Five Criteria for Good Data

Data have the power to transform the way we see the world, from identifying health disparities to strengthening a case for policy change. When determining the best data to support your work, consider the five criteria below: correlation versus causation, credibility, reliability, generalizability and timeliness. Remember, no data are perfect. Use your best judgment and be transparent.

1. **Correlation Versus Causation**

Correlation refers to two findings that are associated. Causation refers to one finding causing another.

- Without statistical testing, do not assume how data are related. For example, number of substantiated abuse and neglect cases and days with ozone levels above standards have both decreased since 1998. They have the same pattern but are not likely related and one does not likely cause the other. As with any two sets of indicators, you cannot make conclusions about a pattern without testing.

*Kidsdata.org does not provide statistical testing to assess how indicators are related. Avoid claims of correlation or causation without statistical testing.*

2. **Credibility**

Credibility refers to the source of the data or who provides the data. Can you trust the entity that produced the data?

- Who paid for, sponsored, or funded the study? Could the data be biased?
- Does the data provider have a stake in a specific finding? Research sponsored by business, religious or political organizations may have missions that influence how they conduct research and interpret findings.
- What is the data provider’s reputation for research? Government and academic institutions are considered credible because research is conducted for the public benefit.

*Kidsdata.org only contains data from credible sources.*

3. **Reliability**

Reliability refers to the accuracy of the data. Can you trust the data?

- Has the research that produced the data been reproduced by other researchers?
- How were the data collected? Did the researchers adhere to ethical research methods?
- If the data come from a survey, is there response bias? For example, did researchers conduct their survey in different languages if they need information about immigrants?

*Kidsdata.org provides only reliable data. See “Data Source” under the “Definition, Source & Notes” section to find more information about data collection methods from the data providers.*

4. **Generalizability**

Generalizability refers to data on a specific population that can be used for other populations.

- Generalizability depends on the way in which the data provider collected the data.
- Understand “who, what, why, when and where” of the data and consider whether the data can apply more broadly. For example, if the data describe Hispanic children, could they also describe Latin American children?
- Be cautious about claiming generalizability and be clear about how populations differ.

Assess data generalizability in kidsdata.org by reviewing the “Definition, Source & Notes” and “Measures of” sections for each indicator.
Five Criteria for Good Data

5 Timeliness

Timeliness refers to when the research was conducted relative to changes in the environment.

- When was the study done - one year ago, three years ago, or over 10 years ago? And, how fast are changes occurring - months or years? Some data may be relevant over a longer period of time than others.
- Often there will be a lag time, especially with big studies such as the American Community Survey (ACS). Most comprehensive surveys will be a few years old by the time findings are published.
- Even if the research seems old, it may be the best source if more recent data are not available. Admit the limitations of the data and supplement it with other closely related research.

Kidsdata.org displays time periods for each indicator above the figures as well as in “Data Source” under the “Definition, Source & Notes” section.