

# Basic Data Interpretation

## Epidemiology 101

**Epidemiologists**

# Summary Measures

- Ratios
- Proportions
- Rates

# Ratio

- Comparison of any two numbers
- Calculated by dividing one quantity by the other
- The numerator and denominator are separate and distinct quantities (neither is included in the other)

# Ratio Example

- What is the ratio of abortions to live births in Montana?
  - 1,842 induced abortions
  - 12,326 live births
  - Ratio =  $1,842/12,326 = .1494$
  - Montana's abortion ratio is 149.4 abortions per 1,000 live births

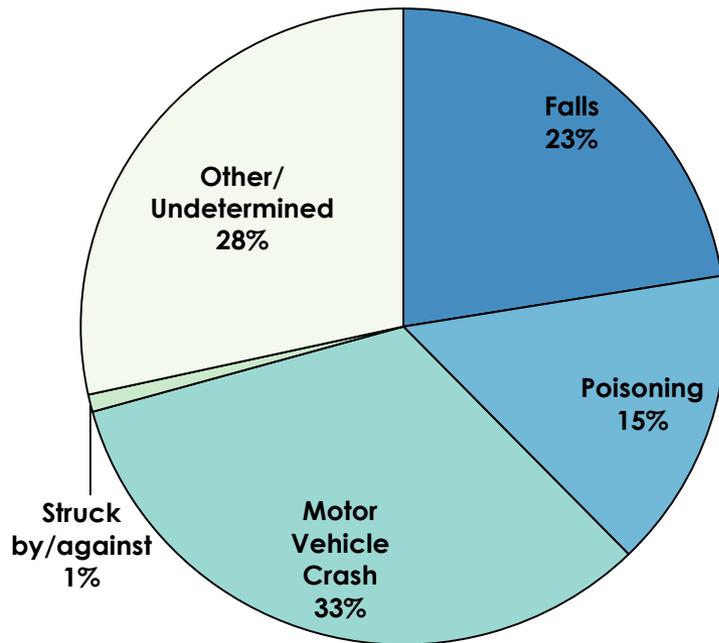


# Proportion

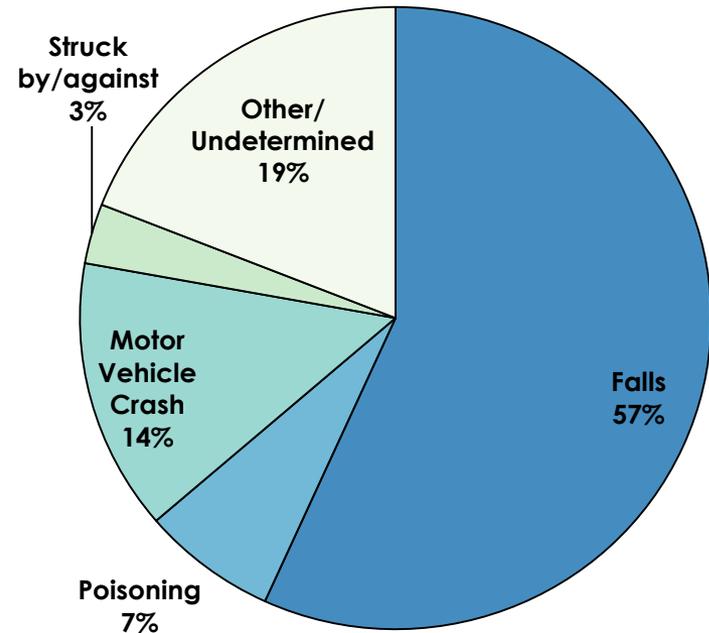
- Comparing two numbers
- Numerator is part of the denominator
- Styles
  - Percent: **10%**
  - Proportion: **.10**
  - Per a constant: **10 per 100**

# Causes of Unintentional Injuries in Montana, 2010-2013

## Injury Deaths



## Injury Hospitalizations



# Rates

- A rate measures the number of events that occur in a defined population, ***with respect to some unit of time.***

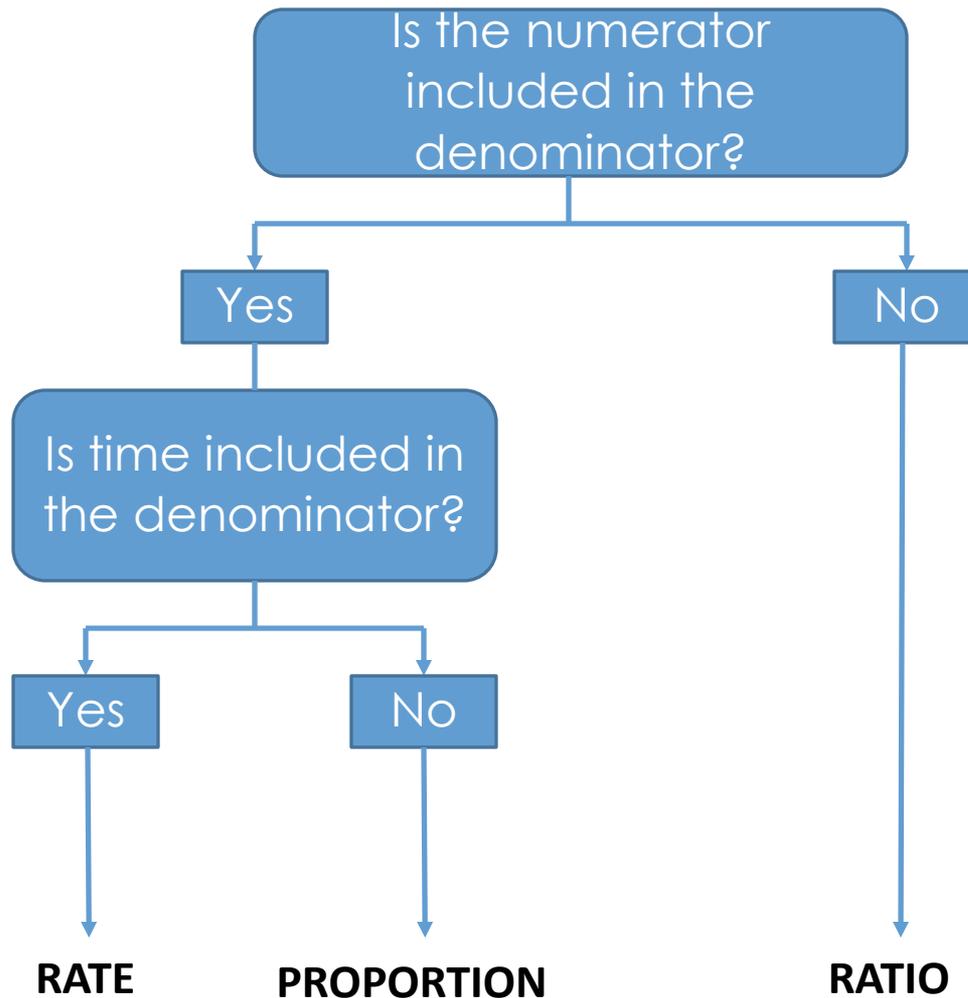
# Crude Rates

- Numerator: count of events in a time period
- Denominator: population at risk during that same time period

# Crude Rates

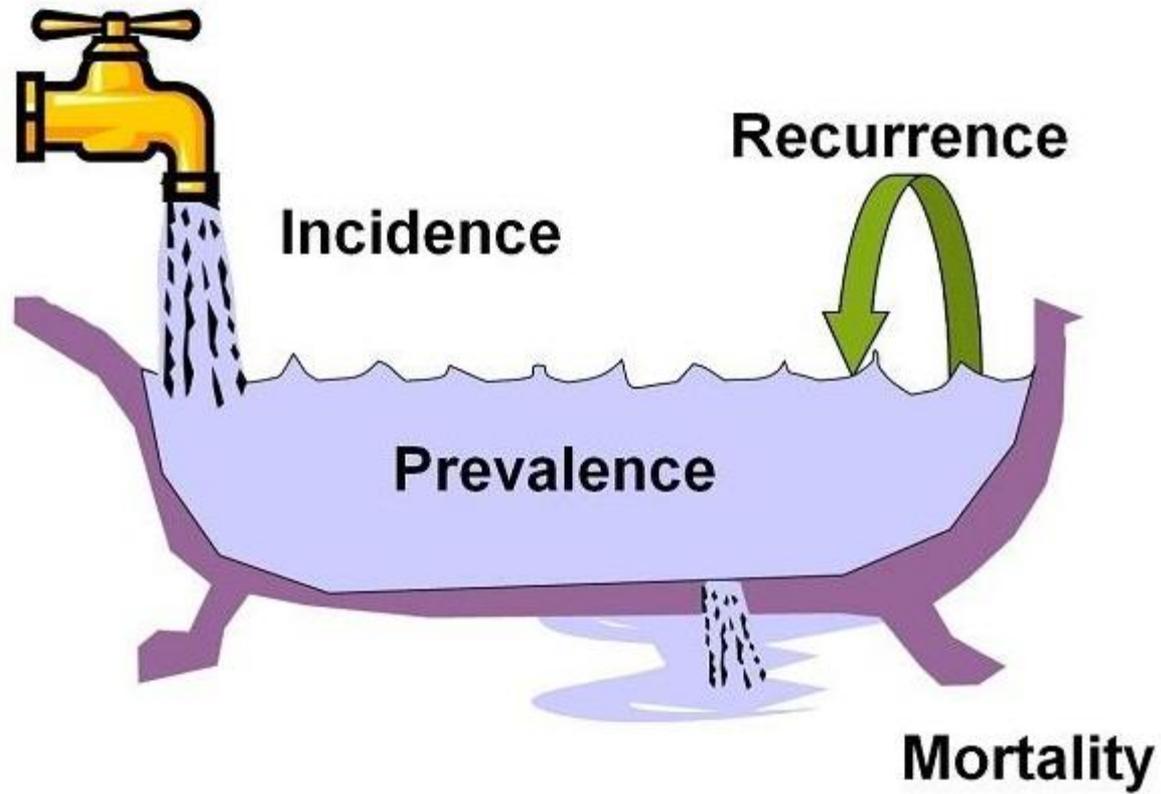
- Asthma Emergency Department (ED) Visits in Montana in 2013
  - Count: 2,188 cases
  - Population at Risk: 84,478 people with asthma
  - Constant: 10,000
- Rate =  $\frac{2188}{84478}$  times 10,000
  - = 259.0 ED visits
  - per 10,000 persons with asthma
  - per year

# Proportions, Rates, and Ratios



# Applying Summary Measures to Disease

- Incidence
- Prevalence
- Mortality



# Incidence

- Measures NEW cases of disease among a population at risk of disease over a period of time
- Examples of diseases/events measured with incidence
  - Cancer
  - Influenza, pertussis, or other notifiable disease
  - Hospitalizations
- Often reported as a **NUMBER** or a **RATE**



# Prevalence

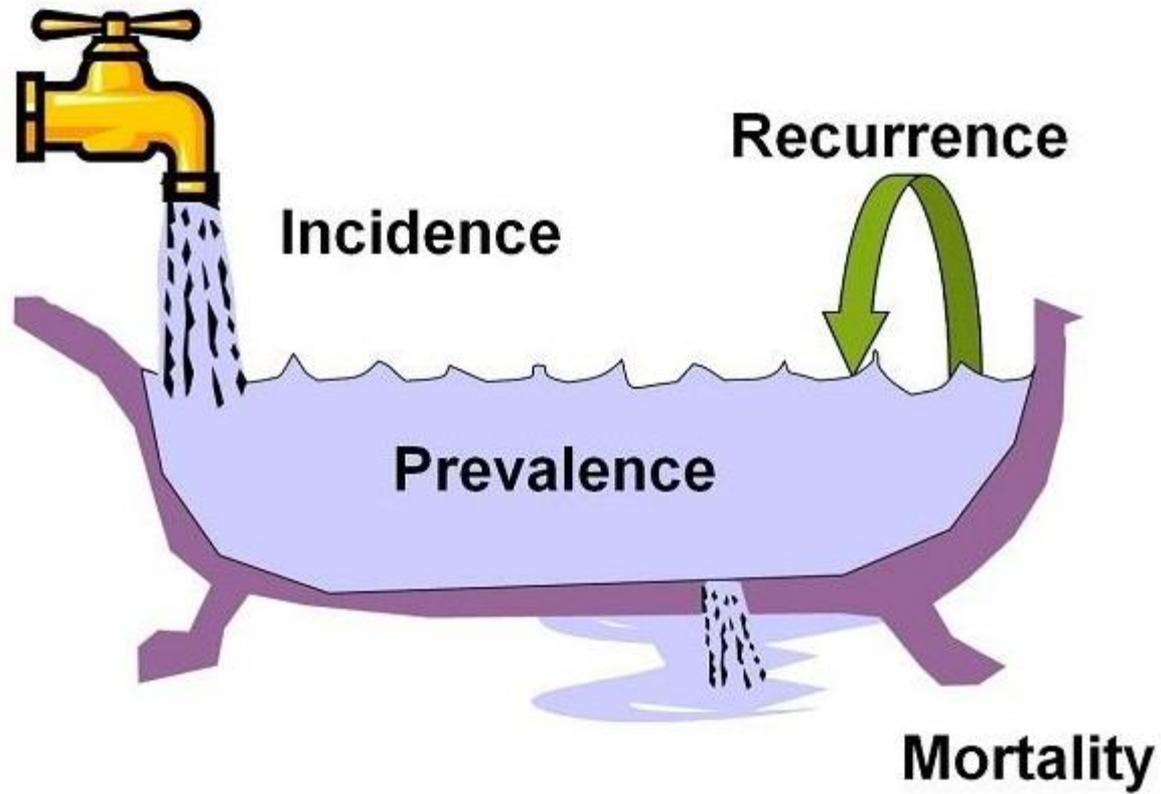
- Measures existing cases of a disease at a particular point in time or over a period of time
- Often reported as a **PERCENTAGE**
- We use **point prevalence**
- Ex: The prevalence of current adult smokers in Montana is 19.0%.



# Mortality

- Measures DEATH due to a particular cause among a population over a period of time.
- Often reported as a **NUMBER** or a **RATE**.





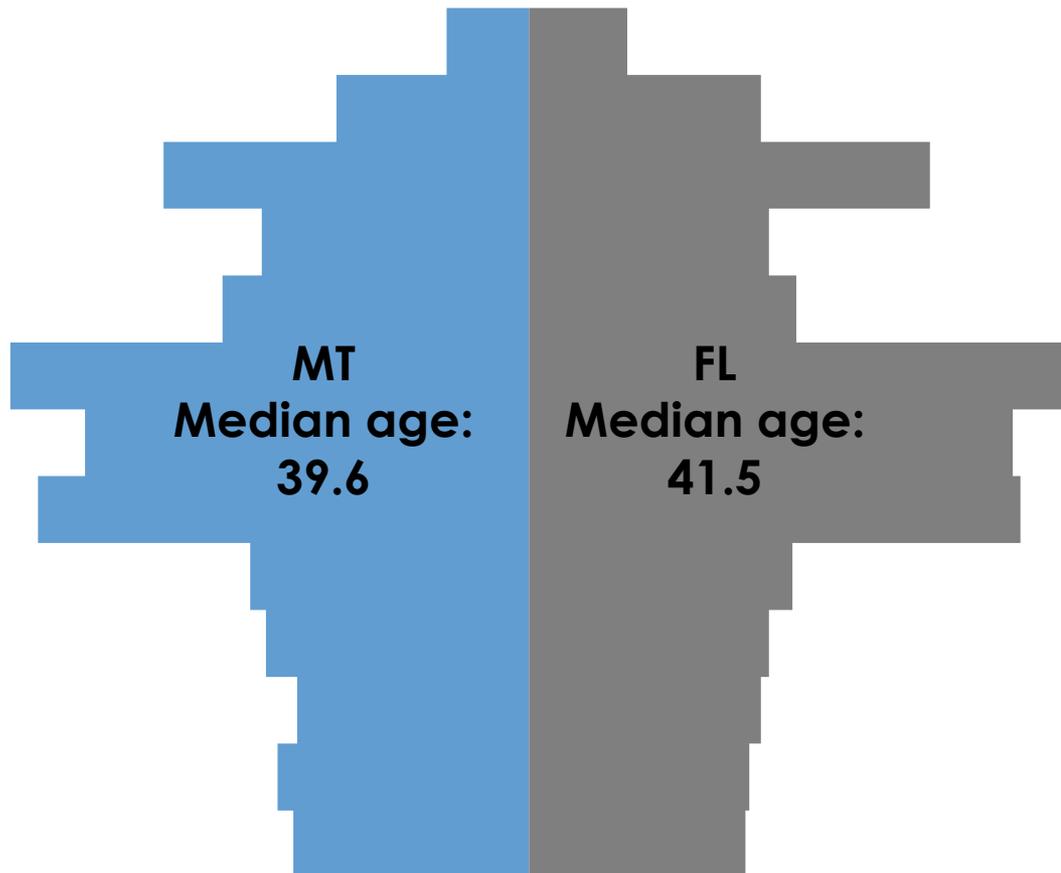
# Age-adjusted rates

- Disease or death is often associated with age.
- Age-adjusted rates are a way to make more fair comparisons between groups with different age distributions.



- Use when:
  - **Comparing geographic areas**  
Ex: MT to U.S.; or MT to counties; or County A to County B
  - **Comparing time periods**  
Ex: 1990-1999 to 2000-2009

# Why age-adjust?



# Why age-adjust?



Number of  
cancer deaths

**1,923**

**41,467**

Mortality rate  
(crude) per  
100,000 people

**194.1**

**220.1**

# Why age-adjust?



Number of  
cancer deaths

**1,923**

**41,467**

Mortality rate  
(crude)

**194.1**

**220.1**

Mortality rate  
(age-adjusted)

**159.8**

**163.9**

# Measures of Uncertainty

- Statistics (from **samples**) estimate the true value of a parameter (from **population**).
- There is unavoidable variability because of this called sampling variability.
- **Confidence intervals** are most common in health data.



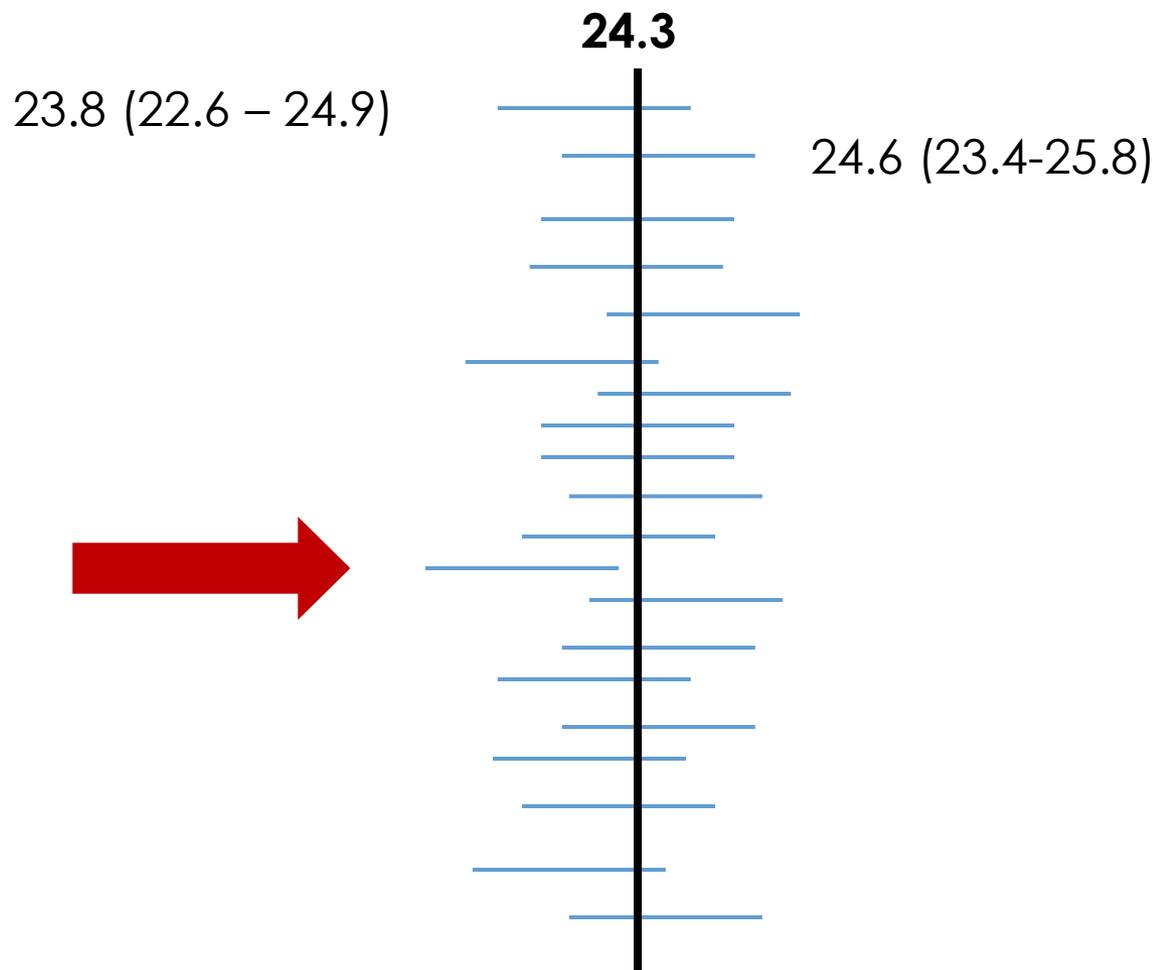
# Confidence Intervals



We can't ask all one million Montanans about their health habits. So we sample.

- BRFSS sample  $n=9,300$
- 2015 – 24.6% obese (23.4 – 25.8)
- If we look at all of the different samples of 9,300 people, and we produced an interval estimate for each sample, then 95% of those intervals would contain the true estimate.

Absolute obesity prevalence = 24.3%  
(If we were able to ask **ALL** Montanans)



# Confidence Intervals

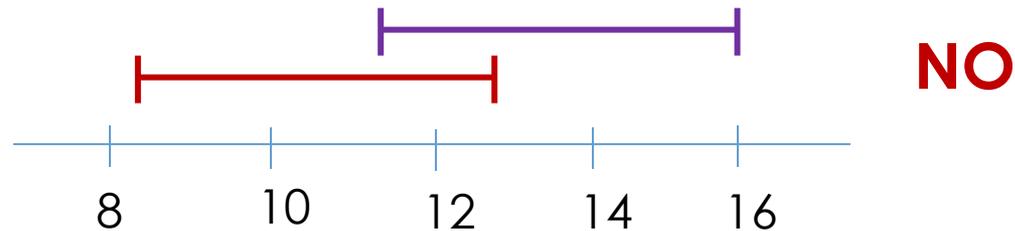
- 2015 – **24.6%** obese
- Margin of error: **1.2%**
- $24.6 - 1.2 = 23.4$   
 $24.6\% + 1.2 = 25.8$
- 95% confidence interval (**23.4% – 25.8%**)

# Confidence Intervals

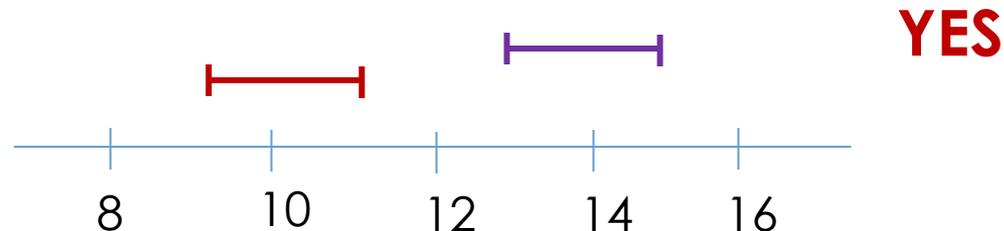
- Confidence intervals are one method to tell if two estimates are **significantly different** or not
- Overlapping confidence intervals indicate that there is **no** significant difference in the two estimates
- Non-overlapping confidence intervals indicate that there **is** a significant difference between the two estimates

# Confidence Intervals

- Are these significantly different:
  - **10.3%** (8.2-12.4) of youth smoke in MT versus **13.9%** (11.8-16.0) of youth nationwide



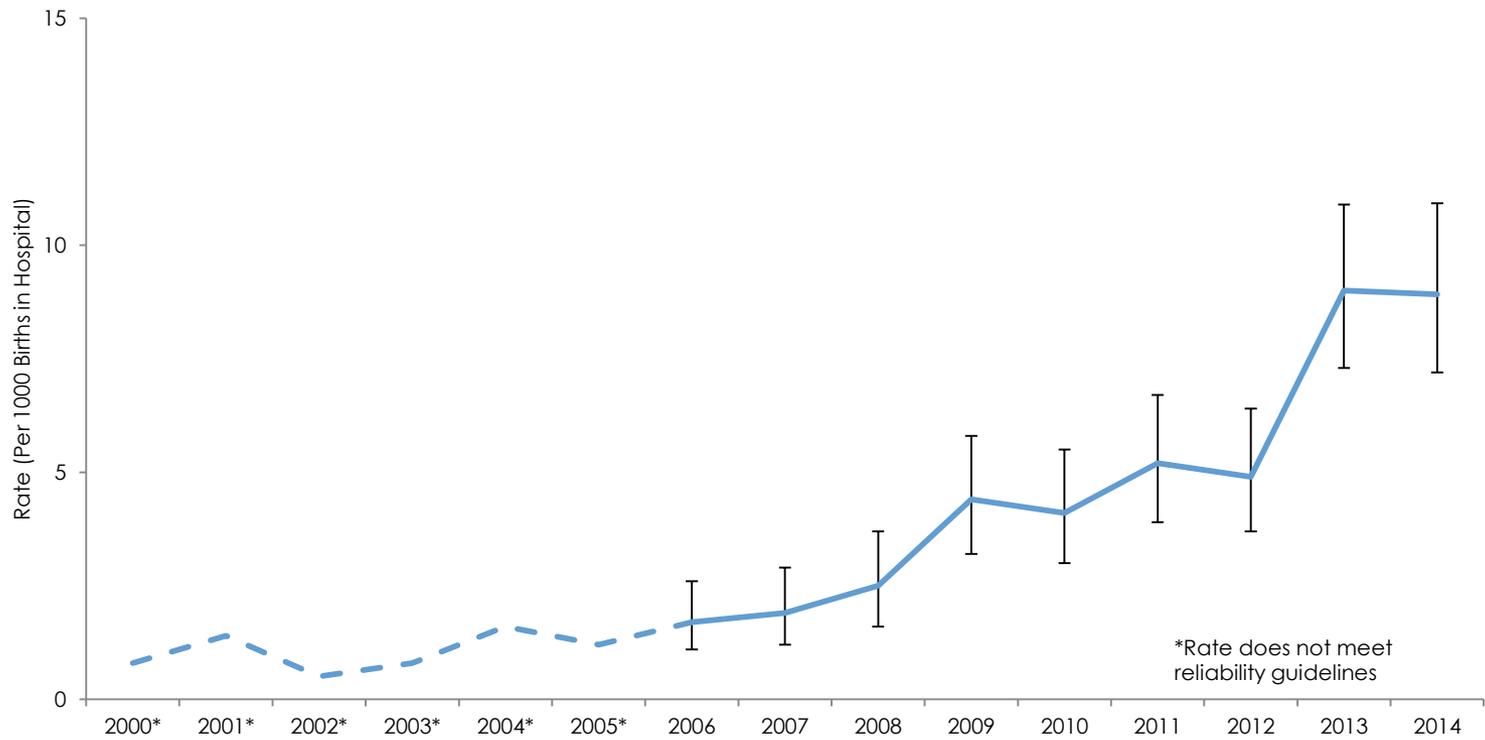
- **10.3%** (9.2-11.4) of youth smoke in MT versus **13.9%** (12.8-15.0) of youth nationwide



# Statistics 101

- Importance of confidence intervals:
  - **9.2%** of adults in Phillips County use smokeless tobacco versus **23.8%** of adults in Sanders County
  - **9.2%** (3.1-17.0) of adults in Phillips County use smokeless tobacco versus **23.8%** (15.9-30.2) of adults in Sanders County
  - **98.6%** of females are exceptional drivers compared to **95.2%** of males
  - **98.6%** (97.8-99.4) of females are exceptional drivers compared to **95.2%** (94.4-96.0) of males

Figure 1. Rate of Newborns with Drug Withdrawl Syndrome (ICD-9-CM: 779.5), Montana Resident Liveborns, 2000-2014



# Confidence Intervals

**The narrower the better**

To  your confidence interval:



**Increase sample size**



decrease standard error



confidence level (from 95% to 90%)

# Questions???

