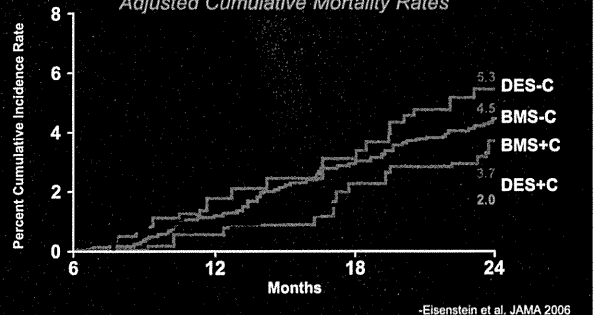
Duke Database: 6-Month Landmark Analysis

Adjusted Cumulative Mortality Rates



-Eisenstein et al. JAMA 2006

Duke Clinical Research Institute

Continuity of Care

Early Clopidogrel Discontinuation of within 30 Days After DES

2498 patients enrolled in PREMIER

1521 patients underwent PCI

189 POBA | 946 BMS | 500 DES

Predictors of Early Stopping
- Older
- Less Educated
- Poorer
- Less D/C Med Instructions

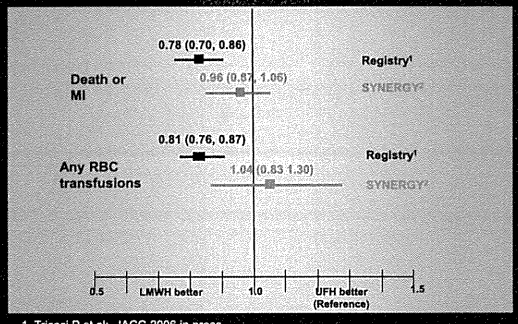
86% taking clopidogrel at 1-month

14% NOT taking clopidogrel at 1-month

Mortality increased nearly 10X

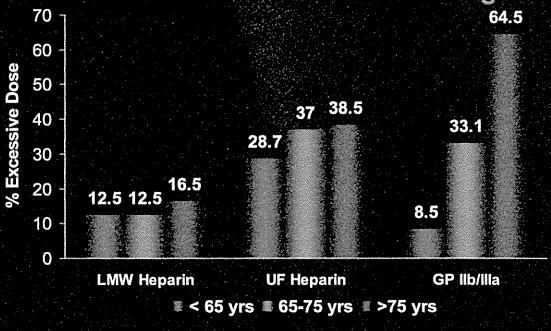
Spertus JA Circ 2006:113: 2803

Comparative Effectiveness: LMHW VS UF Heparin



1. Tricoci P et al: JACC 2006 in press
2. SYNERGY INVESTIGATORS *JAMA* ;2004:45-54

Safe Care? Errors of Commission Excessive Antithrombotic Dosing

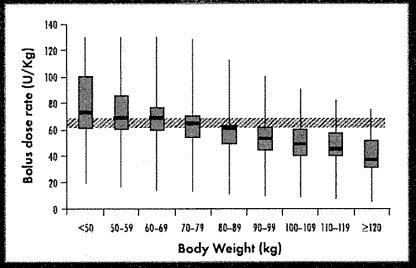


Duke Clinical Research Institute

Alexander KA, et al. *JAMA* 2005;294:3108-3116

Failure to Individualize Care: Weight and Dosing

Unfractionated Heparin Bolus by Weight in Kg

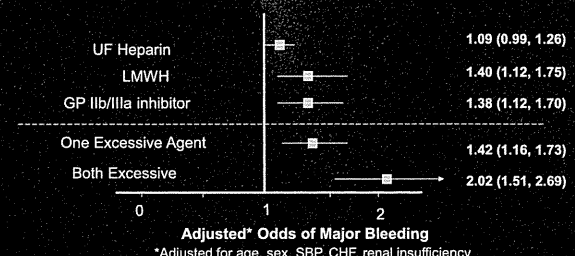


Duke Clinical Research Institute

Melloni, Circulation 2007; Am Heart J 2008

Safety Matters! Major Bleeding Risks with Excess Dosing

Excess v. Recommended



Duke Clinical Research Institute

Alexander KA, JAMA 2005

First Rule of Registries:
Partnering with the Right Partner!

DCRI Mission

"To develop and share knowledge that improves the care of patients around the world through innovative clinical research"

Duke Clinical Research Institute

No Johnny-come-lately to Registries

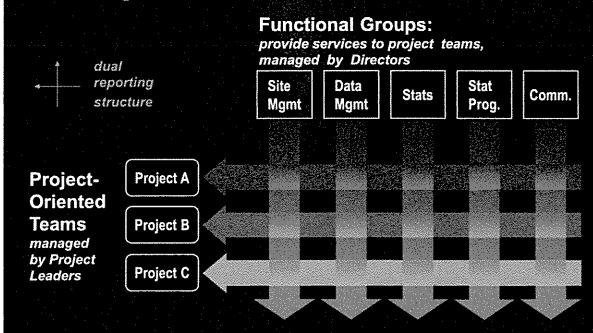
"Chronic diseases can be studied, but not by the methods of the past. If one wishes to create useful data … computer technology must be exploited."

— Eugene Stead, MD

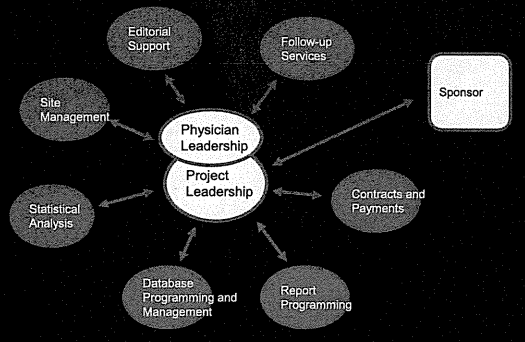
- Led to the concept of "*computerized textbook of medicine*"
- Formed foundation of the Duke Databank for CV Diseases
- Spurred a generation of clinical and quantitative researchers

Duke Clinical Research Institute

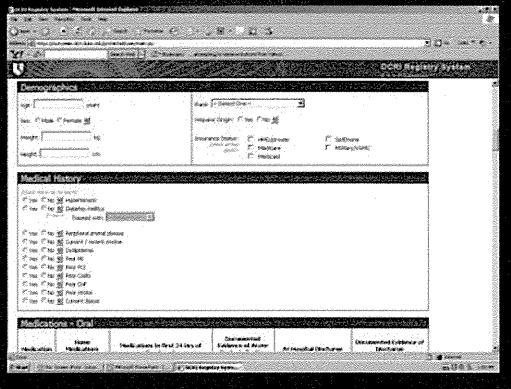
DCRI Registry Operations
Matrix Organizational Structure



Registry Project Team



User-friendly Web-based Data Submission



Outcomes Registry Research: Our Philosophy

- **Dynamic approach to evaluating "real world" care**
  - Health policy application
- **Continuous Quality Improvement**
  - Feedback to sites
  - Patient education
- **Scientific Research**
  - Disparities in care
  - Process of care and associated outcomes
  - Proven success:
  - Elevate analytic methodology
  - Risk models development
  - Hierarchical 'mixed' modeling for site effects
- **Maximize registry scope**

Potential Roles of Registries in the Development and Implementation of Evidence

- Disease Characterization & Surveillance
  - Community-based event rates
- Clinical and comparative 'Effectiveness'
  - Real world look at therapy's benefits, risks and costs
- Quality Measurement
  - Is community adopting and implementing evidence-based care?
- Quality Improvement
  - Use measurement to stimulate practice change
  - Track impact of changes in health policy
- Identify the unmet needs…
  - Defining the 'denominator'

Professional Society CV Clinical Registries

- Society of Thoracic Surgery*: 900+ centers
  - Coronary artery bypass surgery
  - Valve surgery
  - Congenital heart surgery
  - Thoracic surgery
- National Cardiovascular Data Registry*: 1600+ Hospitals
  - Cath/Percutaneous coronary intervention
  - Implantable cardiac defibrillators (ICD)
  - Acute coronary syndromes (ACS)
  - Carotid stenting
  - Ambulatory CV disease (launching)
- AHA-Get With The Guideline Program*: 1500+ hospitals
  - Coronary artery disease (CAD)
  - Heart failure
  - Stroke
  - Ambulatory module (launching)

Duke Clinical Research Institute

*DCRI serves as Data Analytic Center

Pharmaceutical Partnerships

- CRUSADE: **ACS Hospitalizations**
  - Partners Millennium, Schering Plough, BMS, Sanofi
  - 500+ hospitals, 200,000 ACS patients
- MAINTAIN: **Post ACS Longitudinal Study**
  - Partners: Merck- Schering, BMS/Sanofi
  - 40 hospitals, 1500 patients
- AVAIL: **Post Stroke Longitudinal Study**
  - AHA GWTG, BMS
  - 100 hospitals, 3000 patients
- PREVAIL: **Pre-Diabetes Longitudinal Study**
  - Tethys Bioscience
  - 40 centers, 3,000 patients
- ORBIT: **Atrial Fibrillation Longitudinal Study**
  - Ortho-McNeil
  - 400 centers, 10,000 patients (initial)
- TRANSLATE ACS: **ACS Longitudinal Study**
  - Lilly
  - 400 centers, 15,000 patients

Other Clinical Registries Coordinated at DCRI

- Washington State PCI/CABG Registry
- SABG: **Staph A registry**
- ICE: **International Collaboration on Endocarditis**
- CATCH: **Heparin Induced Thrombocytopenia Registry**
- SCVIR: **Uterine Artery Embolism Registry**
- Longitudinal Initiatives
  - **CMS Linkages**
  - **Long-term follow-up for subset of GWTG-Stroke (AVAIL) and CRUSADE (MAINTAIN) patients**

Who Sponsors DCRI Registry Research?

- Government Agencies
  - National Institutes of Health (NIH)
  - Agency for Healthcare Research and Quality (AHRQ)
  - Food and Drug Administration (FDA)
- Professional Societies
  - AHA, ACC, STS
- Private Industry
  - Pharmaceutical companies
  - Biotechnology companies
  - Medical device companies

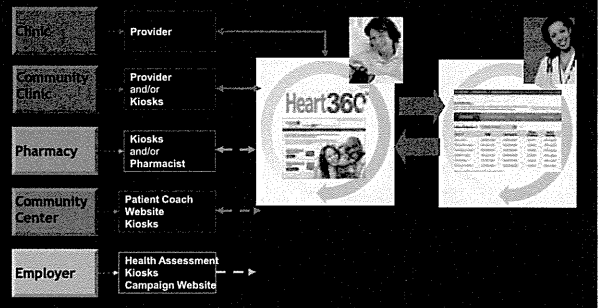
Potential Roles of Registries in the Development and Implementation of Evidence

- Disease Characterization & Surveillance
  - Community-based event rates
- Clinical and comparative 'Effectiveness'
  - Real world look at therapy's benefits, risks and costs
- Quality Measurement
  - Is community adopting and implementing evidence-based care?
- Quality Improvement
  - Use measurement to stimulate practice change
  - Track impact of changes in health policy
- Identify the unmet needs…
  - Defining the 'denominator'

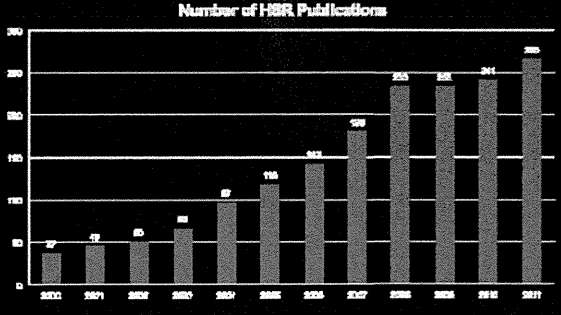
Question 4

- **Will registries be used to directly interact with patients?**
  - A. Yes
  - B. No

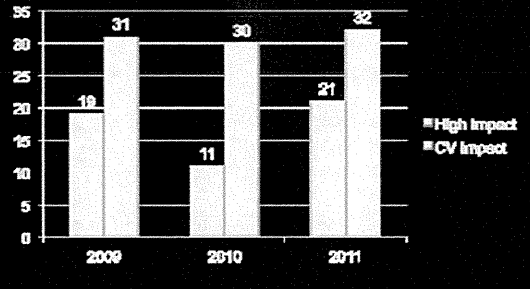
Collaborative Community Care



Health Services Research and Outcomes Publications

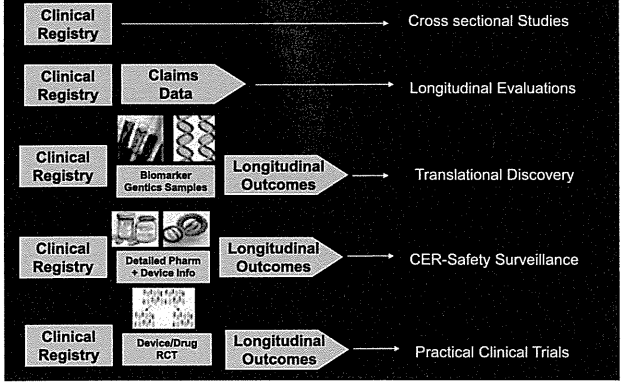


Publication Impact



High Impact: NEJM, JAMA, Annals
CV Impact: Circulation, JACC

Expanding Capacity and Use of Clinical Registries



Conclusions:

Registries have a major role:

- **Evaluating healthcare organization and delivery**
- **Quality of care**
- **Bridging the gap between clinical trials and the real-world**
- **Safety surveillance**
- **Improving patient-centered outcomes**

Questions?

Thank you

Some Examples…

Data Quality Assurance

- **Site Data Collection Training**
  - Virtual Meetings - webinars, teleconferences
  - Multimedia training modules
  - FAQs, newsletters
- **Data Edit Checks**
- **Data Quality Feedback**
  - Iterative process
  - Highly automated
  - Record-specific and aggregate views
- **On-site QA audits**
  - Do-able but costly, not recommended

Follow-up Strategy

entails….

Site-conducted:

- Site training
- Patient return visit
- Site payment for conduct of follow-up
- Site communication
- Site monitoring
- Annual site IRB renewal
- Data QA

Site-Conducted Follow-up

$$$

Centralized Follow-up

Centralized:

- Central personnel training
- Telephone/Web-based follow-up
- Central payment for FU
- Central IRB

Follow-up Alternatives

- **Endpoints collectible without face-face visit**
  - Minimize loss to follow-up
- **Advantages of centralized follow-up?**
  - Consistent data collection
  - Eases site burden
  - Short-term study more appealing to sites
  - Reduces site management/monitoring activities
  - Coordinating center already contacting patients to conduct health outcomes interview

Endpoint Collection/Validation

"Bottom Line: COST"

Collected Endpoints:

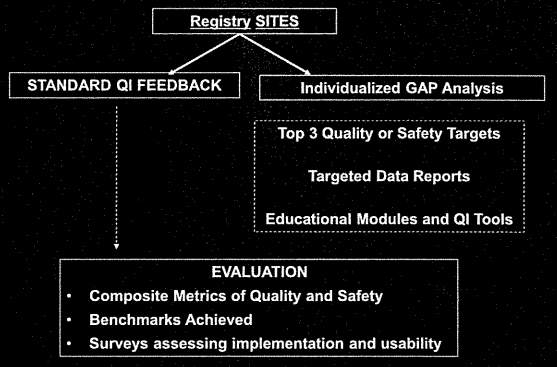
- Death
- Myocardial infarction
- Revascularization
- Bleeding

- **Three potential approaches:**
  - Patient-reported outcomes
  - Obtain medical records with clinical event adjudication
  - Obtain billing data – ICD9 codes
- **Validation**
  - Limited adjudication of early events
  - Annual "check-in" with sites for verification of hospitalization

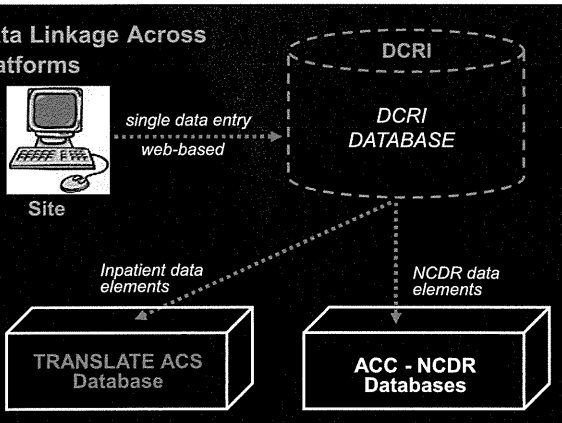
Follow-up and EQOL

- **Integrated with main study follow-up plan**
  - Patient interviews at 6 month intervals
- **Experienced Follow-up Services Group**
  - Duke Databank for Cardiovascular Disease
    - *> 30 years of routine patient contact after cath/revasc*
  - Long-term follow-up for many global trials
    - *Follow-up rates >90%*
- **Health Economics/QOL Group**
  - Centralized data collection
  - Experts in EQOL methodologies
    - *– randomized sampling for QOL study population*

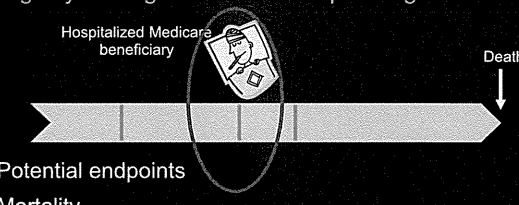
Clinical Trials: Personalized And Targeted QI



Data Linkage Across Platforms



Registry + Longitudinal Follow-up through Medicare



Potential endpoints

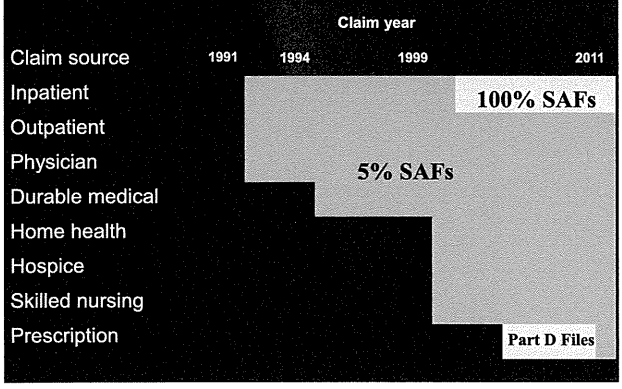
Mortality

Readmission

Procedure

Adverse events (based on coded diagnoses and procedures)

DCRI Medicare Data Warehouse

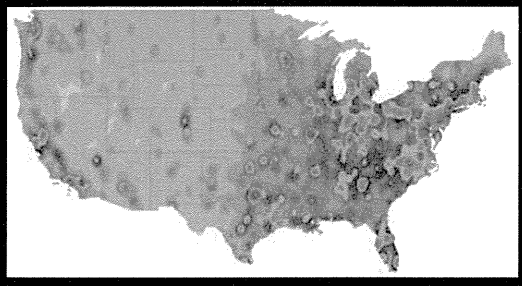


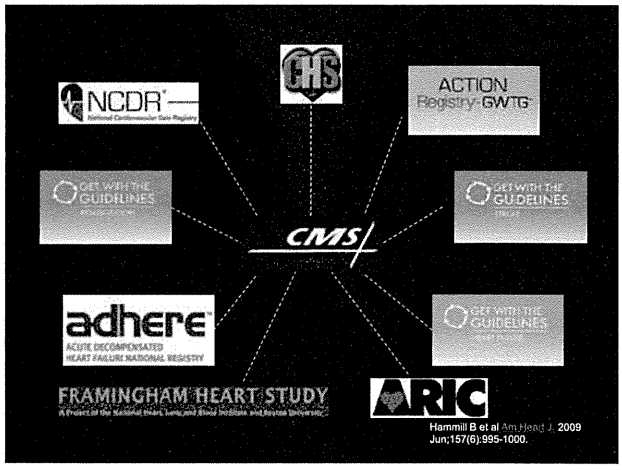
Specialized Medicare Cohorts

- **100% AF Ablation**
- **100% CRT-D**
- **100% Macular degeneration**
- **.....**

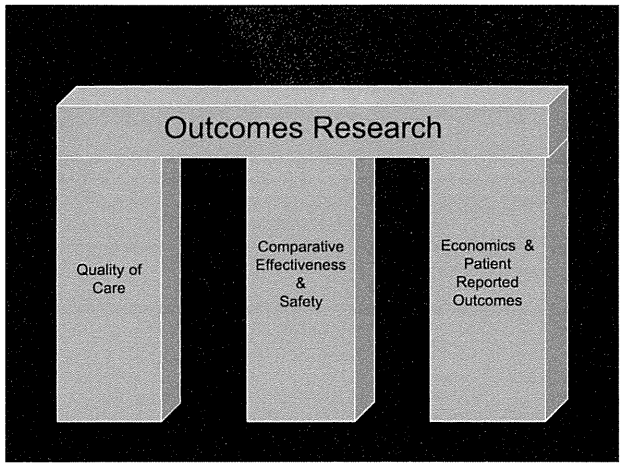
Medicare data warehouse: >78 billion records in >3.8 TB of storage

Patient Access: Heart Failure and VADs





Hammill B et al Am Heart J. 2009 Jun;157(6):995-1000.



Duke Clinical Research Institute

Best Practice in Conducting and Reporting Clinical Studies using Large Databases

International Symposium, Tokyo, Japan

Soko Setoguchi, MD, DrPH, FISPE
Associate Professor of Medicine
Duke Clinical Research Institute and Duke University School of Medicine

Outline

- Databases
- What is the best practice and what is needed?
- What can you do using databases?
- Conducting database studies
  - Important concepts in conducting clinical studies
- Reporting (very briefly)
- Summary

Databases

- Electronic health record data
- Administrative data (claims data)
- Registries
- Cohort study data

Computerization of Health-care Related Records

- Records for billing purposes were standardized/computerized
  - 1970's in pharmacoepidemiology and health service research
- Computerization of health information (e.g., as a form of EHR) is rapidly developing and being used

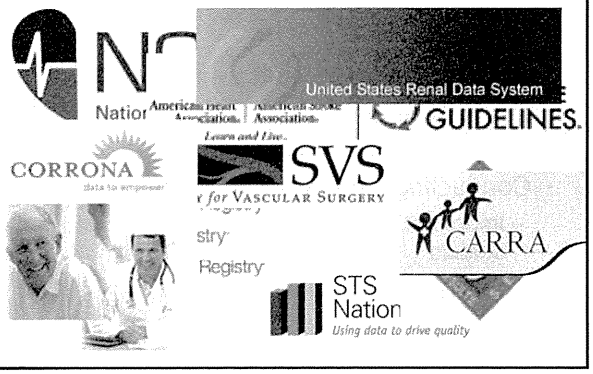
Administrative Databases

- **Claims databases or health care utilization databases'**
- **Examples in North America**
  - Medicaid
  - Medicare + State Pharmacy Assistance Programs or Part D
  - Commercial insurance companies
    - United Health
    - Blue Cross Blue Shield
  - Canadian Provincial claims data
    - Ontario
    - Quebec
    - British Columbia
    - Saskatchewan

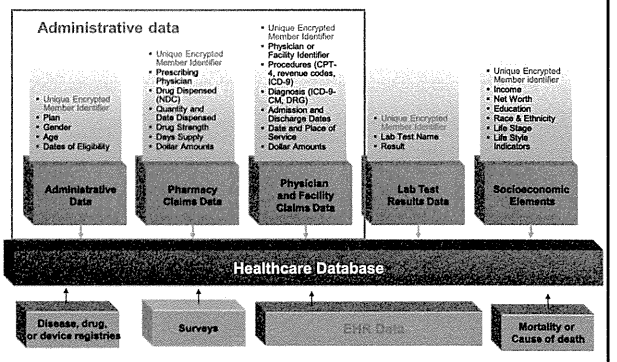
Electronic Health Record (EHR) Database

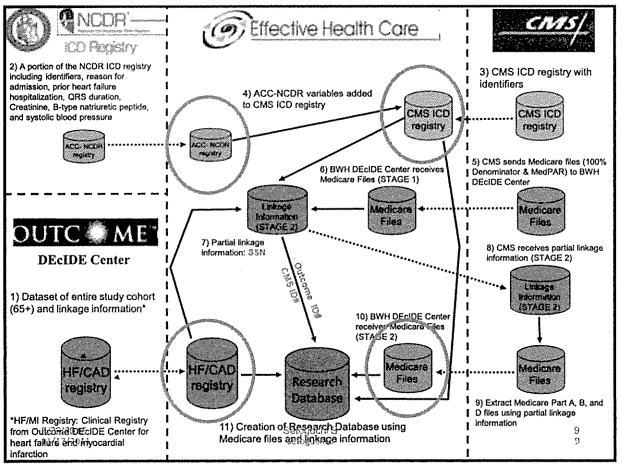
- Examples
  - Single provider
    - DEDUCE (Duke)
    - RPRD (Brigham and Women's Hospital)
  - Multiple providers
    - Geisinger Clinic Electronic Health Records- 41 Clinics covering ~3 million patients
  - EHR combined with administrative data
    - Kaiser Permanente

Registries- Examples



Better Data





Linkage without Unique Identifiers

- Is unique personal identifiers necessary for linkage?
  - SSN, Medicare ID etc
- Hospitalization records can be linked using multiple non-unique identifiers
  - Demographic: date of birth, gender
  - Service information: admission date, discharge date, diagnosis, procedure, provider information
- Key is to use a good combination of variables that makes a record unique

10

Validity of Different Linkage Rules When Patient Unique Identifiers Are Not Available

Linkage Rules	# of 1-1 linkage	Sen.	Spec.	PPV
R1: Gender, DOB, Adm. Date, Provider ID	136,117	95	98	98
R3: Gender, DOB, Adm. Date, **Provider State**	135,537	89	91	93
R4: Gender, DOB, Adm. Date	5	0	100	0
Gold-Standard: SSN, Provider ID, Adm. Date	136,511	ref	ref	ref

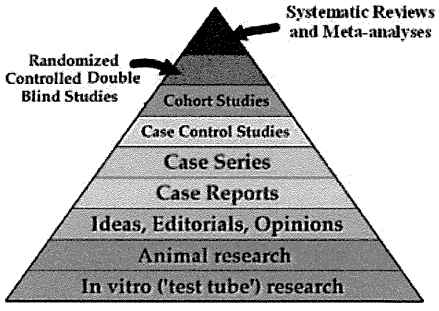
11

Information has to be put together and interpreted in a meaningful way

Anticipated Challenges in Database Studies

- Common problems with databases
  - Inaccurate data (information bias)
  - Missing data
  - Not all important data components may be available in databases to control confounding or selection bias

The Evidence Pyramid



Duke Clinical Research Institute

Quiz 1

- The best evidence come from which of the following types of studies?
  - *Systematic reviews and meta-analyses*
  - *Randomized controlled studies*
  - *Prospective cohort studies*
  - *Observational cohort studies using databases*
  - *Observational case-control studies using databases*
  - *Case reports*

15

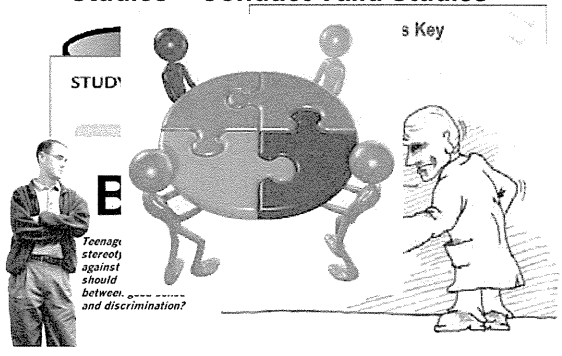
Best Evidence Come From?

- **The best evidence is usually found in clinically relevant research that has been conducted using sound methodology.** (Sackett D, 2002)
- **The forms of studies providing best evidence depends on the nature and types of questions**

Study Design Choices

Explanatory RCT	Large Simple or Pragmatic Trial	Observational Studies
'Gold standard'; selected study population, unusual settings	Randomized ; usual setting of care; non-selected study population	Cohort Study Case-Control Study Case-Cohort Study Case Cross Over Study
Efficacy - Does it work in an ideal situation?	**Effectiveness** - Does it work in the 'real world'?	

Best Practice in Conducting Clinical Studies = Conduct Valid Studies



Best Practice in Conducting Clinical Studies = Conduct Valid Studies

- Knowledge/skills
  - About databases
  - About clinical questions and background
  - Study design (understanding and writing protocol)
  - Bias (selection, confounding, misclassification)
  - Statistical analysis (descriptive statistics, multivariate analysis, propensity score-based methods, instrumental variable analysis, hierarchical modeling)
- Team work

Epidemiology is NOT ONLY All about…

- Measuring incidence and prevalence of diseases
- Understanding disease risk factors
- Prevention of diseases
- Treatment safety or effectiveness

But It is also

- A powerful tool
- Epidemiologic methods can provide approaches and frameworks for valid observational studies including database studies

Epidemiology as A Tool

- Everyone needs a tool to make sense of data
- Epidemiology provide tools and frameworks to deal with and interpret the information
  - So as other disciplines (Health services research, Biostatistics, Econometrics etc)
  - Emphasis on design, measurements, and control of biases

Types of Database Studies

- Disease burden
- Health services research
- Outcome research
- Rick factor studies
- Pharmaco- and device- epidemiology (drug and device safety and effectiveness)
- Comparative effectiveness research

Examples of Studies Using Large Databases

Disease Burden

Trends in Out-of-Hospital Deaths Due to Coronary Heart Disease in Sweden (1991 to 2006)

Kerstin Dudas, PhD; Georg Lappas, BSc; Simon Stewart, PhD; Annika Rosengren, MD, PhD

Background—Case fatality associated with a first coronary event is often underestimated when only those who survive to reach a hospital are considered. Few studies have examined long-term trends in case fatality associated with a major coronary event that occurs out of the hospital.

Methods and Results—Record linkage documented all case subjects 35 to 84 years of age in Sweden during 1991 to 2006 with a first major coronary event (out-of-hospital coronary death or hospitalization for acute myocardial infarction). Of the 384 597 cases identified, 111 319 (28.9%) died out of the hospital, and another 36 552 (9.5%) died in the hospital or within 28 days of hospitaliza ury events declined from 30.5% to 6), however, with a larger decline i 6% confidence interval 5.5% to 6.0 of out-of-hospital deaths to overall to 54 years of age, no more than 1 ale sex (odds ratio 0.85, 95% confi val 0.971 to 0.974 per year) were a ear was associated with increased r

Conclusions—The great majority ng, particularly among younger ind

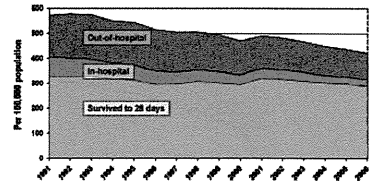
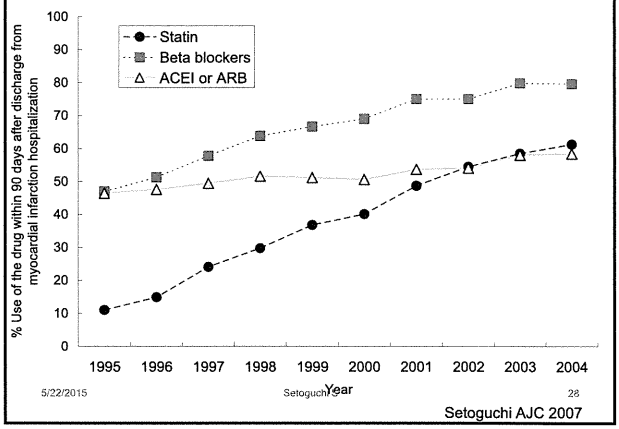


Figure 3. Mortality due to CHD in the hospital (within 28 days) and out of the hospital per 100 000 population 35 to 84 years of age, 1991 to 2006.

Health Services Research

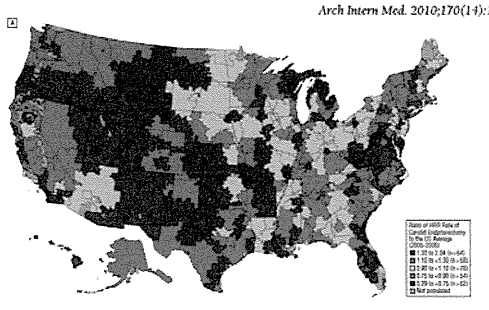


Setoguchi AJC 2007

Geographic Variation in Carotid Revascularization Among Medicare Beneficiaries, 2003-2006

Manesh R. Patel, MD; Melissa A. Greiner, MS; Lisa D. DiMartino, MPH; Kevin A. Schulman, MD; Pamela W. Duncan, PhD, PT; David B. Matchar, MD; Lesley H. Curtis, PhD

Arch Intern Med. 2010;170(14):1218-1225

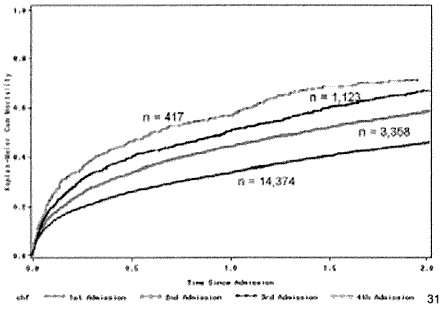


Outcomes Research

Congestive Heart Failure

Repeated hospitalizations predict mortality in the community population with heart failure

Soko Setoguchi, MD, DrPH,* Lynne Warner Stevenson, MD,† and Sebastian Schneeweiss, MD, ScD* *Boston, MA*



Conclusions The simple predictor of mort planning for prognosis.

31

Journal of the American College of Cardiology
© 2008 by the American College of Cardiology Foundation
Published by Elsevier Inc.
Vol. 51, No. 13, 2008
ISSN 0735-1097/08/$34.00
doi:10.1016/j.jacc.2007.10.063

CLINICAL RESEARCH — Acute Myocardial Infarction

Improvements in Long-Term Mortality After Myocardial Infarction and Increased Use of Cardi

A 10-Year Trend

Soko Setoguchi, MD
Murray A. Mittleman
Boston, Massachusetts

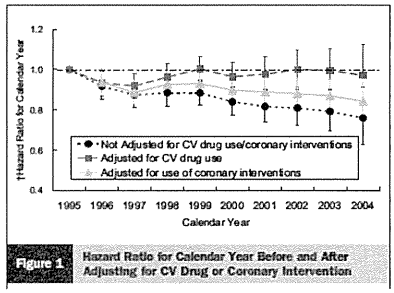


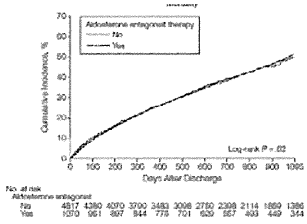
Figure 1 Hazard Ratio for Calendar Year Before and After Adjusting for CV Drug or Coronary Intervention

32

Comparative Effectiveness and Safety Studies

ORIGINAL CONTRIBUTION

Associations Between Aldosterone Antagonist Therapy and Risks of Mortality and Readmission Among Patients With Heart Failure and Reduced Ejection Fraction



Adrian F. Hernandez, MD, MHS
Xiaojuan Mi, PhD
Bradley G. Hammill, MS
Stephen C. Hammill, MD
Paul A. Heidenreich, MD
Frederick A. Masoudi, MD, MSPH

JAMA. 2012;308(20):2097-2107

Statins Prevent Cancer?

Epidemiology

Statins and the Risk of Lung, Breast, and Colorectal Cancer in the Elderly

Soko Setoguchi, MD, DrPH; Robert J. Glynn, PhD, ScD; Jerry Avorn, MD; Helen Mogun, MS; Sebastian Schneeweiss, MD, ScD

Background—Although most randomized trials and meta-analyses suggest a slight or no increase in the risk of cancer in statin users, results from observational studies have been conflicting, and some have even suggested a large protective effect of statins on certain cancers. Long-term statin users tend to be healthier, less frail, and more adherent to therapy than nonusers, however. This could explain such apparent "protective" effects.

Methods and Results—We conducted the present cohort study by linking data from a large state drug benefit program with cancer registry data and Medicare healthcare utilization data. We identified all initiators of statins; initiators of glaucoma medications, another preventive drug, served as a comparison group. Outcomes included all registry-identified cases of colorectal, lung, and breast cancer. Multivariable Cox proportional models were used to adjust for confounding. Patient characteristics were similar in both groups, but statin initiators (n=24 439) were slightly younger and used some services more frequently than glaucoma drug initiators (n=7284). The mean follow-up was 2.9 years, with the longest follow-up being 8.4 years. Incidence rates of colorectal, lung, and breast cancers in both groups were very similar to rates in the general population. Adjusted hazard ratios were 0.96 (95% CI, 0.70 to 1.31) for colorectal cancer, 1.11 (95% CI, 0.77 to 1.60) for lung cancer, and 0.99 (95% CI, 0.74 to 1.33) for breast cancer.

Conclusions—These data from a large population of typical older patients who began using statins indicate that it is unlikely that statins confer a clinically important decrease or increase in the risk of colorectal, lung, or breast cancer o the durations studied. (*Circulation*. 2007;115:27-33.)

35

Timely Assessment

Comparison of Cardiovascular Outcomes in Elderly Patients With Diabetes Who Initiated Rosiglitazone vs Pioglitazone Therapy

Wolfgang C. Winkelmayer, MD, ScD; Soko Setoguchi, MD, DrPH; Raisa Levin, MS; Daniel H. Solomon, MD, MPH

Background: Recent meta-analyses have raised the possibility that rosiglitazone maleate may increase the risk of ischemic cardiovascular events, whereas pioglitazone hydrochloride could not be linked to such a risk. We compared cardiovascular outcomes and mortality between patients initiating pioglitazone vs rosiglitazone therapy.

Methods: We assembled an inception cohort of Medicare beneficiaries older than 65 years with state-sponsored prescription drug benefits who had diabetes mellitus and initiated treatment with rosiglitazone or pioglitazone between January 1, 2000, and December 31, 2005. The study outcomes included all-cause mortality, myocardial infarction, stroke, and hospitalization for congestive heart failure.

Results: Of 28 361 patients selected, 30.3% initiated treatment with pioglitazone and 49.7% with rosiglitazone. Most baseline characteristics were similar between the groups. As preferred in drug safety research, we censored patients at crossover or at 60 days after discontinuation of therapy with their study drug; during 39 060 person-years of follow-up, 1869 patients died. After adjustment for a large number of patient characteristics, Cox regression models revealed 15% greater mortality among patients who initiated therapy with rosiglitazone compared with pioglitazone (95% confidence interval, 5%-26%). Use of rosiglitazone was also associated with a 13% greater risk of congestive heart failure (95% confidence interval, 2%-26%). No differences between the 2 drugs were found in their rates of myocardial infarction or stroke.

Conclusions: Our findings from a large population-based cohort of US seniors are compatible with an increased risk of all-cause mortality and congestive heart failure in patients initiating therapy with rosiglitazone compared with similar patients initiating therapy with pioglitazone. Limitations of this study include residual confounding due to its nonrandomized nature.

Arch Intern Med. 2008;168(22):2268-2275

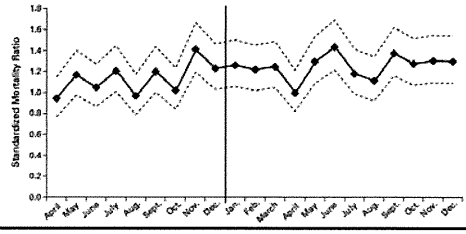
Policy Evaluation

The New England Journal of Medicine

Special Article

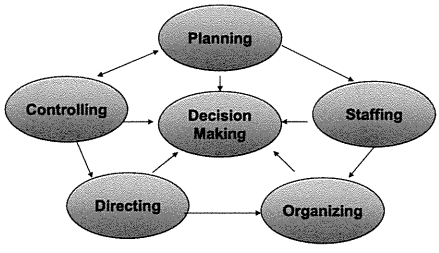
OUTCOMES OF REFERENCE PRICING FOR ANGIOTENSIN-CONVERTING-ENZYME INHIBITORS

Sebastian Schneeweiss, M.D., Alexander M. Walker, M.D., Robert J. Glynn, Ph.D., Malcolm Maclure, Sc.D., Colin Dormuth, M.A., and Stephen B. Soumerai, Sc.D.



Utility of Database Studies in Hospital Setting

Hospital management and Quality of Care



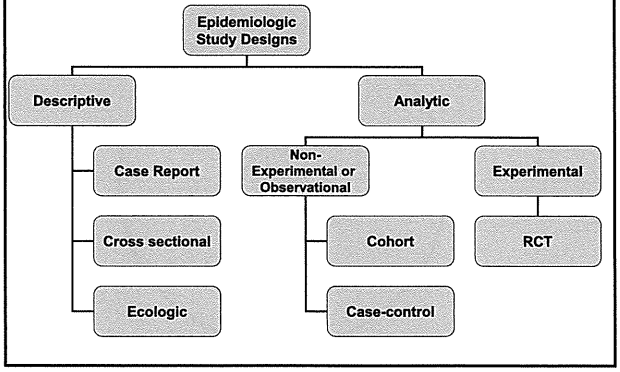
Rakich, Longest, and Darr, 1992

Useful Concepts

Types of Epidemiologic Studies

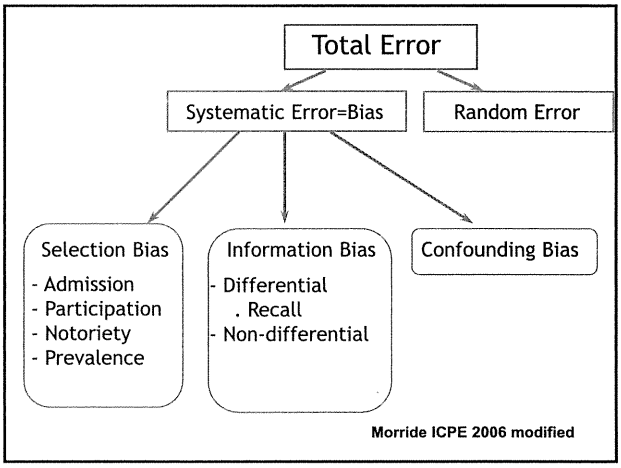
- Descriptive
  - Describe 'what, who, when, where, and how'
- Analytic
  - Understand 'why' (causal inference)
  - *Formulate clear hypotheses as well as causal statements*

Epidemiology Study Designs

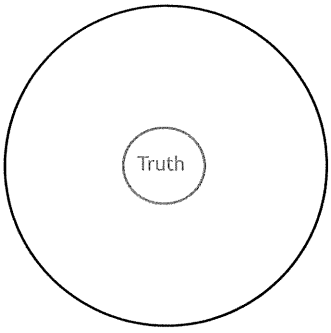


Best Practice = Conduct No Bias Studies

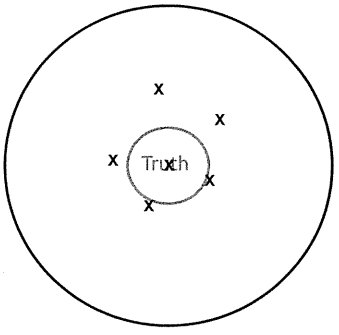
- Maximize the validity= minimize systematic errors (bias)



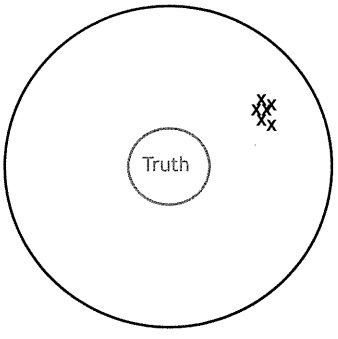
Precision (random error) vs. Validity (bias)



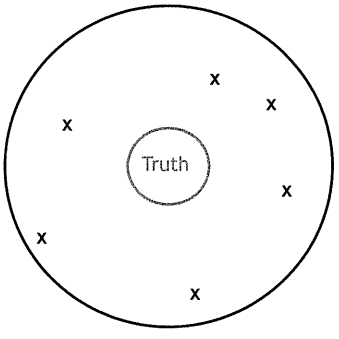
Valid but not precise



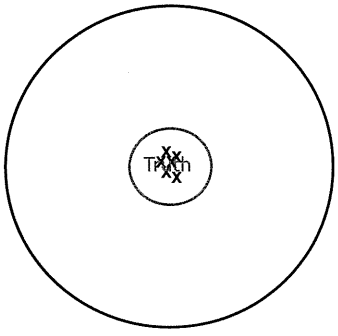
Precise but not valid



Invalid and imprecise



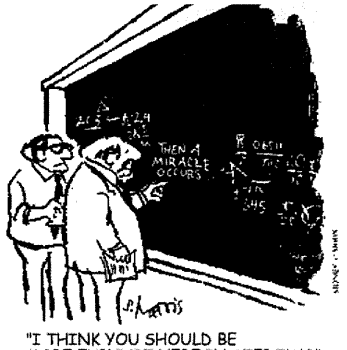
Valid and precise



Internal Validity

- The validity of the inferences drawn for the study subjects
- How can we get the answer from an analysis wrong (bias or distortion in the results)?
  - Information bias or error
  - Confounding (case mix)
  - Selection bias or error

'The devil is in the details…'



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO"

Bias (Systematic Errors)

Inaccurate Information = Information Bias

Quiz 2

- You are interested in assessing geographic distribution in the use of implantable cardioverter defibrillators in elderly (>65 y.o.) patients with heart failure.
- Assume you have a national database for health care utilization data. (e.g., NDB).
- However, you think that heart failure diagnosis in the database is inaccurate and you think the diagnoses for HF will be present who actually do not have clinical heart failure. (e.g., rule out diagnosis).

Duke Clinical Research Institute

Quiz 2

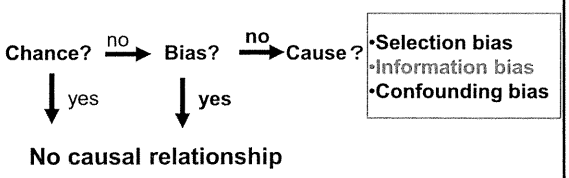
- Using this data, are your estimates for ICD utilization in patients with HF would be inaccurate (biased)? If so, in what way?
  - *My estimates will be accurate and very closer to the true rate!*
  - *My data will likely overestimate the true rate of ICD use*
  - *My data will likely underestimate the true rate of ICD use*
  - *Who knows!*

Calculating Rate of ICD Use in Elderly HF Patients

- Identify patients who are over 65 of age
- Identify and count a subset with diagnosis of HF
- Count the number of ICDs in a specified period in the above population
- Calculate the rate of ICD use

$$\frac{\text{\# of ICDs}}{\text{\# of elderly HF Patients}}$$

Relative Risk > or < 1
Risk Difference > or < 0
Means A Treatment is Effective or Safe?



Information Bias in Safety or Effectiveness Estimates

- Measurement error or classification error on
  - Exposure status
  - Outcome status
  - Confounding

  can cause bias on the effect estimate

Duke Clinical Research Institute

Quiz 3

- Now, you are interested in assessing the effectiveness of ICDs in the elderly HF patients in preventing sudden cardiac death
- Assume that you now have additional information through a linkage of NDB to registries of HF and ICDs and can identify the study population accurately.
- However, you are worried that you cannot accurately identify sudden cardiac death in the linked database.

Expressing 'Accuracy' of Information

- 100% Sensitivity and 100% Specificity
  - Perfectly accurate
- 100% Sensitivity and 50% Specificity
  - Capture all true events but also capture 50% of true non-events as events
- 30% Sensitivity and 100% Specificity
  - Only capture 30% of true events as events but all true non-events are captured as non-events

Numerical Examples of Bias Due to Non-differential Outcome Misclassification

Question: Do ICDs prevent sudden cardiac death in elderly HF patients?

Cohort Study: We followed 10,000 HF patients who received ICDs and 10,000 similar patients who did not and assessed the occurrence of new sudden cardiac death.

	Implantable Defibrillators (ICDs)		
	Exposed	Unexposed	
SCD	80	40	120
No SCD	9920	9960	19880
	10000	10000	20000

Non-differential Outcome Misclassification Bias

No outcome misclassification

		ICDs		
		Yes	No	
SCD	Yes	80	40	120
	No	9920	9960	19880
		10000	10000	20000

Risk Ratio =80/10000 / 40/10000 = 2

Non-differential Outcome Misclassification Bias

No outcome misclassification

		SCD		
		Yes	No	
ICDs	Yes	80	40	120
	No	9920	9960	19880
		10000	10000	20000

Risk Ratio =80/10000 / 40/10000 = 2

100% sensitivity and 50% specificity for outcome definition

	Exposed	Unexposed	
Yes	5040	5020	10060
No	4960	4980	9940
	10000	10000	20000

Risk Ratio =5040/10000 / 5020/10000 = 1.004

Non-differential Outcome Misclassification Bias

No outcome misclassification

Point 1

Non-differential misclassification of outcomes generally cause bias toward the null (direction of the bias is expected toward the null)

Risk Ratio =80/10000 / 40/10000 = 2

100% sensitivity and 50% specificity for outcome definition

	Exposed	Unexposed	
Yes	5040	5020	10060
No	4960	4980	9940
	10000	10000	20000

Risk Ratio =5040/10000 / 5020/10000 = 1.004

Non-differential Outcome Misclassification Bias

No misclassification (100% specificity and 100% sensitivity)

Point 2

Incomplete sensitivity in outcome definition does not cause bias in risk ratio if specificity is 100%

50% sensitivity and 100% specificity

		Exposed	Unexposed	
Outcome	Yes	40	20	60
	No	9960	9980	19940
		10000	10000	20000

Risk ratio =40/10000 / 20/10000 = 2

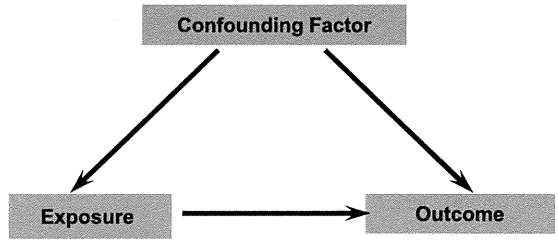
You don't have information in your databases that might distort the association between a therapy and outcome (Confouding)

Confounding Bias

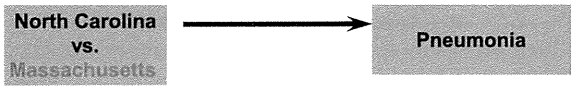
The quantitative association between exposure and outcome is distorted by a third factor with the following characteristics

1. Is a risk factor for the outcome of interest
2. Is a predictor of the exposure of interest
3. Is not an intermediate factor on the causal pathway between exposure and outcome

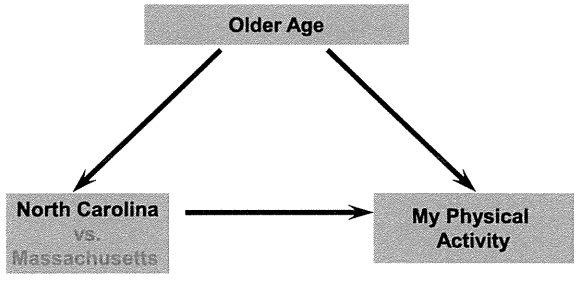
Confounding Bias- 'Famous' Triangle in Epidemiology



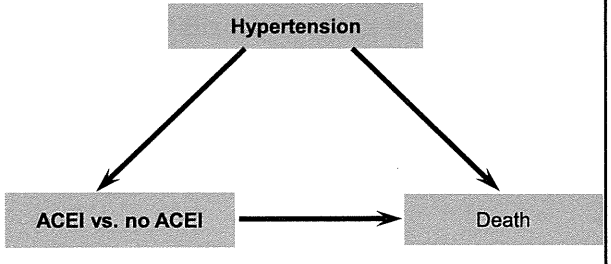
A Real World Example



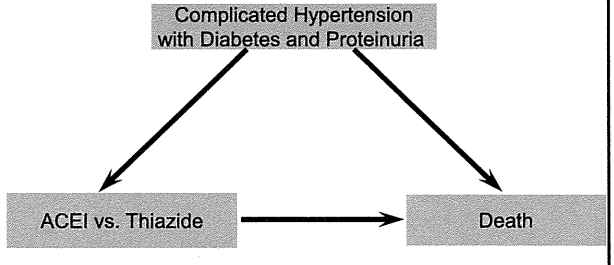
A Real World Example



Confounding by Indication

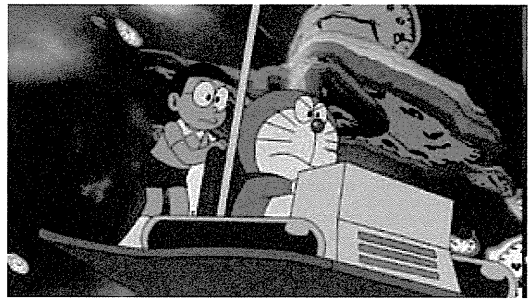


Confounding by Severity in Assessing Effectiveness of ACEIs in Patients with Hypertension



Dream Database Study – No confounding

'ドラえもんがいたら。。'



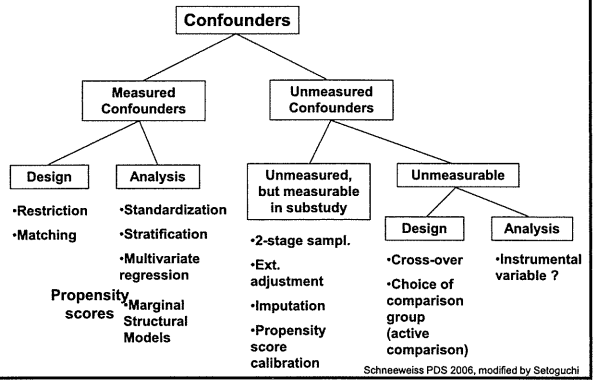
Dream Observational Study

- Time Machine Observational Study
  - No confounding as the treated and non-treated are the same people
- Randomized trials
  - Randomization (if appropriately operationalized) assures the balance in the baseline characteristics between the treated and non-treated patients

Reality in Database Studies

- No Time Machine
  - We cannot observe the same population with and without exposure
- RCT (to compare two groups randomly assigned to exposure) is possible
  - Similar in characteristics on average as a result of randomization
- However, RCT cannot answer every possible question on drug safety or effectiveness
  - Scarce resources
  - Ethical concerns
  - Limitations by design (selective population etc)

Combating Confounding in Observational Studies



Schneeweiss PDS 2006, modified by Setoguchi

Information Bias in Confounding Variables

- Adjustment with a binary non-differentially misclassified confounder reduces the bias and produces a partially adjusted effect estimate that falls between the crude and true effect – Residual confounding Greenland and Robins, AJE 1985
  - "Residual confounding" decreases with increasing sensitivity and specificity of the misclassified confounder Savitz and Baron, AJE 1986
  - Additional assumption – Effect of the confounder on the outcome is in the same direction among the treated and the untreated (ie, there is no qualitative interaction between the treatment and the confounder) Ogburn and VanderWeele, Epidemiology 2012
  - Assumption of no qualitative interaction between treatment and confounder will likely hold in most applications in epidemiology
- Polytomous confounding
  - Conflicting studies Fung and Howe, Int J Epi 1984 Brenner, J Clin Epi 1993