

Mortality and Prevalence Data on NIH's Categorical Spending Page

Description of Data Sources

The disease statistics included in the “Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC)” table are nationally-representative data gathered in 2015 by the National Center for Health Statistics (NCHS) at the Centers for Disease Control & Prevention (CDC). These data reflect two types of disease burdens: mortality and prevalence.

The mortality data are drawn from the [National Vital Statistics System](#) (NVSS), and show the number of deaths in which a particular disease or condition was mentioned on a deceased individual's death certificate. This means that, in some cases, more than one condition may be mentioned, and thus a particular death may count in more than one RCDC category. This method may under-estimate the number of deaths for some conditions, since chronic conditions, such as obesity and atherosclerosis, can greatly contribute to other conditions, such as heart disease, but are sometimes not explicitly mentioned on the death certificate.

The prevalence data are drawn from the [National Health Interview Survey](#) (NHIS) responses. This dataset reports the percentage of respondents who indicated that they were affected by a particular health condition. The survey questions collect different information for different conditions: some ask about occurrences in the last 12 months, while others ask about lifetime experiences; some ask about symptoms, while others ask if a health professional has ever made a particular diagnosis; and questions may be directed at either adults or children under 18. Cases where the posted prevalence value is measured only in children are indicated in the table. In addition, the reported numbers are accompanied by a standard error, which estimates the uncertainty introduced by sampling. In some cases, when the standard error for a particular disease or condition is too large to meet NCHS standards of reliability or precision, the estimate is not reported.

In order to ensure that data for each posted RCDC category was collected in a consistent manner, the posted information comes from two broad data sources provided by NCHS. Other disease statistics may be publicly posted by CDC and other federal sources which are rigorously collected, but are not collected in the same way across a wide range of diseases. In addition, NCHS assembled the posted statistics specifically to match NIH's RCDC categories, and so some numbers on this page which were uniquely assembled to fit a category may not be publicly posted outside of this analysis.

Frequently Asked Questions About This Analysis

(For more information on how the RCDC categories are generated, please see the RCDC FAQ at <https://report.nih.gov/rcdc/faqs.aspx>.)

Q. WHY IS NIH REPORTING THESE DATA?

A. NIH is reporting these data alongside NIH's categorical funding estimates in order to provide the public and policymakers with information that is helpful for understanding the NIH research portfolio

and its relationship to public health needs. NIH chose to report estimates from these two sources because they provide relatively consistent and, within each source, comparable Federal data for a broad range of categories.

Q. ARE THESE DATA THE DEFINITIVE INDICATORS OF PUBLIC HEALTH NEED FOR EACH DISEASE AND CONDITION?

A. No, these data are not meant to be definitive measurements of disease burden, mortality, or prevalence for a given condition in the US. In general, NIH believes that the best way to understand disease burdens is by examining patterns—both within the US and internationally—using multiple methods and measurements, chosen on a case-by-case basis as appropriate for each disease or condition.

Q. WHY DO THESE DATA COVER SOME CATEGORIES AND NOT OTHERS?

A. Some RCDC categories represent areas of research rather than disease (i.e., genetics, neurosciences) and therefore do not map to disease statistics. Further, a number of RCDC categories representing diseases or conditions are not captured by the methods and questions of either data source. While estimates of prevalence or mortality may exist for these categories, it is inadvisable to compare such estimates with the ones listed here, since the methods by which they were collected differ.

Q. HOW WERE THESE DATA GENERATED?

A. NIH and NCHS collaborated to map relevant [International Classification of Disease \(ICD\) codes](#), which allow for standardized reporting of diseases and conditions for epidemiology, health management and clinical purposes, to each RCDC disease or condition category, as appropriate. NCHS then used those codes to generate mortality and prevalence data. In some cases, the appropriate matching between ICD codes and RCDC categories was somewhat subjective, and in such cases these matches represent the best judgement of NIH and NCHS staff.

Q. ARE THE POSTED VALUES UNIQUE TO THIS ANALYSIS? DO THESE FIGURES EXIST ELSEWHERE?

A. All reported figures were derived from NVSS or NHIS, which provide mortality and prevalence data for particular ICD codes. Since these ICD-RCDC matches were created specifically for the purpose of this reporting, some of the *combinations* of ICD codes may be unique to this analysis, and may not have appeared previously.