Advanced Methods for Addressing Selection Bias in Real-World Effectiveness and Cost-Effectiveness Studies

Pre-conference workshop, June 19, 2018

Acknowledgements

- Noemi Kreif (York)
- Jasjeet Sekhon (UC Berkeley)
- Rosalba Radice (Birkbeck)
- Zia Sadique (LSHTM)
- Roland Ramsahai

Funding

- Economic and Social Research Council (Grant no RES-061-25-0434)
- National Institute for Health Research (Senior Research Fellowship, Dr Richard Grieve, SRF-2013-06-016)
- Medical Research Council (Early Career Fellowship in the Economics of Health, Dr Noemi Kreif MR/L012332/1).

Learning Outcomes:

By the end of this session participants should be able to :

- Recognise opportunities for using RWE in effectiveness and cost-effectiveness research
- Consider the fundamental problem of confounding
- Understand the importance of study design
- Be aware of how RWE can be incorporated into decision models

Timetable

12:00 - 12:20 h	Introduction
12:20 - 14:00 h	Statistical methods to address confounding
14:00 - 14:30 h	Lunch
14:30 - 15:30 h	Practical- Genetic Matching
15:30 - 16:45 h	Extensions

Session content

1. RWE in HTA
2. RWE with and without RCTs
3. RWE and decision-making
4. Key Concern: selection bias due to confounding
5. The importance of design
6. The current state of play

Introduction

Comparative effectiveness, cost-effectiveness analysis (CEA)

US Panel (Sanders et al, 2016); NICE (2013); CADTH (2006), PBAC (2008)

- Aim: report relative effectiveness, costs and cost-effectiveness
- Emphasis of Methodological guidelines
  - Relevant comparators
  - Target populations, sub populations interest
  - Appropriate perspective, time horizon
  - Use of appropriate sources of evidence in decision model
- Traditionally published studies reliant on RCTs, relative effectiveness
- No longer the case….which raises new, important issues…

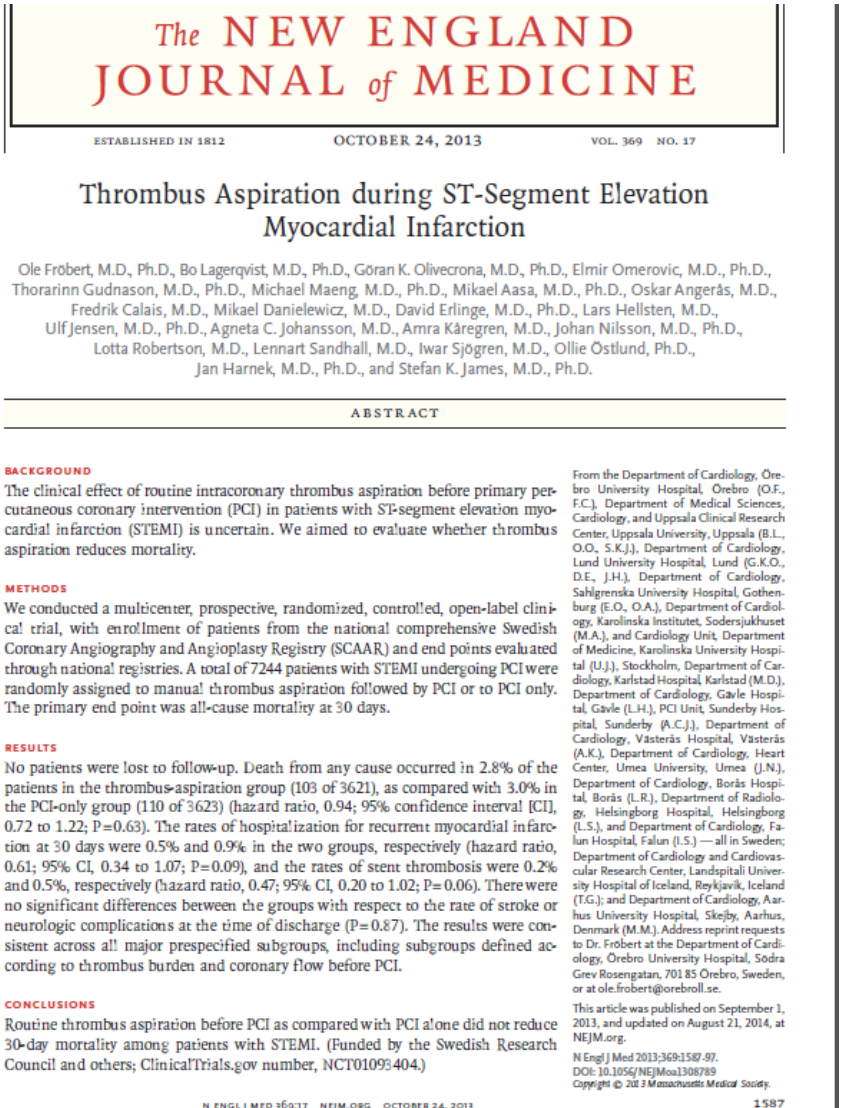
RWE in HTA
What is RWE?

- Disease registries
- Clinical databases
- Administrative databases
- Electronic medical records
- Cost databases
- Evidence that is not from phase III RCTs

HTA that use observational data
i) to complement RCTs

Parameter	Form of observational data	Example
Transition probabilities	Aggregate	Risk equations for cardiovascular disease (CVD) CEA of statins
Costs	Individual patient data (IPD)	Incremental costs for setting of interest (UK) CEA of xigris severe sepsis
Mortality	IPD	Vertebroplasty and kyphoplasty for treating osteoporotic vertebral fractures

RCTs nested within large registries



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 OCTOBER 24, 2013 VOL. 369 NO. 17

Thrombus Aspiration during ST-Segment Elevation Myocardial Infarction

Ole Fröbert, M.D., Ph.D., Bo Lagerqvist, M.D., Ph.D., Göran K. Olivecrona, M.D., Ph.D., Elmir Omerovic, M.D., Ph.D., Thorarinn Gudnason, M.D., Ph.D., Michael Maeng, M.D., Ph.D., Mikael Aasa, M.D., Ph.D., Oskar Angerås, M.D., Fredrik Calais, M.D., Mikael Danielewicz, M.D., David Erlinge, M.D., Ph.D., Lars Hellsten, M.D., Ulf Jensen, M.D., Ph.D., Agneta C. Johansson, M.D., Amra Kåregren, M.D., Johan Nilsson, M.D., Ph.D., Lotta Robertson, M.D., Lennart Sandhall, M.D., Iwar Sjögren, M.D., Ollie Östlund, Ph.D., Jan Harnek, M.D., Ph.D., and Stefan K. James, M.D., Ph.D.

ABSTRACT

BACKGROUND

The clinical effect of routine intracoronary thrombus aspiration before primary percutaneous coronary intervention (PCI) in patients with ST-segment elevation myocardial infarction (STEMI) is uncertain. We aimed to evaluate whether thrombus aspiration reduces mortality.

METHODS

We conducted a multicenter, prospective, randomized, controlled, open-label clinical trial, with enrollment of patients from the national comprehensive Swedish Coronary Angiography and Angioplasty Registry (SCAAR) and end points evaluated through national registries. A total of 7244 patients with STEMI undergoing PCI were randomly assigned to manual thrombus aspiration followed by PCI or to PCI only. The primary end point was all-cause mortality at 30 days.

RESULTS

No patients were lost to follow-up. Death from any cause occurred in 2.8% of the patients in the thrombus-aspiration group (103 of 3621), as compared with 3.0% in the PCI-only group (110 of 3623) (hazard ratio, 0.94; 95% confidence interval [CI], 0.72 to 1.22; P=0.63). The rates of hospitalization for recurrent myocardial infarction at 30 days were 0.5% and 0.9% in the two groups, respectively (hazard ratio, 0.61; 95% CI, 0.34 to 1.07; P=0.09), and the rates of stent thrombosis were 0.2% and 0.5%, respectively (hazard ratio, 0.47; 95% CI, 0.20 to 1.02; P=0.06). There were no significant differences between the groups with respect to the rate of stroke or neurologic complications at the time of discharge (P=0.87). The results were consistent across all major prespecified subgroups, including subgroups defined according to thrombus burden and coronary flow before PCI.

CONCLUSIONS

Routine thrombus aspiration before PCI as compared with PCI alone did not reduce 30-day mortality among patients with STEMI. (Funded by the Swedish Research Council and others; ClinicalTrials.gov number, NCT01093404.)

From the Department of Cardiology, Örebro University Hospital, Örebro (O.F., F.C.), Department of Medical Sciences, Cardiology, and Uppsala Clinical Research Center, Uppsala University, Uppsala (B.L., O.O., S.K.J.), Department of Cardiology, Lund University Hospital, Lund (G.K.O., D.E., J.H.), Department of Cardiology, Sahlgrenska University Hospital, Gothenburg (E.O., O.A.), Department of Cardiology, Karolinska Institutet, Södersjukhuset (M.A.), and Cardiology Unit, Department of Medicine, Karolinska University Hospital (U.J.), Stockholm, Department of Cardiology, Karlstad Hospital, Karlstad (M.D.), Department of Cardiology, Gävle Hospital, Gävle (L.H.), PCI Unit, Sunderby Hospital, Sunderby (A.C.J.), Department of Cardiology, Västerås Hospital, Västerås (A.K.), Department of Cardiology, Heart Center, Umea University, Umea (J.N.), Department of Cardiology, Borås Hospital, Borås (L.R.), Department of Radiology, Helsingborg Hospital, Helsingborg (L.S.), and Department of Cardiology, Falun Hospital, Falun (I.S.) — all in Sweden; Department of Cardiology and Cardiovascular Research Center, Landspitali University Hospital of Iceland, Reykjavik, Iceland (T.G.); and Department of Cardiology, Aarhus University Hospital, Skejby, Aarhus, Denmark (M.M.). Address reprint requests to Dr. Fröbert at the Department of Cardiology, Örebro University Hospital, Södra Grev Rosengatan, 701 85 Örebro, Sweden, or at ole.frobert@orebroll.se.

This article was published on September 1, 2013, and updated on August 21, 2014, at NEJM.org.

N Engl J Med 2013;369:1587-97.
DOI: 10.1056/NEJMoa1308789

Framework combining RCTs with observational data

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Royal Statistical Society Series A

J. R. Statist. Soc. A (2015)
178, *Part* 3, *pp.* 757–778

From sample average treatment effect to population average treatment effect on the treated: combining experimental with observational studies to estimate population treatment effects

Erin Hartman,
University of California at Berkeley, USA
Richard Grieve
London School of Hygiene and Tropical Medicine, UK
Roland Ramsahai
University of Cambridge, UK
and Jasjeet S. Sekhon
University of California at Berkeley, USA

[Received December 2013. Final revision August 2014]

Summary. Randomized controlled trials (RCTs) can provide unbiased estimates of sample average treatment effects. However, a common concern is that RCTs may fail to provide unbiased estimates of population average treatment effects. We derive the assumptions that are required to identify population average treatment effects from RCTs. We provide placebo tests, which formally follow from the identifying assumptions and can assess whether they hold. We offer new research designs for estimating population effects that use non-randomized studies to adjust the RCT data. This approach is considered in a cost-effectiveness analysis of a clinical intervention: pulmonary artery catheterization.

Keywords: Causal inference; Cost-effectiveness studies; External validity; Observational studies; Placebo tests; Randomized controlled trials

1. Introduction

Randomized controlled trials (RCTs) can provide unbiased estimates of the relative effectivene of alternative interventions within the study sample. Much attention has been given to improvi the design and analysis of RCTs to maximize internal validity. However, policy makers requi evidence on the relative effectiveness and cost-effectiveness of interventions for target popul tions that usually differ from those represented by RCT participants (Hoch *et al.*, 2002; Mit and Indurkhya, 2005; Mojtabai and Zivin, 2003; Nixon and Thompson, 2005; Willan *et al.*, 200 Willan and Briggs, 2006). A key concern is that estimates from RCTs and meta-analyses m lack external validity (Allcott and Mullainathan, 2012; Deaton, 2009; Heckman and Urzu

An Approach to Assess Generalizability in Comparative Effectiveness Research: A Case Study of the Whole Systems Demonstrator Cluster Randomized Trial Comparing Telehealth with Usual Care for Patients with Chronic Health Conditions

Adam Steventon, MA, Richard Grieve, PhD, Martin Bardsley, PhD

Background. *Policy makers require estimates of comparative effectiveness that apply to the population of interest, but there has been little research on quantitative approaches to assess and extend the generalizability of randomized controlled trial (RCT)-based evaluations. We illustrate an approach using observational data.* ***Methods.*** *Our example is the Whole Systems Demonstrator (WSD) trial, in which 3230 adults with chronic conditions were assigned to receive telehealth or usual care. First, we used novel placebo tests to assess whether outcomes were similar between the RCT control group and a matched subset of nonparticipants who received usual care. We matched on 65 baseline variables obtained from the electronic medical record. Second, we conducted sensitivity analysis to consider whether the estimates of treatment effectiveness were robust to alternative assumptions about whether "usual care" is defined by the RCT control group or nonparticipants. Thus, we provided alternative estimates of comparative effectiveness by contrasting the outcomes of the RCT telehealth group and matched nonparticipants.* ***Results.*** *For some endpoints, such as the number of outpatient attendances, the placebo tests passed, and the effectiveness estimates were robust to the choice of comparison group. However, for other endpoints, such as emergency admissions, the placebo tests failed and the estimates of treatment effect differed markedly according to whether telehealth patients were compared with RCT controls or matched nonparticipants.* ***Conclusions.*** *The proposed placebo tests indicate those cases when estimates from RCTs do not generalize to routine clinical practice and motivate complementary estimates of comparative effectiveness that use observational data. Future RCTs are recommended to incorporate these placebo tests and the accompanying sensitivity analyses to enhance their relevance to policy making.* ***Key words:*** *causal inference; external validity; generalizability; randomized trials; telehealth; chronic health conditions.* ***(Med Decis Making 2015;35:1023–1036)***

Well-conducted randomized controlled trials (RCTs) can ensure high levels of internal validity because the treatment groups are balanced. However, a major concern with RCT evidence is that criteria and compares the intervention with usual care,[2] the trial may exclude important subgroups of patients and centers.[3] Thus, both observed and unobserved characteristics that modify treatment ef-

RWE to estimate population effects from RCTs

Trial nested within large clinical database

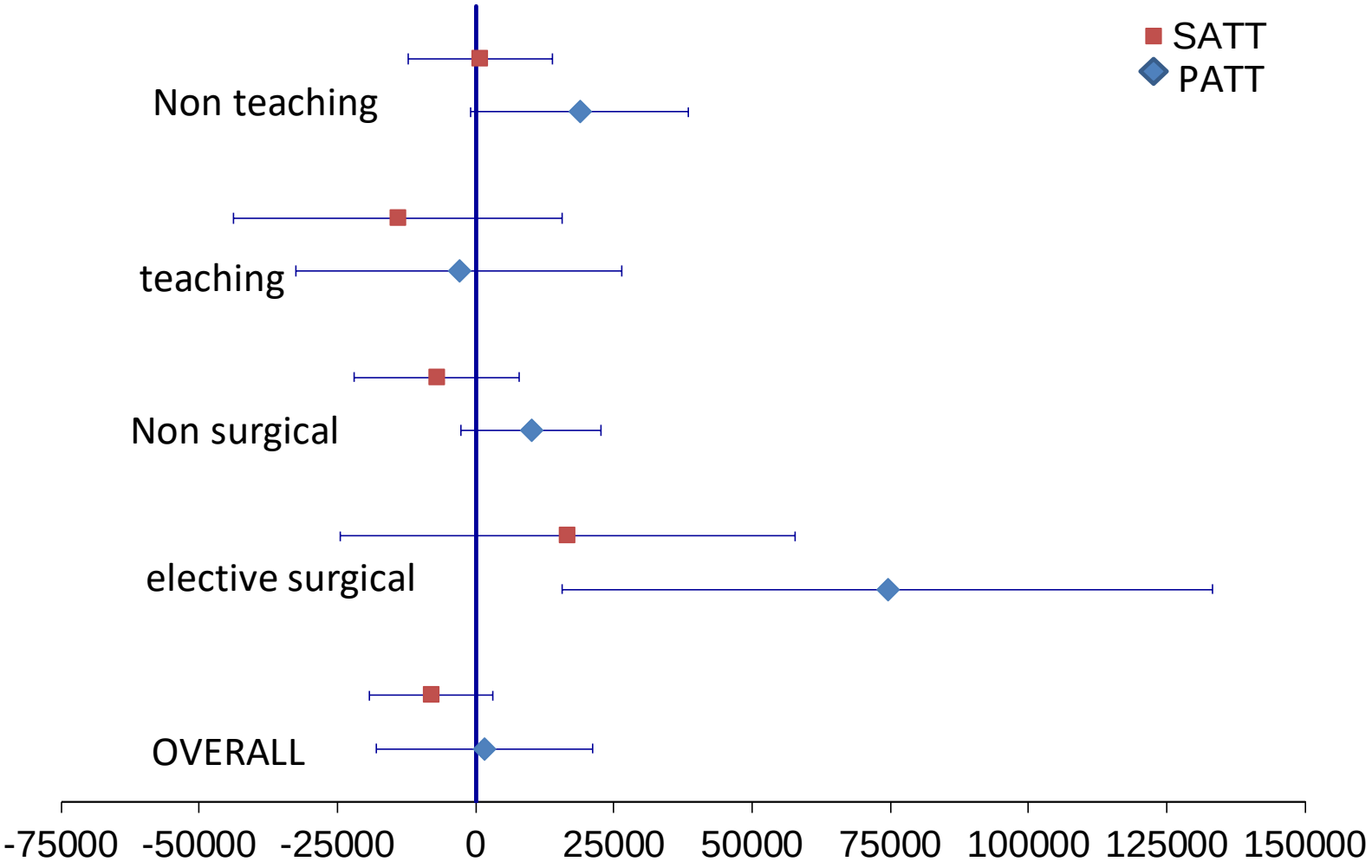
1. Effectiveness within RCT, Sample average treatment effects (SATT)
2. Develop model for trial inclusion to reweight the SATT
3. Reweight to report Population average Treatment effects (PATT)

PATT versus SATT

Incremental net benefits

Pulmonary Artery Catheterization (PAC) vs No PAC

£20,000 per QALY



Observational data when RCTs unavailable..

- To estimate treatment effects
- Effect on mean costs, mean QALYs, time to event
- Traditionally for non-drug interventions
- Health service, health financing, public health interventions
- New medical devices, forms of surgery
- Orphan drugs
- But the floodgates are opening..

Opportunities in observational data initiatives

VIEWPOINT

Artificial Intelligence, Big Data, and Cancer

Hagop Kantarjian, MD
Department of Leukemia, The University of Texas MD Anderson Cancer Center, Houston.

Peter Paul Yu, MD
Department of Hematology-Oncology, Palo Alto Medical Foundation, Palo Alto, California.

Even though the original computers were designed in 1936, computers have become part of the social and professional fabrics of our lives only since the mid-1980s, enhancing workplace and individual productivities. Computers are still evolving, and so are the ways we use them. Artificial intelligence (computers) and big data (incorporating large volumes of information into computers) are gaining wide acceptance in different fields. Consider "Deep Blue," an IBM-trained chess computer that beat Gary Kasparov, the chess world champion.[1] Another IBM-trained Watson System won the *Jeopardy!* game against champions.[2] After 12 years of refinements, poker computers now consistently win against poker world champions.[3]

Having proven the concept in chess, *Jeopardy!*, and poker, this technology is being brought to the real world. This new form of artificial intelligence, or "cognitive computing," learns in ways similar to humans' learning. With proper "training," they address human-like situations and deal with ever-changing data. Cognitive computers are designed to answer questions posed in conversational language with a range of possible accurate answers based on available information. This is exemplified by novel apps that advise, based on previous experiences, choices, and preferences (eg, restaurants, books, clothing) what might be a "best" range of new choices.

This is also the case in medicine and cancer, where there are often no black-and-white answers. The best answers are based on evolving and often ambiguous or even conflicting literature, colored by individual experiences or intuition, what is referred to as "The Art of Medicine." Healthcare and medical research have become a new frontier for artificial intelligence and big data. Endeavors in cancer are incorporating patient information, research publications, and ongoing research into large databases.[4] Cognitive computers can access and analyze these databases and arrive at a "best" range of answers for questions related to research or therapy.

Cancer-oriented cognitive computer systems are developed the same way new cancer specialists are trained. They are designed to "read, remember, recommend, and remind." They can read and remember the evolving body of medical literature. Then, they are trained by expert cancer specialists to weigh a patient's case against existing knowledge and suggest appropriate treatment options, including clinical trials, tailored to the individual patient. They offer evidence supporting the suggestions, allowing physicians to judge its relevance. In other words, cognitive computers do not make decisions; they offer physicians the tools to help tailor the best treatments to an individual patient.

For example, a new system called CanSAR[5] condenses vast quantities of data to help generate new discoveries. Its use may have identified 46 cancer proteins that could be targeted with new or older drugs. The American Society of Clinical Oncology has developed CancerLinQ, which proposes to incorporate data of patients with cancer in the United States into 1 large database. This would capture cancer data on 100% of patients with cancer rather than the 3% who are entered on clinical trials, and thereby accelerate new information, knowledge, and discoveries.[6] Cancer centers like MD Anderson and Memorial Sloan Kettering are launching single-institutional large database endeavors aimed at enhancing cancer care and research. The MD Anderson initiative, referred to as the Oncology Expert Advisor (OEA) will incorporate all information related to the more than 1 million patients with cancer treated over the lifetime of the institution to generate a novel support system for research and patient care.

Novel cognitive computers may include all the clinical and laboratory information available in different cancer populations, worldwide pertinent cancer publications, and novel cutting-edge research discoveries (eg, whole-genome profiling, RNA and proteomics data). Future potential applications of cognitive computers in cancer are limitless. Some are listed in the **Box**.

Box. Future Potential Applications of Artificial Intelligence in Cancer Care and Research

- Develop national, international, and worldwide cancer networks and registries
- Follow national and worldwide epidemiologic trends of cancers (detect causations)
- Detect new associations with particular cancers
- Identify beneficial therapies for rare cancers
- Observe possible differential outcomes or therapeutic benefits of particular cancers by different parameters (geography, regimen, pathways)
- Develop national and international cancer treatment pathways
- Incorporate old and novel patient- and cancer-specific attributes and analyze their associations with cancer etiologies, therapeutic results, and prognosis
- Discover new cancer etiologies (eg, possible cancer peaks in particular geographic areas or with particular habits or diets)
- Incorporate pertinent patient and cancer characteristics into clinic-based uses
- Conduct cost-efficient broad-based cancer trials
- Uncover genomic or molecular events that render subsets of patients more or less sensitive to existing or new treatments
- Analyze therapeutic trends and results in clinical practice on all cancer patients (in contrast to only 3% on clinical trials)
- Elicit new biologic, therapeutic, pathophysiologic, or epidemiologic associations

Corresponding Author: Hagop Kantarjian, MD, Department of Leukemia, MD Anderson Cancer Center, 1400 Holcombe, Unit 428, Houston, TX 77030 (hkantarjian@mdanderson.org).

Drug regulation and reimbursement initiatives

- EMA, Adaptive Pathways. Pilot project on adaptive licensing.

http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/03/WC500163409.pdf

- International committee on Harmonisation (ICH, E9 appendum (august 2017)

http://www.ema.europa.eu/ema/doc_index.jsp?curl=pages/includes/document/document_detail.jsp?webContentId=WC500233916&murl=menus/document_library/document_library.jsp&mid=0b01ac058009a3dc

- Accelerated Access Review

https://www.gov.uk/government/publications/accelerated-access-pathways-for-medical-technologies

- Cancer Drugs Fund (Grieve et al, 2016)

http://www.bmj.com/content/354/bmj.i5090

Real-world evidence and the FDA

While, RWE can

"inform therapeutic development, outcomes research, patient care, research on health care systems, quality improvement, safety surveillance, and well-controlled effectiveness studies…

…

"the confluence of large data sets of uncertain quality and provenance, the facile analytics tools that can be used by nonexperts, and a shortage of researchers with adequate methodologic savvy... could result in poorly conceived study and designs that generate incorrect or unreliable conclusions."

Rob Califf, ex-FDA commissioner, May, 2017

Challenges in CEA use observational data...

Evaluation	Analytical method	Reference
PTCA vs. CABG for Angina	Regression	Griffin et al (2007)
Surgery bladder cancer	Propensity score (PS) methods	Mitra and Indurkhya (2005)
Xigris for severe sepsis	PS matching, Genetic Matching	Sadique et al al (2011)
Alternative types of hip prosthesis	regression, Genetic Matching	Pennington et al (2013)
Alternative surgery for breast cancer	IV methods	Polsky and Basu (2006)
Treatments for psoriasis	Matching-adjusted indirect compar	Signorotvitch et al (2010)
Bosutinib Chronic Myeloid Leukaemia	Naïve comparison	NICE 2015, TA 413

Reformed NICE cancer drugs fund (CDF), decision 1

Bosutinib for chronic myeloid leukaemia

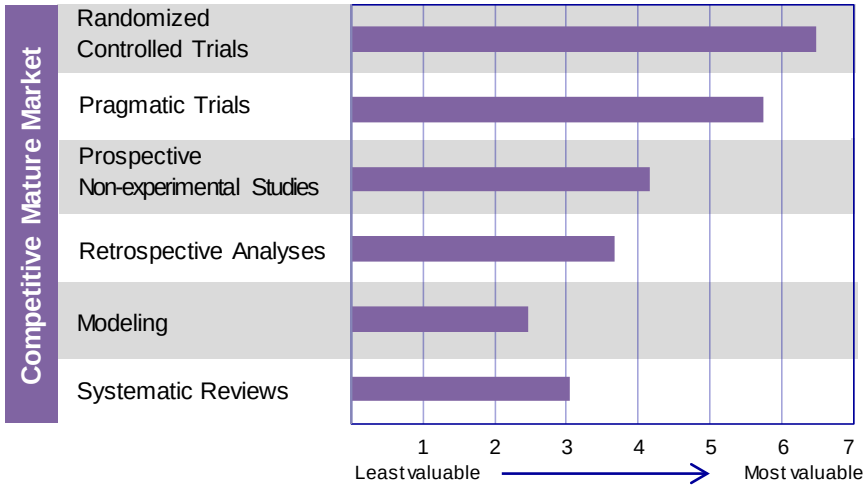
- Previously rejected by NICE and only available through the Cancer Drugs Fund (CDF) (ICERs: 50K to 150K)
- NICE reappraising drugs currently in the CDF.
- Bosutinib appraisal no RCT evidence
- new drug was compared to a small sample of patients who received an alternative drug some years previously.
- The evidence submitted to NICE on whether Bosutinib is effective, did not allow for differences between patients who took the new versus the old drug.
- Bosutinib approved for patients failed previous treatments

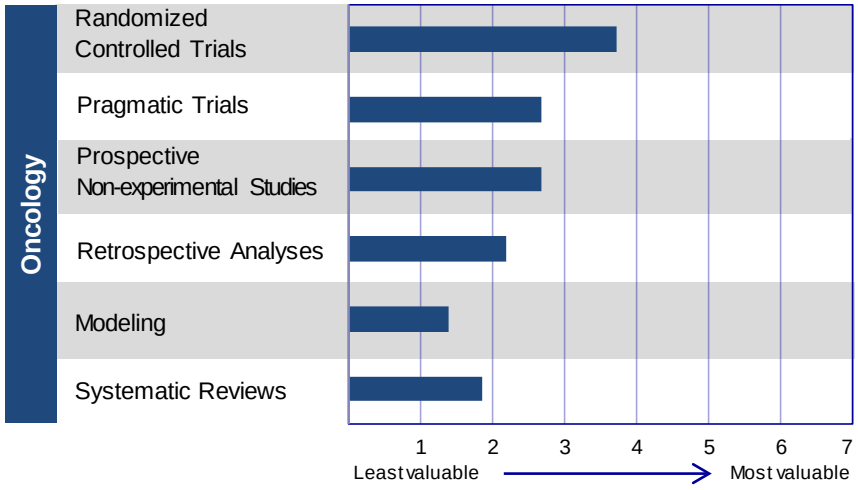
Challenges in use of RWE for decision-making

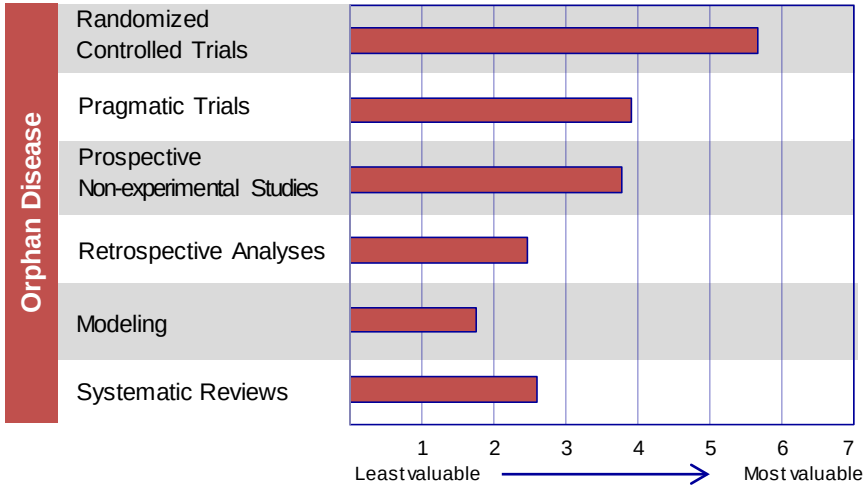
- Beta Interferon for Multiple sclerosis (McCabe et al, 2010)
- Initial RCT reported small levels of clinical effectiveness
- NHS funded drug, conditional on real world evidence of relative effectiveness
- Pharmaceutical company return drug costs if not effective and cost-effective
- Independent, NRS reported smaller effect than original RCT
- Company disputed the findings as based on NRS..
- Key challenge: address the confounding…

Survey by US payers on RWE for comparative effectiveness

Value Assessment for Effectiveness (average ratings)







"Prospective non-experimental studies are the closest to what happens in the practice setting. The biggest challenge are pharma sponsored registries due to biases, so their utility is diminished. If registries are not biased then they would be helpful.." – *National pharmacy director on effectiveness in competitive mature markets*

"Retrospective analysis is sometimes the best we can do. Data of this nature is sometimes published in really good journals, so there's a lot of variation." – *National medical director on effectiveness in competitive mature markets*

"We will manage oncology products to label even if the drug is much more expensive. It's a political time bomb to manage pricing and utilization." – *Regional medical director on effectiveness for oncology*

Key Concern: Confounding

The New York Times

Hormone Replacement Study a Shock to the Medical System

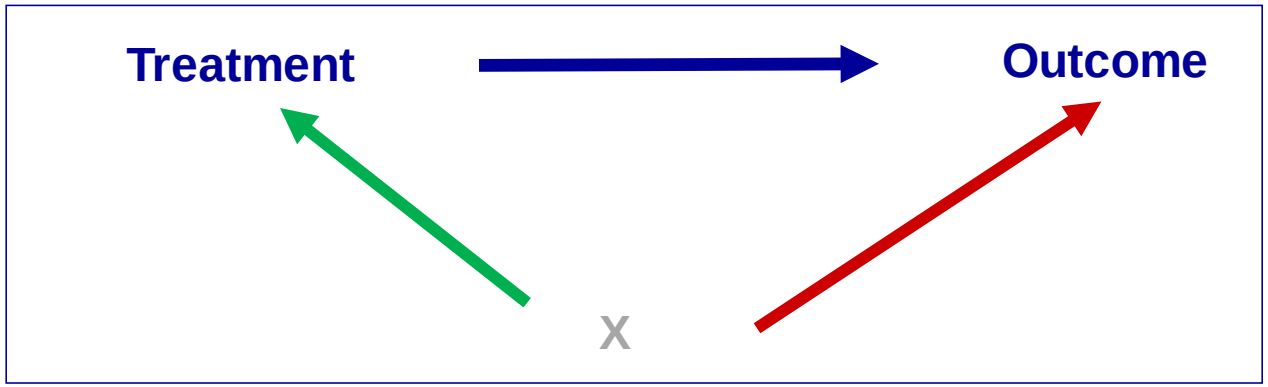
By GINA KOLATA with MELODY PETERSEN

Published: Wednesday, July 10, 2002

"The announcement yesterday that a hormone replacement regimen taken by six million American women did more harm than good was met with puzzlement and disbelief by women and their doctors across the country."

"A rigorous study found that the drugs, a combination of oestrogen and progestin, caused small increases in breast cancer, heart attacks, strokes and blood clots. Those risks outweighed the drugs' benefits: a small decrease in hip fractures and a decrease in colorectal cancer. Many of the 16,000 women in the study, supported by the National Institutes of Health, opened letters yesterday telling them to stop the drugs.."

What is confounding?



1. X associated with treatment assignment
   - Not a consequence of treatment

2. X associated with outcome
   - Independently of treatment (not an intermediary)

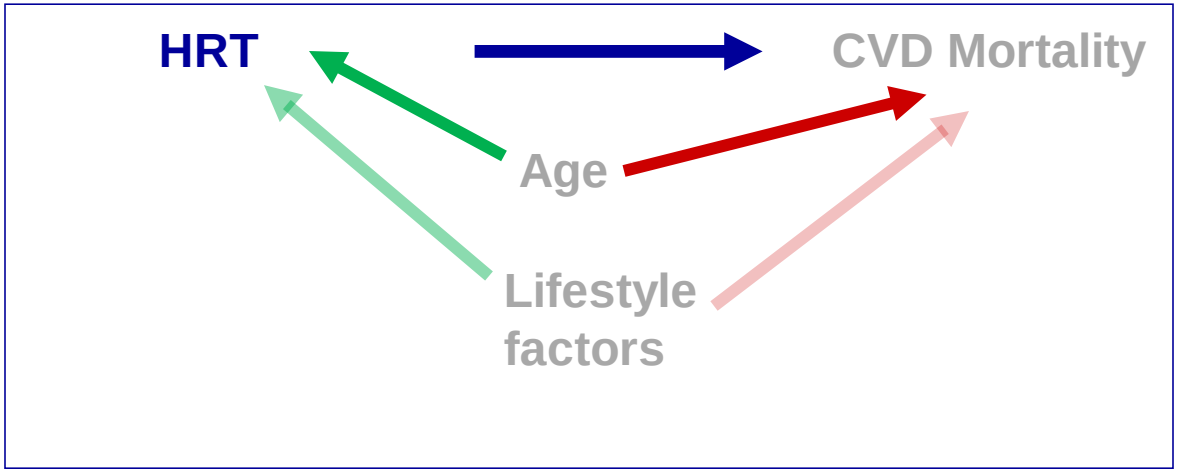
i.e the variable is a common cause

Adjusting for confounding

Want to avoid adjusting for variables that are on the causal pathway

Selection bias from measured and unmeasured confounders

Example of an observed and an unobserved confounding variable



If ignored can lead to selection bias

- Bias from imbalance on unobservables: hidden bias
- Bias from imbalance on observables: overt bias

Statistical Methods for addressing confounding

- **Assume no unobserved confounding**
  - Regression adjustment
  - Matching methods
    - Propensity score matching
    - Genetic Matching

- **Allow for observed and unobserved confounding:**
  - Instrumental variable estimation - Assumes we have a valid instrument!
  - Regression discontinuity design
  - Alternative: sensitivity analysis for unobserved confounding

Instrumental variables (IV)

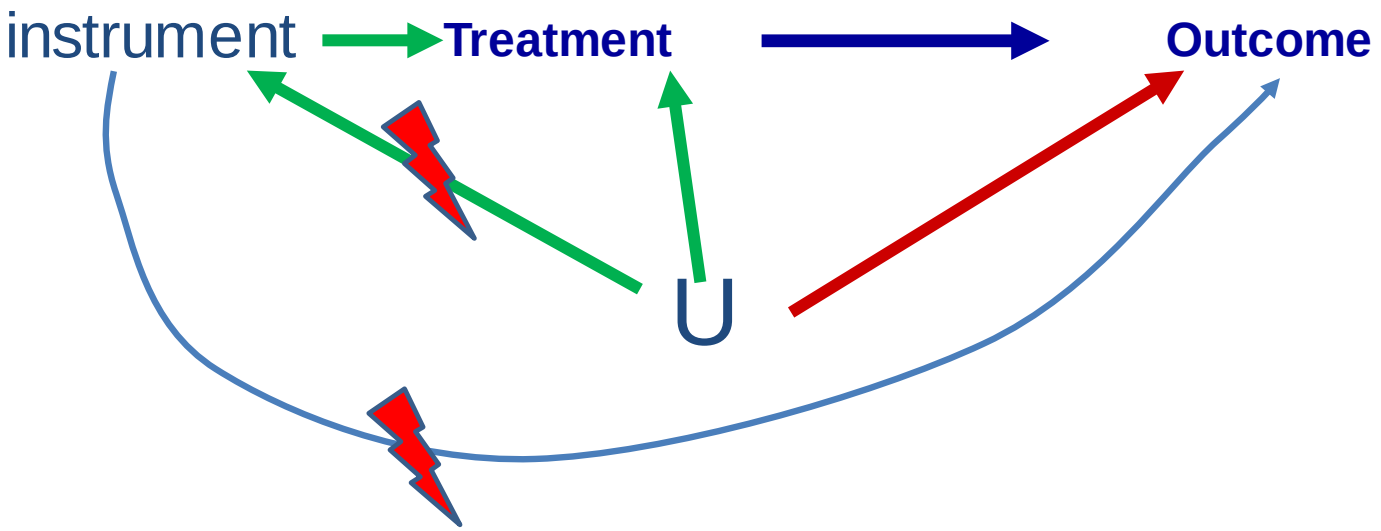
Criteria for IV, often termed Z

1. Correlated with treatment receipt [testable]
2. Independent of unobserved confounder (U) [untestable]
3. Independent of outcome, conditional on U [untestable]

Can effectively randomise between treatment arms

- Both X's and U's equally distributed across treatment arms
- Estimate treatment effects without overt and hidden bias

Criteria for IV



1. Correlated with treatment receipt
2. Independent of U
3. Independent of outcome, conditional on U

Example of IV
Polsky and Basu, 2006

- Breast conserving surgery (BCS) versus mastectomy (M) for breast cancer
- BCS group anticipated to be healthier
- IV- distance to hospital
- Report ATE, mean (SD) incremental costs
  - Unadjusted: $8,593 ($1,522)
  - Regression adjusted: $10,944 ($1,540)
  - IV: $15,417($5,110)

IV: Big challenge

- "The use of instrumental variables replaces the unverifiable assumption of no unmeasured confounding with other unverifiable assumptions. ..Hence the reliance on assumptions that cannot be empirically verified is not solved but shifted to another realm"

- Hernan and Robins, Epidemiology 2006, 17:360-372.

- IV assumptions tenable in some settings e.g. for handling non-compliance in RCTs (see guest lecture)

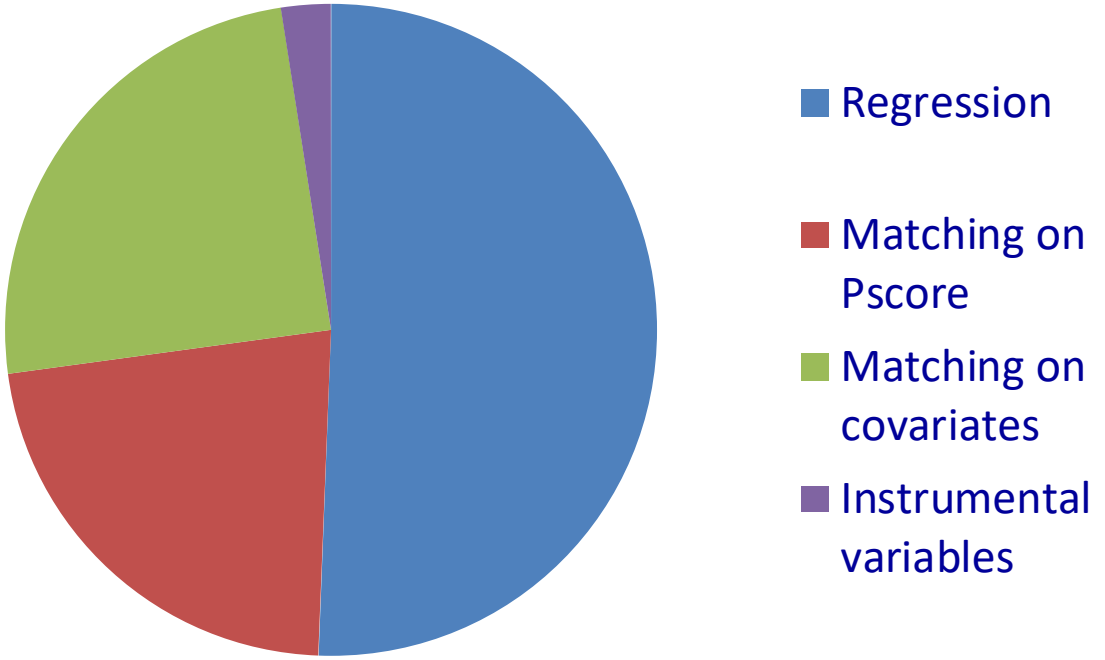
How good are we at addressing confounding in RWE studies?

Systematic review economic evaluation (2000-2011) by Kreif et al (2012)

79 studies

Almost all studies assumed “no unobserved confounding”

Failed to justify this key assumption



How good are we at addressing confounding in observational studies?

- Faria et al. (2015) NICE
  - Technology assessments use matching or regression when using IPD
  - Analysis and reporting not satisfactory
  - Further training in relevant statistical methods essential

Careful design is essential (Rubin 2008)

- Careful definition of intervention and control (counterfactual)
- Measure as many of relevant confounders as possible
- Standard approach assume all relevant confounders observed
- Justify that assumption
- test in sensitivity analyses

Pennington et al (2013 BMJ)

BMJ

BMJ 2013;346:f1026 doi: 10.1136/bmj.f1026 (Published 27 February 2013)

RESEARCH

Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis

OPEN ACCESS

Mark Pennington *lecturer in health economics*[1], Richard Grieve *reader in health economics*[1], Jasjeet S Sekhon *professor*[2], Paul Gregg *professor and consultant orthopaedic surgeon*[3] *vice chairman*[4], Nick Black *professor of health services research*[1], Jan H van der Meulen *professor of clinical epidemiology*[1]

[1]Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London WC1H 9SH, UK; [2]Travers Department of Political Science, Department of Statistics, Center for Causal Inference and Program Evaluation, Institute of Governmental Studies, University of California, Berkeley, CA, USA; [3]James Cook Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK; [4]National Joint Registry for England and Wales, Healthcare Quality Improvement Partnership, London, UK

Abstract

Objective To compare the cost effectiveness of the three most commonly chosen types of prosthesis for total hip replacement.

Design Lifetime cost effectiveness model with parameters estimated from individual patient data obtained from three large national databases.

Setting English National Health Service.

Participants Adults aged 55 to 84 undergoing primary total hip replacement for osteoarthritis.

Interventions Total hip replacement using either cemented, cementless, or hybrid prostheses.

Main outcome measures Cost (£), quality of life (EQ-5D-3L, where 0 represents death and 1 perfect health), quality adjusted life years (QALYs), incremental cost effectiveness ratios, and the probability that each prosthesis type is the most cost effective at alternative thresholds of willingness to pay for a QALY gain.

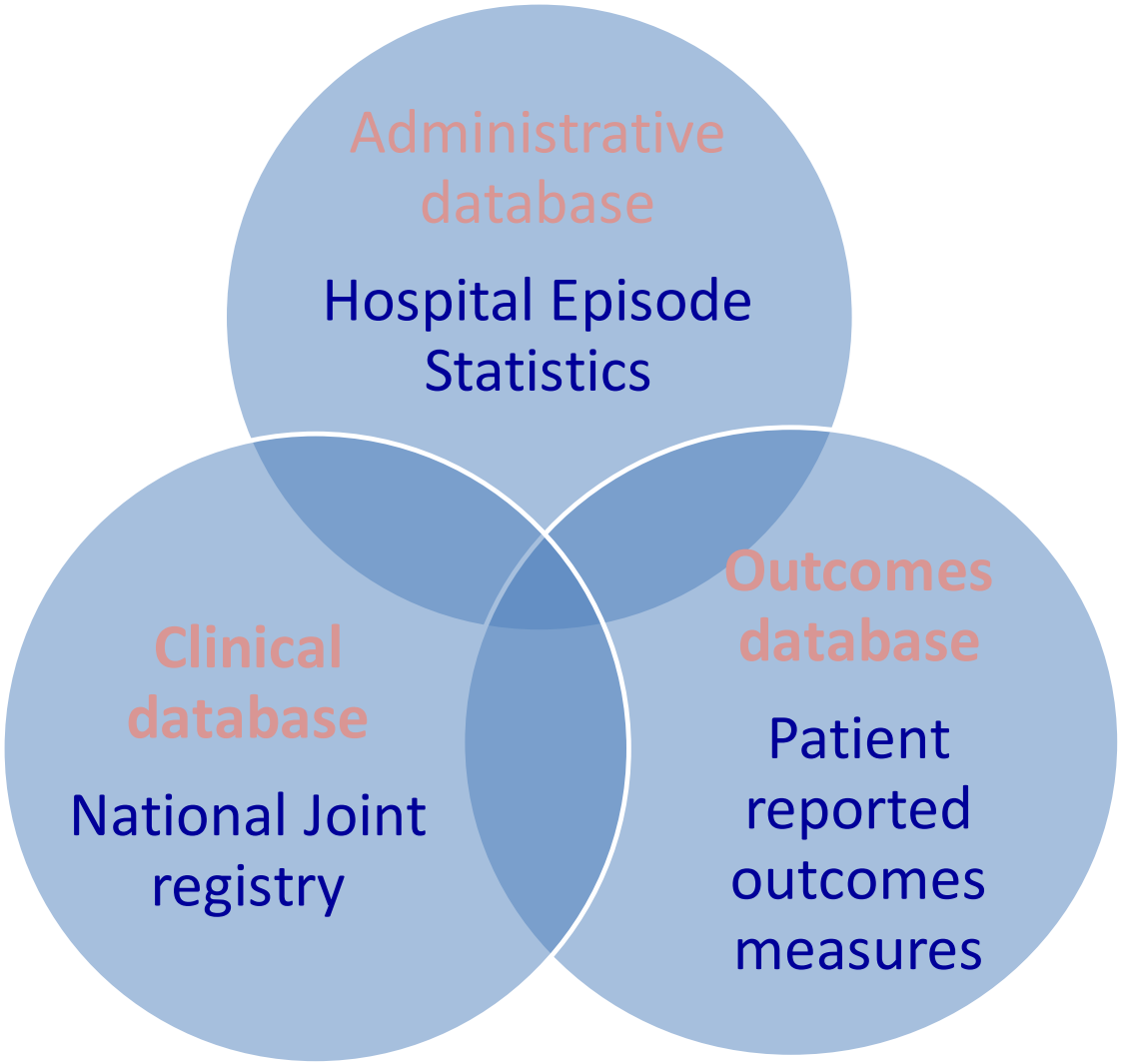
Conclusions Cemented prostheses were the least costly type for total hip replacement, but for most patient groups hybrid prostheses were the most cost effective. Cementless prostheses did not provide sufficient improvement in health outcomes to justify their additional costs.

Introduction

Total hip replacement is one of the most common surgical procedures. In 2010 the global market for hip prostheses was estimated at $4.7b (£3.0b; €3.5b).[1] A large number of different prosthesis designs have been developed and introduced on the market. For example, in England and Wales in 2010 at least 123 different brands of acetabular cups and 146 brands of femoral stems were used.[2] These prosthesis brands are often grouped into cemented, cementless, and hybrid prostheses. Hybrid prostheses consist of cemented stems and cementless cups.

Large linked observational data for decision-making: example of hip prosthesis

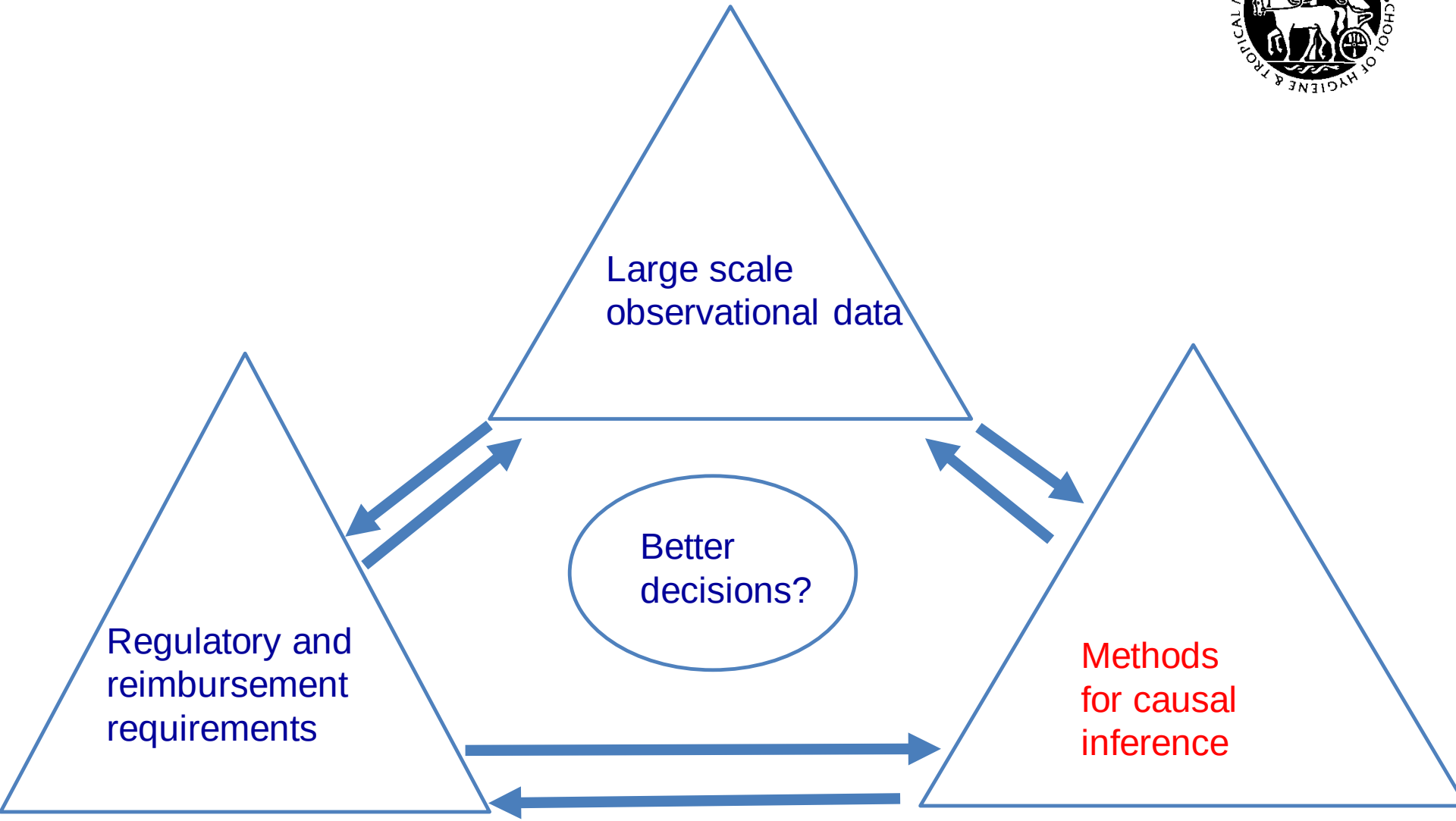
Pennington et al (2013)



Some key design principles
(e.g. Pennington et al, 2013)

Goal	Example
• Careful definition of treatments	• Alternative hip prostheses
• Pre-specification population	• Osteoarthritis, in linked data
• **Pre-specification estimand**	• Average treatment effect, by subgroup
• Covariate measurement	• Baseline QoL etc
• **Good overlap**	• Yes
• Exploit large, linked datasets	• Yes
• Relevant outcomes	• Partial, QoL at 6 months
• Sensitivity analysis	• Partial
• Careful analysis and interpretation	• Partial

Future directions



Summary

- Observational data here to stay, and its improving
- No substitute for RCT
- Addressing confounding is key
- Careful design is necessary first step

Key References

- Faria, R at al (2015): NICE DSU technical support document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data report by the decision support unit http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/TSD17-DSU-Observational-data-FINAL.pdf

- Grieve R, et al. (2016) Cancer Drugs Fund requires further reform: Reliance on "real world" observational data undermines evidence base for clinical practice. *BMJ* 2016;354:i5090

- Pennington et al (2013). Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. *BMJ* Feb 27;346:f1026.

- Rubin DB (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*. 2008 Sep 1:808-40.

- Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence — what is it and what can it tell us? N Engl J Med 2016;375:2293-2297