Finding and using health statistics: A self-study course

Finding and Using Health Statistics:
A Self-Study Course

This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
# Finding and Using Health Statistics

## Finding and Using Health Statistics

### About Health Statistics

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are “Correlates”?</td>
<td>5</td>
</tr>
<tr>
<td>What are “Conditions”?</td>
<td>6</td>
</tr>
<tr>
<td>What Statistics Exist about Health Care?</td>
<td>7</td>
</tr>
<tr>
<td>What Health Costs Can We Measure?</td>
<td>8</td>
</tr>
</tbody>
</table>

### Common Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling</td>
<td>10</td>
</tr>
<tr>
<td>Confidence Intervals</td>
<td>11</td>
</tr>
<tr>
<td>Validity</td>
<td>11</td>
</tr>
<tr>
<td>Dependent and Independent Variables</td>
<td>11</td>
</tr>
<tr>
<td>Age-Adjustment</td>
<td>12</td>
</tr>
</tbody>
</table>

### Sources of Health Statistics

<table>
<thead>
<tr>
<th>Source</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys</td>
<td>14</td>
</tr>
<tr>
<td>Medical Records</td>
<td>15</td>
</tr>
<tr>
<td>Claims Data</td>
<td>17</td>
</tr>
<tr>
<td>Vital Records</td>
<td>18</td>
</tr>
<tr>
<td>Surveillance</td>
<td>19</td>
</tr>
<tr>
<td>Peer-reviewed Literature</td>
<td>20</td>
</tr>
</tbody>
</table>

### Finding Health Statistics

<table>
<thead>
<tr>
<th>Source</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Center for Health Statistics (NCHS)</td>
<td>22</td>
</tr>
<tr>
<td>World Health Organization (WHO)</td>
<td>22</td>
</tr>
<tr>
<td>Agency for Healthcare Research and Quality</td>
<td>23</td>
</tr>
<tr>
<td>Centers for Disease Control and Prevention</td>
<td>24</td>
</tr>
<tr>
<td>Robert Wood Johnson Foundation County Health Rankings</td>
<td>24</td>
</tr>
<tr>
<td>Centers for Medicare and Medicaid Services</td>
<td>24</td>
</tr>
<tr>
<td>Kaiser Family Foundation</td>
<td>25</td>
</tr>
<tr>
<td>United States Census Bureau</td>
<td>25</td>
</tr>
<tr>
<td>HealthData.gov</td>
<td>26</td>
</tr>
<tr>
<td>Dartmouth Atlas of Health Care</td>
<td>26</td>
</tr>
<tr>
<td>Search Strategies</td>
<td>27</td>
</tr>
<tr>
<td>Search by Organization or Resource</td>
<td>27</td>
</tr>
<tr>
<td>Use a Search Engine</td>
<td>27</td>
</tr>
<tr>
<td>Search an Academic Database</td>
<td>28</td>
</tr>
</tbody>
</table>

### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29</td>
</tr>
</tbody>
</table>
Finding and Using Health Statistics

How many Americans die from cardiovascular disease every year? Has the number of Americans with health insurance gone up since the Affordable Care Act was passed? How do the highest causes of death in the United States differ from other countries?

The answers to these questions and many others are in the data collected by government, private, and non-profit organizations. Health statistics are very important for learning about health trends, and planning for health interventions.

This course for librarians and students in health sciences, describes different types of health statistics, how they are collected, and where they can be found.

Course Goals

• Understand what health statistics measure and how we can use them to improve general health
• Gain a basic knowledge of the statistical terms commonly used when reading about health statistics
• Learn different ways health information can be collected, and the pros and cons of each
• Become familiar with a variety of online sources for public health statistics
• Create a set of strategies to find specific health statistics

In this course

• Section 1: About Health Statistics page 3
  Find out what health statistics are, and who uses them.
• Section 2: Common Terms page 8
  Learn what phrases like “statistically valid,” and “sampling” really mean.
• Section 3: Data Sources page 12
  Discover where health statistics come from and why it matters.
• Section 4: Finding Health Statistics page 21
  Start your search for statistics to answer all your questions!
About Health Statistics
Statistics can help us improve everyone’s health. Researchers, public health professionals, policymakers, and physicians use health statistics to understand risk factors for communities, track diseases before they spread, see how effective policies are, and assess the quality, safety, and other aspects of health care.

Health statistics measure four types of information. The types are commonly referred to as the four C’s: Correlates, Conditions, Care, and Costs. We’ll take a look at each type of information and see how statistics can teach us more about them.

About Health Statistics Modules
- **Correlates** page 4
  See how to measure the risk factors and protective factors that impact our health.
- **Conditions** page 5
  Learn to assess how often and how badly diseases impact a community.
- **Care** page 6
  Dig into how health care is delivered to the communities that need it, to treat disease and illness.
- **Costs** page 7
  Get more information on what health care costs, and why.
What are “correlates”?
Health **correlates** are the risk factors that impact our health. They include social and economic factors, like income and education; physical factors like air pollution; and personal behaviors like smoking, exercise, drugs, and alcohol.¹

This graph shows us information about smoking in the United States. The data shows us what percentage of people never smoked, used to smoke, or currently smoke, and whether the current smokers are daily or infrequent smokers.²

![Image 1: A graph showing the percentage distribution of current cigarette smoking status of adults, by sex. Information is from the United States, 2008-2010.](image-url)

1 http://www.who.int/hia/evidence/doh/en/

*This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.*
What are “conditions?”
Health conditions may be the most common type of health statistic. There are two ways to measure health conditions. One way is to look at the prevalence of the disease, or the total number of people with that illness at a specific time. The other is the incidence, or the number of new cases of the illness.

Both numbers can also be shown as percentages, by dividing the new cases by the total number of people at risk. For example, the prevalence of diabetes in the United States in 2012 was about 9.3%, and the incidence was 1.7 million new diagnoses, or a rate of 540 per 100,000 people.

In this chart we can see the incidences of different diseases, going from 1960 to 2010.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diphtheria</td>
<td>3.93</td>
<td>0.51</td>
<td>0.21</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Haemophilus influenzae, invasive</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Hepatitis A</td>
<td>---</td>
<td>27.97</td>
<td>12.84</td>
<td>12.64</td>
<td>4.91</td>
<td>0.86</td>
<td>0.65</td>
<td>0.54</td>
<td>---</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>---</td>
<td>4.08</td>
<td>8.39</td>
<td>8.48</td>
<td>2.95</td>
<td>1.34</td>
<td>1.12</td>
<td>1.10</td>
<td>---</td>
</tr>
<tr>
<td>Lyme disease</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>6.53</td>
<td>11.67</td>
<td>12.71</td>
<td>9.86</td>
</tr>
<tr>
<td>Meningococcal disease</td>
<td>---</td>
<td>---</td>
<td>1.23</td>
<td>1.25</td>
<td>0.99</td>
<td>0.83</td>
<td>0.39</td>
<td>0.32</td>
<td>0.27</td>
</tr>
<tr>
<td>Mumps</td>
<td>---</td>
<td>---</td>
<td>55.55</td>
<td>3.86</td>
<td>2.17</td>
<td>0.13</td>
<td>0.15</td>
<td>0.65</td>
<td>0.85</td>
</tr>
<tr>
<td>Pertussis (whooping cough)</td>
<td>79.82</td>
<td>8.23</td>
<td>2.08</td>
<td>0.76</td>
<td>1.84</td>
<td>2.88</td>
<td>4.40</td>
<td>5.54</td>
<td>8.97</td>
</tr>
<tr>
<td>Poliomyelitis, paralytic</td>
<td>---</td>
<td>1.40</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Rocky Mountain spotted fever</td>
<td>---</td>
<td>---</td>
<td>0.19</td>
<td>0.52</td>
<td>0.26</td>
<td>0.18</td>
<td>0.85</td>
<td>0.60</td>
<td>0.65</td>
</tr>
<tr>
<td>Rubella (German measles)</td>
<td>---</td>
<td>---</td>
<td>27.75</td>
<td>1.72</td>
<td>0.45</td>
<td>0.06</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Rubella (measles)</td>
<td>---</td>
<td>211.01</td>
<td>245.42</td>
<td>23.23</td>
<td>5.96</td>
<td>11.17</td>
<td>0.03</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>Salmonellosis, excluding typhoid fever</td>
<td>---</td>
<td>---</td>
<td>3.85</td>
<td>10.84</td>
<td>14.88</td>
<td>19.54</td>
<td>14.51</td>
<td>16.92</td>
<td>16.18</td>
</tr>
<tr>
<td>Shigellosis</td>
<td>---</td>
<td>15.45</td>
<td>6.94</td>
<td>6.79</td>
<td>8.41</td>
<td>10.89</td>
<td>8.41</td>
<td>7.50</td>
<td>5.24</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>---</td>
<td>30.83</td>
<td>18.28</td>
<td>12.25</td>
<td>10.33</td>
<td>6.01</td>
<td>4.28</td>
<td>3.80</td>
<td>3.64</td>
</tr>
<tr>
<td>Sexually transmitted diseases</td>
<td>---</td>
<td>146.02</td>
<td>68.78</td>
<td>44.80</td>
<td>30.30</td>
<td>54.32</td>
<td>11.20</td>
<td>15.22</td>
<td>14.74</td>
</tr>
<tr>
<td>Syphilis</td>
<td>---</td>
<td>16.73</td>
<td>9.06</td>
<td>10.80</td>
<td>19.00</td>
<td>20.06</td>
<td>7.12</td>
<td>4.44</td>
<td>4.40</td>
</tr>
<tr>
<td>Primary and secondary</td>
<td>---</td>
<td>39.71</td>
<td>10.11</td>
<td>8.00</td>
<td>8.90</td>
<td>22.19</td>
<td>3.25</td>
<td>4.08</td>
<td>4.30</td>
</tr>
<tr>
<td>Early latent</td>
<td>---</td>
<td>70.52</td>
<td>45.91</td>
<td>24.70</td>
<td>9.20</td>
<td>10.32</td>
<td>5.53</td>
<td>6.56</td>
<td>5.70</td>
</tr>
<tr>
<td>Late and late latent</td>
<td>---</td>
<td>368.30</td>
<td>100.70</td>
<td>52.30</td>
<td>7.70</td>
<td>87.95</td>
<td>14.29</td>
<td>10.43</td>
<td>10.01</td>
</tr>
<tr>
<td>Congenital</td>
<td>---</td>
<td>192.50</td>
<td>145.60</td>
<td>294.70</td>
<td>440.10</td>
<td>276.43</td>
<td>129.67</td>
<td>110.75</td>
<td>99.05</td>
</tr>
<tr>
<td>Chlamydia</td>
<td>---</td>
<td>3.24</td>
<td>0.94</td>
<td>0.70</td>
<td>0.30</td>
<td>1.69</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Image 2: A table showing disease rates and the number of new cases per year for select years between 1950 and 2010.
What statistics exist about health care?
Health care statistics can tell us a lot about what health care is available. We can see data on patients, providers, diagnoses, medications, patient satisfaction, who has access to health care, and the quality of the care people get. Data is usually studied about either a group of patients, for example patients of a certain age or race, or about a group of hospitals in a certain location.

Here’s an example of a bar graph, tracking the care patients with colon cancer get at different types of hospitals.

Figure 2.4. Patients with colon cancer who received surgical resection of colon cancer that included at least 12 lymph nodes pathologically examined, by residence location and insurance, 2004-2009

Image 3: Two graphs, side by side, showing the percentage of patients with colon cancer who received surgical resection of colon cancer that included at least 12 lymph nodes pathologically examined. One graph shows the residence locations, and the other shows the types of insurance.
What health costs can we measure?
The fourth type of health statistic measures the **costs** of health care and poor health. Some costs are easy to count, like the price of health insurance or how much the government spends. Others are harder to put numbers to, but can still be measured, like quality of life.

This pie graph breaks down where the United States government spent money for different health care and health needs in 2010.³

![National Health Expenditures, 2010](image)

**Image 4: This pie chart shows examples of different national health expenditures from 2010. All of the items in the chart are examples of measurable health costs.**

Common Terms
To use health statistics, you need to understand some of the basic terms and concepts that are used to collect and present data. This section gives an overview of some of the basics of statistical research as it relates to health care.

At the end of this section, you will be able to take a critical look at data tables and reports, and evaluate what the numbers actually mean. We’ll look at sampling, confidence intervals, validity, dependent and independent variables, direct standardization, and age adjustment.

- **Sampling** page 9
  Take a look at how a small group of people can help us estimate information about a huge population.

- **Confidence Intervals** page 10
  See how to identify how accurate certain data is.

- **Validity** page 10
  Find out how to measure the accuracy of statistics.

- **Dependent and Independent Variables** page 10
  Identify the different ways variables can relate to the people being studied.

- **Age Adjustment** page 11
  Learn how to adapt statistics to account for age differences between communities or groups.
Sampling
How can you possibly study every single American? It’s impossible! Instead, researchers study a random sample of people. Then they create estimates about the total population. Probability sampling, where each person in the group or community has an equal chance of being chosen, is very common.

Of course, researchers also have to be aware of sampling error. Sampling error is the approximate difference between the results from a sample of people from a larger group, and the likely results of studying every single person in that group. In general, the larger the sample size is, the smaller the sampling error. But it’s impossible to know the sampling error exactly, so all sampling errors are approximate.

Image 5: A diagram showing the connection between population and sample size. From the population we take a sampling to create the sample size. Then, from the sample size, we can make inferences as to the population. Source: https://onlinecourses.science.psu.edu/stat506/node/4
Confidence Intervals
How close can studying a sample size get us to the real number? The answer lies in the confidence interval. A confidence interval is how close researchers believe their results are to reality. For example, if a study has a note that results are “50%, plus or minus 3 points,” then the confidence interval is 47-53, the 3 points below and above 50%.

Confidence intervals are usually reported to help explain how reliable, or precise, a result is. If a study is 95% reliable, with a confidence interval of 47-53, that means if researchers did the same study over and over again with samples of the whole population, they would get results between 47 and 53 exactly 95% of the time. Poor reliability can happen with a small population, or if the health event being studied doesn’t happen often, or doesn’t happen at regular times.

Validity
No study can ever be trusted unless we know that the tools we’re using to conduct a study are accurate. The validity of a tool tells us how accurately it measures what it’s supposed to measure. For example, the validity of a study that asks people to write down how much they exercise might have a low validity because people want to write down the “right” answer and not necessarily the true answer.  

Dependent and Independent Variables
When we calculate health statistics, we use two types of variables. The independent variables are the things that we think will influence dependent variables. For example, if we want to find out if people in cold climates are more likely to get the flu, the independent variable is “cold climates.” A dependent variable is the thing we believe happens as a result of the independent variable. In our example, people getting the flu is the dependent variable.

Then there are confounding variables, or confounders. A confounding variable is one that affects the relationship between the independent and dependent variables. For example, if the people in cold climates don’t sleep as much, or don’t wash their hands after sneezing, these would be confounding variables.

When running a study or analyzing statistics, researchers try to remove as many of the confounding variables as possible.
**Finding and using health statistics: A self-study course**

**Age-Adjustment**

Sometimes, health statistics are used to compare how healthy two different groups of people are, or how healthy a certain group is during two different time periods. Since older people are more likely to get ill, and younger people are more likely to injure themselves, we use an age adjustment to make sure our studies are accurate.

Age adjustment is important because age is a *confounding variable*: something that impacts the people being studied, but is not related to the health event being studied. To be able to better compare groups while adjusting for age, we use a process called **direct standardization**.

When we use direct standardization, we assume both groups have the same number of people. Then we calculate the expected number of deaths and death rates in both groups. By doing this, the two populations can be directly compared, independent on the age distribution of each group.

It’s important to remember that age adjusted rates are not the actual rates of death or disease in the population – those are called “crude rates.” Age adjusted rates are *only* useful for comparisons to other populations.

<table>
<thead>
<tr>
<th>Age</th>
<th>White</th>
<th></th>
<th>Black</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rate per 1,000</td>
<td>Population (standard)</td>
<td>Expected no. of deaths</td>
<td>Rate per 1,000</td>
</tr>
<tr>
<td>&lt;1 yr</td>
<td>13.3</td>
<td>15,000</td>
<td>200</td>
<td>19.7</td>
</tr>
<tr>
<td>1-4 yr</td>
<td>0.8</td>
<td>50,000</td>
<td>40</td>
<td>1.7</td>
</tr>
<tr>
<td>5-17 yr</td>
<td>0.5</td>
<td>150,000</td>
<td>75</td>
<td>0.6</td>
</tr>
<tr>
<td>18-44 yr</td>
<td>2.5</td>
<td>325,000</td>
<td>800</td>
<td>3.7</td>
</tr>
<tr>
<td>45-64 yr</td>
<td>18.0</td>
<td>210,000</td>
<td>3,780</td>
<td>25.0</td>
</tr>
<tr>
<td>&gt;65 yr</td>
<td>70.0</td>
<td>200,000</td>
<td>14,000</td>
<td>77</td>
</tr>
<tr>
<td>Total (age-adjusted rate)</td>
<td><strong>24.8</strong></td>
<td><strong>761,000</strong></td>
<td><strong>18,895</strong></td>
<td><strong>29.3</strong></td>
</tr>
</tbody>
</table>

*Image 6: Expected No. of Deaths in Standardized Populations and Age-Adjusted Death Rates*  
This table shows the expected death rate and number of deaths in a standardized population, which provides the age-adjusted death rates.
Sources of Health Statistics
We've looked at what health statistics are, and what some of the common terms for statistics are. Now we'll learn how health and medical information is collected, and where it comes from.

The main sources we’ll review are surveys, administrative and medical records, claims data, vital records, surveillance, disease registries, and peer-reviewed literature. We’ll take a look into each source, and the pros and cons of using it to create health statistics.

- **Surveys** page 13
  Check out the many types of surveys available in the medical community.
- **Medical Records** page 14
  See how medical records help us track information about patients and providers.
- **Claims Data** page 16
  Find out how electronic records can be collected on a large scale.
- **Vital Records** page 18
  Learn about the birth and dead records provide researchers with information on fatal illnesses.
- **Surveillance** page 19
  See how information gathered from state and local governments can help track and control infectious diseases.
- **Peer-Reviewed Literature** page 20
  Find out why peer-reviewed literature is such a high-quality source of data.
Surveys

Surveys are used to collect data from a group of people. Information on a survey might be about patients, providers, or hospitals and doctor’s offices. There are two main types of surveys: population surveys, which are often conducted by mail, telephone, in-person or online interviews, and provider surveys, which are actually a combination of interviews and reviewing existing provider data.

We’ll look at a few of the survey methods that the National Center for Health Statistics and other agencies use.5

- The National Health Interview Survey is an example of an interview-based population survey. Researchers interview people in their homes to learn about how they use health care, insurance, their access to care, and other topics.
- The National Health and Nutrition Examination Survey is another population survey. This survey covers topics like disease conditions, child growth and development, illnesses such as diabetes, hypertension, and cholesterol, and nutrition. This survey uses a mix of personal interviews, physical examinations, and lab tests.
- The National Ambulatory Medical Care Survey is a provider survey. To complete this survey, researchers interview physicians and visit medical centers to learn about patient demographics, diagnoses, provider specialties, and how they use electronic medical records.

Surveys are great because they can be done via mail, email, phone, or in person. They’re not very expensive, and they can help get information from a large sample of the population. However, it can be hard to get detailed information in a survey, and sometimes people choose not to answer difficult questions, or they can’t remember important details. There’s also a bias in surveys. People who don’t like to respond to surveys aren’t included, and people who are homeless or don’t have phones or mailing addresses can’t be included.

This chart was created using data from the National Ambulatory Medical Care Survey6.

---

5 http://www.amstat.org/sections/srms/pamphlet.pdf
Medical Records

Medical records are used to track events and transactions between patients and health care providers. They offer information on diagnoses, procedures, lab tests, and other services. Medical records help us measure and analyze trends in health care use, patient characteristics, and quality of care.

Electronic health records (EHR) is a new type of data collection. It was first introduced in the 1960s, but only became popular recently, in part due to the American Recovery and Reinvestment Act and the Affordable Care Act. EHRs can make it easier for providers to enter information about patients. The data from EHRs is then used for things like comparing how effective providers are, and seeing how patients respond to treatment.

Medical records are helpful in research because they come from the providers, so they’re usually accurate and detailed. The data is automatically collected, so it’s easy and inexpensive to find, and it includes information that patients might...
not think to add or feel comfortable sharing. But, since the information is written
down in a specific context, it can be misinterpreted if taken out of context. And of
course, medical records are (by definition) only available for people who are able
to get medical care.

This chart shows statistics based on information from patient medical records.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>38.2</td>
<td>21.8</td>
<td>53.9</td>
<td>9.0</td>
<td>14.8</td>
<td>5.3</td>
</tr>
<tr>
<td>Asthma</td>
<td>14.9</td>
<td>18.5</td>
<td>36.2</td>
<td>11.4</td>
<td>17.9</td>
<td>7.2</td>
</tr>
<tr>
<td>Digestive system disorders</td>
<td>114.7</td>
<td>29.2</td>
<td>38.1</td>
<td>25.7</td>
<td>*</td>
<td>28.6</td>
</tr>
</tbody>
</table>

* signifies unreliable data. Source: Centers for Disease Control and Prevention. National Center for Health Statistics. Health Data

Image 8: This chart shows the hospital discharges by first-listed diagnosis among children in the United States from 1990-2010. This information is gained from medical records. * signifies unreliable data. Source: Centers for Disease Control and Prevention. National Center for Health Statistics. Health Data
Claims Data

Claims data, also known as administrative data, is another sort of electronic record, but on a much bigger scale. Claims databases collect information on millions of doctors’ appointments, bills, insurance information, and other patient-provider communications.

The good thing about claims data is that it, like other medical records, comes directly from notes made by the health care provider, and happens at the time patient sees the doctor. Also, because of the large sample size of claims data, researchers can analyze groups of patients with rare illnesses and medical conditions. The downside to using claims data is that it is sometimes problematic. The validity can be low due to certain illegal billing practices, like ordering unnecessary tests or billing for things that weren’t provided.

This table was generated using data from the CMS Chronic Conditions Public Use Files. 7

---


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
Vital Records

Vital records are collected by the National Vital Statistics System, and are maintained by state and local governments. Vital records include births, deaths, marriages, divorces, and fetal deaths. They also record information about the cause of death, or details of the birth.

Vital records are useful because their offer very detailed information. They can help us learn about rare disorders that end in death. Unfortunately, because there are so many state and local governments, the records can be inconsistent. Also, vital records are only able to provide information on diseases and illnesses that end in death.

Surveillance
Public health surveillance is a type of record collection that helps us study infectious diseases. The CDC, WHO, and other institutions collect, analyze, and interpret data that they collect from local and state health departments. They use this information to prevent and control the spread of disease.

Disease registries are another type of public health surveillance. Registries are systems that allow people to collect, store, retrieve, analyze, and disseminate information about people with a specific disease or condition. Disease registries let researchers estimate how large a health problem is, determine the incidence of the disease, study trends over time, and evaluate the effects of certain environmental exposures. Registries are kept by governments, hospitals, universities, non-profits, and private groups. They store data from hospital records, lab reports, and other sources.

Surveillance data has a higher validity than surveys, because the data comes from lab tests, diagnoses, and other patient records. Registries also make this data easy to store and analyze. The downside to surveillance data is that, because diseases sometimes change definitions, it makes it difficult to accurately track trends. Also, if hospitals or doctors don’t report some information, there’s no way to access that data.
Peer-reviewed literature

Peer-reviewed journals publish statistics, diagrams, and tables as part of their articles. The research that appears in these journals is reviewed to make sure it is high quality and offers scientific merit. Researchers who publish may have collected their own data using any of the sources we have already reviewed.

Because the research has been reviewed, it's high quality and has high validity. It's also easy to find this data, thanks to online databases. However, it can sometimes be so detailed and complex, it is difficult to understand. It can also often require researchers to subscribe to an expensive journal or database to access the information.
Finding Health Statistics

Are you ready to become a researcher? Now that you know the different types of health statistics and how to find them, let’s take a look at some of the major organizations and government agencies that provide health statistics. That way, you’ll be familiar with some of the resources that can help you find the exact information you’re looking for.

Find the Right Source

- National Center for Health Statistics (NCHS)
  The NCHS has a mission to provide statistical data to improve the health of the American people.
- World Health Organization (WHO)
  The WHO is a part of the United Nations. They act as leaders in worldwide health situations.
- Agency for Healthcare Research and Quality
  The AHRQ is an agency in the United States, dedicated to making American health care higher quality, safer, more efficient, and more effective.
- Centers for Disease Control and Prevention
  The CDC is the primary public health institute of the United States. They provide data on health care and disease.
- Robert Wood Johnson Foundation County Health Rankings
  The Robert Wood Johnson Foundation County Health Rankings is an online application with the health status of almost every county in the United States.
- Centers for Medicare and Medicaid Services
  The CMS is a US government agency that works with individual states to provide Medicare and Medicaid benefits.
- Kaiser Family Foundation
  The Kaiser Family Foundation is a non-profit that offers research, journalism, and communications programs with a focus on health care.
- United States Census Bureau
  The Census Bureau is the best source of current data and statistics on the US population and economy.
- HealthData.gov
  HealthData.gov is an online source for health-related data. It is constantly being updated, and includes thousands of data sets that can be downloaded.
- Dartmouth Atlas of Health Care
  Website with data and resources about health care markets, divided into geographic areas.
National Center for Health Statistics (NCHS)
http://www.cdc.gov/nchs/

The National Center for Health Statistics has a mission to provide statistics and data that can guide public policies and actions. Their goal is to improve the health of Americans. They are the United State’s principal health statistics agency.8

The NCHS website provides access to many health statistics sources, from published reports, to, data briefs on specific topics, and public use data files. For example:

1. Health, United States is a report on national trends in health statistics. It includes detailed tables and charts on selected measures of morbidity, mortality, health status, risk factors, and health care use, among other topics.

2. FastStats is a site within the NCHS website that offers quick and easy access to statistics on specific health topics, from diseases and conditions to health care and insurance.

3. Health Data Interactive is a web-based application giving access to pre-tabulated national data on a vast array of public health indicators over time. This tool allows users to customize tables by age, gender, race/ethnicity, and geographic location.

4. The National Vital Statistics System provides access to the Nation’s official vital statistics, including births, deaths, marriages, and divorces. The webpage features links to statistical reports, an interactive data query tool, and downloadable data files.

World Health Organization (WHO)
http://www.who.int/en/

The World Health Organization (WHO) is an agency within the United Nations. They provide leadership on global health issues. WHO collects and compiles a wide range of data used to monitor and assess health trends worldwide. For example,

- The Global Health Observatory Data Repository contains a list of indicators on priority health topics, including morbidity and mortality

---


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
estimates, child health and nutrition, immunization, health systems, environmental health, HIV/AIDS, tuberculosis, and malaria, and others.

- The World Health Statistics Report, a yearly publication of the WHO, presents the most recent health statistics of its 194 Member States. It provides detailed charts, tables, and figures on nine areas:
  - life expectancy and mortality
  - cause-specific mortality and morbidity
  - selected infectious diseases
  - health service coverage
  - risk factors
  - health systems
  - health expenditures
  - health inequities
  - demographic and socioeconomic statistics

Agency for Healthcare Research and Quality  
http://www.ahrq.gov/

The Agency for Healthcare Research and Quality (AHRQ)’s mission is to make American healthcare higher quality, safer, more efficient, and more effective. AHRQ supports research that helps people make better, more informed decisions, and improve the quality of health care services. For example,

- The Healthcare Cost and Utilization Project (HCUP) is a family of databases and related software tools with the largest collection of nationwide and state-specific hospital care data. Users can create their own tables and statistics using a free system called HCUPnet. They can also see publications and statistical briefs using HCUP reports.

- The Medical Expenditure Panel Survey (MEPS) gives access to data on the cost and use of health care and health insurance coverage. The information is based on surveys of families and individuals, their medical providers, and employers across the country. Users can create summary data tables, download public use data files, or browse through MEPS reports and chart books.

- The National Healthcare Quality Report and the National Healthcare Disparities Report are yearly publications of AHRQ that track trends in care effectiveness, patient centeredness, timeliness of care, patient safety, and efficiency of care. The National Healthcare Disparities Report has information on health care quality and access for racial, ethnic, and income groups, along with other populations.
Centers for Disease Control and Prevention  
http://www.cdc.gov/  

The Centers for Disease Control and Prevention (CDC) is the primary public health institution in the country. The CDC offers many data resources and tools. These include CDC Wonder and the Behavioral Factor Surveillance System.

- **CDC WONDER** (Wide-ranging Online Data for Epidemiologic Research) is a system of searchable databases with access to a wide array of public health indicators. These include measures of chronic and communicable disease, environmental health, disease and injury prevention, and occupational health.

- The **Behavioral Risk Factor Surveillance System** (BRFSS) is a system of telephone surveys about health care, health risks, chronic conditions, and use of preventive services. The BRFSS collects data in all 50 states, Washington, DC, and three U.S. territories. It is the largest continuously conducted health survey system in the world.  

- The CDC website also has a list of interactive database systems with the latest information on many health and disease topics.

Robert Wood Johnson Foundation County Health Rankings  
The Robert Wood Johnson Foundation *County Health Rankings* is an interactive online application. It ranks the overall health status of almost every U.S. county. It offers access to measures of health outcomes like mortality and morbidity, and their determinants, such as health behaviors, clinical care, social and economic factors, and physical environment.

Centers for Medicare and Medicaid Services  
http://www.countyhealthrankings.org/  

The Centers for Medicare and Medicaid Services (CMS) is the U.S. federal agency that works with state governments to manage the Medicare program, and administer Medicaid and the Children’s Health Insurance program. They offer many great resources for researchers who are looking for health data. For example:

- **CMS Statistics** is a yearly reference booklet that people can download on the CMS website. It has summary information about health care expenses

---

9 [http://www.cdc.gov/brfss/about/brfss_faq.htm#1](http://www.cdc.gov/brfss/about/brfss_faq.htm#1)

This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
Finding and using health statistics: A self-study course

and use.

- The Medicare and Medicaid Statistical Supplement has detailed statistics on Medicare, Medicaid, and other CMS programs. It has 115 tables and 67 charts that detail health expenditures for the entire U.S. population.¹⁰

- Hospital Compare is an online tool created by CMS that helps users find information about the quality of care at over 4,000 Medicare-certified hospitals across the United States.

- The CMS Data Navigator lets users search across all CMS programs using a menu-driven search application. Users can specify a particular type of data, or search for all available data types.

Kaiser Family Foundation
http://kff.org/

Kaiser is a non-profit foundations that does research, journalism, and communications programs. They focus on major health care issues in the United States. Their goal is to be an unbiased source of facts, information, analysis, and journalism for the general public, as well as for policy makers, the media, and the health care community.¹¹ Their resources include:

1. State Health Facts has more than 800 health indicators for all 50 states, Washington, D.C., United States Territories, and other locations. Users can view maps, relative health rankings, trends, and download data on demographics, health costs, health coverage, minority health, providers and service use, and more.

2. Global Health Facts has more than 100 indicators on HIV/AIDS, tuberculosis, malaria and other key measures of health and socioeconomic status, divided by country.

United States Census Bureau
http://www.census.gov/

The Census Bureau is the best source for current data and statistics about the United States’ population and economy. The Census Bureau collects and gives

¹¹ http://kff.org/about-us/
out information about health insurance coverage for adults and children. Their resources include:

- The **American FactFinder**, an interactive online application that provides statistics from many Census Bureau data sources. This includes the Economic Census, the American Community Survey, and the 2010 Census.

- The Census Bureau website, which also hosts a [Health Insurance page](http://healthdata.gov/) with statistical reports, tables, and other data on health insurance coverage and medical services utilization.

**HealthData.gov**
http://healthdata.gov/

HealthData.gov brings together high-value datasets, tools, and applications. They offer data about health and health care to help researchers solve problems. They have thousands of health-related data sets from Health and Human Services agencies that can be downloaded, and they are always adding more.

**Dartmouth Atlas of Health Care**
http://www.dartmouthatlas.org/

The Dartmouth Atlas of Health Care was created in 1993 as a study of health care markets. Their website now has data tables, maps, publications, and other resources. They focus on variations in health care resources and how they are used in different geographic regions.

---

12 [http://www.hhs.gov/open/discussion/welcome_healthdata.html](http://www.hhs.gov/open/discussion/welcome_healthdata.html)

*This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.*
Search Strategies

Search by Organization or Resource

1. Identify the type of data you need to answer your research question. Do you need:
   a. A numerical statistic or fact, such as prevalence rates, mortality rates, or other health indicators?
   b. A complete data set for independent analysis?
   c. A chart or figure?

2. What is the geographic scope of your data needs?
   a. Are you looking for county, state, nationwide, or global health data?

3. Identify one or more organizations/agencies that are likely to collect or share the information you are interested in. (ex: The National Center for Health Statistics compiles statistics on virtually all aspects of health and health care, from chronic conditions and risk factors, to insurance and utilization. This is often a good place to start for U.S. health statistics.)
   a. Look for a tab or link that says “Data”, “Statistics”, “Research”, etc. on the organization’s homepage.

4. Alternatively, identify a specific resource (such as any of those listed in this section) that answers your particular research need. (ex: You want to compare various measures of health status between several states. Visit Kaiser’s State Health Facts.)

Use a Search Engine

Once you have clearly identified your specific data needs, use a search engine (Google, Yahoo!, etc.) to find the desired statistics. In order to locate the most relevant information, make sure to enter detailed search terms (what, where, when).

For example, you are writing a report on the magnitude of the diabetes epidemic in the United States. Below are some examples of possible search terms:

- “Diabetes mortality rate United States”
- “Diabetes prevalence United States”
- “Diabetes hospitalizations United States 2012”
Search an Academic Database

If you are in search of more complex statistical analyses for a particular health topic, enter detailed search terms into an academic database (described in Module III). Examples include:

- PubMed
- ScienceDirect
- Google Scholar
Finding and using health statistics: A self-study course

Glossary
Unless otherwise noted, the source of the definitions is the text of the tutorial.

Access: The ability to get needed medical care and services.

Affordable Care Act (ACA): The health care reform law enacted in March 2010 in two parts: the Patient Protection and Affordable Care Act and the Health Care and Education Reconciliation Act.13

Age Adjustment: A process of direct standardization designed to account for age as a confounding variable.

Age-Specific Mortality Rates: The number of deaths of residents of a specified age in a specified geographic area, divided by the population.14

Agency for Healthcare Research and Quality (AHRQ): An operating division of the U.S. Department of Health and Human Services (HHS) that supports research that helps people make better, more informed decisions and improve the quality of health care services in the U.S.

American FactFinder: An interactive online application from the U.S. Census Bureau that provides statistics from many Census Bureau data sources. The Census Bureau is a bureau of the U.S. Department of Commerce.

Behavioral Risk Factor Surveillance System (BRFSS): A system of telephone surveys about health care, health risks, chronic conditions, and use of preventive services from the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS). The BRFSS collects data in all 50 states, Washington, D.C., and three U.S. territories. It is the largest continuously conducted health survey system in the world.15

CDC WONDER (Wide-ranging Online Data for Epidemiologic Research): A system of searchable databases with access to a wide array of public health indicators, including measures of chronic and communicable disease, environmental health, disease and injury prevention, and occupational health. CDC WONDER is a resource of the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS).

Centers for Disease Control and Prevention (CDC): The primary federal public health agency in the U.S., offering many data resources and tools to state and local health departments and the general public. CDC is an operating division of the U.S. Department of Health and Human Services (HHS).

Centers for Medicare & Medicaid Services (CMS): The federal agency that administers Medicare and collaborates with states to administer Medicaid, the Children’s Health Insurance Program (CHIP), and the Health Insurance Marketplace. CMS is an operating division of the U.S. Department of Health and Human Services (HHS).

Chronic Conditions Public Use Files (PUF): A resource of the Centers for Medicare & Medicaid Services (CMS), an operating division of the U.S. Department of Health and Human Services (HHS). Chronic Conditions PUFs contain information from Medicare claims. Each record is a profile defined by the characteristics of Medicare beneficiaries.¹⁶

Claim: A request for payment for services and benefits received.¹⁷

CMS Data Navigator: Lets users search across all Centers for Medicare and Medicaid Services (CMS) programs using a menu-driven search application. Users can specify a particular type of data, or search for all available data types. CMS is an operating division of the U.S. Department of Health and Human Services (HHS).

CMS Statistics: A yearly reference booklet that people can download on the Centers for Medicare and Medicaid Services (CMS) website. It has summary information about health care expenses and use. CMS is an operating division of the U.S. Department of Health and Human Services (HHS).

Confidence Interval: Describes the amount of uncertainty associated with a sampling method. Confidence intervals are usually reported to help explain how reliable, or precise, a result is.

Confounding Variable (or Confounder): Something that impacts the people or condition being studied, but is not related to the health event being studied. Confounding variables affect the relationship between independent and dependent variables.

Dartmouth Atlas of Healthcare: A website created in 1993 as a study of health care markets which has data tables, maps, publications, and other resources. The data focus on variations in health care resources and how they are used in different geographic


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
regions. The Dartmouth Atlas of Healthcare is a resource of the Dartmouth Institute for Health Policy and Clinical Practice at Dartmouth University in New Hampshire.

**Dependent Variable(s):** The response or factors that are measured during a study. The dependent variable responds to the independent variable.

**Direct Standardization:** An adjustment procedure used during data analysis designed to remove an extraneous source of variation to allow for comparison between groups with unlike characteristics.\(^\text{18}\) In direct adjustment, a common structured population is used as standard so as to compare two populations.\(^\text{19}\)

**Electronic Health Records (EHRs):** Electronic versions of patients’ health information including medical history, diagnoses, medications, and other information that can be shared among providers for collaborative treatment, quality improvement, health planning, and clinical research.

**FastStats:** A website within the website of the National Center for Health Statistics’ (NCHS) of the Centers for Disease Control and Prevention (CDC) that offers quick and easy access to statistics on specific health topics, from diseases and conditions to health care and insurance. The CDC is an operating division of the U.S. Department of Health and Human Services (HHS).

**Global Health Facts:** From the Kaiser Family Foundation provides data on more than 100 indicators on HIV/AIDS, tuberculosis, malaria and other key measures of health and socioeconomic status, organized by country. The Kaiser Family Foundation is a non-profit foundation that runs programs in policy analysis, journalism, and communications.

**Global Health Observatory (GHO) Data Repository:** A data repository that contains a list of global indicators on priority health topics, including morbidity and mortality estimates, child health and nutrition, immunization, health systems, environmental health, HIV/AIDS, tuberculosis, and malaria, and more. The GHO Data Repository is a resource of the World Health Organization (WHO), an intergovernmental organization in the United Nations (UN) system.


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
Health Correlates: The risk factors that impact health, including social and economic factors, such as income and education; physical factors like air pollution; and personal behaviors like smoking, exercise, drugs, and alcohol.20

Health Data Interactive: A web-based application from the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS). Health Data Interactive gives access to pre-tabulated national data on a vast array of public health indicators over time.

Health Insurance Page, United States Census Bureau: The U.S. Census Bureau website hosts a Health Insurance page with statistical reports, tables, and other data on health insurance coverage and medical services utilization. The U.S. Census Bureau is a bureau of the U.S. Department of Commerce.

Health, United States: A report from the National Center for Health Statistics (NCHS) on national trends in health statistics in the U.S. Health, United States includes detailed tables and charts on selected measures of morbidity, mortality, health status, risk factors, and health care use, among other topics. NCHS is a part of the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS).

Healthcare Cost and Utilization Project (HCUP): A family of databases and related software tools with the largest collection of nationwide and state-specific hospital care data from the Agency for Healthcare Research and Quality (AHRQ), an operating division of the U.S. Department of Health and Human Services (HHS).

Healthdata.gov: A website managed by the Department of Health and Human Services (HHS) that brings together high-value datasets, tools, and applications, offering data about health and health care to help researchers solve problems.21

Hospital Compare: An online tool created by the Centers for Medicare and Medicaid Services (CMS) that helps users find information about the quality of care at over 4,000 Medicare-certified hospitals across the U.S. CMS is an operating division of the U.S. Department of Health and Human Services (HHS).

Incidence: The number of new cases of an illness.


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
**Independent Variable(s):** A variable that stands alone and isn’t affected by other variables that the study is trying to understand. Independent variables influence dependent variables.

**Indirect Standardization:** An adjustment procedure during data analysis designed to remove an extraneous source of variation to allow for comparison between groups with unlike characteristics. In indirect adjustment, a common set of specific rates is applied to the populations whose rates are to be standardized. A simple form of indirect adjustment is the standardized mortality ratio (SMR).

**Kaiser Family Foundation (KFF):** A non-profit foundation that runs programs in policy analysis, journalism, and communications. They focus on major health care issues in the U.S. Their goal is to be an unbiased source of facts, information, analysis, and health journalism for the general public, as well as for policy makers, the media, and the health care community.

**Medical Expenditure Panel Survey (MEPS):** A survey from the Agency for Healthcare Research and Quality (AHRQ), an operating division of the U.S. Department of Health and Human Services (HHS), that gives access to data on the cost and use of health care and health insurance coverage. The information is based on surveys of families and individuals, their medical providers, and employers across the country.

**Medicare and Medicaid Statistical Supplement:** Has detailed statistics on Medicare, Medicaid, and other Centers for Medicare and Medicaid Services (CMS) programs. The Supplement has 115 tables and 67 charts that detail health expenditures for the entire U.S. population. CMS is an operating division of the U.S. Department of Health and Human Services (HHS).

**National Ambulatory Medical Care Survey (NAMCS):** A national survey conducted by the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS), that is designed to meet the need for objective, reliable information about the provision and use of ambulatory medical care.

---


*This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.*
services in the U.S. Findings are based on a sample of visits to non-federal employed office-based physicians who are primarily engaged in direct patient care.26

**National Center for Health Statistics (NCHS):** A Center of the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS), which has a mission to provide statistics and data to guide public policies and actions.27

**National Health and Nutrition Examination Survey (NHANES):** A program of studies designed to assess the health and nutritional status of adults and children in the U.S. The survey is unique in that it combines interviews and physical examinations and is conducted by the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS).28

**National Health Interview Survey (NHIS):** A survey from that collects data on a broad range of health topics through personal household interviews. For over 50 years, the U.S. Census Bureau has been the data collection agent for the NHIS.29 The U.S. Census Bureau is a bureau of the U.S. Department of Commerce.

**National Healthcare Quality and Disparities Reports:** Yearly publications of the Agency for Healthcare Research and Quality (AHRQ), an operating division of the U.S. Department of Health and Human Services (HHS), that track trends in care effectiveness, patient centeredness, timeliness of care, patient safety, efficiency of care, and health equity.

**National Survey of Children’s Health (NSCH):** A survey sponsored by the Health Resources and Services Administration (HRSA), an operating division of the U.S. Department of Health and Human Services (HHS) that touches on multiple, intersecting aspects of children’s lives. The survey includes physical and mental health status, access to quality health care, as well as information on the child’s family, neighborhood and social context.30

---


This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.
**National Vital Statistics System:** Provides access to the official vital statistics of the U.S., including births, deaths, marriages, and divorces. The webpage features links to statistical reports, an interactive data query tool, and downloadable data files. The National Vital Statistics System is a resource of the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS).

**Population:** In research, the population is the entire set of individuals that are of interest to the researcher.

**Precision (or Reliability):** A measure of how often researchers would get the same answer if they conducted the same study over and over.

**Prevalence:** The total number of people with a certain illness at a specific time.

**Public Health Surveillance:** A type of record collection that helps us study infectious diseases. The Centers for Disease Control and Prevention (CDC), World Health Organization (WHO), and other institutions collect, analyze, and interpret data that they collect from local and state health departments. They use this information to prevent and control the spread of disease.

**Quality:** How well a health plan keeps its members healthy or treats them when they are sick. Good quality health care means doing the right thing at the right time, in the right way, for the right person and getting the best possible results.  


Finding and using health statistics: A self-study course

The Robert Wood Johnson Foundation County Health Rankings: An interactive online application from the Robert Wood Johnson Foundation. It ranks the overall health status of almost every U.S. County. It offers access to measures of health outcomes like mortality and morbidity, and their determinants, such as health behaviors, clinical care, social and economic factors, and physical environment.

Sample: A subset of the population that is actually used in research. One common method for selecting a sample is called probability sampling. In probability sampling, each person in the group or community has an equal chance (probability) of being chosen.

Sampling Error: The approximate difference between the results from a sample of people from a larger group, and the likely results of studying every single person in that group. In general, the larger the sample size is, the smaller the sampling error.

State Health Facts: From the Kaiser Family Foundation, State Health Facts has more than 800 health indicators for all 50 states, Washington, D.C., U.S. Territories, and other locations. Users can view maps, relative health rankings, trends, and download data on demographics, health costs, health coverage, minority health, providers and service use, and more.

Statistical Validity (or Accuracy): Refers to the soundness of research design and methods, and whether researchers are actually measuring what they intend to study.

United States Census Bureau: The best source for current data and statistics about the population and economy of the U.S. The Census Bureau collects and gives out information about health insurance coverage for adults and children. The Census Bureau is a bureau of the U.S. Department of Commerce.

Vital records: Include births, deaths, marriages, divorces, and fetal deaths. They also record information about the cause of death, or details of the birth. Vital Records are collected by the National Vital Statistics System within the National Center for Health Statistics (NCHS), Center for Disease Control and Prevention (CDC), an operating division of the U.S. Department of Health and Human Services (HHS). Vital records data are provided through contracts between NCHS and vital registration systems operated in the various jurisdictions legally responsible for the registration of vital events.

World Health Organization (WHO): An agency within the United Nations (UN) that provides leadership on global health issues. WHO collects, compiles, and disseminates a wide range of data used to monitor and assess health trends worldwide.
**World Health Statistics Report:** A yearly publication of the World Health Organization (WHO) of the United Nations (UN) which presents the most recent health statistics of its 194 Member States.

**Youth Risk Behavior Surveillance System (YRBSS):** From the Centers for Disease Control and Prevention (CDC) which monitors six types of health-risk behaviors that contribute to the leading causes of death and disability among youth and adults. The CDC is an operating division of the U.S. Department of Health and Human Services (HHS).³⁴

---


*This work was conducted by AcademyHealth for the National Information Center for Health Services Research (NICHSR) at the National Library of Medicine (NLM), under the auspices of the Health Services Research Projects in Progress (HSRProj) Professional and Research Support contract # HHSN276201400014C.*