Infectious Diseases Society of America.* https://watermark-silverchair-com.ezp-prod1.hul.harvard.edu/ciaa062.pdf?token=
AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAn0wggJ5BgkqhkiG9w0BBwagggJqMIICZgIBADCC
EQx-LXUgTgLIiPAgEQgIICMC-7gICW1kwD-HanETepgOj5x6deCq5Fr3n9WXriHm-9kI1nEaZorUoyzw9fFR28xBPraP8ad
x9ke4fdLGmXZ6KXr9G06vH0T45mov4AIT93pZ0u8a1WCqluFe92iBuye2CAUOl0ZOOHmI8fTSPxhwTFSZdu6Uncelk9wrR
IfiALtxGG5zMzG9jlEnpSaWEmsBeYMmr7IHFRmuKvFkseVcPMtAgtO3N9Wjo7qzWYW48WLME18g2uFRaidSMtzVMCf
DKiIiFGvq7LbF3z0MNRaohlFQ2WpZagYsnscG41hy5OUQJME4K7VpCChcM0V6H159z1wZf4ooVo65ipAjiWs1tgbeXssyjuk

	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
ext3	39.298**	
	(11.594)	
‘ext3(fit)‘		84.072***
		(20.209)
Observations	357	357
R^2	0.870	0.802
Adjusted R^2	0.846	0.765
Residual Std. Error (df = 299)	7,195.770	8,887.805
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Model 1 Deaths	Model 2 Deaths
2011	20136	43079
2012	21218	45393
2013	24225	51827
2014	32521	69575
2015	40023	85623
2016	47610	101854
2017	53079	113554

4elwIf0YBWCouCB-vWRbtfJRZnWm4etPfARlgtTKLD4DL4kxW6fes-609BTXekgqTMc0qTqoo4StQUMxGL55SsTayp1EFI
rxb-ASlH8-TZlAKAiAW6RneBSD3ClYw2twDMGAZyEsxNON6veX7tgL3_6WcvJ_zgCUPhBkEkJS22-
03MXt6XY0bOflFasTqjYgRBTAj4pbvHAYpRdRckKWft2iuSPaY06a7zF1P1yah5vWnoPbDvZjlzYC22.

Khary K. Rigg, Melody N. Chavez, Shannon M. Monnat. 2018. *Opioid-Related Mortality in Rural America: Geographic Heterogeneity and Intervention Strategies.* International Journal of Drug Policy. https://doi.org/10.1016/j.drugpo.2018.04.011.

Lindsay A. Pearce, Micah Piske, Jeong Eun Min, and Bohdan Nosyk. 2020. *Opioid Agonist Treatment and Risk of Mortality During Opioid Overdose Public Health Emergency: Population Based Retrospective Cohort Study. BMJ.* https://www.bmj.com/content/368/bmj.m772.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

R. Matt Gladden, Bruce A. Goldberger, Julie O'Donnell, and Mbabazi Kariisa. 2019. *Notes from the Field: Opioid-Involved Overdose Deaths with Fentanyl or Fentanyl Analogs Detected — 28 States and the District of Columbia, July 2016–December 2018.* US National Library of Medicine National Institutes of Health. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7075253/.

What Is Fentanyl? n.d. Centers for Disease Control; Prevention. https://www.cdc.gov/drugoverdose/opioids/fentanyl.html.

Zoorob, Michael. 2019. *Fentanyl Shock: The Changing Geography of Overdose in the United States. International Journal of Drug Policy.* https://www.sciencedirect.com/science/article/abs/pii/S0955395919301136.