When calculating ratios, don’t round off the numerator and denominator too much; it can cause the result to vary quite a bit.

\[
\begin{align*}
0.1199 & \text{ rounded to } 0.1 = 1.0 \\
0.1449 & \text{ } 0.1
\end{align*}
\]

Keep 3-4 decimal places in the numerator & denominator. Round off the final result.

\[
\begin{align*}
0.1199 & = 0.827 = 0.8 \\
0.1449 & \text{ } 0.1449
\end{align*}
\]
<table>
<thead>
<tr>
<th></th>
<th>Diseased</th>
<th>No Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>7</td>
<td>124</td>
</tr>
<tr>
<td>Non-exposed</td>
<td>1</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>202</td>
</tr>
</tbody>
</table>

\[
I_{\text{exposed}} = 0.0534 \\
I_{\text{not exp.}} = 0.0127 \\
\text{Risk Ratio} = 4.22
\]

But if I round off before dividing:

\[
0.053/0.013 = 4.07
\]
Reference Group for Multiple Levels of Exposure

Using a reference group for comparison:

You can do this for both cohort studies and case-control studies.
## Reference (Comparison Group) for a Cohort Type Study

<table>
<thead>
<tr>
<th>BMI</th>
<th>Events</th>
<th>Person-Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;21</td>
<td>41</td>
<td>177,356</td>
</tr>
<tr>
<td>21-23</td>
<td>57</td>
<td>194,243</td>
</tr>
<tr>
<td>23-25</td>
<td>56</td>
<td>155,717</td>
</tr>
<tr>
<td>25-29</td>
<td>67</td>
<td>148,541</td>
</tr>
<tr>
<td>&gt;29</td>
<td>85</td>
<td>99,573</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># MIs (non-fatal)</th>
<th>Person-years of observation</th>
<th>Rate of MI per 100,000 P-Yrs. (incidence rate)</th>
<th>Rate Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>177,356</td>
<td>23.1</td>
<td>1.0</td>
</tr>
<tr>
<td>57</td>
<td>194,243</td>
<td>29.3</td>
<td>1.3</td>
</tr>
<tr>
<td>56</td>
<td>155,717</td>
<td>36.0</td>
<td>1.6</td>
</tr>
<tr>
<td>67</td>
<td>148,541</td>
<td>45.1</td>
<td>2.0</td>
</tr>
<tr>
<td>85</td>
<td>99,573</td>
<td>85.4</td>
<td>3.7</td>
</tr>
</tbody>
</table>

**BMI**: weight (kg) / height squared (m²)
### D’Souza Case-Control Study

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Patients with Oropharyngeal Cancer (N = 100)</th>
<th>Control Patients (N = 200)</th>
<th>Unadjusted Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oral hygiene</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tooth loss</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>62 (62)</td>
<td>163 (82)</td>
<td>1.0</td>
</tr>
<tr>
<td>Some</td>
<td>16 (16)</td>
<td>20 (10)</td>
<td>2.1 (1.0–4.4)</td>
</tr>
<tr>
<td>Complete</td>
<td>22 (22)</td>
<td>17 (8)</td>
<td>3.4 (1.7–6.8)</td>
</tr>
<tr>
<td>Daily toothbrushing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>90 (90)</td>
<td>196 (98)</td>
<td>1.0</td>
</tr>
<tr>
<td>No</td>
<td>10 (10)</td>
<td>4 (2)</td>
<td>5.4 (1.7–17.8)</td>
</tr>
</tbody>
</table>
### Table: Oral Hygiene and Oropharyngeal Cancer

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Patients with Oropharyngeal Cancer (N=100)</th>
<th>Control Patients (N=200)</th>
<th>Unadjusted Odds Ratio (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number (percent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral hygiene</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tooth loss</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>62 (62)</td>
<td>163 (82)</td>
<td>1.0</td>
</tr>
<tr>
<td>Some</td>
<td>16 (16)</td>
<td>20 (10)</td>
<td>2.1 (1.0–4.4)</td>
</tr>
<tr>
<td>Complete</td>
<td>22 (22)</td>
<td>17 (8)</td>
<td>3.4 (1.7–6.8)</td>
</tr>
<tr>
<td>Daily toothbrushing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>90 (90)</td>
<td>196 (98)</td>
<td>1.0</td>
</tr>
<tr>
<td>No</td>
<td>10 (10)</td>
<td>4 (2)</td>
<td>5.4 (1.7–17.8)</td>
</tr>
</tbody>
</table>

#### Cancer Case and Control

<table>
<thead>
<tr>
<th>Tooth Loss</th>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>62</td>
<td>163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tooth Loss</th>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>62</td>
<td>163</td>
</tr>
</tbody>
</table>
Evaluating Random Error
(Sampling Error)
In analytic studies one enrolls subjects from a population and groups them in some way to make comparisons that test association between risk factors and outcomes.
What is a “Random Sample”? 

When selection of one subject doesn’t affect the chance of other subjects being chosen, i.e. selection is be independent (uninfluenced by other factors). To achieve this we should select by a process such that selection is independent of other factors… an independent selection process.

In reality, we usually take convenience samples... we take the subjects that are available.

Example: BWHS recruiting via subscribers to a magazine.
1) **Discrete Variables**: Variables that assume only a finite number of values:
   - Dichotomous variables (male/female, alive/died)
   - Categorical variables (nominal variables, e.g., race)
   - Ordinal variables: ordered categories (Excellent/Good/Fair/Poor)

2) **Continuous Variables**: Can take on any value within a range of plausible values, e.g., serum cholesterol level, weight & systolic blood pressure. (Also called quantitative or measurement variables.)
Case fatality rate for H1N1?

Is H1N1 vaccine effective?

Suppose you were reading two studies: one that addressed the first question and another that addressed the second question… what would be the study design of each?

What was being estimated in each?
One Group:

- Estimating a prevalence, risk, or rate in one group (e.g., case-fatality rate with bird flu)
- Estimating mean (e.g. average body weight)

(How precise are the estimates?)

Comparing 2 or more groups:

- Comparing risks, rates
- Comparing means (body weight in smokers versus non-smokers)

(Are differences just due to sampling error?)
(If we calculate a measure of association, how precise is it?)
Four Situations

Estimates in \textit{one} group

- Mean body weight of a population? \textbf{[measurement]}
- Case-fatality rate from bird flu (\% of cases that die)? \textbf{[proportion]}

Comparing \textit{two or more} groups

- Do smokers weigh less than non-smokers? \textbf{[measurement difference]}
- Increased risk of wound infection? \textbf{[relative risk]}

Are the groups different? How precise is measure of association?
Estimates in a Single Group:

How Precise?
Estimates in *one* group

- Mean body weight of a population? [measurement]
- Case-fatality rate from bird flu (% of cases that die)? [proportion]

Precision?

- Are the groups different?
- How precise is measure of association?

Comparing *two or more* groups

- Do smokers weigh less than non-smokers? [measurement difference]
- Increased risk of wound infection? [relative risk]

Four Possibilities
If we measured everyone, we would know the true mean, of the population, but this isn’t usually feasible.

We take samples to estimate the mean, but the estimated means from samples may be misleading as a result of sampling variability (random error), i.e., chance of “the luck of the draw.”
Estimated Mean = 205

Sample #1, N=3
Sample #2, N=3

Estimated Mean = 140
Sample #3, N=3

Estimated Mean = 265
Distribution of Body Weights in a Population

Number of people with each body weight

Total population

Body Weight

100 120 140 160 180 200 220 240 260 280 300 320

 Estimates may vary from sample to sample.

 Estimates will be more precise with larger samples, i.e., more consistent, and they are more likely to be close to the true mean.

Each sample provides an estimate of mean weight in the population.
The range within which the true value lies with a specified degree of confidence.

The estimated mean was 202. With 95% confidence, the true mean lies in the range of 184-220.
Estimates in **one** group

- Case-fatality rate from bird flu (% of cases that die)? [proportion]

Four Situations

- Mean body weight of a population? [measurement]
- Increased risk of wound infection? [relative risk]
- Are the groups different? How precise is measure of association?
- Do smokers weigh less than non-smokers? [measurement difference]

Comparing **two or more** groups

Measurements

- [measurement]

Frequencies

- [proportion]
Initially, 8 are identified as having bird flu and are studied; 4 died.

What is the case-fatality rate in this sample of 8 subjects?

No comparisons; just a single estimate in a group of cases.
What type of measure of disease frequency are we estimating?
Interpretation: The estimated case-fatality rate was 50% (4/8). With 95% confidence the true CFR lies in the range 21.5-78.5%.
95% Confidence Interval for a Proportion

The range within which the true proportion lies, with 95% confidence.

4/8 = 0.5 = point estimate

It is very unlikely (less than 5% probability) that the true value is outside this range.
Suppose it had been 32 deaths out of 64 cases, i.e., same proportion but larger sample?

The confidence interval would be....?

1. Narrower
2. Wider
3. Same width
4. Can’t say
Precision Depends on Sample Size

The greater the sample size...

... the narrower the confidence interval (and the more precise the estimate).
Weymouth, MA conducted an extensive health survey several years ago. One question asked, “Have you experienced violence because you or someone else was drinking alcohol?”

<table>
<thead>
<tr>
<th>Grade</th>
<th># who said yes</th>
<th># respondents</th>
<th>%</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>92</td>
<td>270</td>
<td>34.1</td>
<td>(28-40%)</td>
</tr>
</tbody>
</table>

What kind of study was this?

What type of measure of disease frequency are we estimating?
Hypothesis Testing:

Comparing Two or More Groups
Estimates in *one* group

- Mean body weight of a population? [measurement]
- Case-fatality rate from bird flu (% of cases that die)? [proportion]

Four Situations

- Precision?
- Measurements
- Frequencies

Comparing *two or more* groups

- Do smokers weigh less than non-smokers? [measurement difference]
- Increased risk of wound infection? [relative risk]

Are the groups different?
How precise is measure of association?
Null Hypothesis ($H_0$): the groups *are not* different

$$H_0: \text{Mean Wgt.}_{\text{smokers}} = \text{Mean Wgt.}_{\text{Non-smokers}}$$

Alternative Hypothesis ($H_A$): the groups *are* different

$$H_A: \text{Mean Wgt.}_{\text{Smokers}} \neq \text{Mean Wgt.}_{\text{Non-smokers}}$$
Provide guidance regarding whether the observations are consistent with the null hypothesis. If the null hypothesis were true, what would be the probability of getting a difference this great (or greater) due to random error?

**Could be consistent with** $H_0$

- 190 lbs.
- 195 lbs.

**Could be consistent with** $H_0$

- 190 lbs.
- 205 lbs.

**Probably not consistent with** $H_0$

- 180 lbs.
- 215 lbs.

**Probably not consistent with** $H_0$

- 190 lbs.
- 208 lbs.
Conclusions about whether the observations are compatible with the null hypothesis are based on both the magnitude of difference in the means and the precision of the estimates (variability).

There is no certainty here; we are assessing likelihood that the observations are consistent with the null hypothesis ($H_0$).

We are interested in whether there is strong evidence that the observations are not consistent with $H_0$. If so, we will reject $H_0$ and instead accept the alternate conclusion, i.e., that the two groups are probably different ($H_A$).

The key question is “If the null hypothesis were true, what would be the probability of observing differences this greater or greater?”
P-Values

“p” is for probability.

Definition of a p-value:
The probability of seeing a difference this big or bigger, if the groups were *not* different (meaning, if the null hypothesis were true) just as a result of sampling variability (random error).

If there really were no difference between the groups, there would be a very low probability of observing sample differences this great. Therefore, $H_0$ is probably not correct, so we reject it & conclude that the groups are probably different.

P-values are generated by performing a statistical test. The particular statistical test used depends on study type, type of measurement, etc.
P-Values

- Range between 0 and 1.
- **Small p-values** (≤ 0.05) mean a **low** probability that the observed difference occurred just by chance (so they’re likely to be different).

<table>
<thead>
<tr>
<th>p-value</th>
<th>Probability of observing sample differences this large if H₀ is true</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.55</td>
<td>55%</td>
</tr>
<tr>
<td>0.25</td>
<td>25%</td>
</tr>
<tr>
<td>0.12</td>
<td>12%</td>
</tr>
<tr>
<td>0.05</td>
<td>5%</td>
</tr>
<tr>
<td>0.01</td>
<td>1 in 100 (1%)</td>
</tr>
</tbody>
</table>

Observed differences may be due to sampling variability (random error).

Random error is an unlikely explanation for the observed differences, so the difference is “statistically significant”.

\[ p < \text{or } = 0.05 \]
If “p” \leq 0.05 it is unlikely that the difference is due to chance.

If so, we reject the null hypothesis in favor of the alternative hypothesis, i.e. the groups probably are different.

**BUT…**

- This criterion for significance ($\leq 0.05$) is arbitrary.
- The p-value is a probability, not a certainty.
If we wanted to compare two groups on measurement data (e.g. Is the mean weight of smokers less than that of non-smokers?), we might use a statistical test known as a “t-test” to determine whether there was a “statistically significant” difference.

(There is a worksheet that describes “t-tests” in Epi_Tools.XLS, but you are not responsible for this.)
Evaluating the Role of Chance With Categorical Data
Estimates in one group

- Mean body weight of a population? [measurement]
- Case-fatality rate from bird flu (% of cases that die)? [proportion]

Precision?

Four Situations

- Measurements
- Frequencies

Comparing two or more groups

- Do smokers weigh less than non-smokers? [measurement difference]
- Increased risk of wound infection? [relative risk]

Are the groups different? How precise is measure of association?
Is the risk of wound infection really different?

Or could the differences be due to sampling error?

<table>
<thead>
<tr>
<th>Wound Infection</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>7</td>
<td>124</td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>78</td>
</tr>
</tbody>
</table>

210 Subjects

Cumulative Incidence

\[
RR = \frac{7}{131} = \frac{1}{79} = 5.3 = 4.2
\]

Had Incidental Appendectomy

131 5.3% (7/131)

79 1.3% (1/79)

Is the risk of wound infection really different?
What would the null hypothesis be for the study on wound infections after incidental appendectomy?
Null Hypothesis \((H_0)\): the groups are not different

\[ H_0: \text{Incid. Infection}_{\text{exposed}} = \text{Incid. infection}_{\text{not exposed}} \]

or \( H_0: \text{RR} = 1.0 \) (or \( OR = 1 \))

or \( H_0: \text{RD} = 0 \)

or \( H_0: \text{Attrib. fraction} = 0 \)

Alternative Hypothesis \((H_A)\): the groups are different

\[ H_A: \text{RR} \text{ or } OR \neq 1.0 \]

or \( H_A: \text{RD} \neq 0 \)

or \( H_A: \text{Attrib. fraction} \neq 0 \)
### Observed Data:

<table>
<thead>
<tr>
<th></th>
<th>Infection</th>
<th>No Inf.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appy</td>
<td>7</td>
<td>124</td>
<td>131</td>
</tr>
<tr>
<td>No Appy</td>
<td>1</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>202</td>
<td>210</td>
</tr>
</tbody>
</table>

- 131 (62%)
- 79 (38%)

### Expected if No Association (Null Hypothesis)

<table>
<thead>
<tr>
<th></th>
<th>Infection</th>
<th>No Inf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appy</td>
<td>5</td>
<td>126</td>
</tr>
<tr>
<td>No Appy</td>
<td>3</td>
<td>76</td>
</tr>
</tbody>
</table>

- 5 (62% of 8)
- 3 (38% of 8)

126 total

How big a difference was there between what you expected under the null hypothesis and what was observed?

If the null hypothesis were true, what is the probability of seeing a difference this big just due to sampling variability (the luck of the draw)?
Chi Squared Test Assesses Difference Between Observed & Expected Under $H_0$.

**Observed Data:**

<table>
<thead>
<tr>
<th>Infection</th>
<th>No Inf.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appy</td>
<td>7</td>
<td>124</td>
</tr>
<tr>
<td>No Appy</td>
<td>1</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>202</td>
</tr>
</tbody>
</table>

**Expected if No Association (Null Hypothesis):**

<table>
<thead>
<tr>
<th>Infection</th>
<th>No Inf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appy</td>
<td>5</td>
</tr>
<tr>
<td>No Appy</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(62% of 8)</td>
</tr>
</tbody>
</table>

\[
\chi^2 = \sum \frac{(O-E)^2}{E} = 2.24
\]

*If the null hypothesis were true, there would be a 13% probability of seeing a difference this big (or bigger) due to random error. This is not strong enough evidence to reject the null hypothesis. (i.e. no significant difference).*
### Observed Data:

<table>
<thead>
<tr>
<th>No Inf.</th>
<th>Appy</th>
<th>Total (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>78</td>
<td>79 (38%)</td>
</tr>
<tr>
<td>5</td>
<td>126</td>
<td>131 (62%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No Inf.</th>
<th>Appy</th>
<th>Total (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>76</td>
<td>79 (38%)</td>
</tr>
</tbody>
</table>

### Expected if No Association (Null Hypothesis)

<table>
<thead>
<tr>
<th>Infection</th>
<th>No Inf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>126</td>
</tr>
<tr>
<td>3</td>
<td>76</td>
</tr>
</tbody>
</table>

### Caution:

Chi-Squared Test Exaggerates Significance with Small Samples. Instead of Chi-Squared Test, use Fisher’s Exact Test for small samples, i.e. if expected number in any cell is <5).

---

**NOTE:** Fisher’s requires an iterative calculation that is tough for a spreadsheet. See [http://www.langsrud.com/fisher.htm](http://www.langsrud.com/fisher.htm) for a 2x2 table that performs Fisher’s Exact Test.
Fisher's Exact Test

http://www.matforsk.no/ola/fisher.htm

TABLE = [ 7, 124, 1, 78 ]
Left : p-value = 0.9788955643076613
Right : p-value = 0.12866897889848644
2-Tail : p-value = 0.2631744093886173

P = 0.26
The Limitations of p-Values

1. There is a tendency for p-values to devolve into a conclusion of "significant" or not based on the p-value.

2. If an effect is small and clinically unimportant, the p-value can be "significant" if the sample size is large. Conversely, an effect can be large, but fail to meet the p<0.05 criterion if the sample size is small.

3. When many possible associations are examined using a criterion of p≤ 0.05, the probability of finding at least one that meets the “critical point” increases in proportion to the number of associations that are tested.

4. Statistical significance does not take into account the evaluation of bias and confounding.
Is the risk of wound infection really different, or were the observed differences just due to “the luck of the draw”?

<table>
<thead>
<tr>
<th>Had Incidental Appendectomy</th>
<th>Wound Infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes: 7</td>
</tr>
<tr>
<td>No</td>
<td>1</td>
</tr>
</tbody>
</table>

210 Subjects

RR = \frac{7}{131} = 5.3 = 4.2

How precise is this estimated measure of association?

How confident are we that this is a good estimate?
The “null” value; i.e., no difference.

Could the observed data be consistent with the null hypothesis?

Estimated RR = 4.2

Possible Values of Risk Ratio, Odds Ratio, or Rate Ratio
The range in which the true RR or OR lies, with 95% confidence.

We’re 95% confident that the true value is in this range, so there is <5% probability that the null hypothesis is correct.

Here, the **null value** is outside the interval, so there is <5% chance that the null is the true value. The null is excluded, so p<0.05.
The true RR could be 1.

Here, the null value is inside the interval, so there is >5% chance that the null is the true value. The null could be true, so p>0.05.
The 95% confidence interval for relative risk gives the range of RR values that are compatible with your sample.

If the 95% CI includes the null value, then the findings are compatible with RR=1.0, so there is no significant difference.

But, if the 95% CI does NOT include the null value, RR=1.0 is unlikely (p<0.05).
Examples

Statistically significant decreased risk in exposed group. (p<0.05)

Statistically significant increased risk in exposed group. (p<0.05)

No statistically significant difference in risk (p>0.05)

Possible Values of Risk Ratio, Odds Ratio, or Rate Ratio
A 95% Confidence Interval Puts Non-significant Results in Perspective

Neither of these results are statistically significant, but…

Do you view them differently?

Possible Values of Risk Ratio, Odds Ratio, or Rate Ratio
Having a larger sample size does not ensure that you will get a “statistically significant” difference.

However, with a larger sample…

- The estimate of RR or OR will be more precise.
- If there really is a difference between groups, you will have a better chance of detecting it (better statistical power).
I used the cohort study sheet in Epi_Tools.XLS to calculate the 95% confidence interval for the estimated risk ratio in the incidental appendectomy study:

Estimated RR= 4.2
95% CI: 0.53 – 36.5

What do you think?

Possible Values of Risk Ratio, Odds Ratio, or Rate Ratio
It’s possible that incidental appendectomy really *does* carry an increased risk of wound infection, but the sample size was too small to demonstrate a significant difference.

If the true difference were really 5.3% vs. 1.3% how many subjects would it have taken to demonstrate a significant difference (with 90% power)?

What if the true difference were really 10% vs. 1.3% how many subjects would it have taken to demonstrate a significant difference (with 90% power)?

You can use the Sample Size sheet (Part II) in “Epi_Tools.XLS” to calculate how many subjects would have been needed, assuming equal numbers in each group.
A cohort study examined the association between DES and breast cancer & found RR=1.4 (p=0.006).

1. 40% of women on DES got breast cancer, but this was not ‘statistically significant.’
2. 40% of women on DES got breast cancer, & this was ‘statistically significant.’
3. Women on DES had 1.4 times the risk of getting breast cancer, but this was not statistically significant because the p-value was less than 0.05.
4. Women on DES had 1.4 times the risk of developing breast cancer, and this was statistically significant because the p-value was less than 0.05.
A RCT compared % pain relief from glucosamine + chondroitin vs. placebo in osteoarthritis patients. RR=1.1 (p=0.10).

1. Those on G+C were 10% more likely to report pain relief, but it was *not* ‘statistically significant.’
2. Those on G+C were 10% more likely to report pain relief, and it *was* ‘statistically significant.’
3. Those on G+C had 1.1 times the risk of pain, but it was not statistically significant.
4. Those on G+C had 1.1 times the risk of pain, & it was statistically significant.
The small difference could have been due to random error (sampling error).
A study compared 35 diabetics to a comparable group of 35 non-diabetics and found that the diabetics had a higher incidence of femoral arterial occlusion. The risk ratio was 7.0, and the 95% CI was **1.2 to 14.3**.

How would you interpret these findings?

1. **Not** statistically significant, since the RR was inside the 95% confidence interval.

2. **Not** statistically significant, since the null was outside the 95% confidence interval.

3. Was statistically significant, since the RR was inside the 95% confidence interval.

4. Was statistically significant, since the null was outside the 95% confidence interval.
...comparing 450 diabetics to a comparable group of 900 non-diabetics and found that the diabetics had a higher incidence of femoral arterial occlusion.

The relative risk was 3.0, and the 95% CI was 2.0 - 4.2.

How would you interpret these findings?
Association of Coffee Drinking with Total and Cause-Specific Mortality

Methods
“We examined the association of coffee drinking with subsequent total and cause-specific mortality among 229,119 men and 173,141 women in the National Institutes of Health–AARP Diet and Health Study who were 50 to 71 years of age at baseline. Participants with cancer, heart disease, and stroke were excluded. Coffee consumption was assessed once at baseline.”
Results
“During 5,148,760 person-years of follow-up between 1995 and 2008, a total of 33,731 men and 18,784 women died. In age-adjusted models, the risk of death was increased among coffee drinkers. However, coffee drinkers were also more likely to smoke, and, after adjustment for tobacco-smoking status and other potential confounders, there was a significant inverse association between coffee consumption and mortality. Adjusted hazard [risk] ratios for death among men who drank coffee as compared with those who did not were as follows: 0.99 (95% confidence interval [CI], 0.95 to 1.04) for drinking less than 1 cup per day, 0.94 (95% CI, 0.90 to 0.99) for 1 cup, 0.90 (95% CI, 0.86 to 0.93) for 2 or 3 cups, 0.88 (95% CI, 0.84 to 0.93) for 4 or 5 cups, and 0.90 (95% CI, 0.85 to 0.96) for 6 or more cups of coffee per day (P<0.001 for trend); the respective hazard ratios among women were 1.01 (95% CI, 0.96 to 1.07), 0.95 (95% CI, 0.90 to 1.01), 0.87 (95% CI, 0.83 to 0.92), 0.84 (95% CI, 0.79 to 0.90), and 0.85 (95% CI, 0.78 to 0.93) (P<0.001 for trend). Inverse associations were observed for deaths due to heart disease, respiratory disease, stroke, injuries and accidents, diabetes, and infections, but not for deaths due to cancer.”