ISPE Mid Year Meeting 2011
Introduction to pharmacoepidemiology

Cohort studies
Principles
Measures of disease occurrence
Cohort studies in pharmacoepidemiology

Frank Andersohn, MD, PhD
frank.andersohn@charite.de

Aka: Almut Winterstein, BSPharm, PhD
almut@cop.ufl.edu
What is a cohort?

Cohort = basic tactical unit of a Roman legion
What is a cohort?

In Epidemiology

• Well-defined group of subjects

• Followed over time

• Observed for outcome(s) of interest

• Aim: To study how health status of cohort members change over time

• For instance: How many subjects develop diabetes within one year of “follow-up”? 
What is a cohort?

- Can focus on one similar group of subjects (e.g. treated with drug A)
- Most often: Comparison of (at least) two groups of subjects:

Treated with Drug A

- 20 with new diabetes
- 80 without

Treated with Drug B

- 10 with new diabetes
- 90 without
What is a cohort?

Treated with Drug A

20 with new diabetes
80 without

Treated with Drug B

10 with new diabetes
90 without

Including diabetics in this study does not make sense

=>

Include only subjects „at risk“

= subjects without outcome of interest at study start
What is the main outcome measures of a cohort study?

Cohort = study design to measure disease incidence

**Incidence** = measure of the **risk** of developing a certain (medical) condition

**Two important concepts of incidence**

1. Incidence proportion
2. Incidence density (=incidence rate)
Two types of incidence

Patient 1  10 years
Patient 2  4 years
Patient 3  7 years
Patient 4  8 years
Patient 5  10 years

■ Start of follow-up
■■ Coronary event

Incidence proportion = \( \frac{1}{5} = 0.2 = 20\% \)

Incidence density = \( \frac{1}{10+8+7+4+10} \) person years

(Incidence rate)

= \( \frac{1}{39} \ PY = 0.0256 \) / 1 PY = **25.6 per 1000 PY**
Incidence proportion

Figure of a randomized controlled trial (RCT) – it’s also a cohort study!
Incidence proportion

\[
\text{Incidence proportion} = \frac{\text{Number of subjects with event}}{\text{Number of subjects at risk}}
\]

4 / 10 = 0.4

7 / 10 = 0.7
Incidence proportion

4 / 10 = 0.4

7 / 10 = 0.7

What are problems with this type of incidence?
Incidence proportion

- Incidence proportion is 100% in both scenarios
- But risk of death is obviously higher in left scenario...
- Reason: needs to be defined – must be fixed across all subjects
- Incidence density more flexible

Adapted from Rothman 2008. Modern Epidemiology.
Incidence proportion

Problem 2 with incidence proportion: competing causes

• Study: Very effective drug to prevent death in patients with pancreatic cancer
• RCT vs. placebo, Follow-up 10 years
• Outcome = Incidence proportion of dementia
• What will the study show?

⇒ Increased “risk” (incidence proportion) for drug vs. placebo
⇒ Reason: Many people treated with placebo die before they can develop dementia
⇒ Most often we are not interested in these kind of “effects”
⇒ Solution: Consider time at risk
Problem 3 with incidence proportion: losses-to-follow-up & changes in exposure status
Incidence density

Usually only first outcome of interest is considered

→ follow-up ends after first event

Number of outcomes of interest

Incidence density = \[\frac{\text{Number of outcomes of interest}}{\text{Person time at risk}}\]

(or incidence rate)
## Two types of incidence

<table>
<thead>
<tr>
<th>Incidence proportion</th>
<th>Number of subjects with event</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--------------------------------</td>
</tr>
<tr>
<td></td>
<td>Number of subjects at risk</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incidence density</th>
<th>Number of outcomes of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--------------------------------</td>
</tr>
<tr>
<td>(Incidence rate)</td>
<td>Person time at risk</td>
</tr>
</tbody>
</table>

### Why incidence density is the more appropriate measure in most situations

- Accounts for time to event
- Accounts for competing causes
- Accounts for differences in follow-up times
- Helps to account for changes in exposure status
Incidence rate ratio (IRR)

- Ratio of two incidence rates
- Measure of increased / decreased risk (as compared to a reference)

\[
\text{IRR} = \frac{\text{Incidence rate in exposed}}{\text{Incidence rate in unexposed}}
\]
Cohort studies – classification of exposure

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>PTime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>127</td>
</tr>
</tbody>
</table>
Cohort studies – classification of exposure

<table>
<thead>
<tr>
<th>PTime</th>
<th>Current</th>
<th>Non</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>SUM</td>
<td>127</td>
<td>36</td>
</tr>
</tbody>
</table>
Often used in pharmacoepidemiology: Classification according to timing of use, e.g.
- Current Use
- Recent Use
- Past Use

• Drug effect?
• Withdrawal effect?

➔ Don’t know!
Cohort studies – classification of exposure

Sometimes, adding an additional amount of person time after end of (observed) exposure makes sense (e.g. long half life of a drug)

- Based on prior knowledge / hypotheses
- Should be evaluated in sensitivity analyses (Results stable? If not, check your hypotheses...
Cohort studies – classification of exposure

Classification according to dose and duration of use

- Current low
- Current high
- Past

- Current low
- Past

- Current high
- Past

- Current high
- Current high
- Past

- Current ≤1 year
- Current 1-3 years
- Current >3 years

Time
### Cohort studies – classification of exposure

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>PTIme</th>
<th>Current</th>
<th>Past</th>
<th>Non</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>10</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>8</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
</tbody>
</table>

**SUMME** 127 36 9 82
## Cohort studies – classification of exposure

<table>
<thead>
<tr>
<th>Category</th>
<th>Event Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current Use</strong></td>
<td>2 events in 36 years = 2/36 = 55.6 / 1000 years</td>
</tr>
<tr>
<td><strong>Past Use</strong></td>
<td>1 events in 9 years = 1/9 = 111.1 / 1000 years</td>
</tr>
<tr>
<td><strong>Non Use</strong></td>
<td>1 event in 82 years = 1/82 = 12.2 / 1000 years</td>
</tr>
</tbody>
</table>

### IRR (Current vs. Non)

\[
55.6 / 12.2 = 4.6
\]
Cohort studies

To-Do-List for a (simple) cohort study

1. Define the population of interest (e.g. User of Drug A or Drug B)
2. Define how these subjects are sampled from the underlying source population
3. Define in- and exclusion criteria
4. Explicitly (!) define date of cohort entry and cohort exit
5. Classify person time according to exposure
6. Sum-up person-time of the different exposure categories
7. Assign number of events to these categories
8. Calculate incidence(s) and incidence rates
Cohort studies

1. Define the population of interest (e.g. User of Drug A or Drug B)
2. Define how these subjects are sampled from the underlying source population
3. Define in- and exclusion criteria

• Similar to a randomized trial
• In- and exclusion criteria:
  • Well-defined population
  • Focus on population of interest (e.g. Patients with ADHD)
  • Can be used to reduce bias from confounding ("restriction")
  • Increase study efficiency (high background rate)
  • Etc.
Cohort studies

1. Define the population of interest (e.g. User of Drug A or Drug B)
2. Define how these subjects are sampled from the underlying source population
3. Define in- and exclusion criteria

Too narrow inclusion / exclusion criteria: Disadvantages

- Like in a RCT
- Conclusions for excluded populations not possible
- Not possible to investigate interaction between exclusion factor (e.g. ADHD) and exposure
  - Interaction = Relationship between exposure and outcome is modified by other factor
Cohort studies

4. Explicitly (!) define date of cohort entry and cohort exit

When to start follow-up? E.g.
- Calendar day (Jan 1, 2001)
- Event (Day of ADHD diagnosis; first prescription of central nervous stimulant)
- Time (365 days after inclusion into a registry)
- etc.

When to stop follow-up?
- End of study period (Dec 31, 2008)
- Death / Loss-To-Follow-Up
- Day of first occurrence of outcome of interest

(which ever comes first)
Cohort studies – start of follow-up

Exposure-based cohorts

- often used in pharmacoepi studies
- looks like a RCT (Cave!)
- Follow-up starts with exposure $\rightarrow$ efficient
- Interpretation more straightforward
- Problems might occur in defining start of follow-up in unexposed subjects ("immortal time bias")
- Prevalent users $\rightarrow$ might introduce bias; new user design superior
Cohort studies – start of follow-up

Exposure based cohorts – incident or prevalent user?

Study period

01.01.1998

prevalent user

prevalent user

incident user

incident user

Study period
Cohort studies – hazard function

Deep venous thrombosis

Cancer

Hazard (instantenous risk)

time
Cohort studies – hazard function

Incident user of Drug A

\( \text{IRR} = 1 \) (correct)

Incident user of Drug B
Cohort studies – hazard function

Prevalent user of Drug A

Incident user of Drug B

⇒ IRR >> 1 (Bias!)

Also known as:
- Healthy survivor bias
- Depletion of susceptible
Cohort studies – start of follow-up

Exposure based cohorts – incident or prevalent user?

- Prefer incident user over prevalent user
- Long enough “Look-book” not always possible → include sensitivity analyses to study “depletion-of-susceptible” bias
Cohort studies – the comparator

What is the reference (comparator)?

- Ideal: Head-To-Head with very similar drug
  - Reduces confounding
  - Similar groups of patients (indication of use, comorbidity, etc.)
  - New-user design preferred

- Head-To-Head with inactive comparator drug
  - For outcome acute MI → glaucoma medication (no association with outcome)
  - Easy to define cohort entry
  - It is ensured that all patients are under medical observation

- Comparison with non-use
  - Non-user often substantially different than user → confounding ↑↑
  - If feasible → Interpretation “straight forward”
- Possible to study rare exposures
- Multiple outcomes can be studied (smoking --> lung cancer, COPD, larynx cancer)
- Exposure is assessed before outcome (no recall bias as in case control studies if prospective design)
- Time-to-event analysis is possible
- Possible to estimate absolute risks (incidences)
Cohort studies - Disadvantages

- Long duration (if prospective)
- Expensive
- Not efficient for rare outcomes
- Often not possible to study multiple exposures
- Might be problematic for diseases with long latency
- Environmental changes over time can influence exposure-outcome association
COX-2 selective non-steroidal anti-inflammatory drugs and risk of serious coronary heart disease

Wayne A Ray, C Michael Stein, James R Daugherty, Kathi Hall, Patrick G Arbogast, Marie R Griffin

Results of premarketing and postmarketing trials have raised doubts about the cardiovascular safety of the non-steroidal anti-inflammatory drug (NSAID) rofecoxib, especially at doses greater than 25 mg. Between Jan 1, 1999, and June 30, 2001, we did a retrospective cohort study of individuals on the expanded Tennessee Medicaid programme (TennCare), in which we assessed occurrence of serious coronary heart disease (CHD) in non-users (n=202,916) and in users of rofecoxib and other NSAIDs (rofecoxib n=24,132, other n=151,726). Participants were aged 50–84 years, lived in the community, and had no life-threatening non-cardiovascular illness. Users of high-dose rofecoxib were 1.70 (95% CI 0.98–2.95, p=0.058) times more likely than non-users to have CHD; among new users this rate increased to 1.93 (1.09–3.42, p=0.024). By contrast, there was no evidence of raised risk of CHD among users of rofecoxib at doses of 25 mg or less or among users of other NSAIDs.

The results of VIGOR,1 a large clinical trial, indicated that individuals who took rofecoxib 50 mg were five times more likely to have a myocardial infarction than were those who took naproxen 1000 mg.2 Furthermore, findings of subsequent observational studies3,4 of naproxen 1000 mg have not shown a protective effect sufficient to explain this difference. Therefore, adverse cardiovascular events could be a side-effect of the selective inhibitor of cyclo-oxygenase 2 (COX-2) and non-steroidal anti-inflammatory drug (NSAID) rofecoxib. Our aim was to compare risk of acute myocardial infarction (AMI) and fatal coronary heart disease (CHD) in users of rofecoxib, with risk in users of other frequently prescribed NSAIDs in the expanded Tennessee Medicaid programme, TennCare. Our study included an a priori analysis of rofecoxib at doses greater than 25 mg because such doses, not presently recommended for long-term (>5 days) use,5 could be uniquely associated with adverse cardiovascular effects.6,7

Lancet 2002; 360: 1071–73


Example

**Cohort entry and cohort exit**

We included individuals in the cohort from the first study day of current use, and stopped follow-up at the end of study eligibility, after 365 days of no NSAID use, or at time of switching from one NSAID to another.

**Classification of exposure**

We classified every person-day of cohort membership, according to NSAID use, as current (date prescription filled through end of days’ supply), former (use during past 365 days), or non-use.

**Analysis**

We estimated adjusted incidence rate ratios (IRR) for NSAID exposure groups from a Poisson regression model.
<table>
<thead>
<tr>
<th></th>
<th>Non-user (n=202 916)</th>
<th>Ibuprofen (n=59 007)</th>
<th>Naproxen (n=70 384)</th>
<th>Celecoxib ≤25 mg (n=22 337)</th>
<th>Rofecoxib &gt;25 mg (n=3887)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (mean, SD) (years)</strong></td>
<td>61.8 (9.0)</td>
<td>60.4 (8.1)</td>
<td>50.4 (8.1)</td>
<td>63.7 (8.9)</td>
<td>60.6 (8.1)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td>127 458 (63%)</td>
<td>40 661 (69%)</td>
<td>48 592 (69%)</td>
<td>16 280 (73%)</td>
<td>14 830 (73%)</td>
</tr>
<tr>
<td><strong>White</strong></td>
<td>151 568 (75%)</td>
<td>40 065 (68%)</td>
<td>49 626 (71%)</td>
<td>16 246 (73%)</td>
<td>15 561 (77%)</td>
</tr>
<tr>
<td><strong>TennCare enrolment, uninsured‡</strong></td>
<td>74 718 (37%)</td>
<td>18 247 (31%)</td>
<td>23 054 (33%)</td>
<td>5 780 (26%)</td>
<td>5 884 (29%)</td>
</tr>
<tr>
<td><strong>Treatment for cardiovascular problems in past year‡</strong></td>
<td>155 681 (77%)</td>
<td>49 684 (84%)</td>
<td>58 884 (84%)</td>
<td>19 778 (89%)</td>
<td>17 618 (87%)</td>
</tr>
<tr>
<td><strong>Major cardiovascular disease§</strong></td>
<td>69 150 (34%)</td>
<td>23 213 (39%)</td>
<td>27 011 (38%)</td>
<td>9 625 (43%)</td>
<td>8 507 (42%)</td>
</tr>
<tr>
<td><strong>Cardiovascular drug¶</strong></td>
<td>150 846 (74%)</td>
<td>48 183 (82%)</td>
<td>57 186 (81%)</td>
<td>19 375 (87%)</td>
<td>17 243 (85%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Person-years</th>
<th>Events</th>
<th>Rate/1000</th>
<th>Adjusted IRR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-user</strong></td>
<td>237 975</td>
<td>3085</td>
<td>13.0</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>New user during study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>4319</td>
<td>52</td>
<td>12.0</td>
<td>1.01 (0.77–1.33)</td>
</tr>
<tr>
<td>Naproxen</td>
<td>6489</td>
<td>72</td>
<td>11.1</td>
<td>0.92 (0.73–1.16)</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>4509</td>
<td>55</td>
<td>12.2</td>
<td>0.88 (0.67–1.16)</td>
</tr>
<tr>
<td>Rofecoxib ≤25 mg</td>
<td>3430</td>
<td>47</td>
<td>13.7</td>
<td>1.02 (0.76–1.37)</td>
</tr>
<tr>
<td>Rofecoxib &gt;25 mg</td>
<td>500</td>
<td>12</td>
<td>24.0</td>
<td>1.93 (1.09–3.43)</td>
</tr>
</tbody>
</table>
Analyses of cohort studies

1. Crude estimates (Events / person time for different exposure categories)

2. Stratified analyses (e.g. by age, sex)

3. Multivariate modeling
   Appropriate methods for person-time based analyses (incidence density)
   - Poisson-Regression (models impact of measured variables on events per person-time)
   - COX-proportional hazard regression (models impact of measured variables on hazard ratio)
   If follow-up time is the same for all subjects
   - Can use logistic regression (models probability of developing outcome of interest given the measured variables)
FAQs

• Can one match in cohort studies?
• Can one vary exposure status in cohort studies?
• Can one study multiple exposures?
• Why are there “retrospective cohort studies”? 
• Can one study multiple outcomes?
• Can subjects enter the cohort at different points in time?
• How do I determine when I should start to measure risk and when to stop?
I was lured into epidemiology by a friend in environmental engineering.

“But don't worry,” he assured me. “You don't have to take any classes or anything. It's not a real science like chemistry or physics.”

Kaufman JS, Epidemiology 2009
Thank you.