

Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags

Lauren M. Rossen, PhD, MS

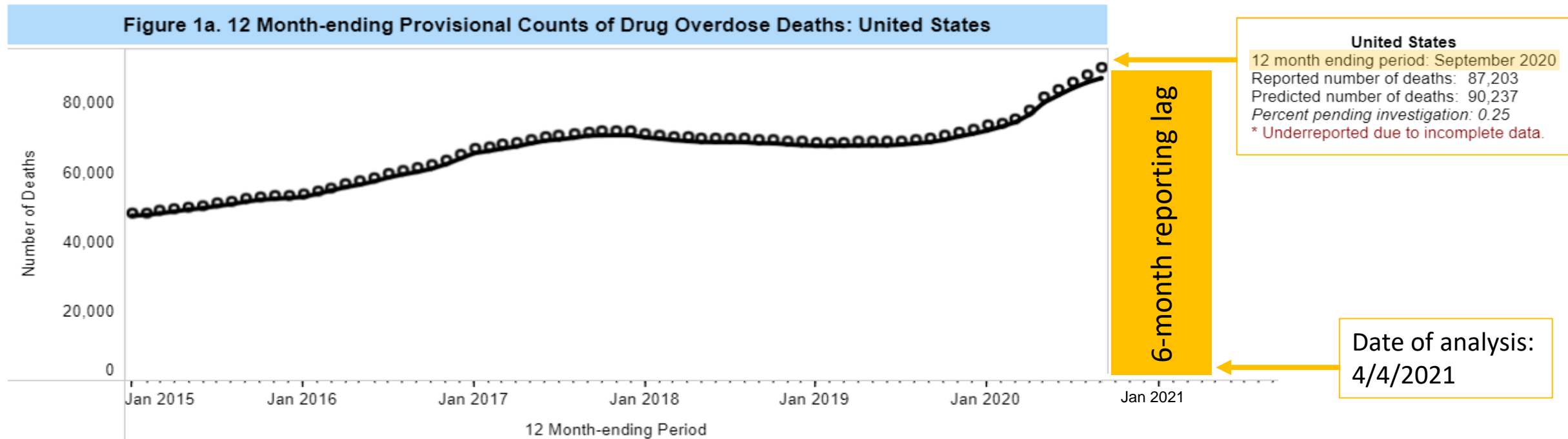
Division of Research and Methodology

National Center for Health Statistics

Centers for Disease Control and Prevention

Monthly provisional drug overdose death counts from the Vital Statistics Rapid Release Program

- 12 month-ending provisional counts of drug overdose deaths are published with a 6-month lag
- Need to ensure counts are reasonably complete and accurate



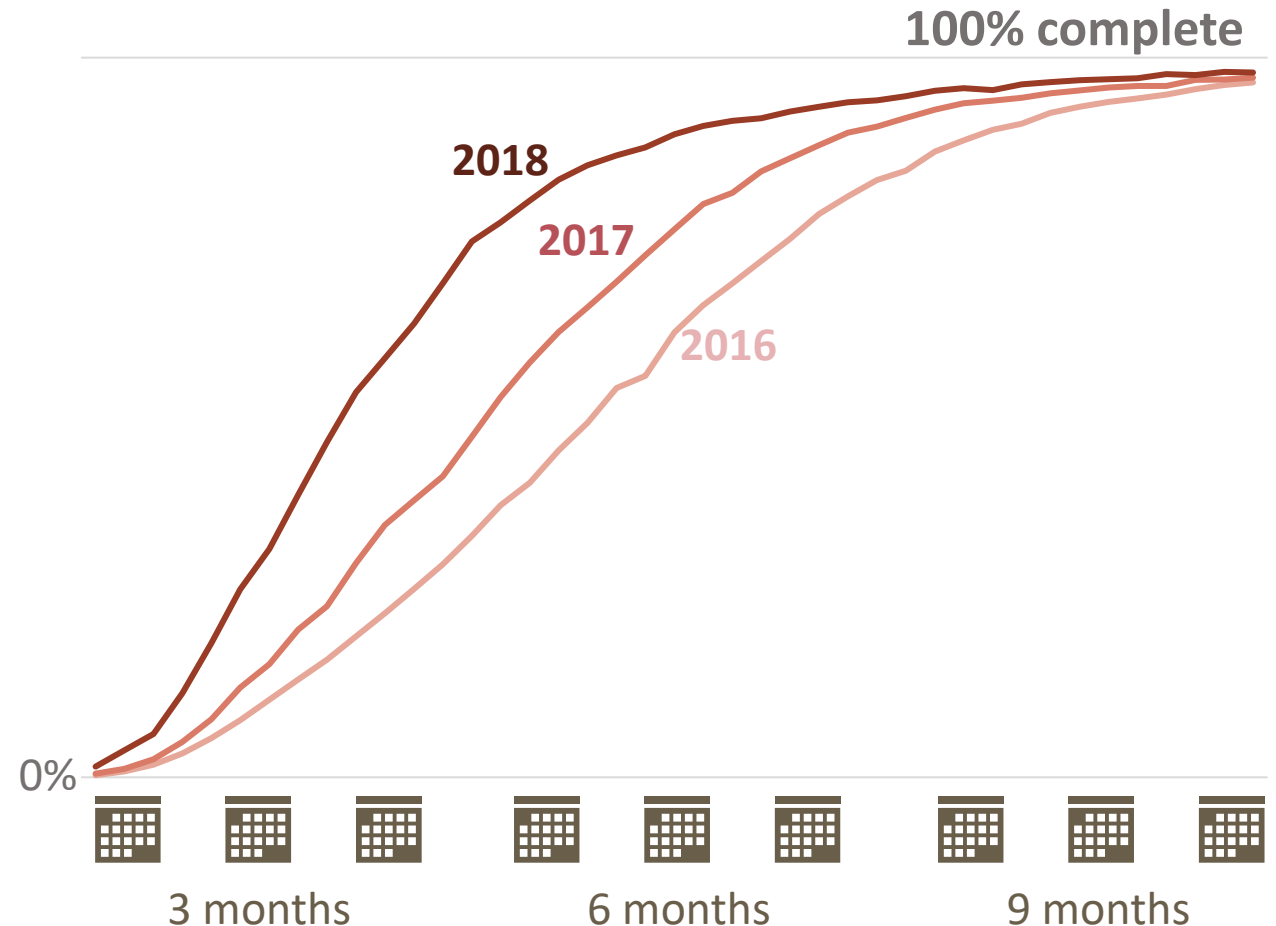
Deaths due to injury-related causes often undergo lengthy investigations and testing to determine the cause and manner of death

- About 60% of death certificate records *overall* are received by NCHS within 10 days of death
 - External causes of death take substantially longer (9 months or more)



Timeliness of drug overdose death data is improving

- Drug overdose deaths are being reported, coded, and processed faster than in years past
- As a result, provisional drug overdose data are more complete than in prior years



What has the impact of the COVID-19 pandemic been on injury-related causes of death?

Problems:

- Need weekly estimates
 - Existing provisional data are rolling annual estimates
- Need timely estimates, account for reporting lags
 - Existing provisional data released with a 6-month lag
- Need to account for improvements in timeliness
 - Unclear how timeliness and reporting may have changed since the pandemic (e.g., overburdened systems, vital statistics offices shifting to remote work, IT changes)

Nowcasting or “predicting the present”

- Used weekly ‘snapshots’ of provisional data (from 2018-2019) to model underreporting relative to final data
- Use models to predict completeness of provisional data within 2 weeks of death, 4 weeks, 6 weeks, ... , 39 weeks
- Create weights based on the inverse of predicted completeness
 - For example, if provisional mortality data in 2019 was 50% complete within 12 weeks of death, then the weight for that jurisdiction would be 2 for data presented with a 12-week lag
- Model the weighted counts, generate predicted numbers of deaths
- Validate the models by running on older provisional data and compare predictions with the reported counts 6-10 months later

Multi-stage hierarchical modeling process

Stage 1: Develop Weights

- **Model completeness**
 - Provisional 2019 data by jurisdiction and lag
- **Predict completeness**
 - 3 scenarios:
 - No change in timeliness
 - Moderate improvement
 - Extreme improvement
- **Create weights**
 - For each jurisdiction and lag, weight = 1/completeness

Stage 2: Model Outcomes

- **Model weighted counts**
 - Under 3 timeliness scenarios
- **Predict number of deaths**
 - Weekly number of deaths per jurisdiction from each model

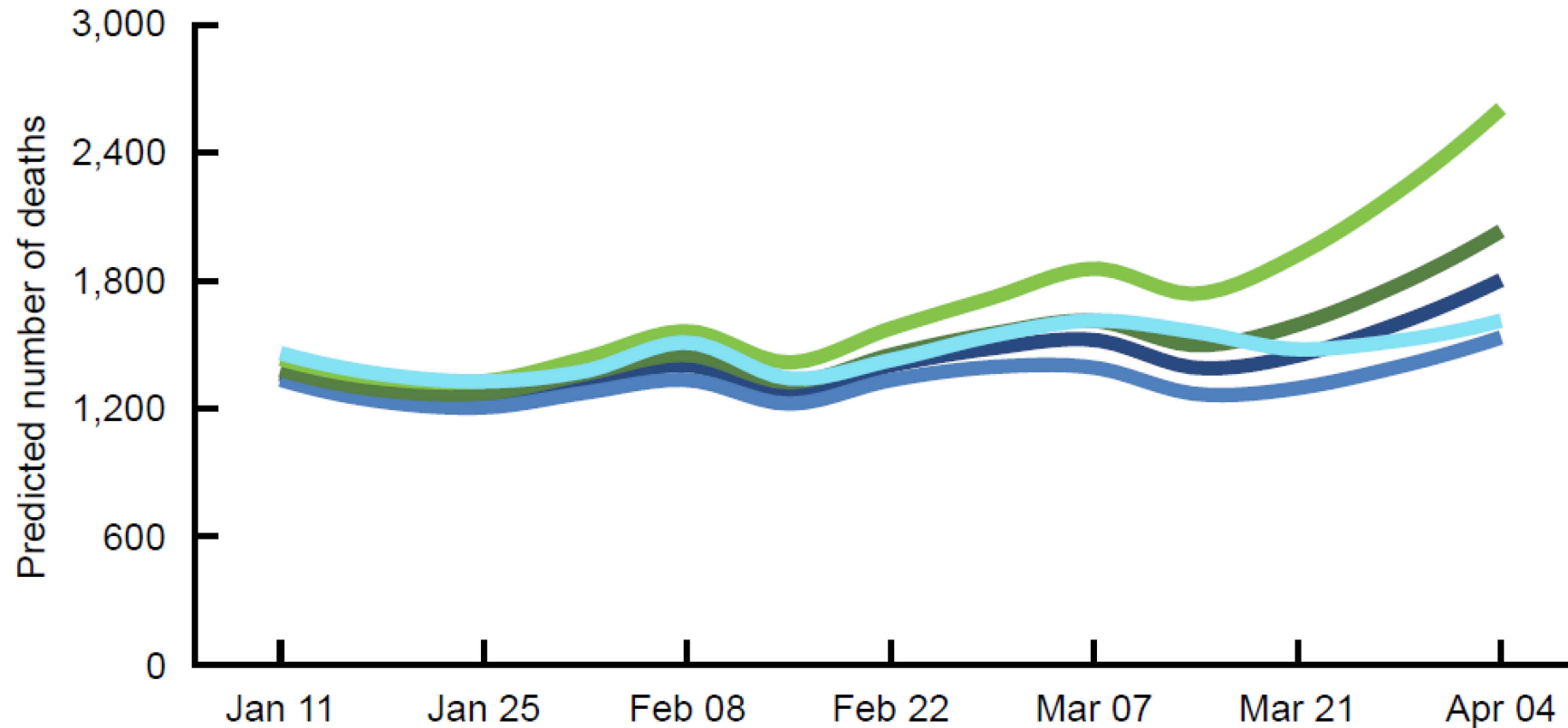
Validation

- **Fit stage 2 models**
 - Using older provisional data (May, 2020)
- **Predict number of deaths**
 - Under 3 timeliness scenarios
- **Compare predicted values to reported values from recent (nearly complete) data**

Validation Results: US*

Counts from November 22 Median timeliness 2019 Moderate improvements Extreme improvements Jurisdiction-specific timeliness

Predicted drug overdose trends



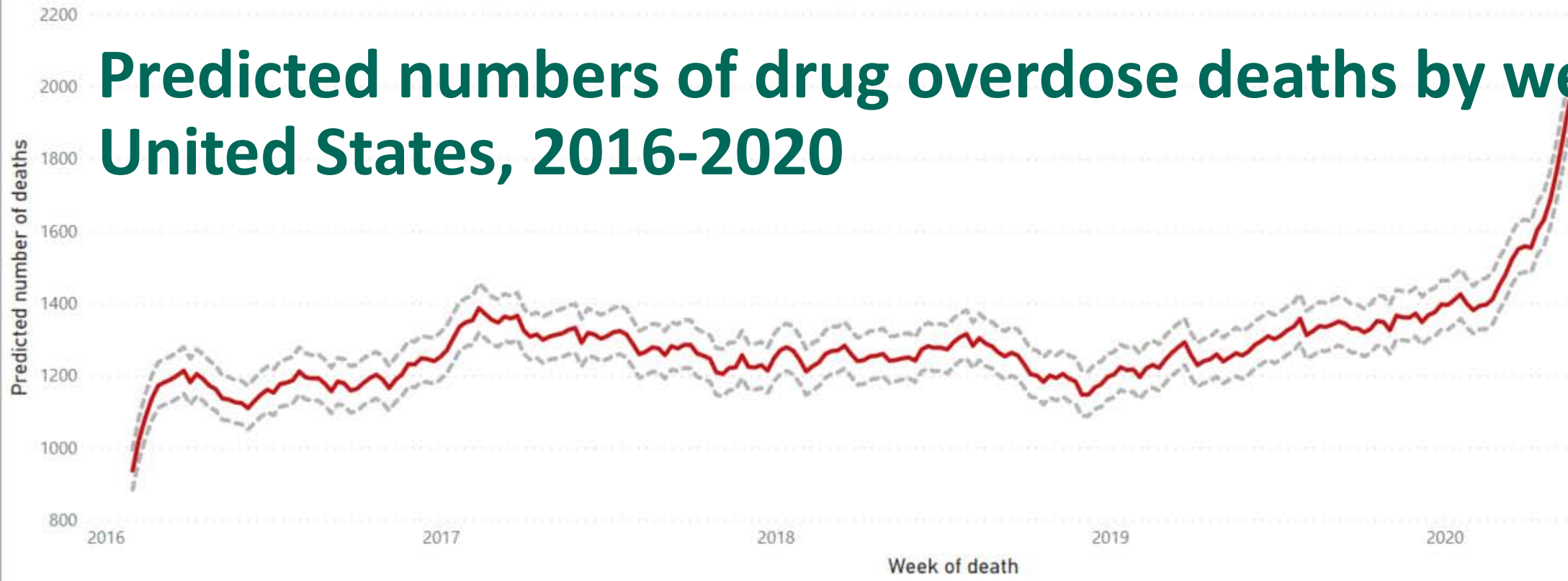
*excludes Connecticut and North Carolina

Figure 1. Predicted numbers of drug overdose, suicide, or transportation-related deaths by week and jurisdiction, 2016-2021

Select outcome:

Select a jurisdiction:

● 95% Credible Interval Lower Bound (rolling 4-week average) ● Posterior predicted median number (rolling 4-week average) ● 95% Credible Interval Upper Bound (rolling 4-week average)



Predicted numbers of drug overdose deaths by week United States, 2016-2020

Existing provisional data: 12 month-ending counts with a 6-month lag

Nowcasting → shorter lag (1 month), weekly counts

NOTES:

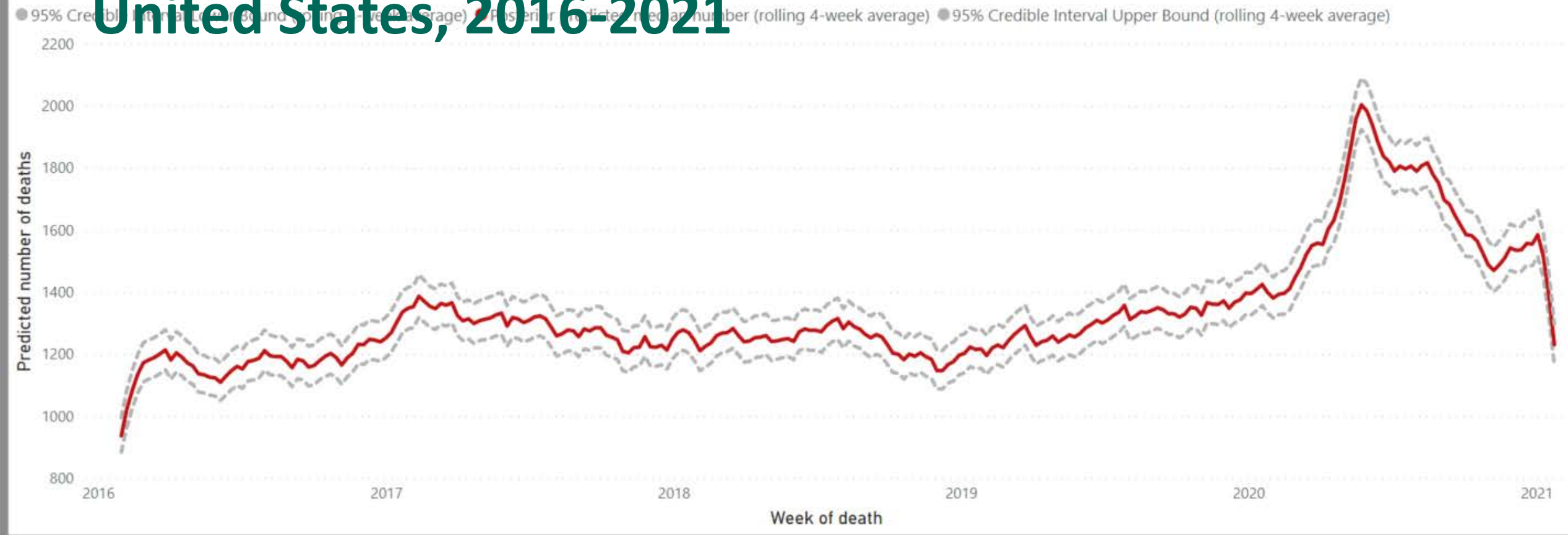
These early provisional estimates were generated using a set of hierarchical Bayesian models to account for reporting lags in injury-related causes of death. These models rely on assumptions about how the reporting lags have changed in recent months, relative to 2019. The resulting estimated numbers of deaths may be too high or too low, depending on how accurate these assumptions are, and the accuracy of the predicted estimates may vary by jurisdiction and over time. The true patterns in reporting lags will not be known until data are finalized, typically 11-12 months following the end of the data year. Estimates for the US exclude Connecticut and North Carolina, and are therefore not nationally representative. The figure above shows four-week rolling averages, and as such, the first few weeks of 2016 are not shown.

For more detail, see: Rossen LM, Hedegaard H, Warner M, Ahmad FB, Sutton PD. Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags. Vital Statistics Rapid Release; no 11. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: <https://doi.org/10.15620/cdc:101132>.

Figure 1. Predicted numbers of drug overdose, suicide, or transportation-related death, by week and jurisdiction, 2016-2021

Predicted numbers of drug overdose deaths by week, United States, 2016-2021

Select outcome: Drug Overdose
Select a jurisdiction: United State



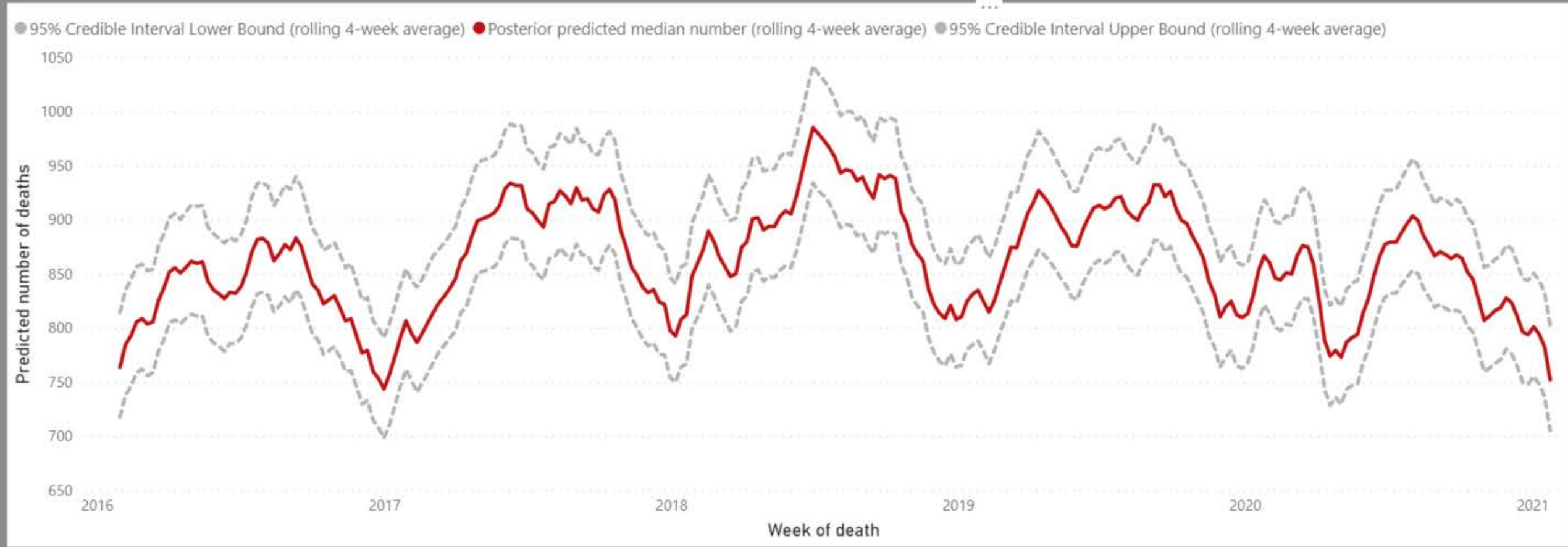
NOTES:

These early provisional estimates were generated using a set of hierarchical Bayesian models to account for reporting lags in injury-related causes of death. These models rely on assumptions about how the reporting lags have changed in recent months, relative to 2019. The resulting estimated numbers of deaths may be too high or too low, depending on how accurate these assumptions are, and the accuracy of the predicted estimates may vary by jurisdiction and over time. The true patterns in reporting lags will not be known until data are finalized, typically 11-12 months following the end of the data year. Estimates for the US exclude Connecticut and North Carolina, and are therefore not nationally representative. The figure above shows four-week rolling averages, and as such, the first few weeks of 2016 are not shown.

For more detail, see: Rossen LM, Hedegaard H, Warner M, Ahmad FB, Sutton PD. Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags. Vital Statistics Rapid Release; no 11. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: <https://doi.org/10.15620/cdc:101132>.

Figure 1. Predicted numbers of drug overdose, suicide, or transportation-related deaths by week and jurisdiction, 2016-2021

Select outcome: Select a jurisdiction:



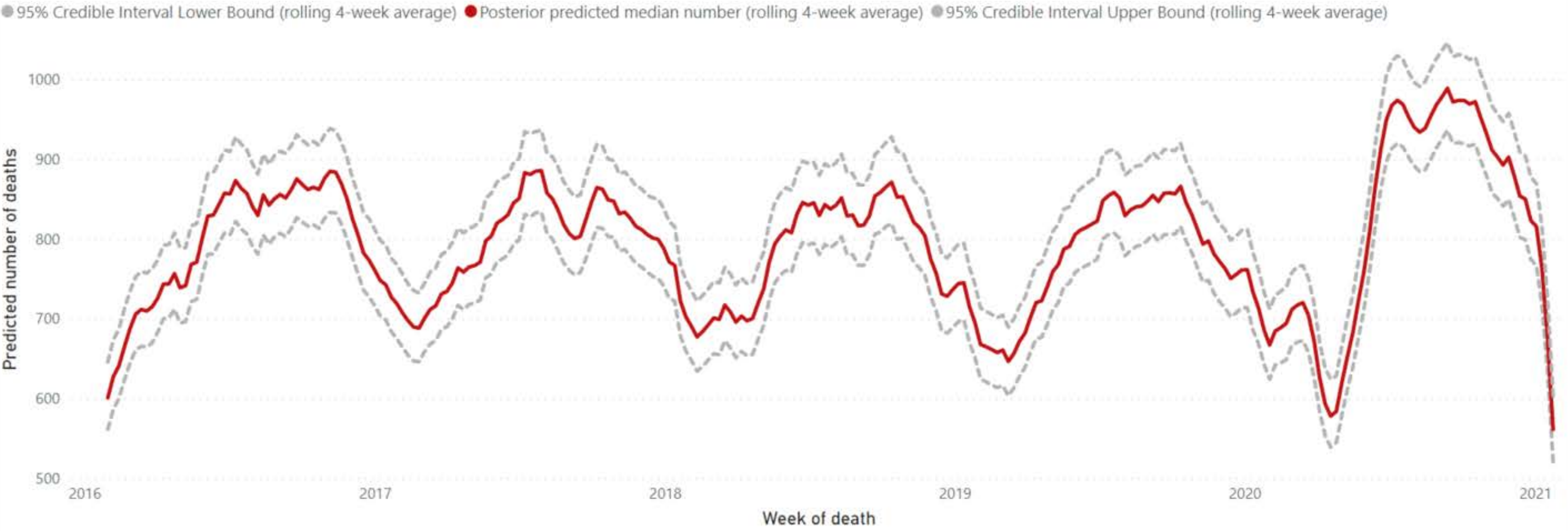
NOTES:

These early provisional estimates were generated using a set of hierarchical Bayesian models to account for reporting lags in injury-related causes of death. These models rely on assumptions about how the reporting lags have changed in recent months, relative to 2019. The resulting estimated numbers of deaths may be too high or too low, depending on how accurate these assumptions are, and the accuracy of the predicted estimates may vary by jurisdiction and over time. The true patterns in reporting lags will not be known until data are finalized, typically 11-12 months following the end of the data year. Estimates for the US exclude Connecticut and North Carolina, and are therefore not nationally representative. The figure above shows four-week rolling averages, and as such, the first few weeks of 2016 are not shown.

For more detail, see: Rossen LM, Hedegaard H, Warner M, Ahmad FB, Sutton PD. Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags. Vital Statistics Rapid Release; no 11. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: <https://doi.org/10.15620/cdc:101132>.

Figure 1. Predicted numbers of drug overdose, suicide, or transportation-related deaths by week and jurisdiction, 2016-2021

Select outcome: Select a jurisdiction:



NOTES: These early provisional estimates were generated using a set of hierarchical Bayesian models to account for reporting lags in injury-related causes of death. These models rely on assumptions about how the reporting lags have changed in recent months, relative to 2019. The resulting estimated numbers of deaths may be too high or too low, depending on how accurate these assumptions are, and the accuracy of the predicted estimates may vary by jurisdiction and over time. The true patterns in reporting lags will not be known until data are finalized, typically 11-12 months following the end of the data year. Estimates for the US exclude Connecticut and North Carolina, and are therefore not nationally representative. The figure above shows four-week rolling averages, and as such, the first few weeks of 2016 are not shown.

For more detail, see: Rossen LM, Hedegaard H, Warner M, Ahmad FB, Sutton PD. Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags. Vital Statistics Rapid Release; no 11. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: <https://doi.org/10.15620/cdc:101132>.

Limitations to the current approach

- Timeliness may have improved more at shorter lags than longer
- Timeliness may have changed throughout the pandemic
 - Hard to quantify, since injury-related death data is not complete for months
- Weights may be too large in some jurisdictions where timeliness is improving, and too small where timeliness has been a challenge
- Better account for uncertainty around timeliness?

Where do we go from here?

- Data Modernization Initiative project focused on further development, automation, expansion
 - Continue to refine methods
 - Expand to additional outcomes and data sources
 - Examine smaller geographies (e.g., county)

Thank you



Contact info:

Lauren Rossen
LRossen@cdc.gov