# Bios 6648: Design & conduct of clinical research

Section 0 - Motivation, context, and objectives

Lecture outline (28 August 2013)

## Previously

<table>
<thead>
<tr>
<th>0.1 Examples:</th>
</tr>
</thead>
<tbody>
<tr>
<td>★ Historic approach to coronary artery disease</td>
</tr>
<tr>
<td>★ Neural tube defects</td>
</tr>
</tbody>
</table>

## Today

<table>
<thead>
<tr>
<th>0.1 Examples (con’t):</th>
</tr>
</thead>
<tbody>
<tr>
<td>★ Selenium supplementation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0.2 Public health objective for clinical research</th>
</tr>
</thead>
</table>
Case study 2: Selenium for cancer prevention

Motivating example (Clark LC, Nutr Cancer. 1984;6(1):13-21)

- Case-control study: Plasma selenium and skin neoplasms:
  - 142 cases (basal cell epithelioma or squamous cell carcinoma); 103 noncancer controls.
  - Odds ratio = 4.39: lowest vs highest selenium decile (cases vs controls)

Abstract

Although experimental studies in animals show that selenium may prevent cancer, case-control studies of internal human cancers have been difficult to interpret because neoplastic tissue sequesters selenium. We therefore conducted a case-control study to examine the association between plasma selenium level and skin cancer, a neoplasm with minimal tumor mass at the time of diagnosis. The mean selenium level among patients with either basal cell epithelioma (N = 142), squamous cell carcinoma (N = 48), or both (N = 50), was 0.141 micrograms/g. This was significantly lower than the mean plasma selenium level of the 103 control subjects, which was 0.155 micrograms/g. The noncancer control groups were drawn from current clinic patients and past clinic patients. The logistic estimate of the odds ratio for the lowest versus the highest decile of selenium for all cases combined versus the group of current patient controls was 4.39; for all cases combined versus the past patient controls, the logistic estimate of the odds ratio was 5.81.
Case study 2: Selenium for cancer prevention

Follow-up clinical trial (Clark, JAMA 1996; 276:1957-1963)

Original Contributions

Effects of Selenium Supplementation for Cancer Prevention in Patients With Carcinoma of the Skin

A Randomized Controlled Trial

Larry C. Clark, MPH, PhD; Gerald F. Combs, Jr, PhD; Bruce W. Turnbull, PhD; Elizabeth H. Slate, PhD; Dan K. Chalker, MD; James Chow, MD; Loretta S. Davis, MD; Renee A. Glover, MD; Gloria F. Graham, MD; Earl G. Gross, MD; Arnon Krongrad, MD; Jack L. Lesher, Jr, MD; H. Kim Park, MD; Beverly B. Sanders, Jr, MD; Cameron L. Smith, MD; J. Richard Taylor, MD; for the Nutritional Prevention of Cancer Study Group

- Design: RCT (double-blind placebo-controlled; 1983-1991)
  - Dietary supplement: oral selenium (200µg) vs placebo
  - Patients with history of basal or squamous cell skin cancer
  - 1312 patents in seven dermatology clinics in eastern US
Case study 2: Selenium for cancer prevention

Clark trial results

<table>
<thead>
<tr>
<th>Cancer Sites, No.</th>
<th>Selenium</th>
<th>Placebo</th>
<th>RR (95% CI)*</th>
<th>P Value</th>
<th>HR (95% CI)†</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung‡</td>
<td>17</td>
<td>31</td>
<td>0.54 (0.30-0.98)</td>
<td>.04</td>
<td>0.56 (0.31-1.01)</td>
<td>.05</td>
</tr>
<tr>
<td>Prostate‡</td>
<td>13</td>
<td>35</td>
<td>0.37 (0.18-0.71)</td>
<td>.002</td>
<td>0.35 (0.18-0.65)</td>
<td>.001</td>
</tr>
<tr>
<td>Colorectal‡</td>
<td>8</td>
<td>19</td>
<td>0.42 (0.18-0.95)</td>
<td>.03</td>
<td>0.39 (0.17-0.90)</td>
<td>.03</td>
</tr>
<tr>
<td>Head and neck</td>
<td>6</td>
<td>8</td>
<td>0.74 (0.21-2.43)</td>
<td>.58</td>
<td>0.77 (0.27-2.24)</td>
<td>.64</td>
</tr>
<tr>
<td>Bladder</td>
<td>8</td>
<td>6</td>
<td>1.32 (0.40-4.61)</td>
<td>.62</td>
<td>1.27 (0.44-3.67)</td>
<td>.66</td>
</tr>
<tr>
<td>Esophageal</td>
<td>2</td>
<td>6</td>
<td>0.33 (0.03-1.84)</td>
<td>.15</td>
<td>0.30 (0.06-1.49)</td>
<td>.14</td>
</tr>
<tr>
<td>Breast</td>
<td>9</td>
<td>3</td>
<td>2.88 (0.72-16.5)</td>
<td>.09</td>
<td>2.95 (0.80-10.9)</td>
<td>.11</td>
</tr>
<tr>
<td>Other specific carcinomas</td>
<td>5</td>
<td>9</td>
<td>0.55 (0.14-1.82)</td>
<td>.27</td>
<td>0.54 (0.18-1.62)</td>
<td>.27</td>
</tr>
<tr>
<td>Total carcinomas‡$</td>
<td>59</td>
<td>104</td>
<td>0.55 (0.40-0.77)</td>
<td>&lt;.001</td>
<td>0.54 (0.39-0.75)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Melanomas</td>
<td>8</td>
<td>8</td>
<td>0.97 (0.32-2.96)</td>
<td>.91</td>
<td>0.92 (0.34-2.45)</td>
<td>.87</td>
</tr>
<tr>
<td>Leukemia/lymphomas</td>
<td>8</td>
<td>5</td>
<td>1.58 (0.46-6.14)</td>
<td>.41</td>
<td>1.50 (0.49-4.60)</td>
<td>.48</td>
</tr>
<tr>
<td>Other specific noncarcinomas</td>
<td>3</td>
<td>3</td>
<td>0.99 (0.13-7.37)</td>
<td>.98</td>
<td>0.99 (0.20-4.94)</td>
<td>.99</td>
</tr>
<tr>
<td>Total noncarcinomas</td>
<td>19</td>
<td>16</td>
<td>1.17 (0.57-2.44)</td>
<td>.65</td>
<td>1.16 (0.60-2.27)</td>
<td>.65</td>
</tr>
<tr>
<td>Total cancer†$</td>
<td>77</td>
<td>119</td>
<td>0.63 (0.47-0.85)</td>
<td>.001</td>
<td>0.61 (0.46-0.82)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*RR indicates relative risk; CI, confidence interval. P values derived from log rank tests.

- Lung cancer results (incident cases):
  - Selenium: 17 cases; Placebo: 31 cases
  - RR: 0.54 (95%CI: 0.30-0.98; p = 0.04)

- Prostate cancer results (incident cases):
  - Selenium: 13 cases; Placebo: 35 cases
  - RR: 0.37 (95%CI: 0.18-0.71; p = 0.002)
ECOG 5597

Phase III Chemoprevention Trial of Selenium Supplementation in Persons with Resected Stage I Non-Small Cell Lung Cancer

- Design: RCT (double-blind placebo-controlled)
  - Dietary supplement: oral selenium (200µg) vs placebo
  - Patients with resected stage I NSC lung cancer
  - 1522 patients from ECOG-participating clinics from 2000-2009.

- Results (interim analysis in 2009):
  - 5-year risk of recurrence or death: Selenium: 72%; Placebo 78%
  - Trial stopped early: “not an effective chemoprevention agent.”
Case study 2: Selenium for prostate cancer prevention

Effect of Selenium and Vitamin E on Risk of Prostate Cancer and Other Cancers
The Selenium and Vitamin E Cancer Prevention Trial (SELECT)

SELECT trial; JAMA. 2009 301(1): 39-51

- Randomized 35,533 men to 4 treatment groups (2 × 2 factorial):
  - Selenium + Vit E placebo
  - Selenium placebo + Vit E
  - Selenium + Vit E
  - Selenium placebo + Vit E placebo
- Follow-up for 4.17-7.33 years over 12 years
SELECT trial results

Hazard ratios and 99% CI for prostate cancer:

- Vit E: 1.13 (0.95 to 1.35)
- Selenium: 1.04 (0.87 to 1.24)
- Selenium + Vit E: 1.05 (0.88 to 1.25)
The selenium story represents:
- Excellent demonstration of careful evaluation of a hypothesis illustrating:
- Interplay between careful epidemiology and clinical trials in a range of diseases:
  - Epidemiology as foundation for major intervention trials.
  - Demonstrates the importance of confirmatory trials for subgroup effects in large trials.
  - Large RCT’s of the same hypothesis in multiple diseases
  - The question has been answered??
Formulating the public health objective

- **Ultimate objectives:**
  - Discover things that are true
  - Develop the science in order to provide public health benefit (therapies, prevention, etc...)
  - Want high prevalence of truly beneficial therapies/practices among all things (therapies or public health recommendations) that are adopted in practice.

- **These objectives are quantified as the positive predictive value (PPV) of clinical research**
  - Review of PPV: cervical cancer screening
  - PPV in clinical trials
    - Illustration of practices to increase (or decrease) PPV.
PPV Example: Cervical cancer screening in New Zealand

New Zealand National Cervical Screening Program (NSCP)

- Over 70% participation
- Two screening tests (circa 2000)
  - Pap smear (∼$5): funded by NSCP
  - ThinPrep (∼$20): offered by some physicians for $15 fee

ThinPrep versus Pap (Stein 2003)

- Pap smear (Papanicolaou test)
  - Cervical swab on slide for pathologist evaluation
  - Sensitivity ∼ 50% (up to 68%)  
  - Specificity ∼ 98% (up to 79%)
- “ThinPrep”: liquid-based cytology screening test
  - Cervical swab rinsed in tube with liquid preservative
  - Sensitivity ∼ 80%
  - Specificity ∼ 90%
### PPV Example: Cervical cancer screening in New Zealand

**NCSP and equitably (circa 2000)**

- Lower SES communities unable to pay for ThinPrep
- *IF* superior should NCSP adopt ThinPrep?
- Key questions:
  - Is ThinPrep really more accurate than pap?
  - What are the potential cost impacts?
    - ThinPrep costs $15 more
    - A positive screening test is referred for colposcopy ($200).
    - Lower specificity might overwhelm budget with unnecessary colposcopies.
Impact of sensitivity and specificity on the NCSP

▶ Suppose:
  ▶ 1,000,000 women are screened
  ▶ Prevalence of high grade lesions is 1%:
    ▶ 10,000 with high grade lesion
    ▶ 990,000 without high grade lesion
  ▶ Each positive test is sent for colposcopy
**PPV Example: Cervical cancer screening in New Zealand**

### Impact of sensitivity and specificity on the NCSP

- Suppose NCSP uses pap:
  - Sensitivity = 50%
  - Specificity = 98%
- Results of screening:
  - Number of positive tests:
    - True positive tests: $10,000 \times 0.50 = 5000$
    - False positive tests: $990,000 \times 0.02 = 19,800$
    - PPV: $\frac{5000}{24,800} = 0.20$
- Cost:
  - Cost of tests: $5.00M$
  - Cost of colposcopy: $4.96M$
  - Total: $\sim 10M$
Impact of sensitivity and specificity on the NCSP

- Suppose NCSP uses ThinPrep:
  - Sensitivity = 80%
  - Specificity = 95%

- Results of screening:
  Number of positive tests:
    - True positive tests: \(10,000 \times 0.80 = 8000\)
    - False positive tests: \(990,000 \times 0.05 = 49,500\)
  
  PPV:
  \[
  \frac{8000}{57,500} = 0.14
  \]

- Cost:
  - Cost of tests: \(\$20M\)
  - Cost of colposcopy: \(\$11.5M\)
  - Total: \(\sim \$31.6M\)
**PPV Example: Cervical cancer screening in New Zealand**

### Impact of sensitivity and specificity on the NCSP

- Suppose NCSP uses ThinPrep:
  - Sensitivity = 80%
  - Specificity = 90%
- Results of screening:
  - Number of positive tests:
    - True positive tests: \(10,000 \times 0.80 = 8000\)
    - False positive tests: \(990,000 \times 0.1 = 99,000\)
  - PPV: \(\frac{8000}{99,900} = 0.075\)

### Cost:
- Cost of tests: \(\$20.0M\)
- Cost of colposcopy: \(\$21.4M\)
- Total: \(\sim \$41.4M\)
PPV Example: Cervical cancer screening in New Zealand

**Summary remarks: public health objective**

- Rare diseases:
  - High risk for false positive
  - Important to control *specificity*

- Consequences of a false positive
  - Costs to healthcare system
  - Anxiety costs for women

- Clearly:
  - Weigh costs against risk/consequences of false negative

- Public health objective:
  - Highest PPV for lowest total cost
(b) Formulating the public health objective

**PPV as the objective in public health research**

- **So what is the right answer?**
  - Diagnostic testing
    - Identify people with disease who can benefit from care
    - Identify people who should not be treated
  - Public health research?
    - Identify hypotheses that are in fact true
    - Identify hypotheses that are not worthy of further exploration

- **What are the consequences of a wrong answer?**
  - Diagnostic testing?
    - People do not receive beneficial treatment
    - People receive non-beneficial treatment
  - Public health research?
    - Populations do not receive beneficial practice/care
    - Populations receive non-beneficial practice/care

- **Objective:**
  - Maximize the proportion of right answers (PPV)
Positive predictive value of research

PPV in research

- A Statistical hypothesis tests can be viewed as a test for beneficial treatments.
  - $\alpha$-level: probability of observing a positive (statistically significant) test in absence of a true treatment effect:
    - Level of significance is $1 - \text{specificity}$.
    - Choosing $\alpha = 0.05$ gives 95% specificity.

- Statistical power ($\beta$): Probability of observing a positive (statistically significant) test when there is a true treatment effect:
  - Power is sensitivity.
  - Common choice of 80% sensitivity (not usually recommended by me).

- Prevalence ($\pi_0$): the percentage of effective treatments among all tested treatments.
Positive predictive value of research

**PPV in research**

- Positive predictive value: probability that a statistically significant trial indicates a truly effective treatment.

\[
PPV = \frac{\beta \pi_0}{\beta \pi_0 + \alpha (1 - \pi_0)}
\]

- The probability that our public health recommendation is in fact beneficial.
Example: The Amgen experience

CG Begley and LM Ellis: "Raise the standards for preclinical cancer research" Nature 483:531-533; 2012

* Over the past decade Amgen scientists tried to confirm the results of 53 ‘landmark’ studies

* Only 6/53 (11%) of these results were confirmed

* “The scientific process demands the highest standards of quality, ethics, and rigour.”

* JK perspective:
  ▶ Agree: High standards are an absolute requirement
  ▶ But lack of reproducibility is not surprising if initial false-positive risk is high
The Public Health Objective

Clinical trials as diagnostic tests

- We routinely consider power ($\beta = \text{sensitivity}$) and type I error ($\alpha = 1 - \text{specificity}$).

- What is the prevalence ($\pi_0$)?
  - NCI Developmental Therapeutics Program:
    - Over 400,000 candidate compounds since 1955 (over 80,000 since 1990).
    - NCI sponsors about 1500 trials involving 25,000 patients/year.
  - 10% of treatments entering phase I trials are positive in subsequent phase III trials (Von Hoff, 1998)
  - Results of NCI-sponsored trials 1955-2006 (Djulbegovic, 2008)
    - 743 randomized comparisons, 176 (24%) are significant for new treatments
    - 116 (15%) discover ‘breakthrough interventions’.
  - Results of phase II cancer trials (J Lee, 2005)
    - 266 randomized phase II trials: 39 (15%) led to phase III.
  - Prevalence of truly beneficial treatments entering phase II trials is probably less than 10%.
The Public Health Objective
How do clinical trials determine PPV?

Example: Phase II studies as screening tests

Consider the following approaches to evaluating new treatments:

1. Study every treatment in a large definitive experiment.
2. Perform small screening tests, and perform large definitive experiments only in those treatments that pass the screening tests.

Suppose that we want to evaluate the efficiency of these strategies. Assume:

- 10% of all treatments actually work.
- Level of significance = 0.05 (specificity = 0.95).
- 1,000,000 subjects are available for clinical trials.
- Power for a clinically important difference:
  - 1000 subjects → 97.5% power
  - 500 subjects → 80% power
  - 50 subjects → 15% power
The Public Health Objective
How do clinical trials determine PPV?

Example: Phase II studies as screening tests

- Scenario 1 (only large trials):
  - Suppose we evaluate 1000 new treatments (100 effective and 900 ineffective).
  - On average we have positive tests for:
    - 98 of the 100 effective treatments ($0.975 \times 100 \approx 98$).
    - 45 of the 900 ineffective treatments ($0.05 \times 900 = 45$).
  - PPV: $\frac{98}{45 + 98} = 0.69$; that is, only 69% of the 143 treatments identified actually work.
The Public Health Objective
How do clinical trials determine PPV?

Example: Phase II studies as screening tests

- Scenario 2 (preliminary screening trials):
  (a) Suppose we first screen 12,500 new treatments (1,250 effective and 11,250 ineffective).
  - Using 50 subjects in the screening trials (625,000 total) with 15% power.
  - On average the screening trials give positive tests for:
    - 187 of the 1,250 effective treatments (0.15 × 1250 ≈ 187).
    - 562 of the 11,250 ineffective treatments (0.05 × 11250 ≈ 562).
  - PV+ for the screening phase: 187/(187 + 562) = 0.25.

(b) Now evaluate the 749 treatments (187 effective and 562 ineffective) from the screening trials.
  - Using 500 subjects per trial (374,500 total) with 80% power.
  - On average these confirmatory trials give positive tests for:
    - 150 of the 187 effective treatments (0.8 × 187 ≈ 150).
    - 28 of the 562 ineffective treatments (0.05 × 562 ≈ 28).
  - PV+ for confirmatory trials: 150/178 = 0.84.
The Public Health Objective
How do clinical trials determine PPV?

Example: Phase II studies as screening tests

- Comparison of scenarios:
  - Scenario 1 (large trials only):
    - Use 1,000,000 subjects
    - Screen 1,000 new treatments
    - Adopt 98 effective treatments
    - Adopt 45 ineffective treatments
    - PPV = 98/143 = 0.69
  
  - Scenario 2 (screening studies followed by large trials):
    - Use 999,500 subjects
    - Screen 12,500 new treatments
    - Adopt 150 effective treatments
    - Adopt 28 ineffective treatments
    - PPV = 150/178 = 0.84
The Public Health Objective
How do clinical trials determine PPV?

Example: Phase II studies as screening tests

- Bottom line:
  - Using the same number of subjects, phase II studies increase the predictive value of a positive study. A greater number of effective treatments are identified due in part to the greater number of treatments screened.
  - (Different choices of statistical power in screening and confirmatory trials can be used to optimize the strategy for a particular setting.)
The Public Health Objective
How do clinical trials determine PPV?

PPV is increased through good experimental practice

\[ PPV = \frac{\beta \pi_0}{\beta \pi_0 + \alpha (1 - \pi_0)} \]

* Increase \( \pi_0 \):
  - Careful planning of preliminary studies
  - Avoid "novel" and "innovative" ideas
  - Careful specification of hypothesis-driven research (avoid "science by hunch")

* Increase \( \beta \):
  - Good practice (no missing data, low variation in outcome assessment, good adherence, etc.)
  - Increase sample size.

* Reduce \( \alpha \):
  - Pre-specify outcomes
  - Pre-specify all analyses
  - Avoid multiple comparisons
  - Avoid surrogate outcomes.
  - Avoid subgroups
The Public Health Objective
How do clinical trials determine PPV?

Sensitivity to $\pi_0$ (how likely is it that the new treatment works?)

3. Trial of an ‘incremental’ advance for a known compound:
   ▶ $\pi_0 = 0.20; \alpha_2 = 0.05; \beta_2 = 0.15; \alpha_3 = 0.05; \beta_3 = 0.80$
   ▶ Results:

<table>
<thead>
<tr>
<th>Trials</th>
<th>True</th>
<th>False</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 2</td>
<td>11765</td>
<td>353</td>
<td>471</td>
</tr>
<tr>
<td>Phase 3</td>
<td>824</td>
<td>282</td>
<td>24</td>
</tr>
</tbody>
</table>

4. Trial of a novel and innovative therapy:
   ▶ $\pi_0 = 0.01; \alpha_2 = 0.05; \beta_2 = 0.15; \alpha_3 = 0.05; \beta_3 = 0.80$
   ▶ Results:

<table>
<thead>
<tr>
<th>Trials</th>
<th>True</th>
<th>False</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 2</td>
<td>13245</td>
<td>20</td>
<td>656</td>
</tr>
<tr>
<td>Phase 3</td>
<td>675</td>
<td>16</td>
<td>33</td>
</tr>
</tbody>
</table>
The Public Health Objective
How do clinical trials determine PPV?

Sensitivity to $\beta_3$ (ultimate sensitivity for effective therapies)

5. Sufficiently powered phase III ($\beta_3 = 0.975$)
   - $\pi_0 = 0.10; \alpha_2 = 0.05; \beta_2 = 0.15; \alpha_3 = 0.05; \beta_3 = 0.975$
   - Results:
     | Trials | True Pos | False Pos | PPV |
     |--------|---------|-----------|-----|
     | Phase 2 | 9091    | 136       | 409 | 0.25 |
     | Phase 3 | 545     | 133       | 20  | 0.87 |

6. Underpowered phase III:
   - $\pi_0 = 0.10; \alpha_2 = 0.05; \beta_2 = 0.15; \alpha_3 = 0.05; \beta_3 = 0.50$
   - Results:
     | Trials | True Pos | False Pos | PPV |
     |--------|---------|-----------|-----|
     | Phase 2 | 15385   | 231       | 692 | 0.25 |
     | Phase 3 | 923     | 115       | 35  | 0.77 |
The Public Health Objective
How do clinical trials determine PPV?

### Sensitivity to \( \alpha \) (false positive risk; specificity)

#### 7. Relax phase II alpha \((\alpha_2 = 0.20)\)
- \( \pi_0 = 0.10; \alpha_2 = 0.20; \beta_2 = 0.15; \alpha_3 = 0.05; \beta_3 = 0.80 \)
- Results:

<table>
<thead>
<tr>
<th>Trials</th>
<th>True Pos</th>
<th>False Pos</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 2</td>
<td>6780</td>
<td>102</td>
<td>1220</td>
</tr>
<tr>
<td>Phase 3</td>
<td>1322</td>
<td>81</td>
<td>61</td>
</tr>
</tbody>
</table>

#### 8. Relax both phase II and III alpha \((\alpha_2 = 0.2, \alpha_3 = 0.10)\):
- \( \pi_0 = 0.10; \alpha_2 = 0.20; \beta_2 = 0.15; \alpha_3 = 0.10; \beta_3 = 0.80 \)
- Results:

<table>
<thead>
<tr>
<th>Trials</th>
<th>True Pos</th>
<th>False Pos</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 2</td>
<td>6780</td>
<td>102</td>
<td>1220</td>
</tr>
<tr>
<td>Phase 3</td>
<td>1322</td>
<td>81</td>
<td>122</td>
</tr>
</tbody>
</table>
The Public Health Objective
How do clinical trials determine PPV?

Summary: PPV as a function of $\pi_0$, $\alpha$, and $\beta$

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$\pi_0$</th>
<th>$\alpha_2$</th>
<th>$\beta_2$</th>
<th>$\alpha_3$</th>
<th>$\beta_3$</th>
<th>Drugs Evaluated</th>
<th>True Pos</th>
<th>False Pos</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.10</td>
<td>*</td>
<td>*</td>
<td>0.05</td>
<td>0.800</td>
<td>1000</td>
<td>98</td>
<td>45</td>
<td>0.685</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td>0.800</td>
<td>12500</td>
<td>150</td>
<td>28</td>
<td>0.842</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td>0.800</td>
<td>11765</td>
<td>282</td>
<td>24</td>
<td>0.923</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td>0.800</td>
<td>13265</td>
<td>16</td>
<td>33</td>
<td>0.327</td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td>0.975</td>
<td>9091</td>
<td>133</td>
<td>20</td>
<td>0.867</td>
</tr>
<tr>
<td>6</td>
<td>0.10</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td>0.500</td>
<td>15385</td>
<td>115</td>
<td>35</td>
<td>0.769</td>
</tr>
<tr>
<td>7</td>
<td>0.10</td>
<td>0.20</td>
<td>0.15</td>
<td>0.05</td>
<td>0.800</td>
<td>6780</td>
<td>81</td>
<td>61</td>
<td>0.571</td>
</tr>
<tr>
<td>8</td>
<td>0.10</td>
<td>0.20</td>
<td>0.15</td>
<td>0.10</td>
<td>0.800</td>
<td>6780</td>
<td>81</td>
<td>122</td>
<td>0.400</td>
</tr>
</tbody>
</table>
The Public Health Objective
PPV and good science?

Summary remarks: How to get high PPV with fewer trials:
See also paper: EfficiencyForTargetedTX.pdf on website.

- Design for scientifically informative negative trials
  * All trials (positive or negative) must reduce the number of viable hypotheses.
- Accept that no means no
  * Example 2 (never give up): Avoid inflating $\alpha$ with
    - multiple endpoints
    - subgroup analyses
    - surrogate endpoints
  * Example 3 (try try again): Avoid recycling ideas
- Assure power ($\beta$)
  - Good practice reduces variability
  - Good recruitment/retention
  - Adequate sample size
- Avoid development programs with low pre-test probability ($\pi_0$)
  - "Novel" and "innovative" approaches have low $\pi_0$. 

PPV as the public health objective
Summary remarks

- A wide range of situations/therapies are studied in clinical trials.
- Globally, clinical trials need to assure:
  - Scientific credibility
  - Ethical experiments
  - Efficient experiments:
    - Minimize time
    - Minimal number of extra subjects
    - Minimize cost
  - A high prevalence of truly beneficial therapies among all therapies used in routine care.
## 0. Motivation and context

### Summary remarks

- Empirical science is about using observations on a sample to make inference about a population.
- Achieving the public health objective requires:
  - A meaningful population characteristic (appropriate “endpoint" or “outcome") *(Sec 2.1)*
  - A representative sample *(Sec 3)*
  - An unbiased and efficient estimate from the data *(Sec 2)*
- In this introduction we have seen examples of several types of studies
  - Descriptive studies
  - Analytic-Observational studies (case-control and cohort studies)
  - Analytic-Experimental studies (RCT's)
- Summary of study structure and examples so far...
0. Motivation and context: Summary remarks

Empirical science and the research process:

Research Process

Underlying Population

\[ \mu \text{ denotes unknown center} \]

Sample

Inference about \( \mu \)

Statistics

Sample summary measure: \( \bar{X} \)
0. Motivation and context: Summary remarks

Example: descriptive epidemiology

- Risk of NTD:
  - Sample: representative sample of pregnant women
  - Data:
    \[ \bar{X} = \text{proportion of births with NTD} \]
  - Inference:
    \[ \mu = \text{risk of NTD in the population} \]
Schematic representation: Analytic-observational studies

Underlying Population A
\( \mu_A \) - Sample A
Statistics A \( \bar{X}_A \)

Inferential Question:
\( \mu_A = \mu_B \)

Underlying Population B
\( \mu_B \) - Sample B
Statistics B \( \bar{X}_B \)
Example: Analytic-observation studies (Case-control study)

- **Participants**: Individuals with disease (cases); Individuals without disease (controls).
- **Comparison**: Compare risk factors in cases and controls.
- **Advantages**: Good for rare diseases (and common exposures); relative quick.
- **Disadvantages**: Time sequence between exposure and disease initiation is not always apparent; confounding.

**Example**: Case-control study of dietary folate and NTD risk:

- **Population A**: Women with NTD birth ("cases")
  
  \[ \mu_A = \text{true risk of low folate in NTD-birth population} \]
  \[ \bar{X}_A = \text{proportion with low folate in NTD-birth sample} \]

- **Population B**: Women without NTD-birth ("controls")
  
  \[ \mu_B = \text{true risk of low folate in normal-birth population} \]
  \[ \bar{X}_B = \text{proportion with low folate in normal-birth sample} \]

- **Inference is usually about the odds ratio**:
  
  \[ \theta = \frac{\mu_A/(1 - \mu_A)}{\mu_B/(1 - \mu_B)} \]
0. Motivation and context: Summary remarks

Example: Analytic-observation studies (Cohort study)

- **Participants**: Exposed individuals; unexposed individuals.
- **Comparison**: Compare disease incidence in exposed and unexposed groups.
- **Advantages**: Good for rare exposures (and common diseases); prospective design gives correct time sequence (exposure precedes outcome) and lower risk of confounding.
- **Disadvantages**: Potential for imbalances make it difficult to infer causality.
- **Example**: Cohort study of dietary folate and NTD risk:
  - Population A: Women with low dietary folate
    \[
    \mu_A = \text{true risk of NTD birth in low folate population} \\
    \bar{X}_A = \text{proportion with NTD birth in low folate sample}
    \]
  - Population B: Women with normal folate
    \[
    \mu_B = \text{true risk of NTD birth in normal folate population} \\
    \bar{X}_B = \text{proportion with NTD birth in normal folate sample}
    \]
  - Inference is usually about risk difference or relative risk:
    \[
    \theta = \mu_A - \mu_B \quad \text{risk difference} \\
    \theta = \mu_A / \mu_B \quad \text{relative risk}
    \]
Schematic representation: Analytic-experimental studies

Underlying Population:

μ_A: Hypothetical mean, treatment A
μ_B: Hypothetical mean, treatment B

Inferential Question:

μ_A = μ_B

Sample

Treatment A

Statistics:

X_A
X_B

Treatment B
0. Motivation and context: Summary remarks

Example: Analytic-experimental study (Randomized trial – RCT)

- **Participants**: Sample of relevant study population; intervention (treatment vs control) assigned as part of study.
- **Comparison**: Outcome compared across groups.
- **Advantages**: Best design for studying causality; exposure occurs before outcome; confounding risk is minimized.
- **Disadvantages**: Not always ethical or practical.
- **Note**: RCT is a cohort study in which exposure is assigned by investigators. This is a defining distinction.

**Example**: MRC vitamin trial:

- **Population A**: Women randomized to folate supplement
  \[ \mu_A = \text{true risk of NTD in population given folate supplement} \]
  \[ \bar{X}_A = \text{proportion with NTD in folate supplement group} \]

- **Population B**: Women randomized to control supplement
  \[ \mu_B = \text{true risk of NTD in population without supplement} \]
  \[ \bar{X}_B = \text{proportion with NTD in control group} \]

- Inference is usually about risk difference (or relative risk):
  \[ \theta = \mu_A - \mu_B \quad \text{risk difference} \]
  \[ \theta = \mu_A / \mu_B \quad \text{relative risk} \]
0. Motivation and context

0.1 Examples
Case Study 1: Neural tube defects
Case Study 2: Selenium supplementation

0.2 Public health objective
Cervical cancer screening example
PPV as the public health objective

0. Motivation/context: Summary remarks

Outline – where to from here (next week)

0. Motivation, context, and objectives (Ch 1-4)

1. Specifying the study setting and objectives
   1.1 Defining the study population (Ch 12)
   1.2 Defining the study type (Ch 6)
   1.3 Defining the interventions (Ch 4, 6)