The Perspective from Clinical Translational Researchers

February 2012

Brad H. Pollock, MPH, PhD
Department of Epidemiology and Biostatistics
University of Texas Health Science Center at San Antonio
San Antonio, TX

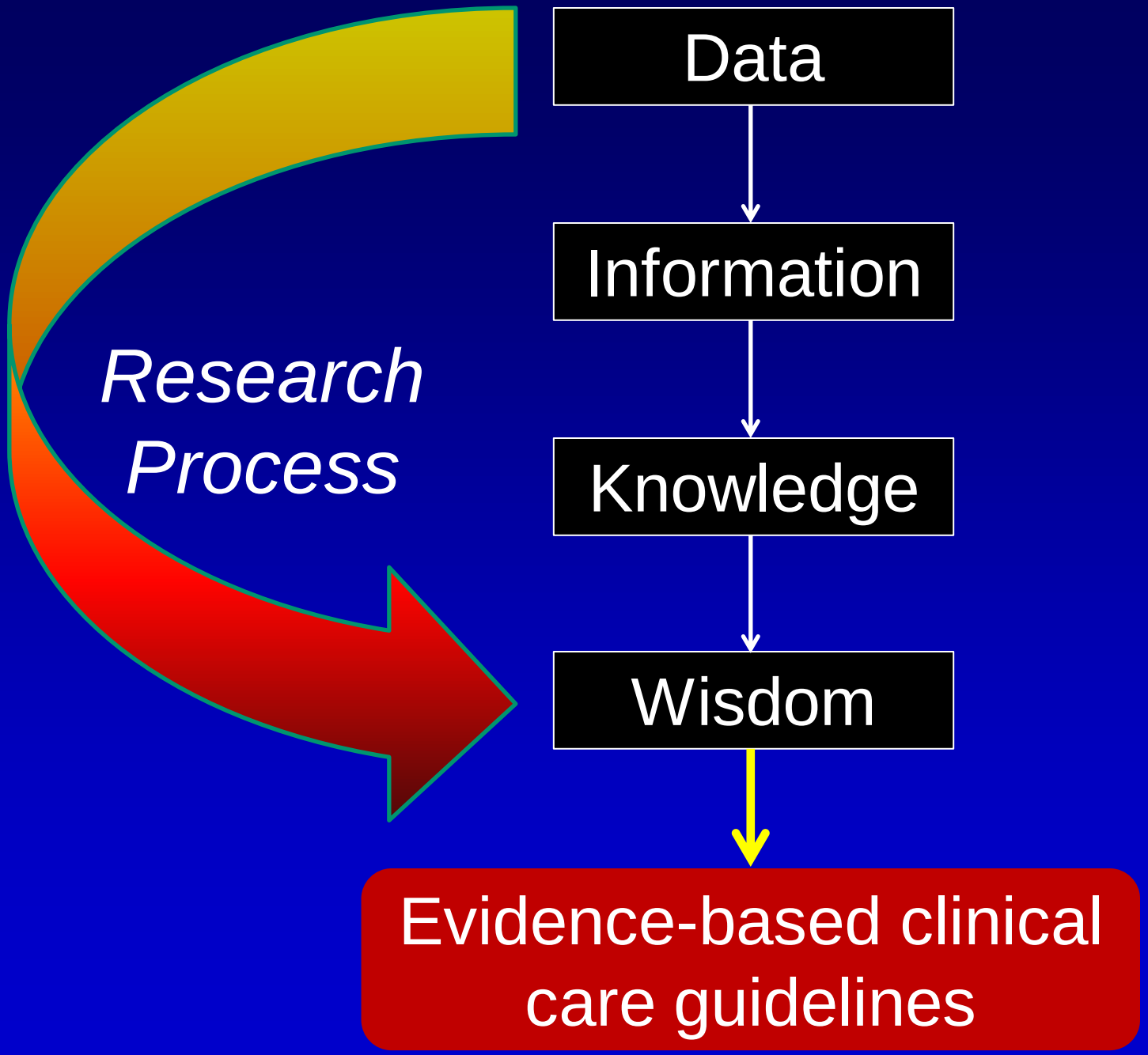
SCHOOL OF MEDICINE
UT HEALTH SCIENCE CENTER™
SAN ANTONIO

Disclosures

- No financial conflicts of interest
- *Director*, Biostatistics and Informatics Shared Resource, Cancer Therapy & Research Center, University of Texas Health Science Center at San Antonio (P30 CA054174)
- *Principal Investigator*, Children's Oncology Group (COG) Community Clinical Oncology Program (CCOP) Research Base (U10 CA095861)
- *Vice Chair and Chair-Elect*, Biostatistics, Epidemiology, Research Design (BERD) Key Function Committee, Clinical Translational Science Award (CTSA) Consortium (UL1 RR025767)

Informatics Needs and Challenges in Cancer Research Planning Committee Welcome Note

"…At a time of huge scientific opportunity, cancer researchers are finding it challenging to collect, aggregate, integrate, analyze, interpret, and exchange biomedical data, and the problem is growing steadily worse with the exponential increase in complex, multidimensional data that they confront."


Data
Information
Knowledge
Wisdom
Research Process
Evidence-based clinical care guidelines

Research Process

- Hypothesis formation
- Study design
- Data
- Statistical modeling and analysis
- Inferences and translation

Research Process

- Hypothesis formation
- Study design
- **Data**
- Statistical modeling and analysis
- Inferences and translation

Research Process

- Hypothesis formation
- Study design
- Data
- Statistical modeling and analysis
- Inferences and translation

Anything we can do to promote or improve these components at any level is welcome!

Hypothesis Formation

A paradigm shift is occurring with the ever-increasing generation and availability of data:

From: **Hypotheses** in search of **Data**

To: **Data** in search of **Hypotheses**

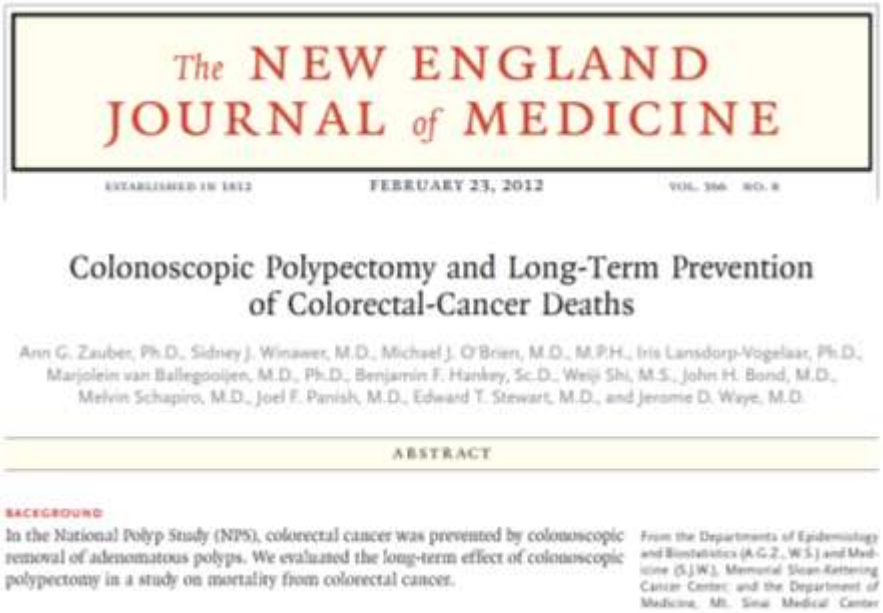
Hypothesis-generating work is important, but the most novel clinical oncology discoveries have been made using a traditional hypothesis-driven research framework.

Study Design

- From my perspective in the trenches, study design deserves a lot of attention
- Good designs → Efficient use of data
- Selection of a study design has profound implications…

NEJM, 2012, vol. 366 (No. 8)
(February 23, 2012)

Prospective Cohort Study



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 FEBRUARY 23, 2012 VOL. 366 NO. 8

Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths

Ann G. Zauber, Ph.D., Sidney J. Winawer, M.D., Michael J. O'Brien, M.D., M.P.H., Iris Lansdorp-Vogelaar, Ph.D., Marjolein van Ballegooijen, M.D., Ph.D., Benjamin F. Hankey, Sc.D., Weiji Shi, M.S., John H. Bond, M.D., Melvin Schapiro, M.D., Joel F. Panish, M.D., Edward T. Stewart, M.D., and Jerome D. Waye, M.D.

ABSTRACT

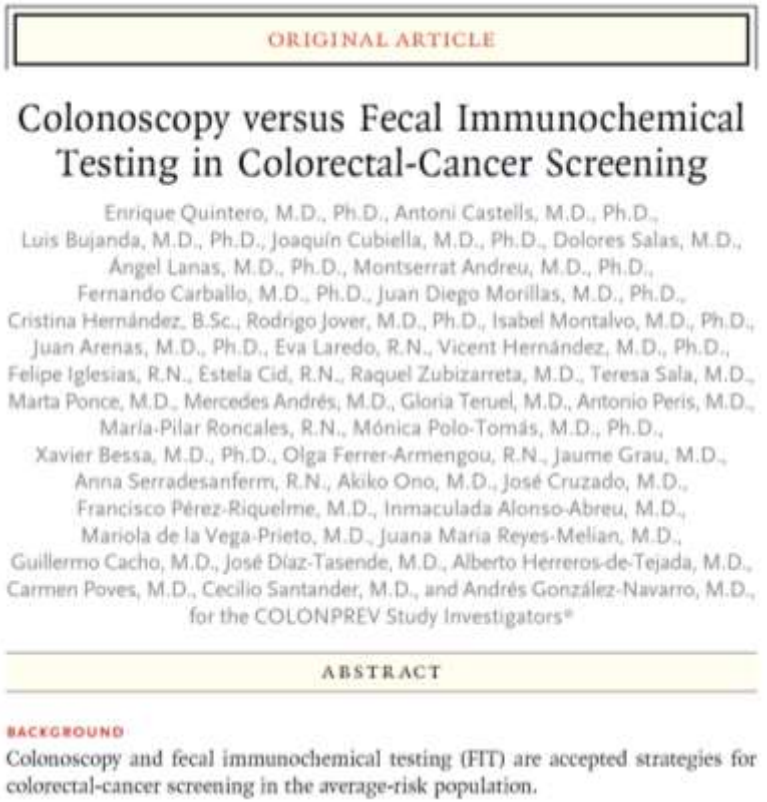
BACKGROUND

In the National Polyp Study (NPS), colorectal cancer was prevented by colonoscopic removal of adenomatous polyps. We evaluated the long-term effect of colonoscopic polypectomy in a study on mortality from colorectal cancer.

From the Departments of Epidemiology and Biostatistics (A.G.Z., W.S.) and Medicine (S.J.W.), Memorial Sloan-Kettering Cancer Center, and the Department of Medicine, Mt. Sinai Medical Center

- The National Polyp Study cohort was compared to the SEER populaiton that had higher mortality from all causes
- Not a screening trial

RCT



ORIGINAL ARTICLE

Colonoscopy versus Fecal Immunochemical Testing in Colorectal-Cancer Screening

Enrique Quintero, M.D., Ph.D., Antoni Castells, M.D., Ph.D., Luis Bujanda, M.D., Ph.D., Joaquín Cubiella, M.D., Ph.D., Dolores Salas, M.D., Ángel Lanas, M.D., Ph.D., Montserrat Andreu, M.D., Ph.D., Fernando Carballo, M.D., Ph.D., Juan Diego Morillas, M.D., Ph.D., Cristina Hernández, B.Sc., Rodrigo Jover, M.D., Ph.D., Isabel Montalvo, M.D., Ph.D., Juan Arenas, M.D., Ph.D., Eva Laredo, R.N., Vicent Hernández, M.D., Ph.D., Felipe Iglesias, R.N., Estela Cid, R.N., Raquel Zubizarreta, M.D., Teresa Sala, M.D., Marta Ponce, M.D., Mercedes Andrés, M.D., Gloria Teruel, M.D., Antonio Peris, M.D., María-Pilar Roncales, R.N., Mónica Polo-Tomás, M.D., Ph.D., Xavier Bessa, M.D., Ph.D., Olga Ferrer-Armengou, R.N., Jaume Grau, M.D., Anna Serradesanferm, R.N., Akiko Ono, M.D., José Cruzado, M.D., Francisco Pérez-Riquelme, M.D., Inmaculada Alonso-Abreu, M.D., Mariola de la Vega-Prieto, M.D., Juana María Reyes-Melian, M.D., Guillermo Cacho, M.D., José Díaz-Tasende, M.D., Alberto Herreros-de-Tejada, M.D., Carmen Poves, M.D., Cecilio Santander, M.D., and Andrés González-Navarro, M.D., for the COLONPREV Study Investigators*

ABSTRACT

BACKGROUND

Colonoscopy and fecal immunochemical testing (FIT) are accepted strategies for colorectal-cancer screening in the average-risk population.

- Randomized, controlled population-based screening trial

Strength of Evidence

- At the study level, the gold standard is the randomized clinical trial (RCT); however, an RCT may not be feasible in all situations.
- Beware, there may be more methodologic hurdles to jump through in choosing observational designs…

- Statins for CVD
- Assessed the effect of the inclusion of prevalent cases vs. incident cases
- Can prevalent vs. incidence cases be distinguished from the EHR?

American Journal of Epidemiology

Vol. 175, No. 4
DOI: 10.1093/aje/kwr301
Advance Access publication:
January 5, 2012

Systematic Reviews and Meta- and Pooled Analyses

Bias in Observational Studies of Prevalent Users: Lessons for Comparative Effectiveness Research From a Meta-Analysis of Statins

Goodarz Danaei*, Mohammad Tavakkoli, and Miguel A. Hernán

* Correspondence to Dr. Goodarz Danaei, Department of Global Health and Population, Harvard School of Public Health, 665 Huntington Avenue, Bldg. 1, Room 1107, Boston MA 02115 (e-mail: gdanaei@hsph.harvard.edu).

Initially submitted March 7, 2011; accepted for publication August 3, 2011.

Randomized clinical trials (RCTs) are usually the preferred strategy with which to generate evidence of comparative effectiveness, but conducting an RCT is not always feasible. Though observational studies and RCTs often provide comparable estimates, the questioning of observational analyses has recently intensified because of randomized-observational discrepancies regarding the effect of postmenopausal hormone replacement therapy on coronary heart disease. Reanalyses of observational data that excluded prevalent users of hormone replacement therapy led to attenuated discrepancies, which begs the question of whether exclusion of prevalent users should be generally recommended. In the current study, the authors evaluated the effect of excluding prevalent users of statins in a meta-analysis of observational studies of persons with cardiovascular disease. The pooled, multivariate-adjusted mortality hazard ratio for statin use was 0.77 (95% confidence interval (CI): 0.65, 0.91) in 4 studies that compared incident users with nonusers, 0.70 (95% CI: 0.64, 0.78) in 13 studies that compared a combination of prevalent and incident users with nonusers, and 0.54 (95% CI: 0.45, 0.66) in 13 studies that compared prevalent users with nonusers. The corresponding hazard ratio from 18 RCTs was 0.84 (95% CI: 0.77, 0.91). It appears that the greater the proportion of prevalent statin users in observational studies, the larger the discrepancy between observational and randomized estimates.

bias (epidemiology); comparative effectiveness research; confounding factors (epidemiology); meta-analysis; prospective studies; selection bias

Study Planning

- Careful planning is needed to ensure that meaningful results can be obtained feasibly.
- *NEJM* now requests protocols for all clinical trials that reach the stage of statistical review.
- Vanderbilt is implementing a new university-wide policy requiring Pre-study Statistical Plans for clinical trials and observational studies.

Where do you start when designing a research investigation?

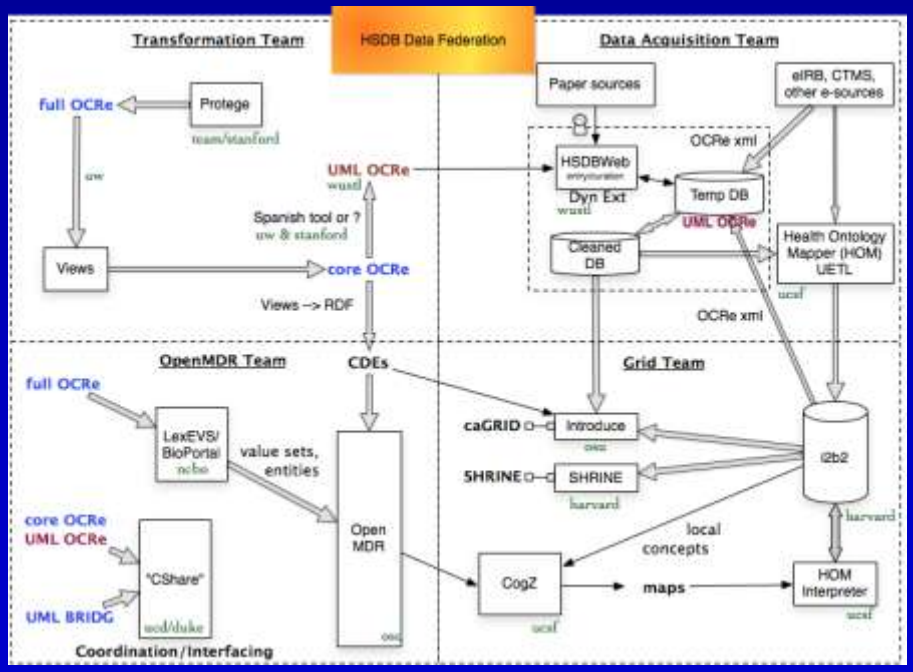
- Comb the literature
  - Publication bias
  - Latency between findings and publication
- *ClinicalTrials.gov*
  - Clinical trials only
  - Information is not at the level of detail, rigor, or standardization necessary for scientific computation
- We need accessible meta-study data
  - For all study designs

Human Studies Database Project
(http://hsdbwiki.org/)

- Goal is to create a federated, CTSA-wide database of past and ongoing human studies
  - Interventional and observational
- This will enable computational reuse of human studies data for:
  - Systematic reviews
  - Planning future studies
  - Scientific portfolio analysis
  - Research networking

Human Studies Database Project

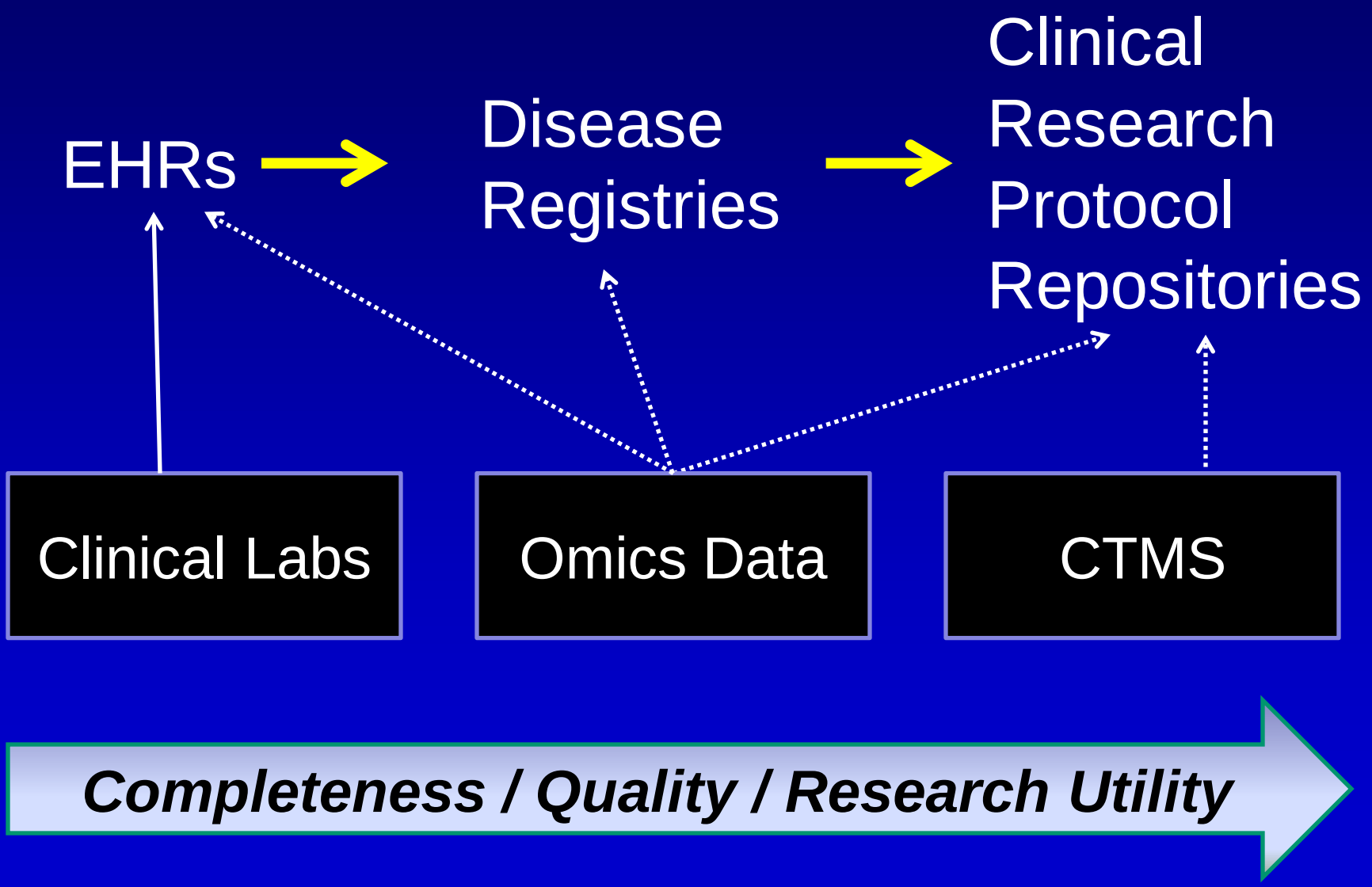
- Ontology of Clinical Research (OCRe):
  - Study design type
  - Interventions/exposures
  - Participants
  - Outcomes
  - Statistical analyses
- Piloting a federation model



HSDBgrid data sharing architecture

DATA CONSIDERATIONS

Research Data Sources



Limitations of "Existing" Non-Research Data

- Missing or inappropriately structured data elements, e.g. EHR
  - Dose schedule/intensity/AUC
  - Dose modifications
  - Reasons for stopping
  - Patient PK/PD/pharmagenomic characteristics
- Systematic biases in massive clinical data stores only get amplified, not attenuated, in analysis
- *In silico* research depends on having complete and valid information:

  Garbage In → Garbage Out

Current Challenges

- Lack of harmonization
  - Different EHRs in use at the same AHC
  - Different research databases in use at the same NCI-designated cancer center
- Clinical data systems and research data systems still do not routinely interoperate
- CTMS commercial vs. open source
  - Sustainability
  - Costs

Regulatory Barriers For Public Use Data

- Across-state permission heterogeneity for multi-state cancer registry-based studies
  - CDC-Westat survey showed 1–4 levels of approval required
- Within-state registry linkage restrictions
  - Hospital discharge data and Texas Cancer Registry

Tools / Technology

- Explosion in omnic analysis tools
- Imaging informatics
  - Highly dimensional, increasingly dimensional; e.g., real-time functional imaging
- Decision support tools require large scale validation and constant updating
  - Adjuvant! Online
  - PCPT Prostate Cancer Risk Calculator
- Data storage and networking
  - Human Genome Sequencing at BCM
  - Texas Advanced Computing Center
  - LEARN: Lonestar Education and Research Network

“Big Data” Advantages

- Big data can facilitate study planning with cohort queries to:
  - Generate hypotheses
  - Assess feasibility of planned investigations
- Big data can lower the cost of conducting clinical translational research
  - More pre-collected data
  - More automation
  - More interconnectivity

Challenges

- We need to bring the clinical data up to the same standards as high quality “research data”
- Extend statistical methodologies to deal with clinical data and design compromises
- Develop better data mining and filtering approaches for sorting through massive datasets

Challenges (continued)

- Like it or not, research process will have to evolve to accommodate changes in technology.
- Increased access to more data does not automatically lead us where we want to go:

More Data ≠ More Discovery

I'm a believer…

- We should connect genomic and molecular data with clinical data
- We should structure clinical data
- But these processes should be guided in a way that is compatible with a research framework