Statistical Methods: Suppression of Rates and Counts | U.S. Cancer Statistics Data Visualizations Tool Technical Notes | CDC

7/1/23, 7:53 AM

CDC | Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

United States Cancer Statistics (USCS)

United States Cancer Statistics (USCS) Home

Suppression of Rates and Counts

Suppression for Reliability

When the numbers of cases or deaths used to compute rates are small, those rates tend to have poor reliability.[1] Therefore, to discourage misinterpretation or misuse of rates or counts that are unstable, incidence and death rates and counts are not shown in tables and figures when the case or death counts are below 16. A count of fewer than 16 results in a standard error of the rate that is approximately 25% (or more) larger than the rate itself. Similarly, a case count below 16 results in the width of the rate's 95% confidence interval being at least as large as the rate itself. These relationships were derived under the assumption of a Poisson process and with the standard population age distribution assumed to be similar to the observed population age distribution. A suppressed rate does not necessarily mean that the rate was low.

Suppression for Confidentiality

Another important reason for using a threshold value for suppressing cells is to protect the confidentiality of patients whose data are included in a report by reducing or eliminating the risk of disclosing their identity.[2] The cell suppression threshold value of 16, which was selected to reduce misuse and misinterpretation of unstable rates and counts in this report, is more than sufficient to protect patient confidentiality. [PDF-324KB]

Suppression for Other Reasons

While data meet the U.S. Cancer Statistics (USCS) publication criteria, a central cancer registry may suppress its data for various reasons. For example, a state may have racial and ethnic groups (American Indian or Alaska Native, Asian or Pacific Islander, Hispanic) where the algorithms to correct for unknown race or ethnicity may not function properly. In these circumstances, data are suppressed upon the state's request.

U.S. State and County Data

Cancer incidence rates are presented for each county or county equivalent as available over the most recent 5-year period. County data are not available from Kansas and Minnesota because state legislation and regulations prohibit the release of county-level data to outside entities. County data from Virginia are suppressed due to incomplete data.

Statistical Methods: Suppression of Rates and Counts | U.S. Cancer Statistics Data Visualizations Tool Technical Notes | CDC

7/1/23, 7:53 AM

Additional data suppressions are applied in accordance with the rules outlined above.

Total United States

Cancer incidence rates for the United States are aggregate rates based on cancer cases reported from central cancer registries that met the USCS publication criteria and are the best estimates of the U.S. cancer burden available that are based on observed data. Case counts for the U.S. incidence rates for all ages combined are presented. Puerto Rico data are not included in the total U.S. counts and corresponding rate calculations.

References

1. Brillinger DR. The natural variability of vital rates and associated statistics. ↗ *Biometrics* 1986;42(4):693–734.

2. Doyle P, Lane JI, Theeuwes JM, Zayatz LM. *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies.* Amsterdam, The Netherlands: Elsevier Science; 2001.

Last Reviewed: June 8, 2023

Source: Division of Cancer Prevention and Control, Centers for Disease Control and Prevention