OHDSI: Drawing reproducible conclusions from observational clinical data

George Hripcsak, MD, MS
Biomedical Informatics, Columbia University
Medical Informatics Services, NewYork-Presbyterian
August 2010: “Among patients in the UK General Practice Research Database, the use of oral bisphosphonates was not significantly associated with incident esophageal or gastric cancer”

Sept 2010: “In this large nested case-control study within a UK cohort [General Practice Research Database], we found a significantly increased risk of oesophageal cancer in people with previous prescriptions for oral bisphosphonates”
Observational Health Data Sciences and Informatics (OHDSI, as “Odyssey”)

Mission: To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care

A multi-stakeholder, interdisciplinary, international collaborative with a coordinating center at Columbia University

Aiming for 1,000,000,000 patient data network

http://ohdsi.org
OHDSI’s global research community

- >200 collaborators from 25 different countries
- Experts in informatics, statistics, epidemiology, clinical sciences
- Active participation from academia, government, industry, providers
- Over a billion records on >400 million patients in 80 databases

http://ohdsi.org/who-we-are/collaborators/
Why large-scale analysis is needed in healthcare

All health outcomes of interest
Patient-level predictions for personalized evidence require big data.

2 million patients seem excessive or unnecessary?

- Imagine a provider wants to compare her patient with other patients with the same gender (50%), in the same 10-year age group (10%), and with the same comorbidity of Type 2 diabetes (5%).
- Imagine the patient is concerned about the risk of ketoacidosis (0.5%) associated with two alternative treatments they are considering.
- With 2 million patients, you’d only expect to observe 25 similar patients with the event, and would only be powered to observe a relative risk > 2.0.

Aggregated data across a health system of 1,000 providers may contain 2,000,000 patients.
OHDSI’s approach to open science

- Open science is about sharing the journey to evidence generation
- Open-source software can be part of the journey, but it’s not a final destination
- Open processes can enhance the journey through improved reproducibility of research and expanded adoption of scientific best practices
Evidence OHDSI seeks to generate from observational data

• Clinical characterization
  – Natural history: Who has diabetes, and who takes metformin?
  – Quality improvement: What proportion of patients with diabetes experience complications?

• Population-level estimation
  – Safety surveillance: Does metformin cause lactic acidosis?
  – Comparative effectiveness: Does metformin cause lactic acidosis more than glyburide?

• Patient-level prediction
  – Precision medicine: Given everything you know about me, if I take metformin, what is the chance I will get lactic acidosis?
  – Disease interception: Given everything you know about me, what is the chance I will develop diabetes?
How OHDSI Works

Source data warehouse, with identifiable patient-level data

Standardized, de-identified patient-level database (OMOP CDM v5)

ETL

Standardized large-scale analytics

Analysis results

OHDSI Coordinating Center

Data network support
Analytics development and testing
Research and education

Summary statistics results repository

OHDSI Data Partners

OHDSI.org
Extensive vocabularies
Preparing your data for analysis

**WhiteRabbit**: profile your source data

**RabbitInAHat**: map your source structure to CDM tables and fields

**ATHENA**: standardized vocabularies for all CDM domains

**Usagi**: map your source codes to CDM vocabulary

**CDM**: DDL, index, constraints for Oracle, SQL Server, PostgreSQL; Vocabulary tables with loading scripts

**ACHILLES**: profile your CDM data; review data quality assessment; explore population-level summaries

**OHDSI Tools built to help**

**WhiteRabbit**: profile your source data

**RabbitInAHat**: map your source structure to CDM tables and fields

**ATHENA**: standardized vocabularies for all CDM domains

**Usagi**: map your source codes to CDM vocabulary

**CDM**: DDL, index, constraints for Oracle, SQL Server, PostgreSQL; Vocabulary tables with loading scripts

**ACHILLES**: profile your CDM data; review data quality assessment; explore population-level summaries

**OHDSI Forums**:
Public discussions for OMOP CDM Implementers/developers

http://github.com/OHDSI
## ACHILLES Heel Data Validation

<table>
<thead>
<tr>
<th>Message Type</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERROR</td>
<td>101-Number of persons by age, with age at first observation period; should not have age &lt; 0, (n=848)</td>
</tr>
<tr>
<td>ERROR</td>
<td>103 - Distribution of age at first observation period (count = 1); min value should not be negative</td>
</tr>
<tr>
<td>ERROR</td>
<td>114-Number of persons with observation period before year-of-birth; count (n=851) should not be &gt; 0</td>
</tr>
<tr>
<td>ERROR</td>
<td>206 - Distribution of age by visit_concept_id (count = 7); min value should not be negative</td>
</tr>
<tr>
<td>ERROR</td>
<td>301-Number of providers by specialty concept_id; 224 concepts in data are not in correct vocabulary (Specialty)</td>
</tr>
<tr>
<td>ERROR</td>
<td>400-Number of persons with at least one condition occurrence, by condition_concept_id; 115 concepts in data are not in correct vocabulary (SNOMED)</td>
</tr>
<tr>
<td>ERROR</td>
<td>406 - Distribution of age by condition_concept_id (count = 753); min value should not be negative</td>
</tr>
</tbody>
</table>
ATLAS to build, visualize, and analyze cohorts

- People having any of the following: Add Primary Criteria...
  - a condition occurrence of Delivery
  - occurrence start is: Between 2005-01-01 and 2013-12-31
  - with age Between 18 and 55
  - with a gender of: FEMALE

- with observation at least days prior and days after index
- Limit primary events to: All Events per person.

For people matching the Primary Criteria, include:

- People having All of the following criteria: Add New Criteria...
  - with At Least occurrences of:
    - a condition occurrence of Depression
    - occurring between days Before and days After index
  - and with At Most occurrences of:
    - a condition occurrence of Depression
    - occurring between days Before and days After index

Delete
Add Criterion...
Characterize the cohorts of interest
OHDSI in Action
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Name</th>
<th>Description</th>
<th>Population, millions</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUSOM</td>
<td>Ajou University School of Medicine</td>
<td>South Korea; inpatient hospital EHR</td>
<td>2</td>
</tr>
<tr>
<td>CCAE</td>
<td>MarketScan Commercial Claims and Encounters</td>
<td>US private-payer claims</td>
<td>119</td>
</tr>
<tr>
<td>CPRD</td>
<td>UK Clinical Practice Research Datalink</td>
<td>UK; EHR from general practice</td>
<td>11</td>
</tr>
<tr>
<td>CUMC</td>
<td>Columbia University Medical Center</td>
<td>US; inpatient EHR</td>
<td>4</td>
</tr>
<tr>
<td>GE</td>
<td>GE Centricity</td>
<td>US; outpatient EHR</td>
<td>33</td>
</tr>
<tr>
<td>INPC</td>
<td>Regenstrief Institute, Indiana Network for Patient Care</td>
<td>US; integrated health exchange</td>
<td>15</td>
</tr>
<tr>
<td>JMDC</td>
<td>Japan Medical Data Center</td>
<td>Japan; private-payer claims</td>
<td>3</td>
</tr>
<tr>
<td>MDCD</td>
<td>MarketScan Medicaid Multi-State</td>
<td>US; public-payer claims</td>
<td>17</td>
</tr>
<tr>
<td>MDCR</td>
<td>MarketScan Medicare Supplemental and Coordination of Benefits</td>
<td>US; private and public-payer claims</td>
<td>9</td>
</tr>
<tr>
<td>OPTUM</td>
<td>Optum ClinFormatics</td>
<td>US; private-payer claims</td>
<td>40</td>
</tr>
<tr>
<td>STRIDE</td>
<td>Stanford Translational Research Integrated Database Environment</td>
<td>US; inpatient EHR</td>
<td>2</td>
</tr>
<tr>
<td>HKU</td>
<td>Hong Kong University</td>
<td>Hong Kong; EHR</td>
<td>1</td>
</tr>
</tbody>
</table>
Treatment pathway event flow

INDEX: First exposure

≥1 exposure 121d-240d after index

≥1 exposure 241d-360d after index

≥1 exposure 361d-480d after index

≥1 exposure 481d-600d after index

≥1 exposure 601d-720d after index

≥1 exposure 721d-840d after index

≥1 exposure 841d-960d after index

≥1 exposure 961d-1080d after index

≥1 condition occurrence of disease of interest between all time prior to index and all time after index

≤0 condition occurrence of any excluded diseases between all time prior to index and all time after index

>365 day of prior observation

>1095 days of observation post-exposure

≤0 exposures 365d before index
Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak, Patrick B. Ryan, Jon D. Duke, Nigam H. Shah, Rae Woong Park, Vojtech Huser, 
Marc A. Suchard, Martijn J. Schuemie, Frank J. DeFalco, Adler Perotte, Juan M. Banda, Christian G. Reich, 
Lisa M. Schilling, Michael E. Matheny, Daniella Meeker, Nicole Pratt, and David Madigan

Characterizing treatment pathways at scale using the OHDSI network

Observational research promises to complement experimental research by providing large, diverse populations that would be infeasible for an experiment. Observational research can test its own clinical hypotheses, and observational studies also can contribute to the design of experiments and inform the generalizability of experimental research. Understanding the diversity of populations

Without sufficiently broad databases available in the first stage, randomized trials are designed without explicit knowledge of actual disease status and treatment practice. Literature reviews are restricted to the population choices of previous investigations, and pilot studies usually are limited in scope. By exploiting the ClinicalTrials.gov national trial registry and electronic health
Treatment pathways for diabetes

T2DM: All databases

<table>
<thead>
<tr>
<th>Drug</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metformin</td>
<td>29.42%</td>
</tr>
<tr>
<td>pioglitazone</td>
<td></td>
</tr>
<tr>
<td>sitagliptin</td>
<td></td>
</tr>
<tr>
<td>Glipizide</td>
<td></td>
</tr>
<tr>
<td>Glimepiride</td>
<td></td>
</tr>
<tr>
<td>Gliclazide</td>
<td></td>
</tr>
<tr>
<td>Glyburide</td>
<td></td>
</tr>
<tr>
<td>Rosiglitazone</td>
<td></td>
</tr>
<tr>
<td>Insulin, Glargine, Human</td>
<td></td>
</tr>
<tr>
<td>Exenatide</td>
<td></td>
</tr>
<tr>
<td>Insulin, Aspart, Human</td>
<td></td>
</tr>
<tr>
<td>Liraglutide</td>
<td></td>
</tr>
<tr>
<td>Saxagliptin</td>
<td></td>
</tr>
<tr>
<td>Insulin, Lispro, Human</td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
</tr>
<tr>
<td>Insulin, Isophane, Human</td>
<td></td>
</tr>
</tbody>
</table>

First drug

Second drug

Only drug
Type 2 Diabetes Mellitus

- Metformin
- Glitazone
- Pioglitazone
- Sitagliptin
- Glimepiride
- Alogliptin
- Rosiglitazone
- Glyburide

Hypertension

- Hydrochlorothiazide
- Lisinopril
- Metoprolol
- Amlodipine
- Furosemide
- Losartan
- Atenolol
- Valsartan
- Carvedilol
- Triamterene
- Diltiazem
- Ramipril
- Benazepril
- Olmesartan
- Spironolactone
- Clonidine

Depression

- Citalopram
- Bupropion
- Sertraline
- Escitalopram
- Fluoxetine
- Trazodone
- Venlafaxine
- Duloxetine
- Paroxetine
- Amitriptyline
- Mirtazapine
- Desvenlafaxine
- Nortriptyline
- Doxepin

Population-level heterogeneity across systems, and patient-level heterogeneity within systems.
25% of HTN patients (10% of others) have a unique path despite 250M pop
Monotherapy – diabetes

General upward trend in monotherapy
Academic medical centers differ from general practices.
Monotherapy – diabetes

General practices, whether EHR or claims, have similar profiles.
Conclusions: Network research

• It is feasible to encode the world population in a single data model
  – Over 1,000,000,000 records by voluntary effort
• Generating evidence is feasible
• Stakeholders willing to share results
• Able to accommodate vast differences in privacy and research regulation
Open science

• Admit that there is a problem
• Study it scientifically
  – Define that surface and differentiate true variation from confounding ...
• Total description of every study
• Research into new methods
Take a scientific approach to science

1. Database heterogeneity:
   Holding analysis constant, different data may yield different estimates
   “Evaluating the Impact of Database Heterogeneity on Observational Study Results”

2. Parameter sensitivity:
   Holding data constant, different analytic design choices may yield different estimates
   Madigan D, Ryan PB, Schuemie MJ, Therapeutic Advances in Drug Safety, 2013: “Does design matter?
   Systematic evaluation of the impact of analytical choices on effect estimates in observational studies”

3. Empirical performance:
   Most observational methods do not have nominal statistical operating characteristics

4. Empirical calibration can help restore interpretation of study findings
   Schuemie MJ, Ryan PB, DuMouchel W, et al, Statistics in Medicine, 2013:
   “Interpreting observational studies: why empirical calibration is needed to correct p-values”
Reproducible research

1. Address confounding that is measured
   • Propensity stratification
   • *Systematic* (not manual) variable selection
     • Balance 58,285 variables ("Table 1")

After stratification on the propensity score, all 58,285 covariates have standardized difference of mean < 0.1
Reproducible research

2. Unmeasured (residual) confounding
   - Confidence interval calibration
   - Adjust for all uncertainty, not just sampling
   - Many negative controls
   - Unique to OHDSI (PNAS in press)

After calibration, 4% have $p < 0.05$ (was 16%)
Reproducible research

3. Multiple databases, locations, practice types
   • Exploit international OHDSI network

![Graph showing comparison between Uncalibrated and Calibrated data for CCAE, MDCD, MDCR, and Optum.]
Reproducible research

4. Open: publish all
   - Hypotheses
   - Code
   - Parameters
   - Runs

OURL<1000
Generating evidence for US FDA

Potential Signals of Serious Risks/New Safety Information Identified by the FDA Adverse Event Reporting System (FAERS) between October - December 2015

| Keppra (levetiracetam) tablet, oral solution, injection | Angioedema | FDA is evaluating the need for regulatory action. |

- Protocol completed, code tested, study announced
- 50 viewed protocol, 25 viewed the code, and 7 sites ran the code on 10 databases (5 claims / 5 EHR), 59,367 levetiracetam patients matched with 74,550 phenytoin patients
Generating evidence for US FDA

No evidence of increased angioedema risk with levetiracetam use compared with phenytoin use

“The study is focused, appears well designed, and provides new insight that should be of interest to clinicians and regulators... This is an important contribution to improved pharmacovigilance.”

Add word to title, move diagram from supplement to body
How can we improve the literature
Literature

Effect size (1 = no effect)

Standard error

P = 0.05
Literature

- Effect size
- Standard error
- Not significant
- Protect
- Harm
- P = 0.05
Observational research results in literature

85% of exposure-outcome pairs have $p < 0.05$

29,982 estimates
11,758 papers
Observational research results in literature

29,982 estimates
11,758 papers

Publication bias
➢ Don’t know the denominator of negative studies.

29,982 estimates
11,758 papers
Observational research results in literature

- 29,982 estimates
- 11,758 papers

P-value hacking

29,982 estimates
11,758 papers
Observational research results in literature

• Individuals may produce good research studies
• In aggregate, the medical research system is a data dredging machine
Look at many outcomes at once

Duloxetine vs. Sertraline for these 22 outcomes:

<table>
<thead>
<tr>
<th>Acute liver injury</th>
<th>Hypotension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myocardial infarction</td>
<td>Hypothyroidism</td>
</tr>
<tr>
<td>Alopecia</td>
<td>Insomnia</td>
</tr>
<tr>
<td>Constipation</td>
<td>Nausea</td>
</tr>
<tr>
<td>Decreased libido</td>
<td>Open-angle glaucoma</td>
</tr>
<tr>
<td>Delirium</td>
<td>Seizure</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Stroke</td>
</tr>
<tr>
<td>Fracture</td>
<td>Suicide and suicidal ideation</td>
</tr>
<tr>
<td>Gastrointestinal hemorrhage</td>
<td>Tinnitus</td>
</tr>
<tr>
<td>Hyperprolactinemia</td>
<td>Ventricular arrhythmia and sudden cardiac death</td>
</tr>
<tr>
<td>Hyponatremia</td>
<td>Vertigo</td>
</tr>
</tbody>
</table>
Many treatments at once

<table>
<thead>
<tr>
<th>Type</th>
<th>Class</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug</td>
<td>Atypical</td>
<td>Bupropion</td>
</tr>
<tr>
<td>Drug</td>
<td>Atypical</td>
<td>Mirtazapine</td>
</tr>
<tr>
<td>Procedure</td>
<td>ECT</td>
<td>Electroconvulsive therapy</td>
</tr>
<tr>
<td>Procedure</td>
<td>Psychotherapy</td>
<td>Psychotherapy</td>
</tr>
<tr>
<td>Drug</td>
<td>SARI</td>
<td>Trazodone</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>Desvenlafaxine</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>duloxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>venlafaxine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Citalopram</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Escitalopram</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Fluoxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Paroxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Sertraline</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>vilazodone</td>
</tr>
<tr>
<td>Drug</td>
<td>TCA</td>
<td>Amitriptyline</td>
</tr>
<tr>
<td>Drug</td>
<td>TCA</td>
<td>Doxepin</td>
</tr>
<tr>
<td>Drug</td>
<td>TCA</td>
<td>Nortriptyline</td>
</tr>
</tbody>
</table>
Large-scale estimation for depression

- 17 treatments
- $17 \times 16 = 272$ comparisons
- 22 outcomes
- $272 \times 22 = 5,984$ effect size estimates
- 4 databases (Truven CCAE, Truven MDCD, Truven MDCR, Optum)
- $4 \times 5,984 = 23,936$ estimates
Estimates are in line with expectations

11% of exposure-outcome pairs have calibrated $p < 0.05$
Example

Mirtazapine vs. Citalopram Constipation
Database: Truven MDCD

Calibrated HR = 0.90 (0.70 – 1.12)
Propensity models for all comparisons
(Truven CCAE, one outcome)

<table>
<thead>
<tr>
<th>Amitriptyline</th>
<th>Bupropion</th>
<th>Citalopram</th>
<th>Desvenlafaxine</th>
<th>Doxepin</th>
<th>Duloxetine</th>
<th>Escitalopram</th>
<th>Fluoxetine</th>
<th>Mirtazapine</th>
<th>Nortriptyline</th>
<th>Paroxetine</th>
<th>Psychotherapy</th>
<th>Sertraline</th>
<th>Venlafaxine</th>
<th>Vilazodone</th>
<th>Trazodone</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></td>
</tr>
<tr>
<td>Amitriptyline</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>Bupropion</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>Citalopram</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>Desvenlafaxine</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>Doxepin</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>Duloxetine</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>Escitalopram</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>Fluoxetine</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>Mirtazapine</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>Nortriptyline</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>Paroxetine</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>Psychotherapy</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>Sertraline</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>Venlafaxine</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>Vilazodone</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>Trazodone</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>
</tbody>
</table>
Large-scale estimation for depression

• How do we use it? Troll for effects?

• Professor what should I study this year?
  – Simple, go to Pubmed and find the smallest p-values in the literature; surely those must be the most significant things to study

• Which is safer?
  • Seizure in 0.0000000001 to 0.0000000002 (p=0.00001)
  • Seizure in 0 to 0.2 (p=.45)

• Large-scale studies become the literature
  • Come with hypothesis and ask a question
Large-scale estimation for depression

• Not “data-dredging”!
  – Data-dredging is not about what you do but about what you *throw out*
    • This can’t be done for literature

• One-off studies
  – Wouldn’t it be best to optimize each study?
    • Never get 10 or 100 parameters right
  – Still good to see the distribution

• At the very least, publish every last parameter so it can be reproduced
How often
How often do side effects occur?

• New incidence of any condition for any drug on the world market
  – Show range of answers for disparate databases
• Absolute risk (vs. attributable risk)
  – Not know if it is causal or not: MI with statin
• More complicated than it looks
  – Standard framework for reporting incidence
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.
Summary

• Current observational research is suspect
• Large-scale observational research appears to be possible and more reliable than the current approach
• Need to extend to other areas
• Further research on reproducibility