Journey through Clinical Characterization: Large-Scale Honest Incidence

George Hripcsak, MD, MS
Columbia University Medical Center
NewYork-Presbyterian Hospital

2017 OHDSI Symposium
All-by-All Incidence

• Ask a doctor important side effects of a drug
• Then ask the incidence of that side effect
  – Many side effects are well known, but most clinicians have no idea of the incidence
  – The evidence is sparse
• Start simple
  – Characterization = non-causal rates
  – Tally how often conditions occur in drug therapy
Why start simple?

• If incidence is low, then I am set
• If incidence is high, then need to look out for it even if not caused by drug
• Feasible to execute all-by-all
• Fewer assumptions
• More complicated than it looks, so need to get this one right first

“When I start this drug, what is the chance that I’ll experience a condition in the next year?”
Dissecting the anatomy of incidence

Incidence metrics:

Incidence proportion = \frac{\text{# persons in the target cohort who have new outcome occurrence during the time-at-risk}}{\text{# persons in the target cohort with time-at-risk}^*}

Incidence rate = \frac{\text{# persons in the target cohort who have new outcome occurrence during the time-at-risk}}{\text{person-time at-risk for persons in the target cohort with time-at-risk}^*}
Let target cohort be new users of warfarin.

Incidence we observe:

- Baseline risk of stroke off warfarin

Warfarin causing hemorrhage

Warfarin averting emboli

Attributable effect of warfarin

Incidence we observe: net rate of stroke on warfarin
Incidence rates do not tell causal effect (attributable risk or benefit)

- **Incidence on warfarin**
  - Baseline: 24%
  - Causal additive effect: 22%
  - Causal preventive effect: 20%
  - Total: 18%

- **Benefit of warfarin**

- **Incidence on aspirin**
  - Baseline: 16%
  - Causal additive effect: 14%
  - Causal preventive effect: 12%
  - Total: 10%

- **Benefit of aspirin**
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

• How should the target cohort be defined?
• How should the outcome be defined?
• How should the time-at-risk be defined?
• How to account for patients with incomplete time-at-risk?
• Which statistical metrics should be reported?
• Which data should be used?
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

• How should the target cohort be defined?
  – For a cohort of ‘new users of a drug’, cohort entry can be defined as the date of first exposure
    • Should other inclusion criteria be imposed, such as requiring prior diagnosis of labeled indication? How do these criteria impact the generalizability of this estimate to the target population?
  – What minimum lookback period is required to ensure ‘new user’?
    • Shorter period provides larger (and more generalizable) sample to yield more precise estimate
    • Longer period provides greater confidence that patient is truly ‘newly exposed’ and provides longer prior history to ensure outcome is incident occurrence
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

- **How should the outcome be defined?**
  - Alternative phenotype definitions often represent different sensitivity/specificity tradeoffs, though those operating characteristics are commonly unknown at the time of choosing the definition
  - ‘First diagnosis’ may be more sensitive but less specific than ‘first diagnosis with hospitalization’
  - Outcome cohort can include ‘first ever occurrence’ vs. ‘first occurrence post-exposure’ vs. ‘all occurrences’
  - Phenotype evaluation diagnostics required to quantify potential measurement error and calibrate incidence estimates
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

• **How should the time-at-risk be defined?**

  - **Cohort entry**
  - **Cohort exit**

**Person timeline**

- **Observation period start**
- **Observation period end**

**Time-at-risk**

- 1 day to 30 day after cohort start
- 1 day to 365 day after cohort start

**Outcome occurrence**

- ‘on treatment’: cohort start through cohort end
- ‘intent-to-treat’: cohort start through observation period end
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

- **How to account for patients with incomplete time-at-risk?**

  - Include persons with incomplete follow-up time
    - Assumes unobserved time did not have events
    - Lower bound of true incidence estimate
      \[ \frac{\#\text{observed}_\text{events}}{\#\text{observed}_\text{events} + \#\text{missed}_\text{events}} \]
    - Worsens with increased censoring or more events in censored pts
  
  - Include only persons with full time-at-risk
    - Usually higher than true incidence estimate (if rate is uniform)
      \[ \approx \frac{\#\text{observed}_\text{events}}{\#\text{observed}_\text{events} - \#\text{missed}_\text{events}} \]
    - Worsens with increased censoring (also smaller sample size)
    - Can flip if high rate of events in censored period
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

- **Which statistical metrics should be reported?**
  - Incidence proportion requires a defined time-at-risk
  - Incidence rate allows variable-length time-at-risk, but assumes constant hazard over time-at-risk
  - 95% confidence intervals commonly reported, but only represent sampling variability.
  - Within-source systematic error and between-source heterogeneity represent larger sources of uncertainty that are not adequately quantified in current practice
  - Characterizing the range of estimates across network analysis (e.g. minimum → maximum) may be more reflective of uncertainty than sampling statistics from any given data source
Myriad difficult choices that researchers have to make to produce a ‘simple answer’

• **Which data should be used?**
  – Incidence estimation requires a minimum longitudinal follow-up for the desired time-at-risk
  – Data should be represent patients that are contained within the target population of interest (but not necessarily be a random sample or fully representative of the target population)
  – A network analysis may provide heterogeneity across patients, health systems, geographies and represent different perspectives and health care process biases
Hierarchy of uncertainty

• Biology (genetics)
  – This is signal that you want to measure, not error
• Environment (i.e., its effect on biology)
  – Also signal that you want to measure
• Health care process bias
  – Measurement error
• Extract-transfer-load
  – ETL errors, and ETL interpretations
• Sampling error
  – Sampling error goes to zero with sample size
• Confounding
  – Different confounders in different populations
Problems with current practice

- For a majority of incidence questions of potential interest, there is no readily accessible evidence available.
- When evidence is identified in the literature, it can be difficult to interpret:
  - Incidence metric – ambiguity in what’s reported
  - Unspecified time-at-risk
  - Generalizability of target population
  - Diversity of phenotype definitions
  - Different evidence sources (RCT, systematic reviews, observational studies)
    - Systematic reviews synthesize results from different metrics/time-at-risk/phenotypes
  - Observational data have different sources of systematic error that are rarely quantified or corrected for.
Inspiration from Woody Allen

Two elderly women are at a Catskill mountain resort, and one of them says, "Boy, the food at this place is really terrible." The other one says, "Yeah, I know; and such small portions."
How could OHDSI help?

• Develop a standardized framework for incidence evidence generation and dissemination
• Fill the gaps where there is currently no available evidence
• Augment existing knowledge with new evidence systematically generated across the world’s largest observational data network
  – Demonstrate reliability of current knowledge through replication
  – Reconcile discordant evidence observed in the literature through quantification of uncertainty
  – Apply causal effect estimates to overall incidence to assess attributable risk
“Things we know that we know”

• What we think we know:
  – ACE inhibitors cause angioedema

• What we want to know:
  – Clinical characterization: Incidence of angioedema in patients exposed to ACE inhibitors
  – Population-level effect estimation:
    • Safety surveillance: Strength of association with ACE inhibitor vs. counterfactual
    • Comparative effectiveness: Strength of association with ACE inhibitor, relative to alternative treatments
    • Attributable risk
  – Patient-level prediction: Probability that a patient will experience event, given baseline characteristics
ANGIOEDEMA: Angioedema has been reported in patients receiving lisinopril (0.1%). Angioedema associated with laryngeal edema may be fatal. If angioedema of the face, extremities, lips, tongue, glottis and/or larynx occurs, treatment with lisinopril should be discontinued and appropriate therapy instituted immediately. (See WARNINGS.)
What’s the published evidence?

<table>
<thead>
<tr>
<th>Publication</th>
<th>Person-years</th>
<th>Events</th>
<th>Incidence (per 1000 person-years)</th>
<th>95% CI (Incidence rate per 1000 person-years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller Hypertension 2008</td>
<td>179,088</td>
<td>352</td>
<td>1.97</td>
<td>(1.76-2.17)</td>
</tr>
<tr>
<td>Makani Am J Cardiol. 2012</td>
<td>185,067</td>
<td>394</td>
<td>3.00</td>
<td>(2.80-3.20)</td>
</tr>
<tr>
<td>Toh AIM 2012</td>
<td>753,105</td>
<td>3,301</td>
<td>4.38</td>
<td>(4.23-4.53)</td>
</tr>
</tbody>
</table>

Incidence rate interval estimate predicated on 2 assumptions:
- Observed data represents a random sample of a target population
- Estimator in unbiased, so no systematic error
How does it get distilled to clinicians?

The overall incidence of angioedema related to ACE inhibitors has been estimated between 0.1 percent and 0.7 percent [1-5,14-18]. However, the lower end of this range may overlap with the background rate of angioedema in the general population. In the TRANSCEND trial of ACE inhibitor-intolerant individuals given an angiotensin II receptor blocker (ARB) or placebo, rates of angioedema were 0.07 and 0.1 percent in the ARB and placebo groups, respectively [17].
What if a standardized incidence estimation was consistently applied across the OHDSI network?

Range of incidence proportions from across 8 sources in the OHDSI data network: 0.1% - 0.8%
How does OHDSI evidence compare with prior evidence?
ACE inhibitors have many potential side effects listed on the product label.

Label listed what appears to be ‘attributable’ risk but not absolute risk.
How is risk of cough among ACE inhibitors summarized in UpToDate?

**Cough** — A dry, hacking cough has been described in 5 to 20 percent of patients treated with an ACE inhibitor [21]. The best data come from a meta-analysis of 29 trials in which cough was noted in 9.9 percent of patients treated with ACE inhibitors [22,23]. In the ONTARGET trial, cough sufficiently severe to discontinue the drug was observed in 4.2 percent of the patients treated with ramipril [3]. Cough is much less common with ARBs. (See 'ARBS' below.)
What evidence can we find in the literature?

**Angiotensin-Converting Enzyme Inhibitor Associated Cough: Descriptive Information from the Physicians' Desk Reference**

**RESULTS:** One hundred twenty-five studies that satisfied our inclusion criteria enrolled 198,130 patients. The pooled weighted incidence of cough for enalapril was 11.48% (95% confidence interval [CI], 9.54% to 13.41%), which was ninefold greater compared to the reported rate in the *PDR/drug label* (1.3%). The pooled weighted withdrawal rate due to cough for enalapril was 2.57% (95% CI, 2.40-2.74), which was 31-fold greater compared to the reported rate in the *PDR/drug label* (0.1%). The incidence of cough has increased progressively over the last 2 decades with accumulating data, but it has been reported consistently several-fold less in the *PDR* compared to the RCTs. The results were similar for most other ACE inhibitors.

**CONCLUSION:** The incidence of ACE inhibitor-associated cough and the withdrawal rate (the more objective metric) due to cough is significantly greater in the literature than reported in the *PDR/drug label* and is likely to be even greater in the real world when compared with the data from RCTs. There exists a gap between the data available from the literature and that which is presented to the consumers (prescribing physicians and patients).
What if a standardized incidence estimation was consistently applied across the OHDSI network?

UpToDate range: 5%-20%
Meta-analysis estimate: 9.5% -13.4%
OHDSI range: 3%-14%
Amongst new users of sertraline, how often do **suicidal thoughts and behavior** occur in a given time horizon?
Amongst new users of sertraline, how often does **suicidal thoughts and behavior** occur in a given time horizon?

**Clinical Worsening and Suicide Risk**

Patients with major depressive disorder (MDD), both adult and pediatric, may experience worsening of their depression and/or the emergence of suicidal ideation and behavior (suicidality) or unusual changes in behavior, whether or not they are taking antidepressant medications, and this risk may persist until significant remission occurs. Suicide is a known risk of depression and certain other psychiatric disorders, and these disorders themselves are the strongest predictors of suicide. There has been a long-standing concern, however, that antidepressants may have a role in inducing worsening of depression and the emergence of suicidality in certain patients during the early phases of treatment. Pooled analyses of short-term placebo-controlled trials of antidepressant drugs (SSRIs and others) showed that these drugs increase the risk of suicidal thinking and behavior (suicidality) in children, adolescents, and young adults (ages 18 to 24) with major depressive disorder (MDD) and other psychiatric disorders. Short-term studies did not show an increase in the risk of suicidality with antidepressants compared to placebo in adults beyond age 24; there was a reduction with antidepressants compared to placebo in adults aged 65 and older.
The pooled analyses of placebo-controlled trials in children and adolescents with MDD, obsessive compulsive disorder (OCD), or other psychiatric disorders included a total of 24 short-term trials of 9 antidepressant drugs in over 4400 patients. The pooled analyses of placebo-controlled trials in adults with MDD or other psychiatric disorders included a total of 295 short-term trials (median duration of 2 months) of 11 antidepressant drugs in over 77,000 patients. There was considerable variation in risk of suicidality among drugs, but a tendency toward an increase in the younger patients for almost all drugs studied. There were differences in absolute risk of suicidality across the different indications, with the highest incidence in MDD. The risk differences (drug vs. placebo), however, were relatively stable within age strata and across indications. These risk differences (drug-placebo difference in the number of cases of suicidality per 1000 patients treated) are provided in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Drug-Placebo Difference in Number of Cases of Suicidality per 1000 Patients Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;18</td>
<td>Increases Compared to Placebo</td>
</tr>
<tr>
<td>18–24</td>
<td>14 additional cases</td>
</tr>
<tr>
<td>25–64</td>
<td>Decreases Compared to Placebo</td>
</tr>
<tr>
<td>≥65</td>
<td>1 fewer case</td>
</tr>
<tr>
<td></td>
<td>6 fewer cases</td>
</tr>
</tbody>
</table>

Amongst new users of sertraline, how often does **suicidal thoughts and behavior** occur in a given time horizon?

Table 1 provides estimate of ‘attributable risk’, but does not provide baseline risk
Maybe. Why not measure?

Summary of clinical trial data abstracted from the Medicine and Healthcare products Regulatory Agency’s review of the safety of SSRIs

<table>
<thead>
<tr>
<th>SSRI (conditions included in RCTs; No of trials contributing data)</th>
<th>Active (SSRI) arm</th>
<th>Placebo arm</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of subjects</td>
<td>No of episodes</td>
<td>No of subjects</td>
</tr>
<tr>
<td><strong>(a) Suicides in placebo controlled trials in adults</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citalopram (depression; 9 trials)</td>
<td>1320</td>
<td>1</td>
</tr>
<tr>
<td>Escitalopram (all indications; 34 trials)</td>
<td>2648</td>
<td>1</td>
</tr>
</tbody>
</table>

Fluoxetine; the risk of suicidal thoughts was similar to that for non-fatal self harm (387/100 000 (177 episodes of suicidal thoughts among 45 704 subjects)). As the mean duration of the trials included in the synthesis was eight to 10 weeks, the overall rates of suicidal behaviour and thoughts per person year at risk are likely to be some five times higher than the risks calculated here.
What if a standardized incidence estimation was consistently applied across the OHDSI network?

Meta-analysis estimate: 0.38% (variable-length studies, median 8-10 wks)

OHDSI range: 0.06%-0.99% during 30 days after exposure start
Evaluating the impact of time-at-risk on incidence estimation

OHDSI range: 0.06%-0.99% during 30 days after exposure start

OHDSI range: 0.26%-3.67% during 365 days after exposure start
Amongst new users of sertraline, how often does **gastrointestinal bleeding** occur in a given time horizon?

**Abnormal Bleeding**

SSRIs and SNRIs, including sertraline, may increase the risk of bleeding events. Concomitant use of aspirin, non-steroidal anti-inflammatory drugs, warfarin, and other anticoagulants may add to this risk. Case reports and epidemiological studies (case-control and cohort design) have demonstrated an association between use of drugs that interfere with serotonin reuptake and the occurrence of gastrointestinal bleeding. Bleeding events related to SSRIs and SNRIs use have ranged from ecchymoses, hematomas, epistaxis, and petechiae to life-threatening hemorrhages.
UpToDate summary of SSRI risk of GI bleeding

Amongst new users of sertraline, how often does gastrointestinal bleeding occur in a given time horizon? This gives only OR.

Upper gastrointestinal bleeding — Multiple meta-analyses of observational studies suggest that SSRIs are associated with an elevated risk of upper gastrointestinal bleeding [76-78]; however, the absolute risk is low [75]. As an example, one meta-analysis compared the risk of upper gastrointestinal bleeding in SSRI users with the risk in non-SSRI users, pooling data from 22 observational studies (n >1,000,000 individuals, including more than 56,000 cases of bleeding) [79]. Exposure to SSRIs was associated with an increased risk of bleeding (odds ratio 1.6, 95% CI 1.4-1.8). The risk was even greater in the subgroup of patients who took SSRIs plus NSAIDS (odds ratio 3.7, 95% CI 3.0-4.7). By contrast, a separate subgroup analysis found that the risk of bleeding was comparable for patients who took SSRIs plus NSAIDS plus acid suppressing drugs and for patients who were not exposed to SSRIs. Based upon these findings, some clinicians use non-SSRI antidepressants in patients at high risk for bleeding (eg, prior history of upper gastrointestinal bleeding), or prescribe a proton pump inhibitor when SSRIs are used in conjunction with NSAIDS; however, this is not standard practice.
What if a standardized incidence estimation was consistently applied across the OHDSI network?

OHDSI range: 0.19%-2.33% during 365 days after exposure start
With outcome = diagnosis + hospitalization
Evaluating the impact of outcome definition on incidence estimation

OHDSI range: 0.19%-2.33% during 365 days after exposure start
With outcome = diagnosis + hospitalization

OHDSI range: 0.29%-5.54% during 365 days after exposure start
With outcome = diagnosis
Amongst new users of sertraline, how often does **sexual dysfunction** occur in a given time horizon?

According to the product label, **probably higher than what is actually reported**...
Amongst new users of sertraline, how often does sexual dysfunction occur in a given time horizon?

According to UpToDate, anywhere between ‘1 in 7’ to ‘6 in 7’
What if a standardized incidence estimation was consistently applied across the OHDSI network?

OHDSI range: **only 0.04%-1.17%** during 365 days after exposure start

Consistent (and likely inaccurate) estimate potentially due to common bias of measurement error across sources
Large-scale incidence estimation

• We have developed a standard framework for clinical characterization of outcome incidence

• We demonstrated its reliability across several examples
  – But also highlighted that (as with all observational studies) we cannot assure reliable results for all drugs and outcomes
Caveats to All-by-All Incidence

• Why might rate be high
  – (Recall that indications reduced b/c first occurrence is after exposure)
  – High in the underlying population
  – Indication is a risk
  – Things associated with indication
  – Reversed timing (Drug -> Indication)
  – Or could be causal (attributable risk)

• But if rate is low and side effect is not serious, then side effect may not be important
Caveats to All-by-All Incidence

• Current version based on billing codes
  – Only get side effects reported and worthy of billing

• Not good for discovering side effects
  – Simvastatin’s first 1000 are less interesting (probably associated with indication)
  – But those known to be side effects (e.g., from product label) match the (sparse) literature rates extremely well
All-by-All incidence

• I will probably use it frequently for personal questions, keeping caveats in mind
Demo

**How Often...**

How often do patients get a condition after starting a drug?

Which drug are you interested in?

Lisinopril

Which condition are you interested in?

Angioedema

**What this does**

Use this tool to look up the proportion of people starting a drug who are newly diagnosed with a condition within 1 year of starting the drug. You can search for a specific drug-condition incidence by entering your drug and condition of interest in the fields above. Or you can browse a list of conditions of potential interest by leaving the condition field blank, and you'll be shown conditions listed on the drug's product label.

**What this does not do**

This tool does not demonstrate that a drug causes a condition (i.e., that the condition is a side effect of the drug). Instead, for example, the condition may be part of the reason you are taking the drug, or the condition may just be common in the population.

This tool provides the overall observed risk in a population, but does not provide the attributable risk due to drug exposure. The results provided are raw unadjusted numbers for each diagnosis. The data made available through this site are for informational purposes only and are not a substitute for professional medical advice or services. You should not use this information for comparing drugs or making decisions related to diagnosing or treating a medical or health condition; instead, please consult a physician or healthcare professional in all matters related to your health.
Observations

• Uncertainty assessment
  – The lynchpin of reproducibility and honest evidence
  – Not just sampling variation; time to stop pretending
    • “Just an observational study”
    • At least get upper limit
  – Exploit the network to learn about uncertainty
    • (Although some bias is replicated across sites)
  – Learn to model full uncertainty

• Future steps
  – Target populations, restricted to treatments within specific indications?
  – Incidence risk stratification (e.g. age/gender)
  – When do you transition from clinical characterization to patient-level prediction? (think about this when you see Jenna’s talk)
We need you!

• We have shown proof-of-concept
• But this will only work if everyone contributes
• How can you help?
  – If you have data, run the “all-by-all” incidence analysis and share your results, which will be compiled into the open-source evidence repository
    • The more databases we get, the more honest we think our range of estimates will become
  – If you are a methods researcher, use the open-source evidence repository to develop new models for estimating credible uncertainty ranges
  – If you are an open-source developer, build a better user interface to share this evidence more broadly with all stakeholders, including providers and patients
Acknowledgments

– Supported by the JAYNE KOSKINAS TED GIOVANIS FOUNDATION FOR HEALTH AND POLICY, a Maryland private foundation dedicated to effecting change in health care for the public good.

  • The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and not necessarily those of the JAYNE KOSKINAS TED GIOVANIS FOUNDATION FOR HEALTH AND POLICY, its directors, officers, or staff.

• NLM R01 LM006910 “Discovering and Applying Knowledge in Clinical Databases”
Thanks

OHDSI collaborators
• Marc Suchard
• Martijn Schuemie
• David Madigan
• Jon Duke
• Patrick Ryan

Columbia team
• Ray Chen
• Mark Velez
• Karthik Natarajan
• Jungmi Han
• Peng Jin

OHDSI Infrastructure
• Lee Evans
Join the journey

http://ohdsi.org