Overview of Confounding and Bias: What is Next?

Fadia Tohme Shaya, PhD, MPH

Professor and Vice-Chair of Academic Affairs

Viktor Chirikov and Priyanka Gaitonde , PhD students

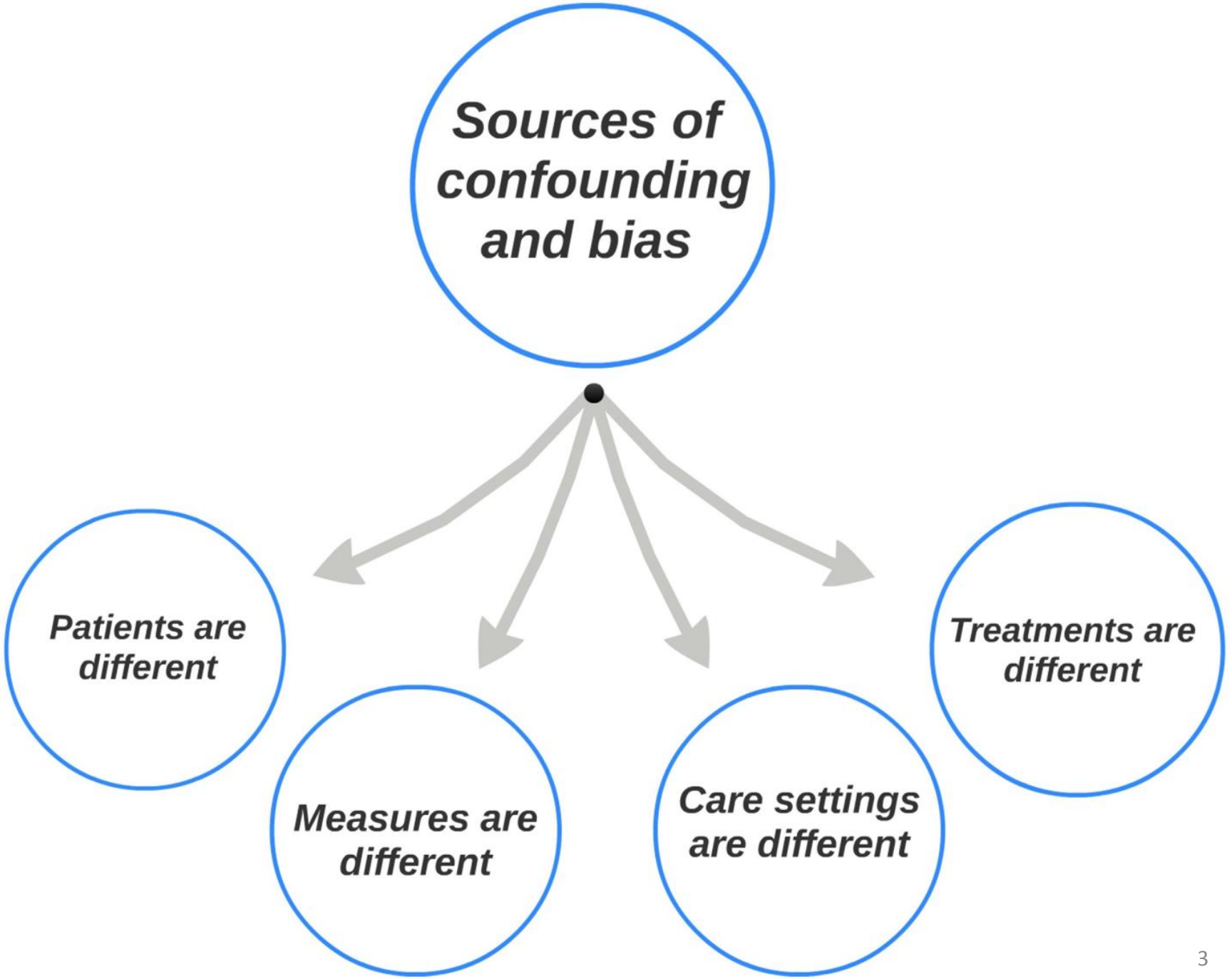
University of Maryland School of Pharmacy

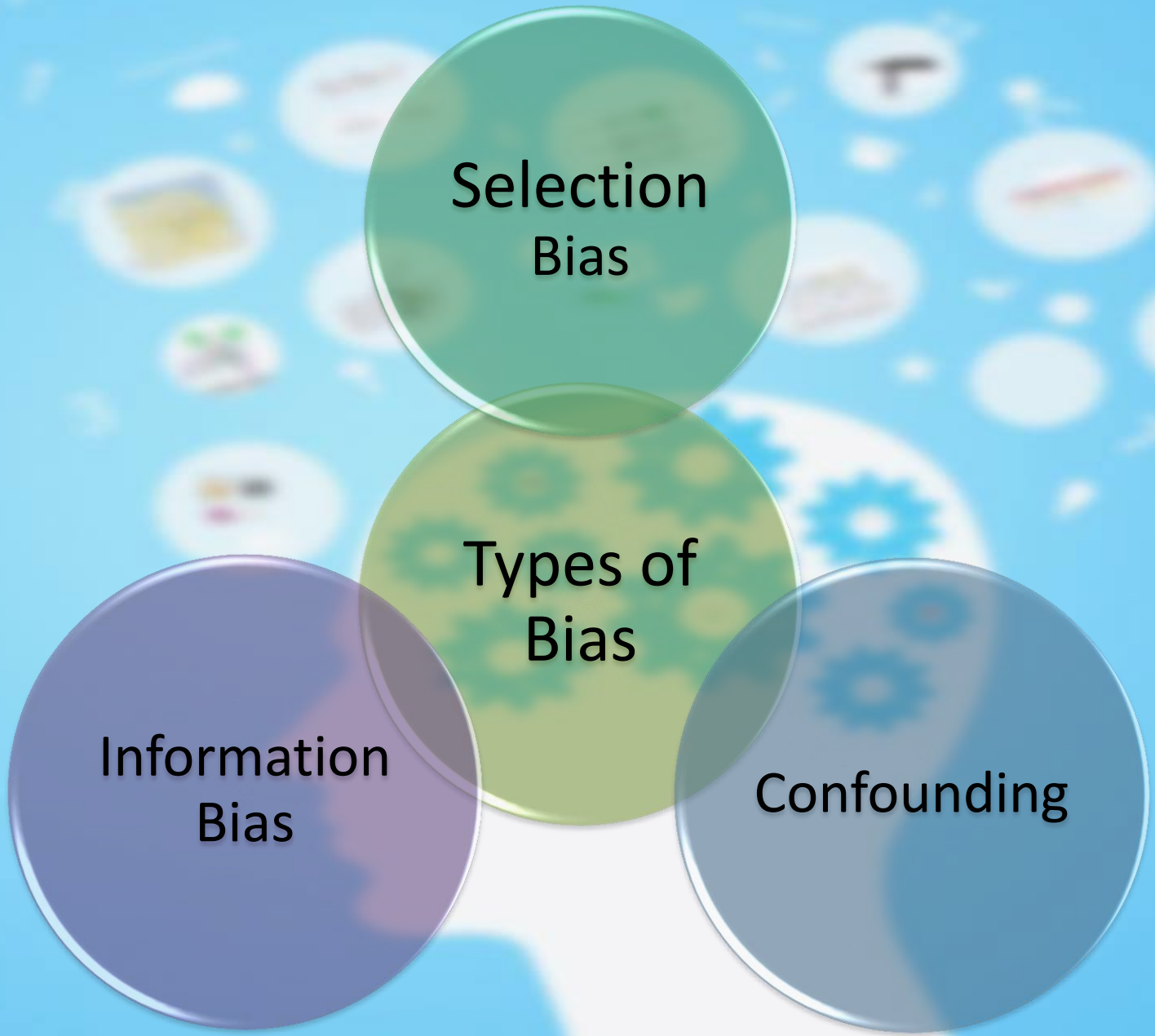
fshaya@rx.umaryland.edu

410-706-5392

OVERVIEW OF CONFOUNDING

by Fadia Shaya


Sources of confounding and bias
Patients are different
Measures are different
Care settings are different
Treatments are different


Selection Bias
Types of Bias
Information Bias
Confounding

Information bias

- Interviewer bias
- Recall Bias
- Observer bias
- Loss to follow-up
- Surveillance bias

Misclassification bias

- Differential misclassification
- Non-differential misclassification

Restrictive Reimbursement Policies: Bias Implications for Claims-Based Drug Safety Studies

Joshua J. Gagne

Published online: 4 September 2014

Abstract Restrictive reimbursement policies—including those based on non-formulary drug status and prior authorizations—can create situations in which patients' use of prescription medications is not fully captured in administrative claims data. This can create bias in drug safety studies that depend solely on these data. An analysis in two Canadian provinces found that primary administrative databases captured only 61 % of dispensations of drugs for which restrictive reimbursement policies were in place. A subsequent simulation study found that, in certain circumstances bias due to exposure misclassification resulting from restrictive reimbursement policies can be quite large in analyses comparing outcomes between drug exposure groups. Investigators need to be knowledgeable about the data they analyze and know whether restrictive reimbursement policies are in place that might affect the capture of drugs of interest. It is also critical to understand the mechanisms by which restrictive reimbursement might cause bias in claims-based drug safety studies, the direction and magnitude of the potential bias, and strategies that could be used to mitigate such bias.

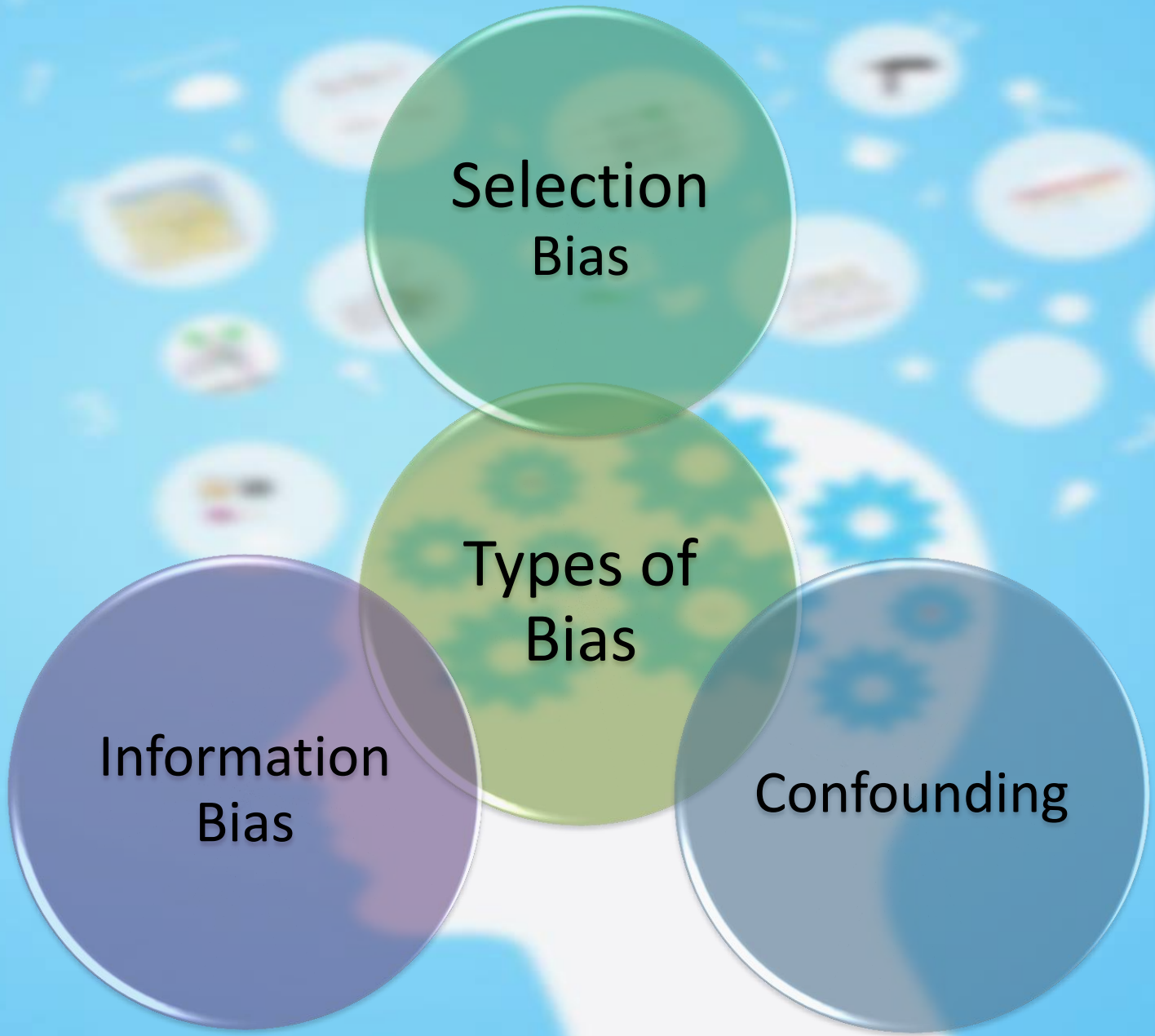
Key Points

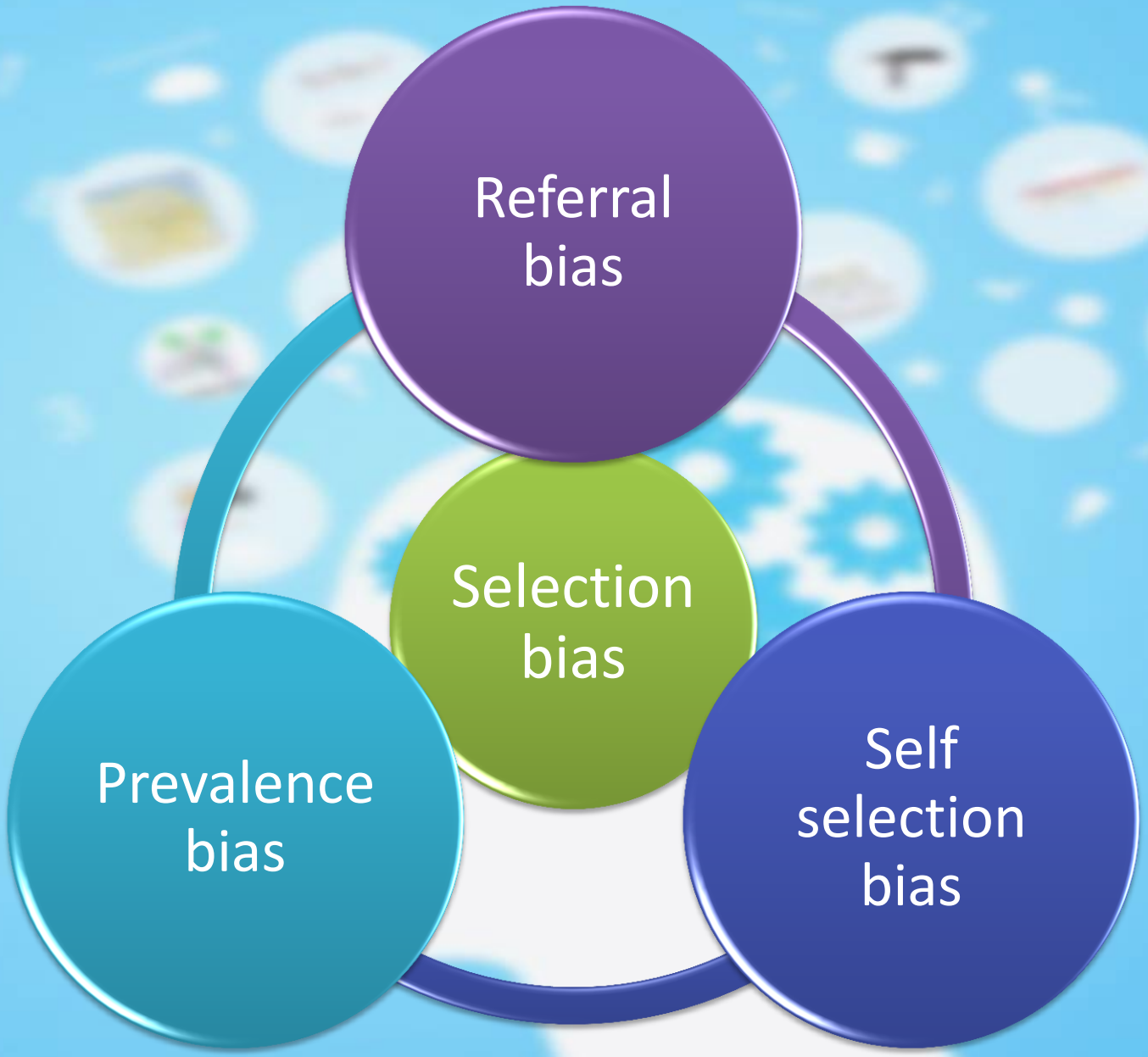
Restrictive reimbursement policies can lead to substantial under-ascertainment of prescription drug use in administrative claims data

Under certain conditions, this missing information can cause or increase biases arising from misclassification and confounding in drug safety studies that rely on claims data

By understanding the mechanisms by which restrictive reimbursement might cause bias, investigators can anticipate the direction and magnitude of the potential bias and implement strategies that could be used to mitigate it

1 Introduction


Selection Bias
Types of Bias
Information Bias
Confounding

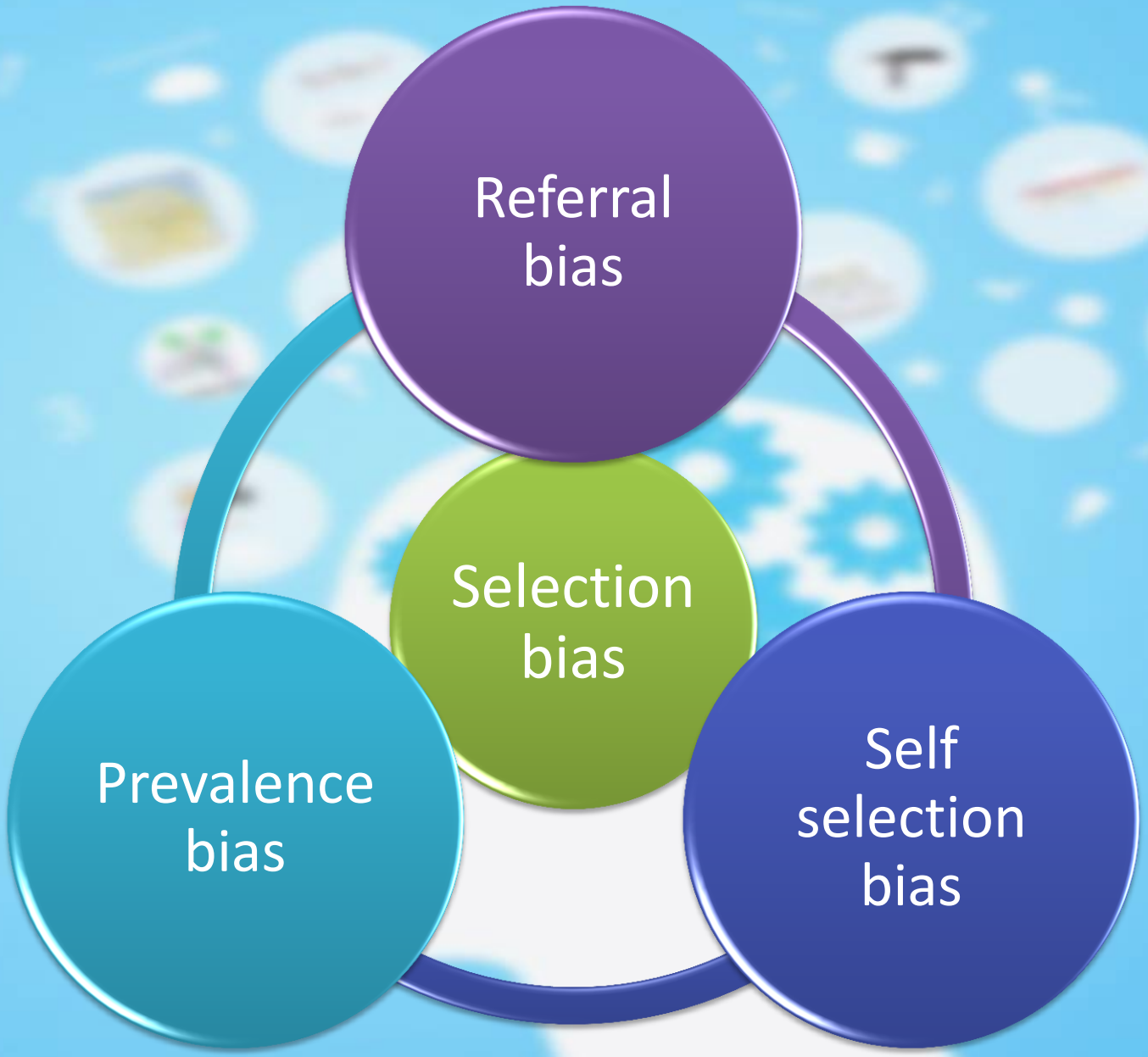

Referral bias
Selection bias
Prevalence bias
Self selection bias

Referral Selection Bias in the Medicare Hospital Mortality Prediction Model: Are Centers of Referral for Medicare Beneficiaries Necessarily Centers of Excellence?

David J. Ballard, Sandra C. Bryant, Peter C. O'Brien, David W. Smith, Michael B. Pine, and Denis A. Cortese

---

Objective. Although the Health Care Financing Administration (HCFA) uses Medicare hospital mortality data as a measure of hospital quality of care, concerns have been raised regarding the validity of this concept. A problem that has not been fully evaluated in these data is the potential confounding effect of illness severity factors associated with referral selection and hospital mortality on comparisons of risk-adjusted hospital mortality. We address this issue.


Referral bias
Selection bias
Prevalence bias
Self selection bias

SPECIAL ARTICLE

Evaluating Medication Effects Outside of Clinical Trials: New-User Designs

Wayne A. Ray[1,2]

[1] Division of Pharmacoepidemiology, Department of Preventive Medicine, Vanderbilt University School of Medicine, Nashville, TN.
[2] Geriatric Research, Education and Clinical Center, Nashville Veteran's Administration Medical Center (WAR), Nashville, TN.

Received for publication November 7, 2002; accepted for publication May 1, 2003.

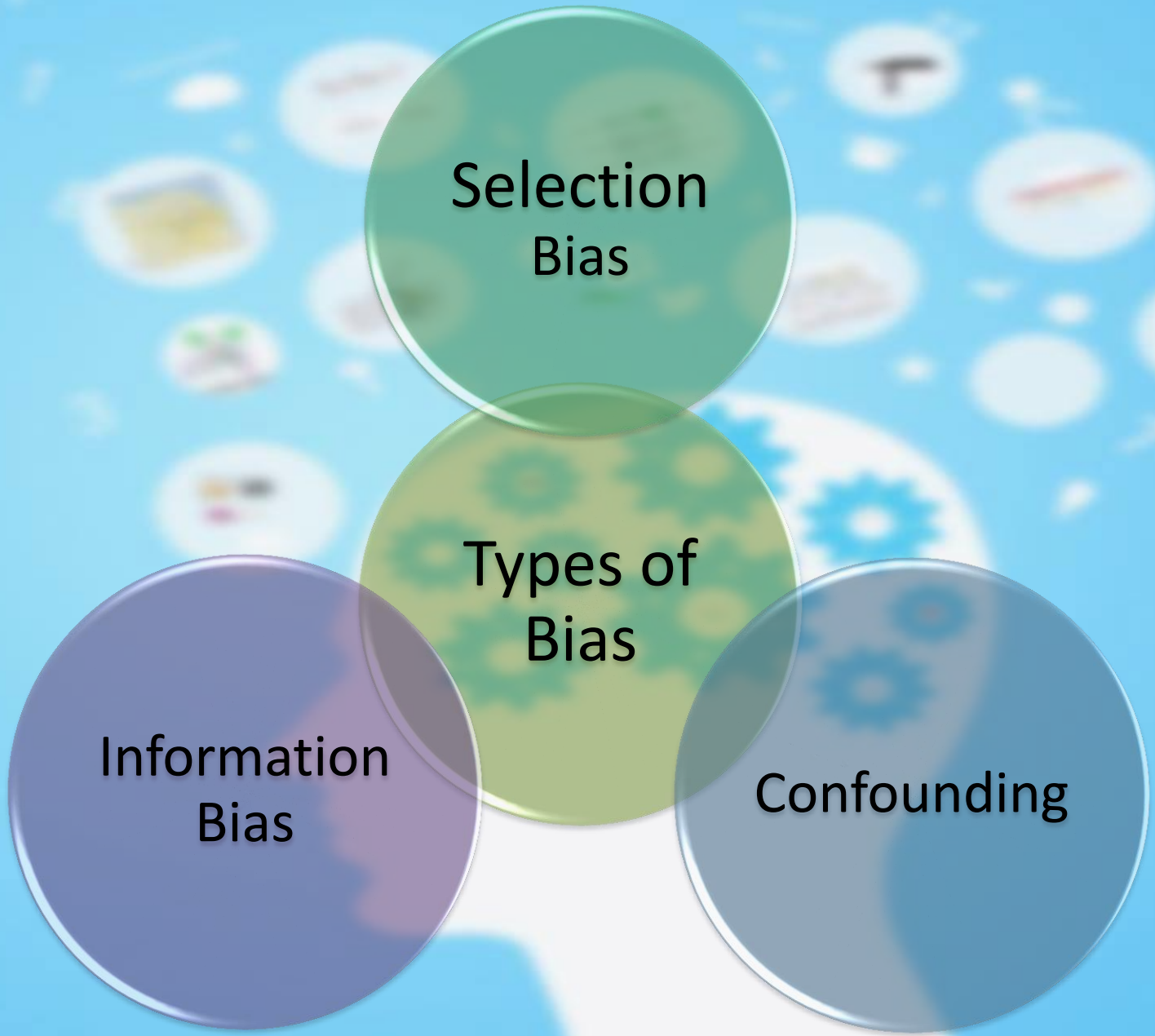
Recent clinical trials demonstrating that hormone replacement therapy (HRT) does not prevent coronary heart disease in women have again raised doubts concerning observational studies. Although much of the explanation probably lies in what might be called the "healthy HRT user" effect, another contributing factor may be that most observational studies included many prevalent users: women taking HRT for some time before study follow-up began. This practice can cause two types of bias, both of which plausibly may have contributed to the discrepancy between observational and randomized studies. First, prevalent users are "survivors" of the early period of pharmacotherapy, which can introduce substantial bias if risk varies with time, just as in studies of operative procedures that enroll patients after they have survived surgery. This article provides several examples of medications for which the hazard function varies with time and thus would be subject to prevalent user bias. Second, covariates for drug users at study entry often are plausibly affected by the drug itself. Investigators often do not adjust for these factors on the causal pathway, which may introduce confounding. A *new-user design* eliminates these biases by restricting the analysis to persons under observation at the start of the current course of treatment. This article thus argues that such designs should be used more frequently in pharmacoepidemiology.

bias (epidemiology); confounding factors (epidemiology); epidemiologic research design; hormone replacement therapy; pharmacoepidemiology; research design

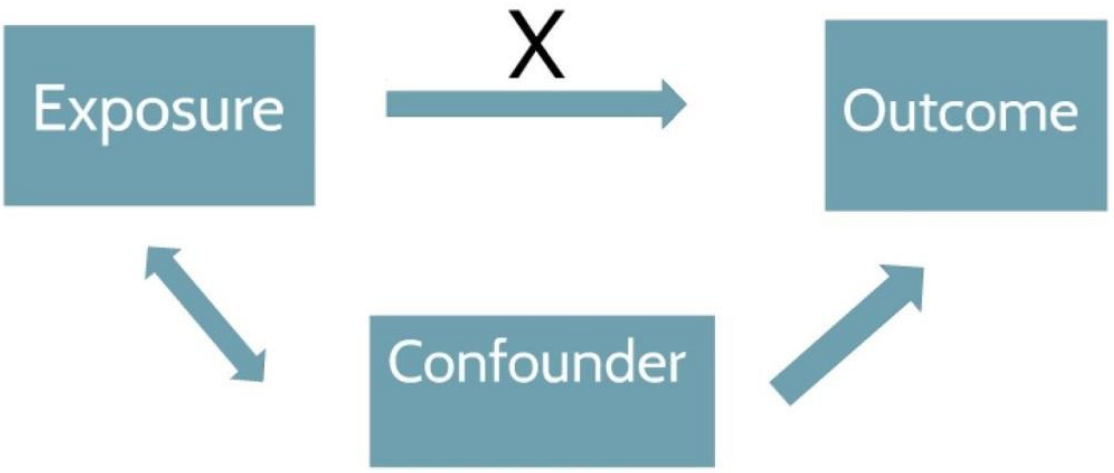
New user design

- Patients with exposure have similar observed and unobserved characteristic
- Clear temporal sequence
- Events occurring on immediate consumption can be ascertained

- Loss of sample by excluding prevalent users
- Loss of precision


Selection Bias
Types of Bias
Information Bias
Confounding

A confounder is an extraneous factor which causes bias or distorts the true association between the exposure and outcome of interest



American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2015. This work is written by (a) US Government employee(s) and is in the public domain in the US.

Vol. 181, No. 7
DOI: 10.1093/aje/kwu324
Advance Access publication:
March 15, 2015

Original Contribution

Is the Inverse Association Between Selenium and Bladder Cancer Due to Confounding by Smoking?

Laura E. Beane Freeman*, Margaret R. Karagas, Dalsu Baris, Molly Schwenn, Alison T. Johnson, Joanne S. Colt, Brian Jackson, G. M. Monawar Hosain, Kenneth P. Cantor, and Debra T. Silverman

* Correspondence to Dr. Laura E. Beane Freeman, Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Room 6E136, MSC 9771, Bethesda, MD 20892 (e-mail: freemala@mail.nih.gov).

Initially submitted April 25, 2014; accepted for publication October 20, 2014.

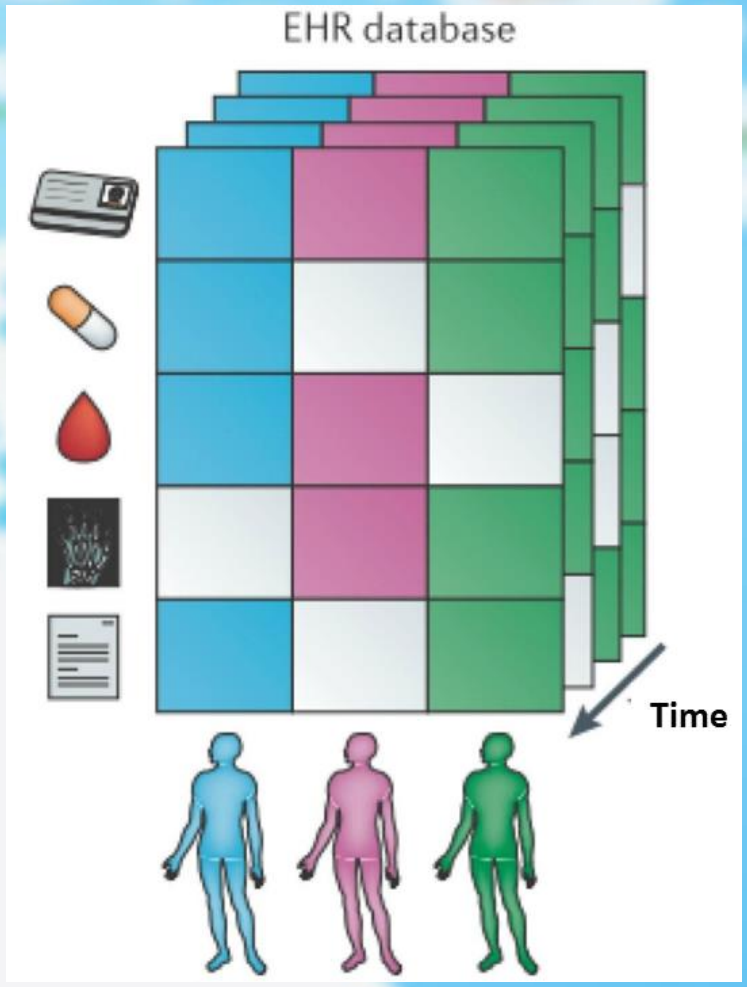
Selenium has been linked to a reduced risk of bladder cancer in some studies. Smoking, a well-established risk factor for bladder cancer, has been associated with lower selenium levels in the body. We investigated the selenium-bladder cancer association in subjects from Maine, New Hampshire, and Vermont in the New England Bladder Cancer Case-Control Study. At interview (2001–2005), participants provided information on a variety of factors, including a comprehensive smoking history, and submitted toenail samples, from which we measured selenium levels. We estimated odds ratios and 95% confidence intervals among 1,058 cases and 1,271 controls using logistic regression. After controlling for smoking, we saw no evidence of an association between selenium levels and bladder cancer (for fourth quartile vs. first quartile, odds ratio (OR) = 0.98, 95% confidence interval (CI): 0.77, 1.25). When results were restricted to regular smokers, there appeared to be an inverse association (OR = 0.76, 95% CI: 0.58, 0.99); however, when pack-years of smoking were considered, this association was attenuated (OR = 0.91, 95% CI: 0.68, 1.20), indicating potential confounding by smoking. Despite some reports of an inverse association between selenium and bladder cancer overall, our results, combined with an in-depth evaluation of other studies, suggested that confounding from smoking intensity or duration could explain this association. Our study highlights the need to carefully evaluate the confounding association of smoking in the selenium-bladder cancer association.

bladder cancer; case-control study; selenium; smoking

CONFOUNDING BY POPULATION ADMIXTURE/ETHNICITY

- Baseline disease risks and genotype frequency vary across ethnicity

- The larger the number of ethnicities involved in an admixed population, the less likely that population stratification can be the explanation for an observation

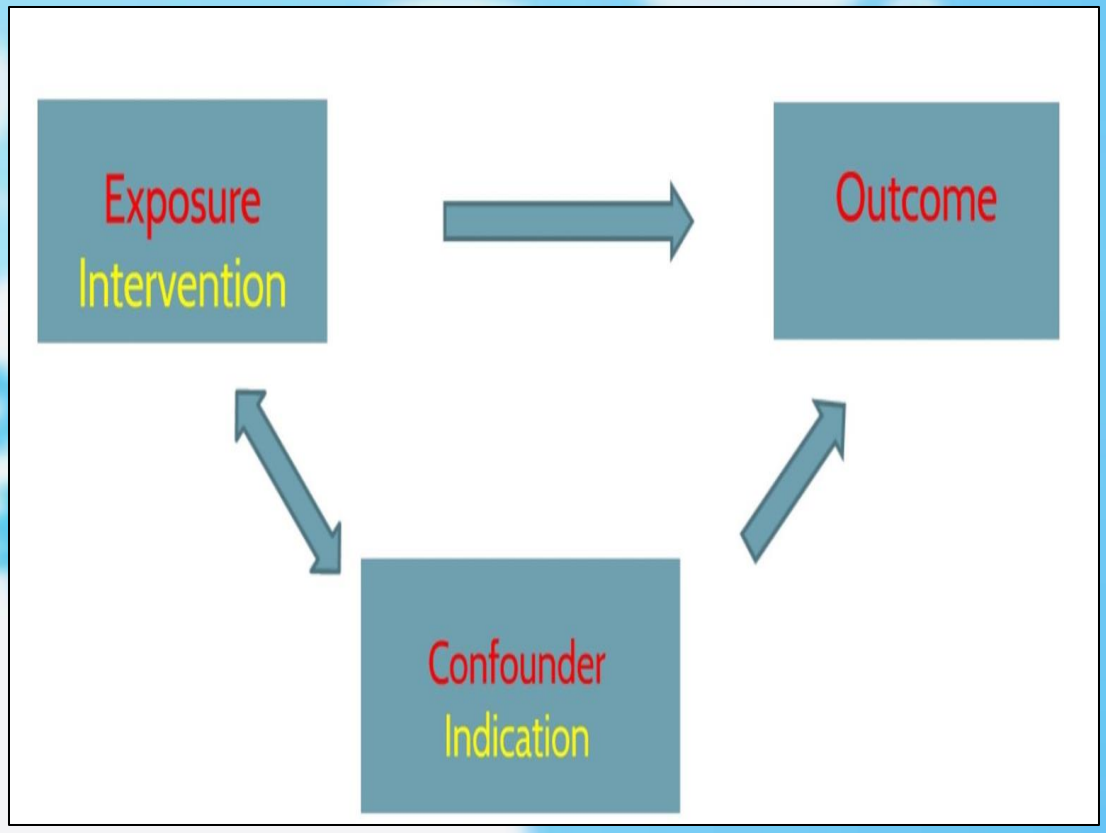


The role of body weight, fat distribution and weight change in ethnic differences in the 9-year incidence of hypertension

Laura R. Grootveld[a], Irene G.M. Van Valkengoed[a], Ron J.G. Peters[b], Joanne K. Ujcic-Voortman[c], Lizzy M. Brewster[a], Karien Stronks[a], and Marieke B. Snijder[a]

CONFOUNDING BY INDICATION

- In situations where the indication itself becomes a confounder
- Frequent risk of bias in observational studies of treatment effect
- Difference in underlying risk profile or baseline prognosis between treated and untreated



Confounding by Indication Probably Distorts the Relationship between Steroid Use and Cardiovascular Disease in Rheumatoid Arthritis: Results from a Prospective Cohort Study

Alper M. van Sijl*, Maarten Boers, Alexandre E. Voskuyl, Michael T. Nurmohamed

Department of Rheumatology, VU University Medical Center, Amsterdam, the Netherlands

Abstract

Objective: To evaluate the risk of cardiovascular disease in patients with rheumatoid arthritis exposed to glucocorticoids.

Methods: Retrospective analysis of exposure to glucocorticoids in a prospective cohort of 353 patients with rheumatoid arthritis followed from June 2001 up to November 2011 for incident cardiovascular disease in a hospital-based outpatient cohort in the Netherlands. Hazard ratios with 95%-confidence intervals were calculated for the association between different types of exposure to glucocorticoids and incident cardiovascular disease. Associations were adjusted for demographics, cardiovascular risk factors and disease related parameters.

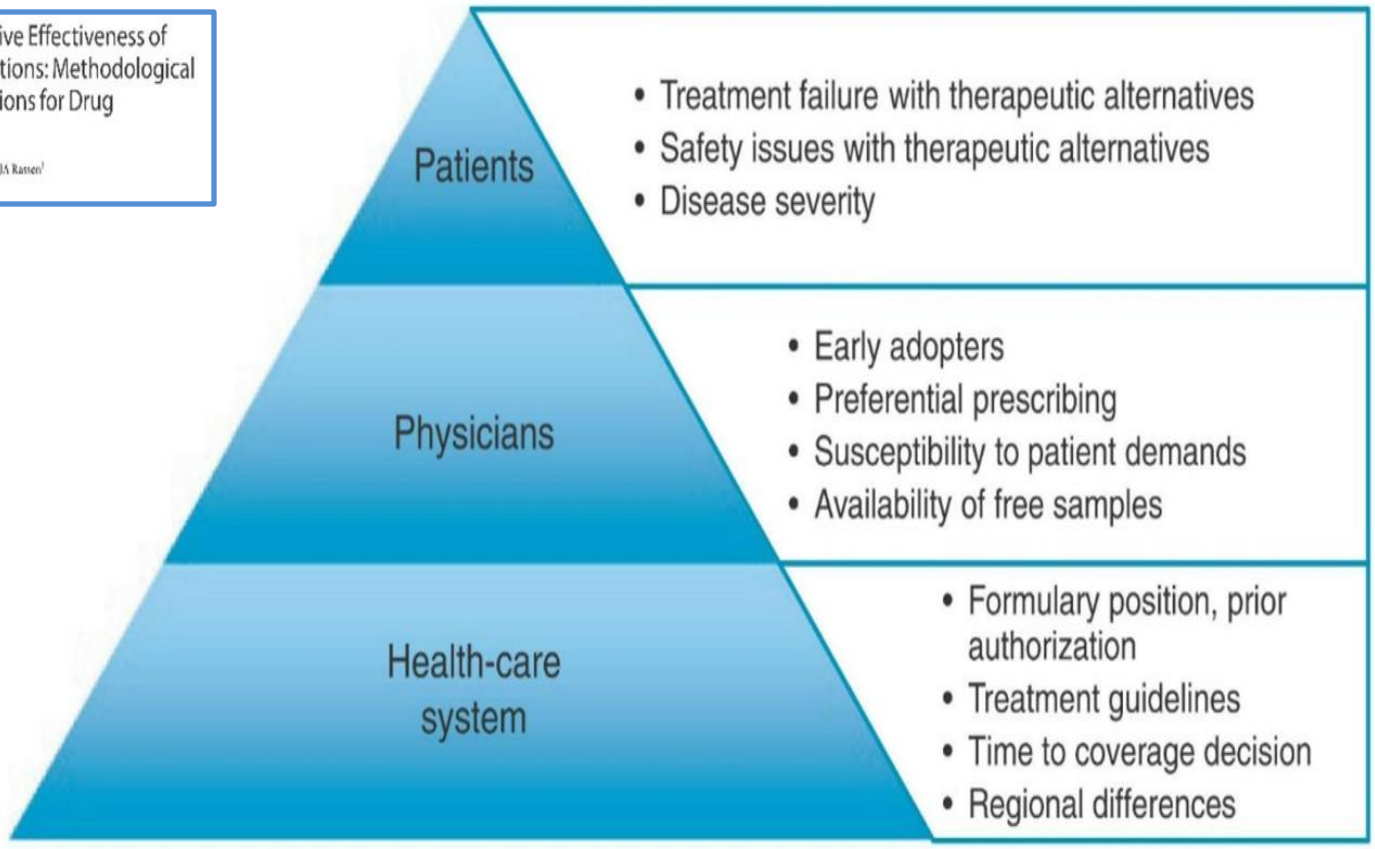
Results: Recent and current exposure to glucocorticoids were associated with incident cardiovascular disease, as was a longer duration of exposure and cumulative exposure to glucocorticoids. Adjustment for disease activity and severity negated the association.

Conclusion: In observational studies the finding of incident cardiovascular disease in patients with rheumatoid arthritis exposed to glucocorticoids is strongly confounded by indication due to high disease activity. The adverse cardiovascular effects of glucocorticoids might be balanced by positive effects working through inflammation control.

CONFOUNDING BY PRESCRIBING

Assessing the Comparative Effectiveness of Newly Marketed Medications: Methodological Challenges and Implications for Drug Development

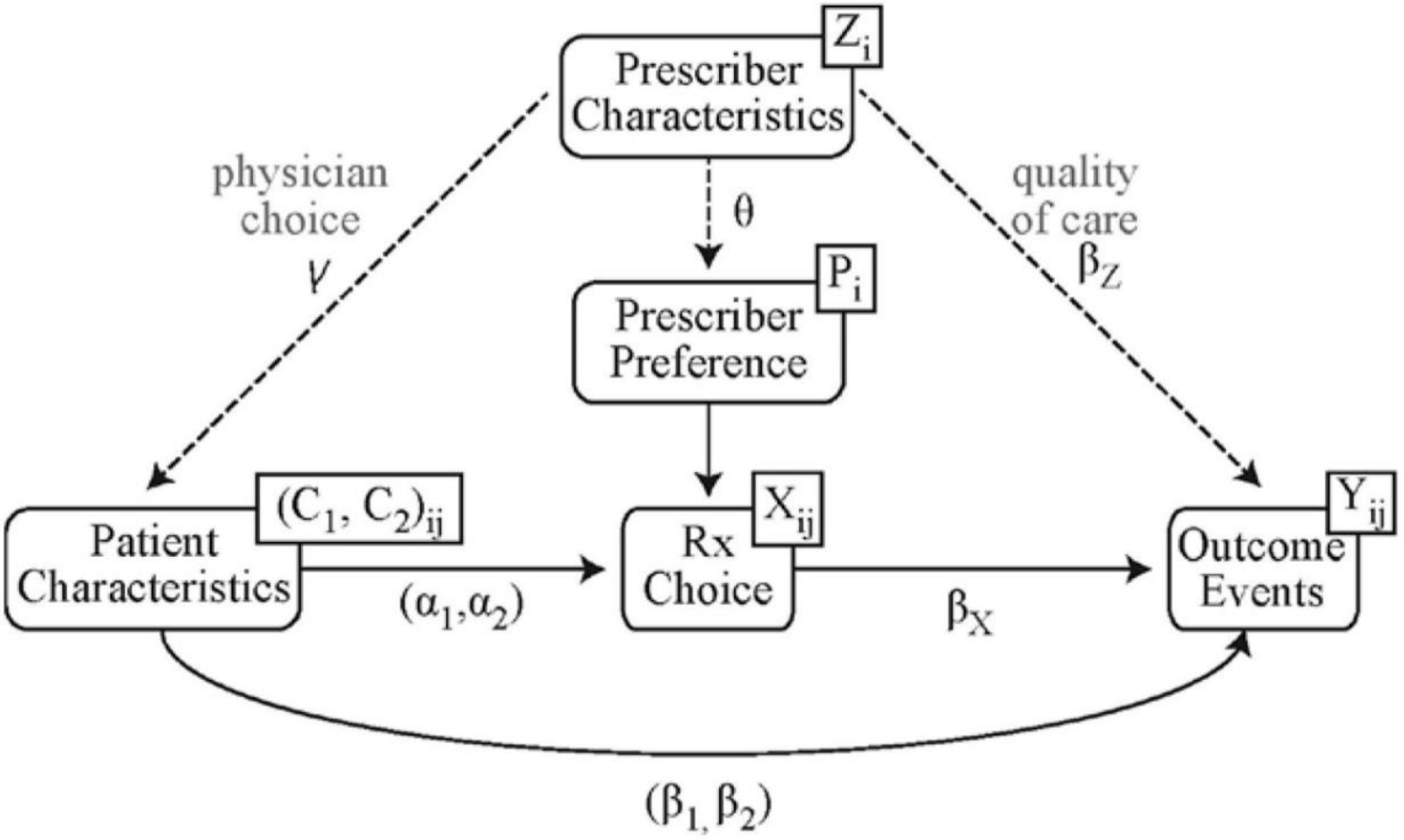
S Schneeweiss[1], JJ Gagne[1], RJ Glynn[1], M Ruhl[1] and JA Rassen[1]



The multilevel nature of factors that determine prescribing of newly marketed medications.

Evaluating Possible Confounding by Prescriber in Comparative Effectiveness Research

Jessica M. Franklin, Sebastian Schneeweiss, Krista F. Huybrechts, and Robert J. Glynn

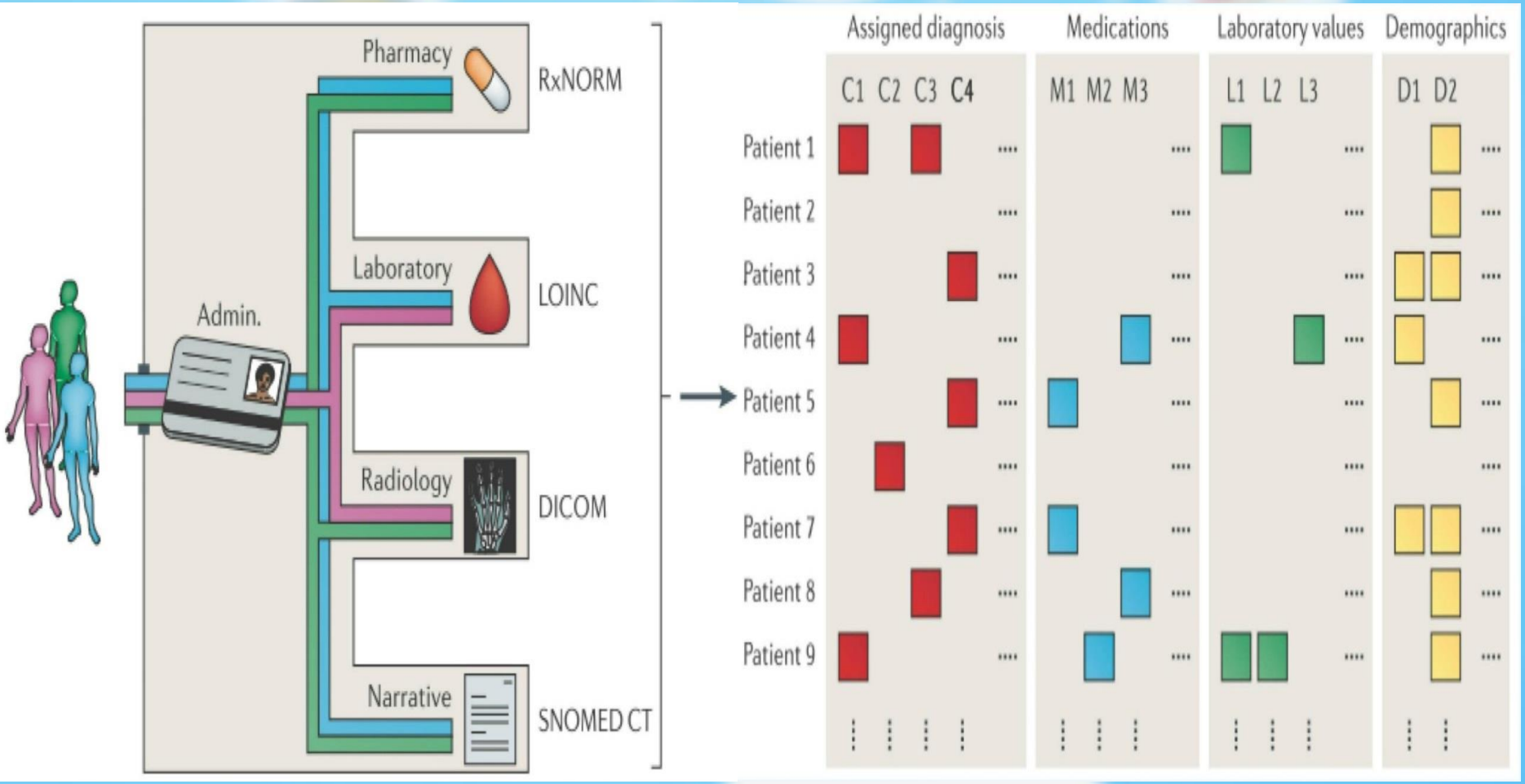


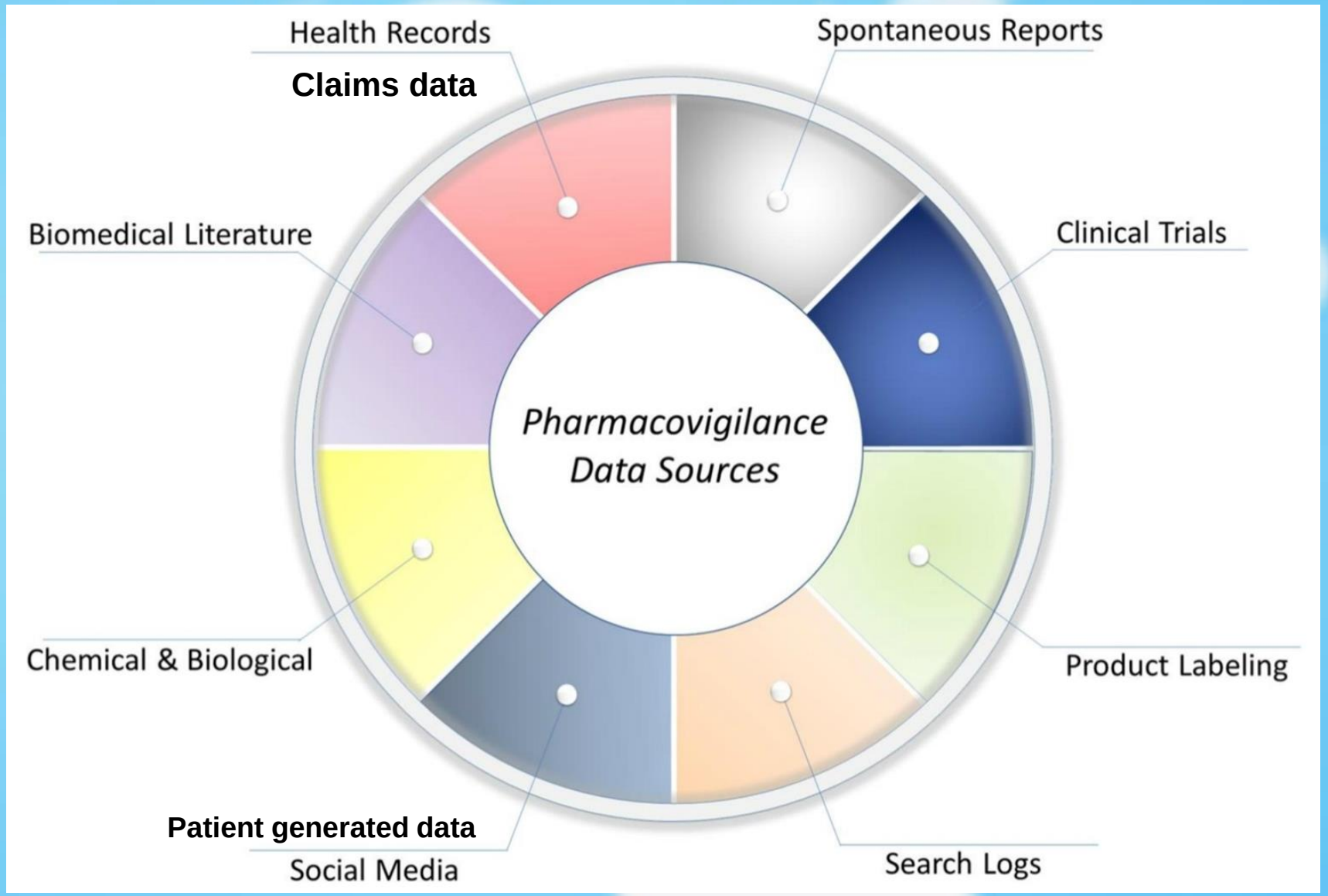
WE MAY NOT LIVE IN A DATA DESERT ANYMORE…

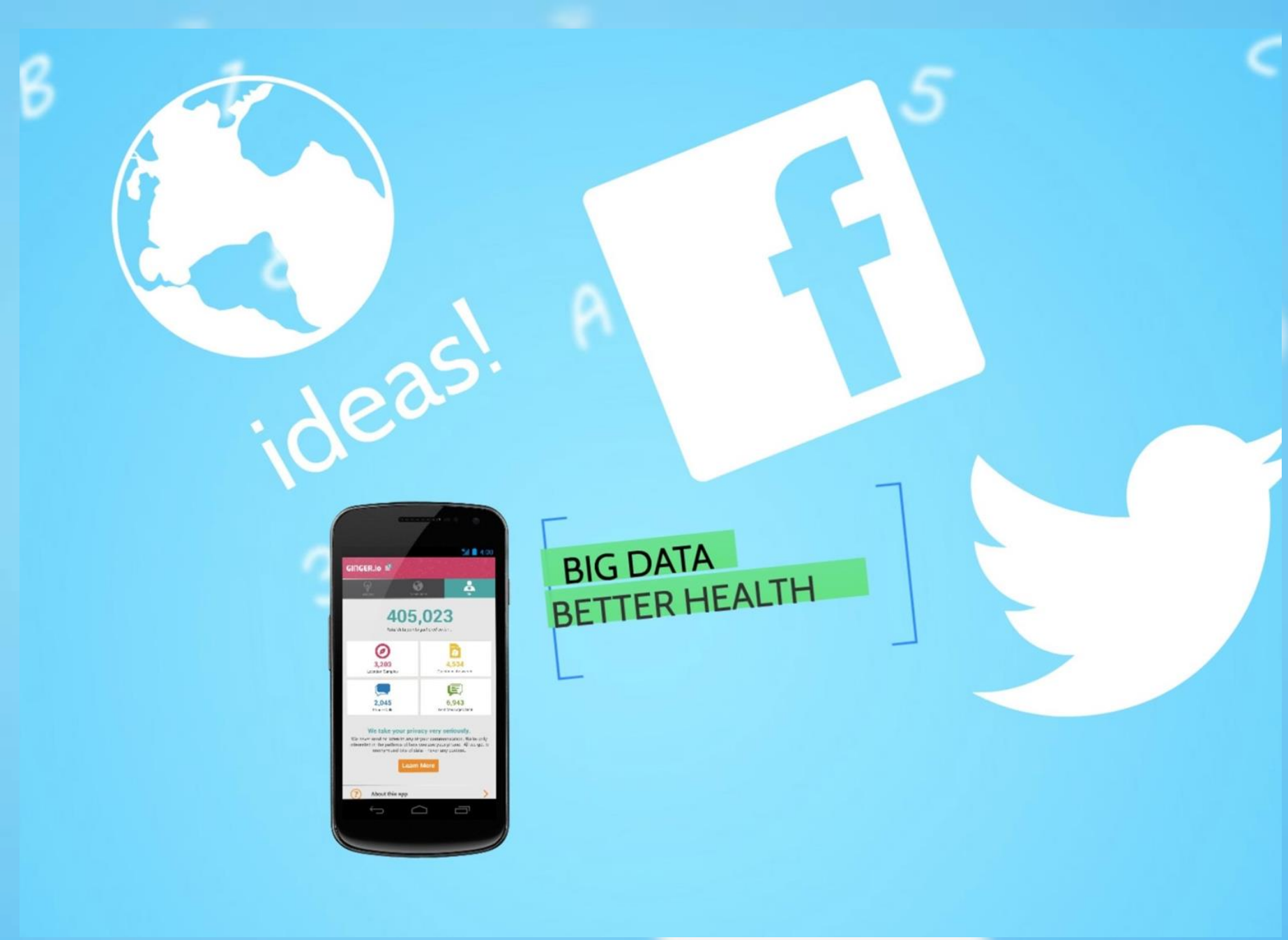
Regular data sources are missing important confounders

Uncertainty about causal relationships

DATA BUILD-UP

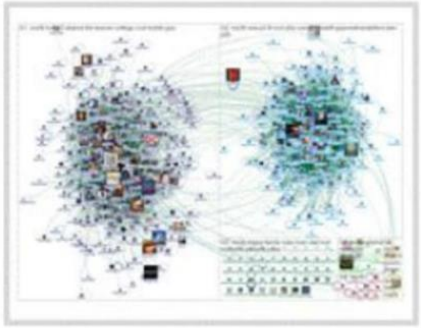



Health Records
Claims data
Spontaneous Reports
Biomedical Literature
Clinical Trials
Pharmacovigilance
Data Sources
Chemical & Biological
Product Labeling
Patient generated data
Social Media
Search Logs


ideas!
BIG DATA
BETTER HEALTH
GINGER.io
405,023
3,203
4,534
2,045
6,943
We take your privacy very seriously.
Learn More
About this app

TYPES OF DATA NETWORK

[Divided]
Polarized Crowds

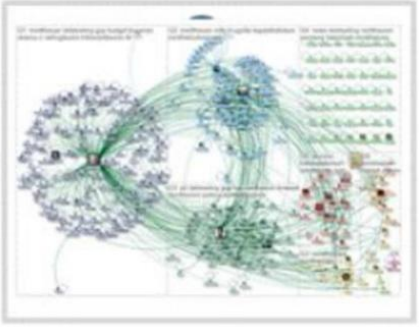


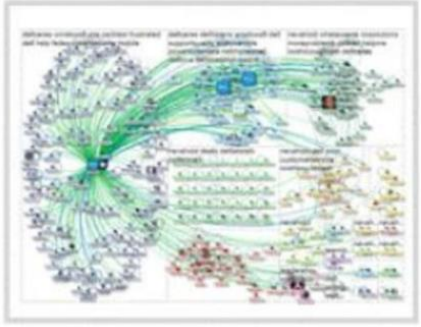
[Unified]
Tight Crowd

[Fragmented]
Brand Clusters

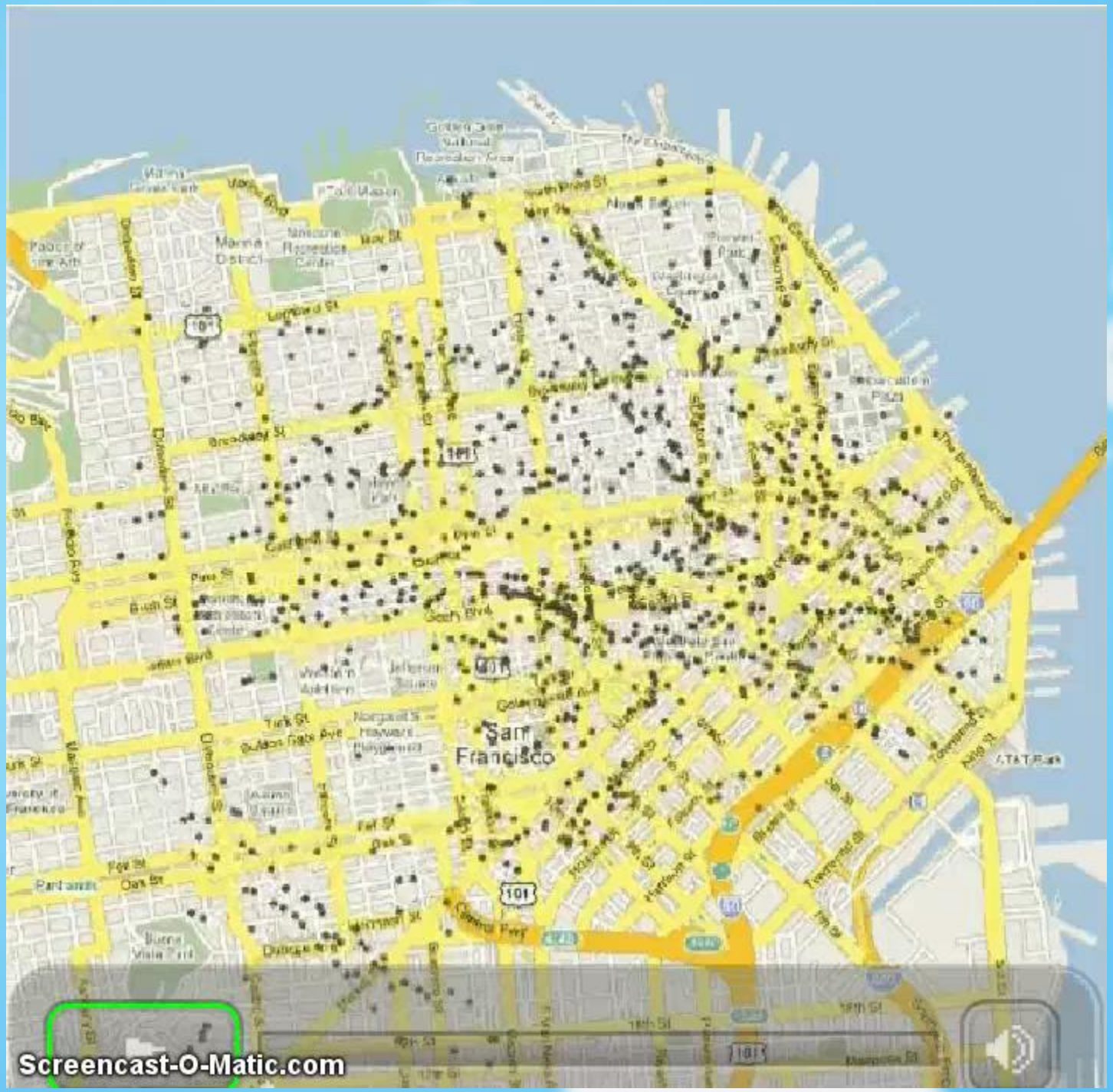
[Clustered]
Community Clusters

[In-Hub & Spoke]
Broadcast Network





[Out-Hub & Spoke]
Support Network


San Francisco
Screencast-O-Matic.com

Mining clinical text for signals of adverse drug-drug interactions

Srinivasan V Iyer, Rave Harpaz, Paea LePendu, Anna Bauer-Mehren, Nigam H Shah

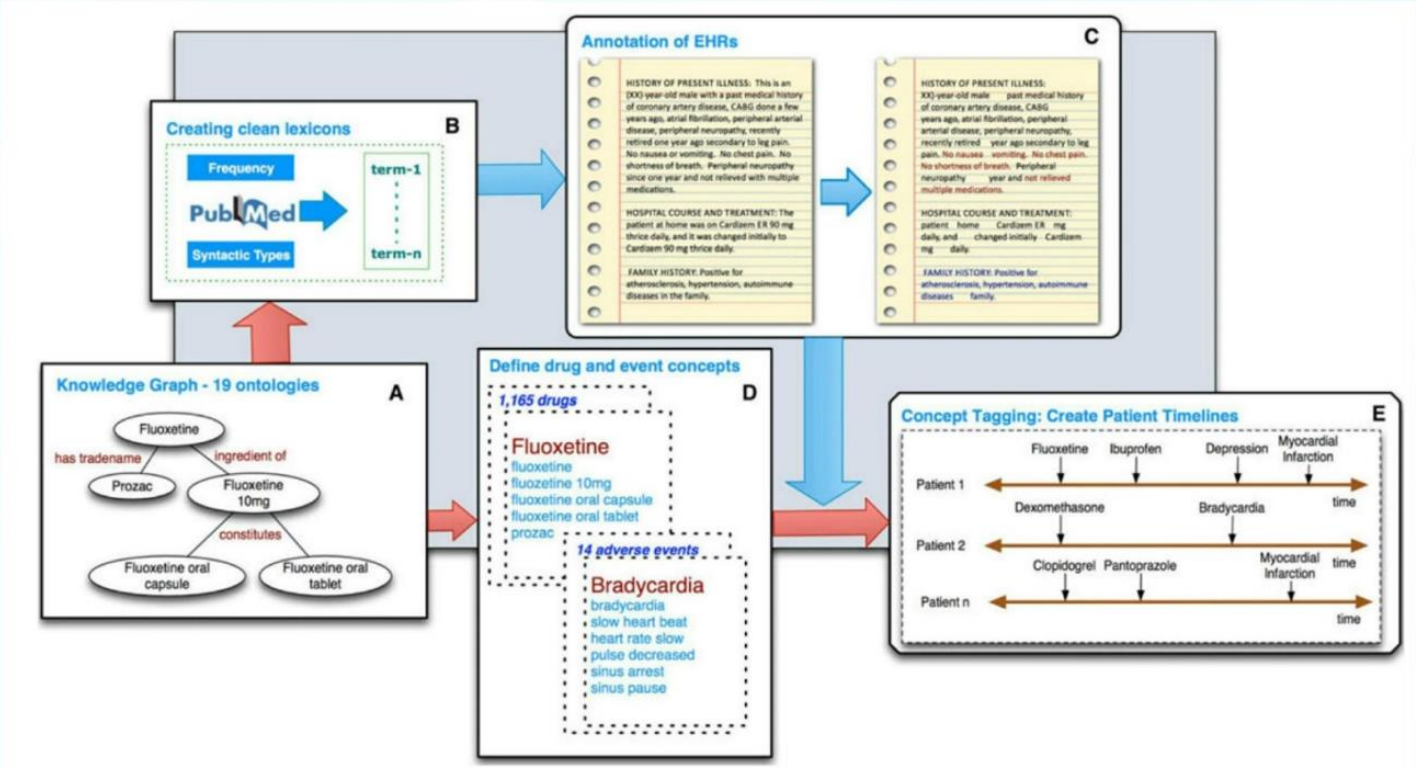


Figure 1 The annotator workflow. (A) The annotator uses a lexicon of approximately 5.6M terms derived from the Unified Medical Language System (UMLS) and BioPortal, as well as trigger terms for NegEx and ConText. (B) It uses term frequency and syntactic type information from Medline to prune the set of strings into a clean lexicon. (C) It then uses the lexicon for exact string matching on the textual notes, followed by negation detection (red) and family history detection (blue). The output is a list of positively mentioned terms recognized in the text. (D) UMLS and BioPortal terms are used to define concepts (a set of terms), making use of the relationships in the ontologies to expand the set. (E) Each note is tagged with a concept if any one of the defining terms appears in the note as a positive mention. The concepts are ordered by the note's timestamp, creating a concept timeline for each patient.

Dose-Specific Adverse Drug Reaction Identification in Electronic Patient Records: Temporal Data Mining in an Inpatient Psychiatric Population

Robert Eriksson · Thomas Werge · Lars Juhl Jensen · Søren Brunak

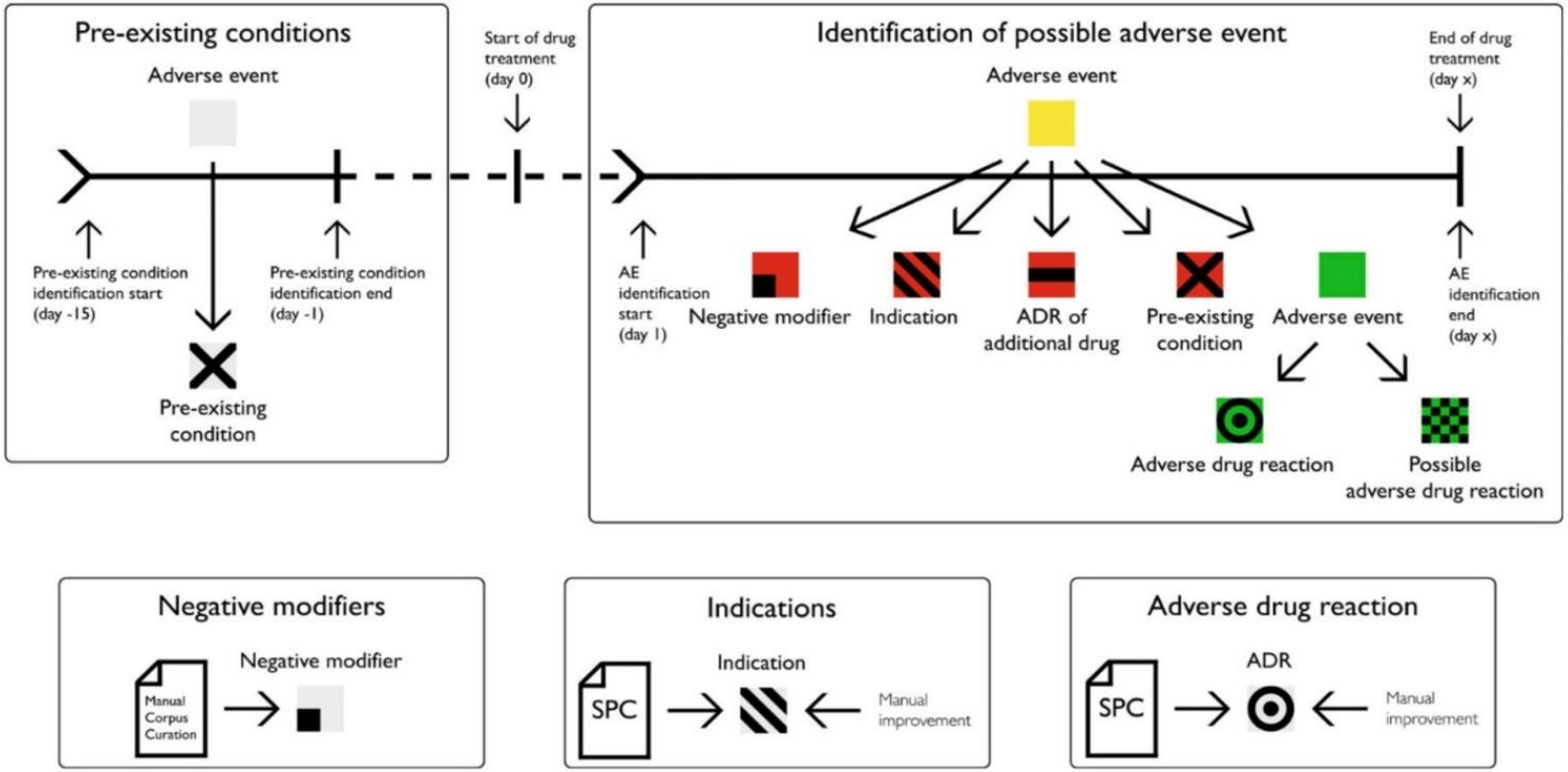


Fig. 1 AE extraction and data integration. AEs were extracted between drug introduction and discontinuation, where we filtered out AEs if the text in the clinical note suggests it did not happen, affected someone else or happened in the past. Additionally, we filtered all indications of the drug and ADRs related to additional drugs. Finally, all pre-existing conditions were removed. Remaining AEs were sorted into ADRs and possible ADRs; the latter was presented for manual review. *ADR* adverse drug reaction, *AE* adverse event, *EPR* electronic patient record, *SPC* Summary of Product Characteristics

Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features

Azadeh Nikfarjam[1], Abeed Sarker[1], Karen O'Connor[1], Rachel Ginn[1], Graciela Gonzalez[1]

re 1: Examples of user-posted drug reviews in Twitter (a) and DailyStrength (b).

a) #***Schizophrenia***$_{indication}$ #*Seroquel did not suit me at all. Had severe* ***tremors***$_{ADR}$ *and* ***weight gain***$_{ADR}$.

b) *I felt awful, it made my* ***stomach hurt***$_{ADR}$*with bad* ***heartburn***$_{ADR}$ *too,* ***horrid taste in my mouth***$_{ADR}$ *tho it does tend to clear up the* ***infection***$_{Indication}$.

Text Mining for Adverse Drug Events: the Promise, Challenges, and State of the Art

Rave Harpaz · Alison Callahan · Suzanne Tamang · Yen Low · David Odgers · Sam Finlayson · Kenneth Jung · Paea LePendu · Nigam H. Shah

Circulation. 2004 May 4;109(17):2068-73. Epub 2004 Apr 19.

NOUN PHRASE — NOUN PHRASE

Relationship between selective cyclooxygenase-2 inhibitors and acute myocardial infarction in older adults.

RELATIONSHIP?

Solomon DH[1], Schneeweiss S, Glynn RJ, Kiyota Y, Levin R, Mogun H, Avorn J.

Author information

Abstract

NOUN PHRASE — VERB PHRASE — NOUN PHRASE

BACKGROUND: Although cyclooxygenase-2 inhibitors (coxibs) were developed to cause less gastrointestinal hemorrhage than nonselective nonsteroidal antiinflammatory drugs (NSAIDs), there has been concern about their cardiovascular safety. We studied the relative risk of acute myocardial infarction (AMI) among users of celecoxib, rofecoxib, and NSAIDs in Medicare beneficiaries with a comprehensive drug benefit.

METHODS AND RESULTS: We conducted a matched case-control study of 54 475 patients 65 years of age or older who received their medications through 2 state-sponsored pharmaceutical benefits programs in the United States. All healthcare use encounters were examined to identify hospitalizations for AMI. Each of the 10 895 cases of AMI was matched to 4 controls on the basis of age, gender, and the month of index date. We constructed matched logistic regression models including indicators for patient demographics, healthcare use, medication use, and cardiovascular risk factors to assess the relative risk of AMI in patients who used rofecoxib compared with persons taking no NSAID, taking celecoxib, or taking NSAIDs. Current use of rofecoxib was associated with an elevated relative risk of AMI compared with celecoxib (odds ratio [OR], 1.24; 95% CI, 1.05 to 1.46; P=0.011) and with no NSAID (OR, 1.14; 95% CI, 1.00 to 1.31; P=0.054). The adjusted relative risk of AMI was also elevated in dose-specific comparisons: rofecoxib < or =25 mg versus celecoxib < or =200 mg (OR, 1.21; 95% CI, 1.01 to 1.44; P=0.036) and rofecoxib >25 mg versus celecoxib >200 mg (OR, 1.70; 95% CI, 1.07 to 2.71; P=0.026). The adjusted relative risks of AMI associated with rofecoxib use of 1 to 30 days (OR, 1.40; 95% CI, 1.12 to 1.75; P=0.005) and 31 to 90 days (OR, 1.38; 95% CI, 1.11 to 1.72; P=0.003) were higher than >90 days (OR, 0.96; 95% CI, 0.72 to 1.25; P=0.8) compared with celecoxib use of similar duration. Celecoxib was not associated with an increased relative risk of AMI in these comparisons.

CONCLUSIONS: In this study, current rofecoxib use was associated with an elevated relative risk of AMI compared with celecoxib use and no NSAID use. Dosages of rofecoxib >25 mg were associated with a higher risk than dosages < or =25 mg. The risk was elevated

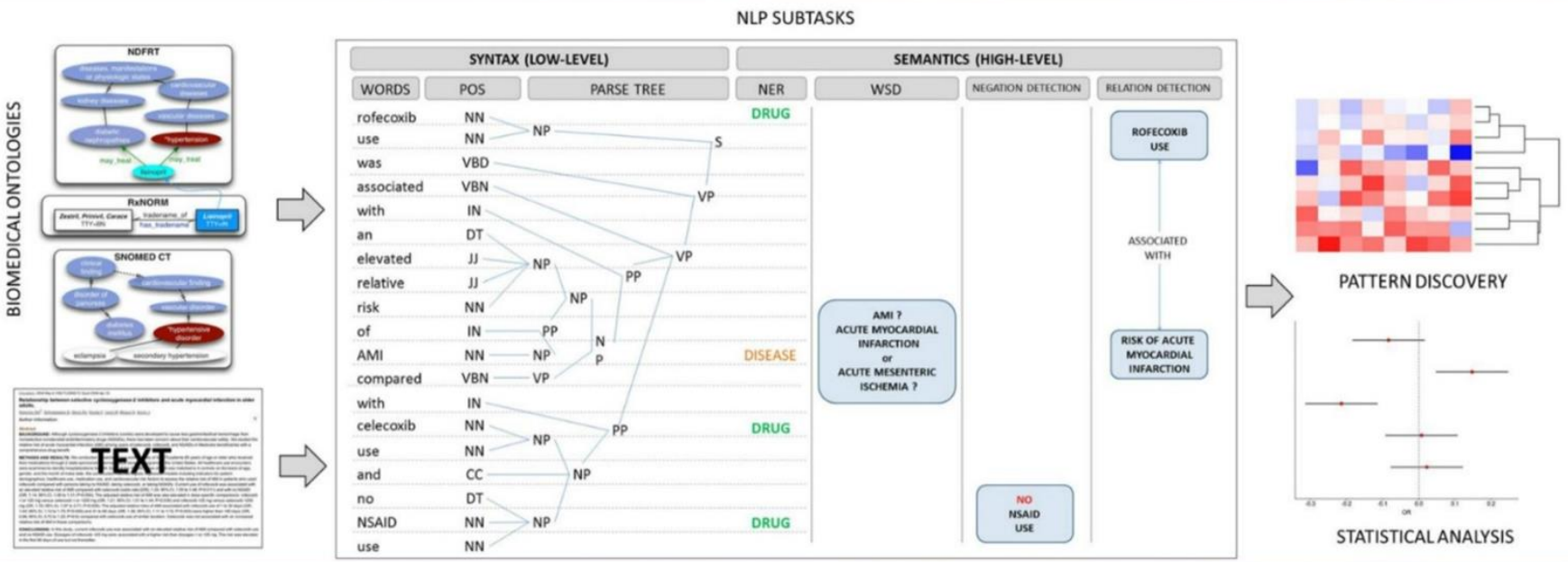
MeSH Terms

- Anti-Inflammatory Agents, Non-Steroidal/adverse
- Case-Control Studies
- Cyclooxygenase 2
- Cyclooxygenase 2 Inhibitors
- Cyclooxygenase Inhibitors/adverse effects*
- Dose-Response Relationship, Drug
- ...
- Myocardial Infarction/chemically induced*
- Myocardial Infarction/epidemiology
- Prostaglandin-Endoperoxide Synthases

Substances

- Anti-Inflammatory Agents, Non-Steroidal
- Cyclooxygenase 2 Inhibitors
- Cyclooxygenase Inhibitors
- Isoenzymes
- Lactones
- Membrane Proteins
- Pyrazoles
- Sulfonamides
- Sulfones
- rofecoxib
- celecoxib
- Cyclooxygenase 2
- PTGS2 protein, human
- Prostaglandin-Endoperoxide Synthases

= SUBHEADING * = MAJOR TOPIC


NLP SUBTASKS
BIOMEDICAL ONTOLOGIES
NDFRT
RxNORM
SNOMED CT
TEXT
SYNTAX (LOW-LEVEL)
SEMANTICS (HIGH-LEVEL)
WORDS
POS
PARSE TREE
NER
WSD
NEGATION DETECTION
RELATION DETECTION
rofecoxib NN
use NN
was VBD
associated VBN
with IN
an DT
elevated JJ
relative JJ
risk NN
of IN
AMI NN
compared VBN
with IN
celecoxib NN
use NN
and CC
no DT
NSAID NN
use NN
DRUG
DISEASE
DRUG
DRUG
AMI ? ACUTE MYOCARDIAL INFARCTION or ACUTE MESENTERIC ISCHEMIA ?
NO NSAID USE
ROFECOXIB USE
ASSOCIATED WITH
RISK OF ACUTE MYOCARDIAL INFARCTION
PATTERN DISCOVERY
STATISTICAL ANALYSIS

THANK YOU

Fadia T. Shaya, PhD, MPH

fshaya@rx.umaryland.edu