



## Data quality considerations when using county-level opioid overdose death rates to inform policy and practice: A reply



To the Editor,

Jones et al. (2019) highlight an important issue in opioid mortality research regarding the reliability of information on specific substances present in drug poisoning cases. Efforts to improve data collection are vital. At the same time, the immediacy of the opioid crisis may argue in favor of cautiously using the existing data with methods that are robust to data limitations. I welcome the opportunity to present further results highlighting the robustness of the Gini coefficient declines reported in McClellan (2019) to the data limitations Jones et al. discuss.

Since the Gini characterizes the entire distribution of the data and depends on the rank-ordering of the distribution (Gastwirth, 2017), the scope of under-reporting must be somewhat large to alter the conclusions of the analysis. To illustrate, consider replacing the opioid mortality rates in the entire bottom 10% of counties with randomly selected values from the top 10% of counties. This simulates a considerable correction of the data in the spirit of the example given by Jones et al. (2019).

In 2000, such a reshuffling of county mortality only moves the Gini to 0.755 from the originally calculated Gini of 0.813. To achieve the observed -0.2 change in the Gini between 2000 and 2014 (McClellan, 2019), the bottom 30.7% of counties must be assigned values on par with the top 10% of counties. For comparison, in 2000, 14.4% of

counties reported only unidentified-substance mortality, while unidentified-substance mortality comprised over half of substance-related deaths in only 22.6% of counties. This under-reporting is likely not extensive enough to explain the observed downward trend.

Other methods can address the data limitations more directly. One such approach, pioneered by Ruhm (2017), imputes opioid-related deaths by modeling those deaths as a function of observed characteristics and predicting opioid presence in unspecified substance-related deaths. Alternatively, researchers could assign an upper bound to opioid mortality by assuming all drug poisonings by unspecified substances are opioid-related. Though these approaches still contain mis-measurement, they help to provide an upper bound for the potential impact of measurement error in the reporting of opioid-related deaths.

Fig. 1 details how the analysis of opioid mortality inequality over time would be affected by (a) applying Ruhm's (2017) method and (b) by assuming all unspecified substance poisonings were the result of opioids. For both new measures, the Gini is approximately 0.1 lower over the entire period than the original estimate, however the pattern of increasing equality in mortality rates among counties remains virtually unchanged. While the issue of undercounting opioid deaths may be large, mortality inequality between counties remains high even after accounting for misreporting. Additionally, better reporting is not responsible for the observed decline in inequality.

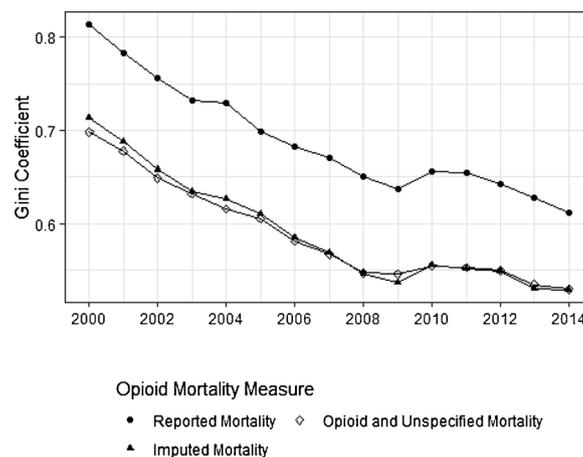


Fig. 1. U.S. Opioid Mortality Gini Coefficients by Imputation Method: 2000-2014.

Caption: The trend lines of Gini Coefficients for each year between 2000 and 2014 for three data series: Reported Mortality, Imputed Mortality, and Opioid and Unspecified Mortality. All three series show a steady downward trend (with a slight reversal in 2010). Reported Mortality Gini is approximately 0.1 points higher than the other Ginis for each year.

Notes: Estimates of the Gini are indicated by points. Reported Mortality are all deaths reporting presence of opioids. Imputed Mortality includes all reported opioid deaths plus unspecified substance related deaths statistically imputed to be related to opioids. Opioids and Unspecified Mortality includes all reported opioid deaths plus unspecified substance deaths.

<https://doi.org/10.1016/j.drugalcdep.2019.107550>

Received 26 August 2019

Available online 05 September 2019

0376-8716/ Published by Elsevier B.V.

While data quality is a very important consideration, it need not be an insurmountable barrier to research – particularly during times of pressing need for policy-related insights. Researchers should carefully consider the implications of data limitations for their analysis and seek innovative ways to overcome poor data while waiting on better data to become available. Jones and colleagues do an important service in raising awareness of the policy relevant limitations of the data that should be a focus of these research efforts and emphasizing the need for cautious interpretation of work with unavoidably problematic data.

#### **Role of funding source**

Nothing Declared

#### **Contributors**

C. McClellan was responsible for all stages of analysis and manuscript preparation.

#### **Declaration of Competing Interest**

No conflict declared.

#### **Acknowledgements**

None.

#### **References**

- Gastwirth, J.L., 2017. Is the gini index of inequality overly sensitive to changes in the middle of the income distribution? *Stat. Public Policy (Phila)*. 4, 1–11.
- Jones, C.M., Warner, M., Hedegaard, H., Compton, W., 2019. Data quality considerations when using county-level opioid overdose death rates to inform policy and practice. *Drug Alcohol Depend* Forthcoming.
- McClellan, C.B., 2019. Disparities in opioid related mortality between united states counties from 2000 to 2014. *Drug Alcohol Depend*. 199, 151–158.
- Ruhm, C.J., 2017. Drug involvement in fatal overdoses. *SSM Popul. Health* 3, 219–226.

Chandler B. McClellan\*  
*Agency for Healthcare Research and Quality, United States*  
*E-mail address: [chandler.mcclellan@ahrq.hhs.gov](mailto:chandler.mcclellan@ahrq.hhs.gov).*

---

\* Corresponding author at: Center for Financing, Access, Cost, and Trends, Agency for Healthcare Research and Quality. 5600 Fishers Lane, Rockville, MD 20852, United States.