Clinical Evidence Development via Real-World Big Data Study

Duk-Woo Park, MD, PhD

Heart Institute, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Definition

What is Evidence-based Medicine?

> Combining quantitative evidence about medical practice with expert judgment in an effort to ensure the provision of medical care with reproducible high quality

Adapted from D Sackett

Why EBM

Evidence-based Medicine

Why should we rely on evidence for medical decision-making?

Because our intuition might be wrong!

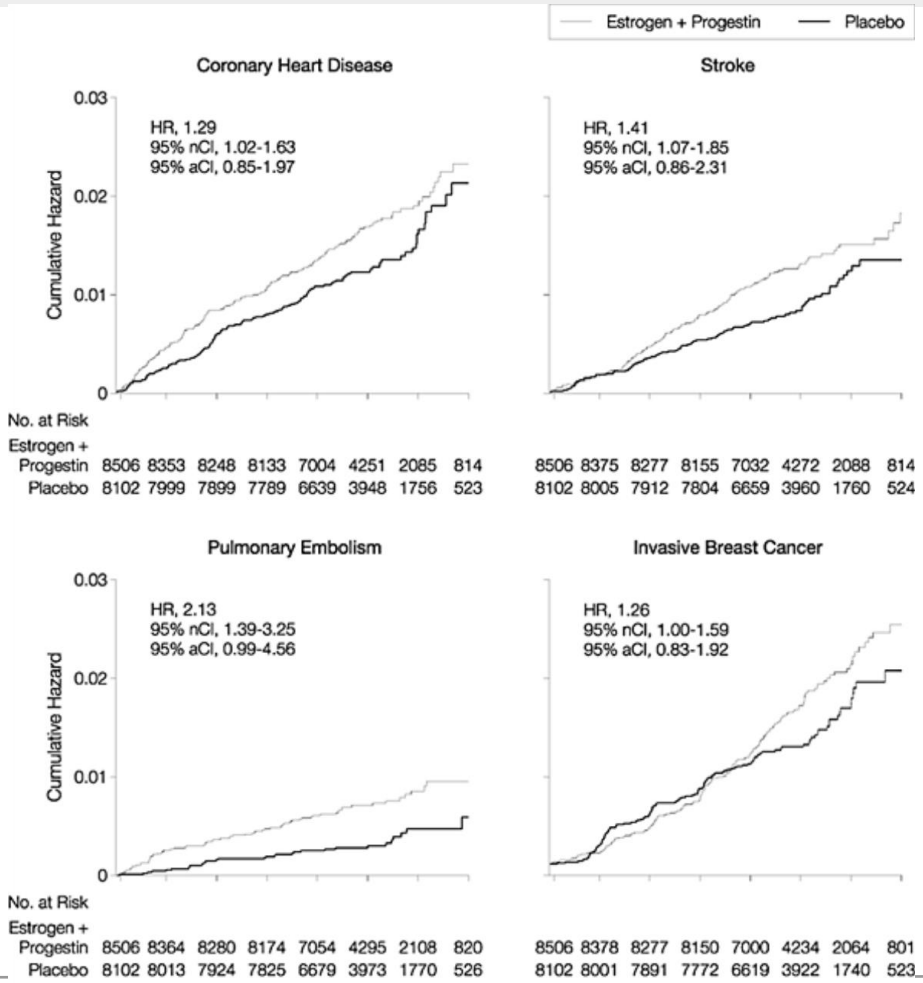
+ Why EBM

Menopause and HRT Use: WHS

- 50 million post-menopausal women in U.S.
  - *1.8 million reach menopause each year*
- ~38% of U.S. menopausal women use HRT
- In 2000:
  - *46 million prescriptions for Premarin*
    - 2nd most frequently prescribed drug in US
  - *22 million prescriptions for Prempro*
    - 6 million users
    - $900 million in sales

From: **Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Controlled Trial**

JAMA. 2002;288(3):321-333. doi:10.1001/jama.288.3.321

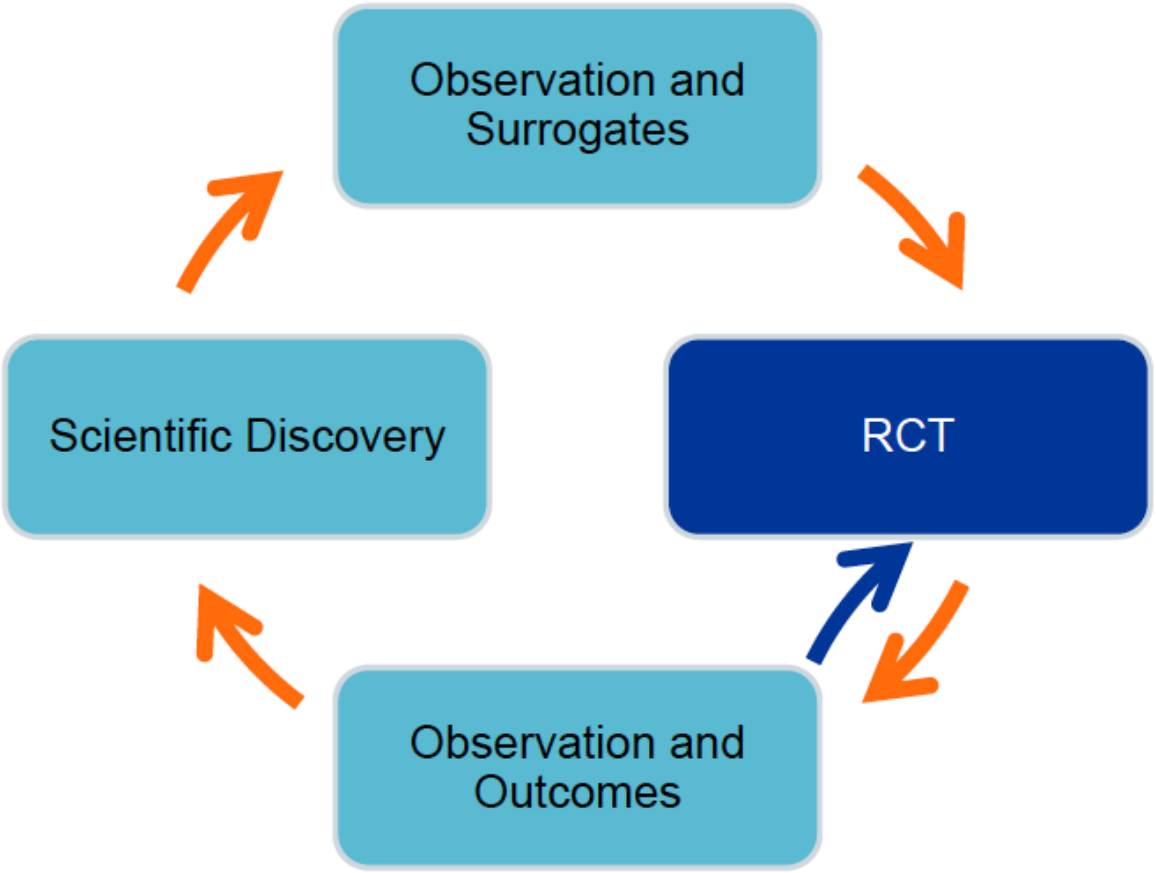


Date of download: 12/7/2013

Copyright © 2012 American Medical Association. All rights reserved.

Clinical Research

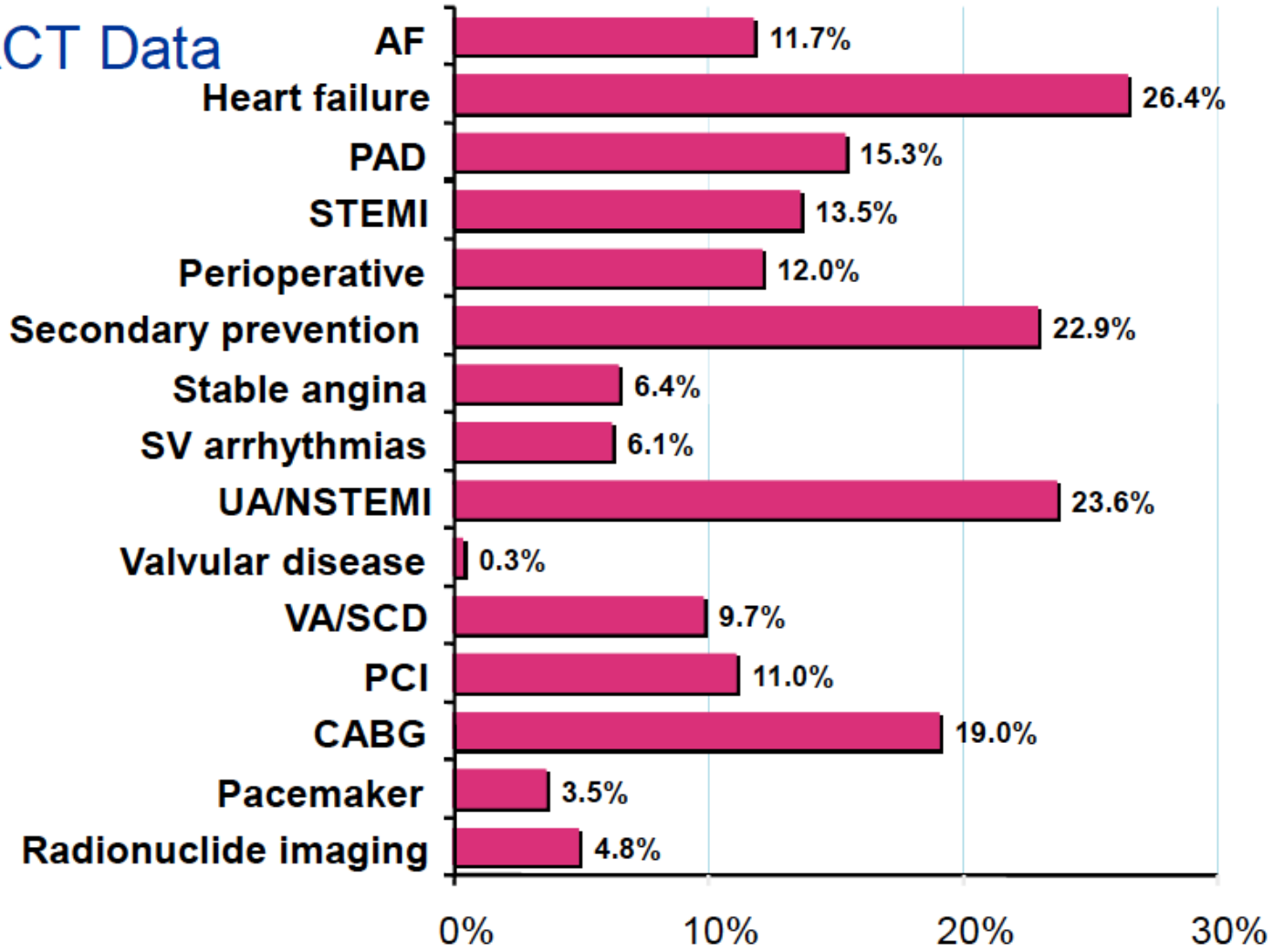
The Cycle of Research



Lack of EBM

Lack of Evidence in Guidelines:

Recommendation
Based on RCT Data

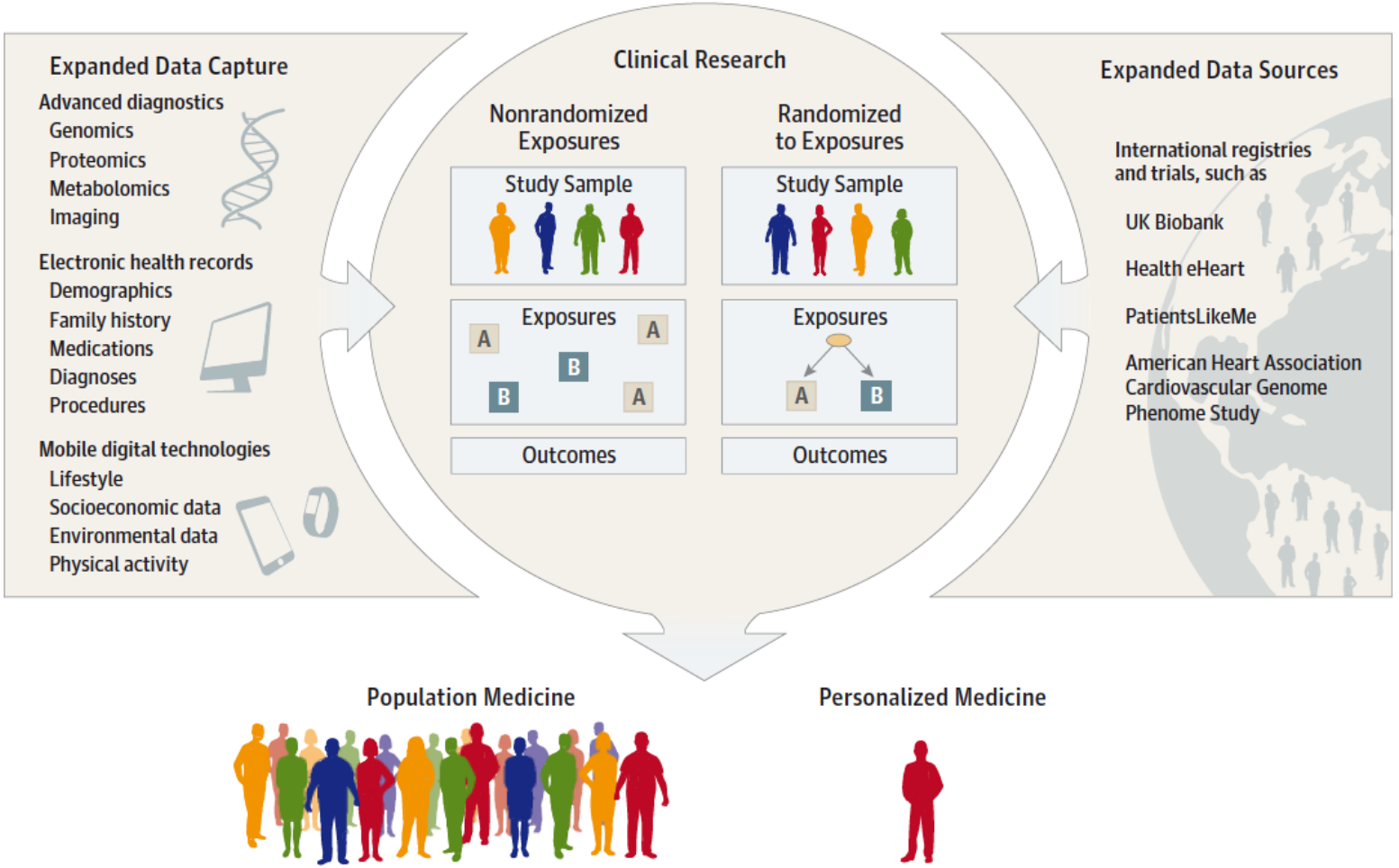


Tricoci P et al
JAMA 2009

VIEWPOINT

Population and Personalized Medicine in the Modern Era

Figure. Tools Being Used in Clinical Research to Understand Population and Personalized Medicine



Current Status of EBM in US vs. Korea?

- RCT/ Registry/Big Data in U.S.
- RCT/ Registry/Big Data in Korea

Duke Clinical Research Institute (DCRI)

The DCRI is the world largest academic clinical research organization. We combine the clinical expertise and academic leadership of a premier teaching hospital with the full-service operational capabilities of a major contract research organization

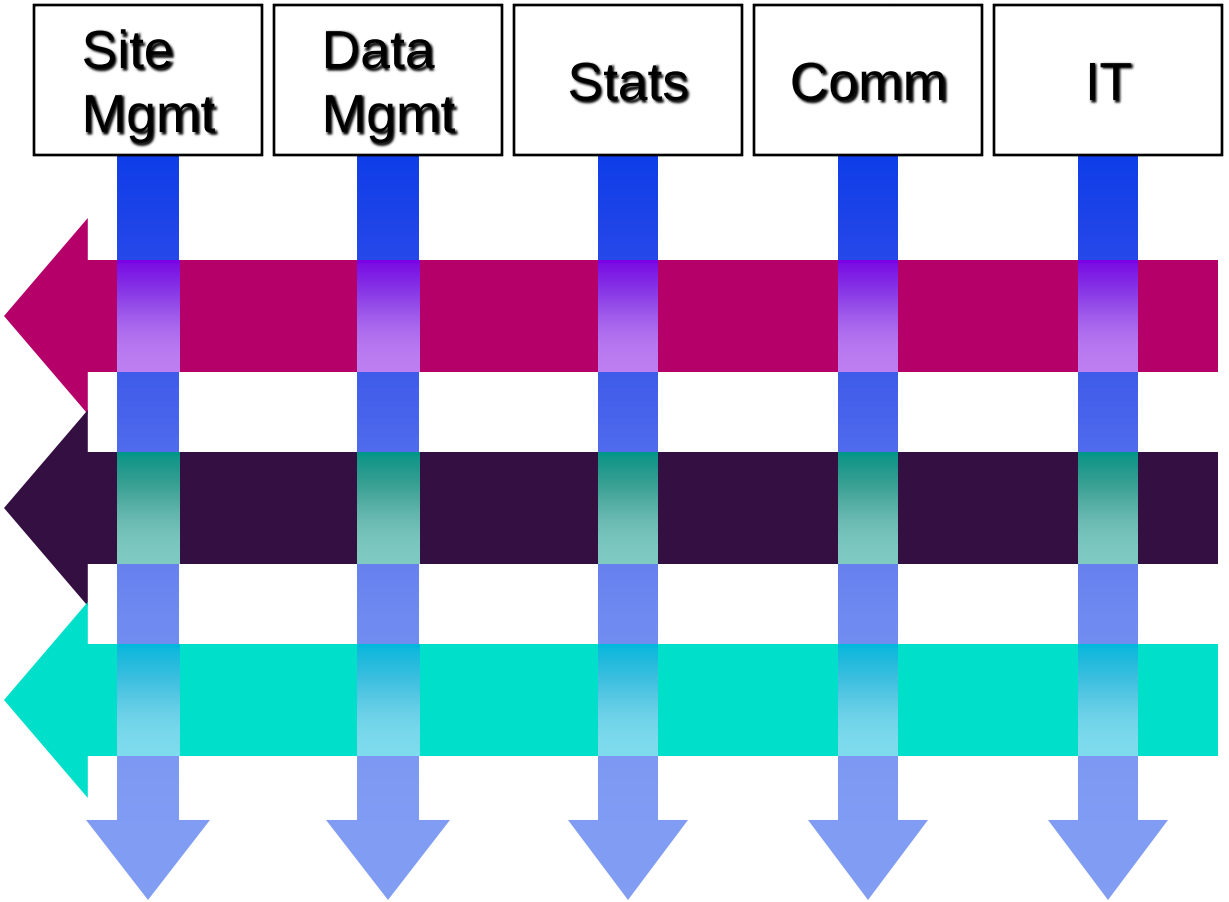
- **Employee; 1500**
- **Faculty; 300**
- **Statistician (faculty); 43**
- **Lawyer; 10**
- **Research Fellows; 30-40**

Project Level Matrix

Dual Reporting Structure

Functional Groups

provide services to Project Teams; managed by Directors



Project-Oriented Teams

managed by Project Leaders

DCRI ARO Advertising

One-Stop Services

- World renowned faculty
- Therapeutic area expertise
- High level operational capability
- Far reaching network experience
- Exceptional scientific technologies
- Publication record impacting clinical practice
- Commitment to public-private partnerships
- Credibility with regulators and medical community

Why Duke Clinical Research Institute (DCRI) ?

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Prasugrel versus Clopidogrel for Acute Coronary Syndromes without Revascularization

Matthew T. Roe, M.D., M.H.S., Paul W. Armstrong, M.D., Keith A.A. Fox, M.B., Ch.B., Harvey D. White, M.B., Ch.B., D.Sc., Dorairaj Prabhakaran, M.D., D.M., Shaun G. Goodman, M.D., Jan H. Cornel, M.D., Ph.D., Deepak L. Bhatt, M.D., M.P.H., Peter Clemmensen, M.D., D.M.Sc., Felipe Martinez, M.D., Diego Ardissino, M.D., Jose C. Nicolau, M.D., Ph.D., William E. Boden, M.D., Paul A. Gurbel, M.D., Witold Ruzyllo, M.D., Anthony J. Dalby, M.D., Darren K. McGuire, M.D., M.H.Sc., Jose L. Leiva-Pons, M.D., Alexander Parkhomenko, M.D., Ph.D., Shmuel Gottlieb, M.D., Gracita O. Topacio, M.D., Christian Hamm, M.D., Gregory Pavlides, M.D., Assen R. Goudev, M.D., Ali Oto, M.D., Chuen-Den Tseng, M.D., Ph.D., Bela Merkely, M.D., Ph.D., D.Sc., Vladimir Gasparovic, M.D., Ph.D., Ramon Corbalan, M.D., Mircea Cinteză, M.D., Ph.D., R. Craig McLendon, R.N., Kenneth J. Winters, M.D., Eileen B. Brown, Ph.D., Yuliya Lokhnygina, Ph.D., Philip E. Aylward, B.M., B.Ch., Ph.D., Kurt Huber, M.D., Judith S. Hochman, M.D., and E. Magnus Ohman, M.B., Ch.B., for the TRILOGY ACS Investigators*

ABSTRACT

BACKGROUND

The effect of intensified platelet inhibition for patients with unstable angina or myocardial infarction without ST-segment elevation who do not undergo revascularization has not been delineated.

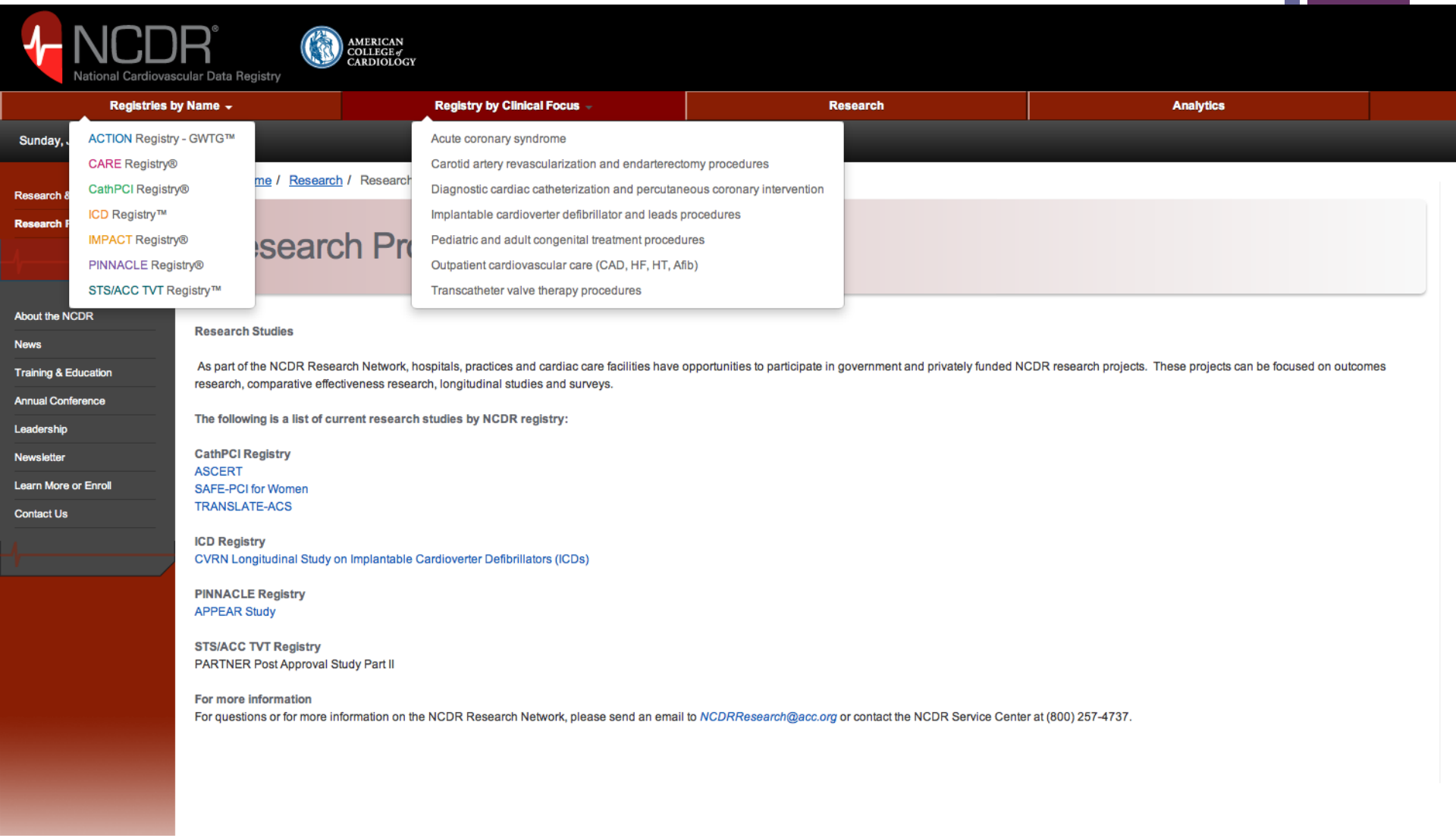
METHODS

In this double-blind, randomized trial, in a primary analysis involving 7243 patients under the age of 75 years receiving aspirin, we evaluated up to 30 months of treatment with prasugrel (10 mg daily) versus clopidogrel (75 mg daily). In a secondary analysis involving 2083 patients 75 years of age or older, we evaluated 5 mg of prasugrel versus 75 mg of clopidogrel.

The authors' affiliations are listed in the Appendix. Address reprint requests to Dr. Roe at Duke Clinical Research Institute, 2400 Pratt St., Rm. 7035, Durham, NC 27705, or at matthew.roe@duke.edu.

*The Targeted Platelet Inhibition to Clarify the Optimal Strategy to Medically Manage Acute Coronary Syndromes (TRILOGY ACS) investigators are listed in the Supplementary Appendix, available at NEJM.org.

NCDR (National Cardiovascular Data Reg[redacted])



GWTG (Get With The Guidelines)

NCDR-CathPCI DB

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Low Diagnostic Yield of Elective Coronary Angiography

Manesh R. Patel, M.D., Eric D. Peterson, M.D., M.P.H., David Dai, M.S., J. Matthew Brennan, M.D., Rita F. Redberg, M.D., H. Vernon Anderson, M.D., Ralph G. Brindis, M.D., and Pamela S. Douglas, M.D.

ABSTRACT

BACKGROUND

Guidelines for triaging patients for cardiac catheterization recommend a risk assessment and noninvasive testing. We determined patterns of noninvasive testing and the diagnostic yield of catheterization among patients with suspected coronary artery disease in a contemporary national sample.

METHODS

From January 2004 through April 2008, at 663 hospitals in the American College of Cardiology National Cardiovascular Data Registry, we identified patients without known coronary artery disease who were undergoing elective catheterization. The patients' demographic characteristics, risk factors, and symptoms and the results of noninvasive testing were correlated with the presence of obstructive coronary artery disease, which was defined as stenosis of 50% or more of the diameter of the left main coronary artery or stenosis of 70% or more of the diameter of a major epicardial vessel.

N Engl J Med 2010;362:886-95

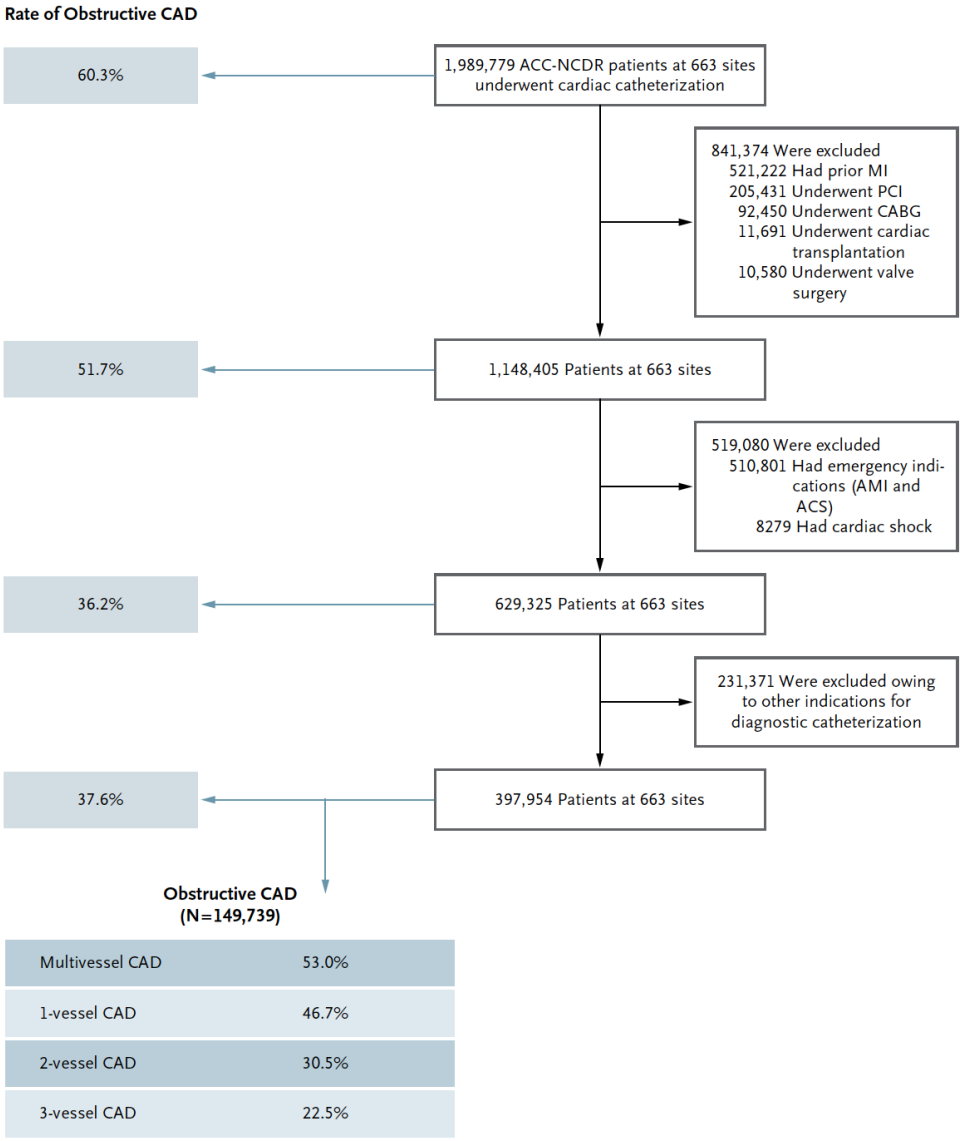


Figure 1. Study Population and Rates of Obstructive Coronary Artery Disease.

ACC-NCDR denotes American College of Cardiology National Cardiovascular Data Registry, ACS acute coronary syndrome, AMI acute myocardial infarction (MI), CABG coronary-artery bypass grafting, CAD coronary artery disease, and PCI percutaneous coronary intervention.

Academic Trends; Top Priority

SPECIAL ARTICLE

Future Directions for Cardiovascular Disease Comparative Effectiveness Research

Report of a Workshop Sponsored by the National Heart, Lung, and Blood Institute

Mark A. Hlatky, MD,* Pamela S. Douglas, MD,† Nakela L. Cook, MD, MPH,‡ Barbara Wells, PHD,‡ Emelia J. Benjamin, MD, SCD,§ Kay Dickersin, PHD, MA,‖ David C. Goff, MD, PHD,¶ Alan T. Hirsch, MD,# Elaine M. Hylek, MD,§ Eric D. Peterson, MD, MPH,† Véronique L. Roger, MD, MPH,** Joseph V. Selby, MD, MPH,†† James E. Udelson, MD,‡‡ Michael S. Lauer, MD‡

Stanford and Oakland, California; Durham and Winston-Salem, North Carolina; Bethesda and Baltimore, Maryland; Boston, Massachusetts; and Minneapolis and Rochester, Minnesota

CER

- CER has recently emerged as a national priority, spurred by healthcare reform and and economic stimulus legislation.
- Congress appropriated **$1.1 billion** for CER.
- PCORI (Patient-Centered Outcomes Research Institute).
- PCORI gave priority for project management to the NIH and the AHRQ (Agency for Health Research and Quality).

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

JULY 23, 2009

Prioritizing Comparative-Effectiveness Research — IOM Recommendations

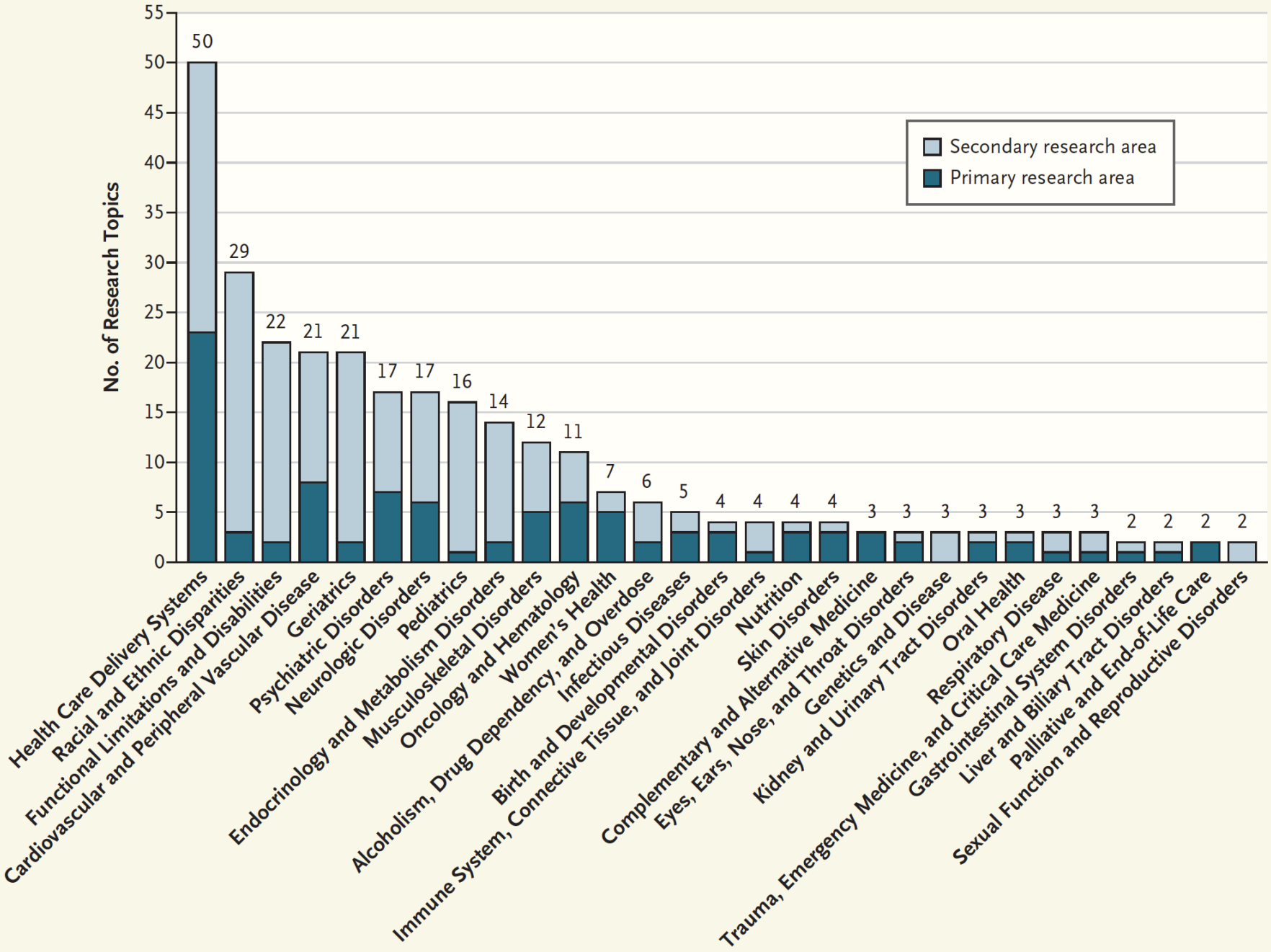
John K. Iglehart

Directed by Congress to rapidly develop a list of broad-based priorities for the Department of Health and Human Services (DHHS) to consider as it implements a new agenda for comparative-

effectiveness research (CER), the Institute of Medicine (IOM) re-

sustainable national CER strategy" and that Congress and the

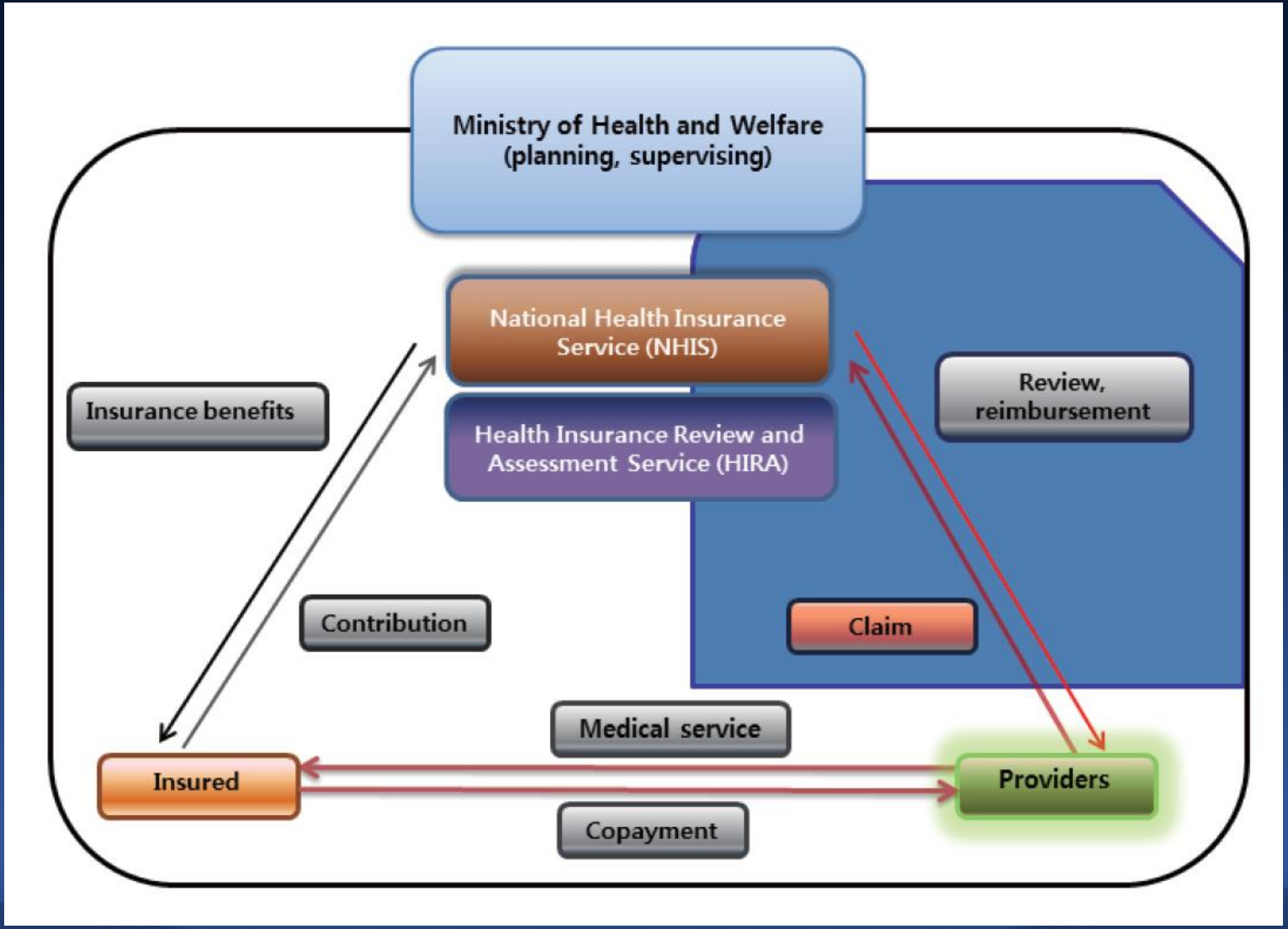
trust in the U.S. research enterprise." The committee began with 1268 CER topics that were nominated by stakeholders and the public and winnowed them down to 82; the other 18 topics were recommended by the committee to fill gaps in the portfolio.


No. of Research Topics
55
50
45
40
35
30
25
20
15
10
5
0
Secondary research area
Primary research area
50
29
22
21
21
17
17
16
14
12
11
7
6
5
4
4
4
4
3
3
3
3
3
3
3
2
2
2
2
Health Care Delivery Systems
Racial and Ethnic Disparities
Functional Limitations and Disabilities
Cardiovascular and Peripheral Vascular Disease
Geriatrics
Psychiatric Disorders
Neurologic Disorders
Pediatrics
Endocrinology and Metabolism Disorders
Musculoskeletal Disorders
Oncology and Hematology
Women's Health
Alcoholism, Drug Dependency, and Overdose
Infectious Diseases
Birth and Developmental Disorders
Immune System, Connective Tissue, and Joint Disorders
Nutrition
Skin Disorders
Complementary and Alternative Medicine
Eyes, Ears, Nose, and Throat Disorders
Genetics and Disease
Kidney and Urinary Tract Disorders
Oral Health
Respiratory Disease
Trauma, Emergency Medicine, and Critical Care Medicine
Gastrointestinal System Disorders
Liver and Biliary Tract Disorders
Palliative and End-of-Life Care
Sexual Function and Reproductive Disorders

Registry-based comparative effectiveness research

How Do We Make A Good Clinical Evidence Using Big Data in Korea

Operational structure of the National Health Insurance program in Korea



Song SO et al. Diabetes Metab J 2014;38:395-403

Components of Claim Data in the National Health Insurance in Korea

20T	30T	40T	60T
Payment specification	Consultation statement	Diagnosis statement	Detail statement of prescription
Personal identification	Medical examination and treatment such as:	Principal diagnosis from 1st to 9th additional diagnoses	Name of drug
Health and medical care institution			Date
Principal diagnosis	Medical care		Filled days
1st additional diagnosis	In-hospital administration of medicine		Supply
Days of medical care	Procedure		Quantity dispensed
Commencement date of medical care	Surgery		Price of each drug
No. of visiting days			
Insurer and deduction payment			

20T, 20 table, consist of unique number delimiter; 30T, 30 table, consist of unique number delimiter; 40T, 40 table, consist of unique number delimiter; 60T, 60 table, consist of unique number delimiter.

Song SO et al. Diabetes Metab J 2014;38:395-403

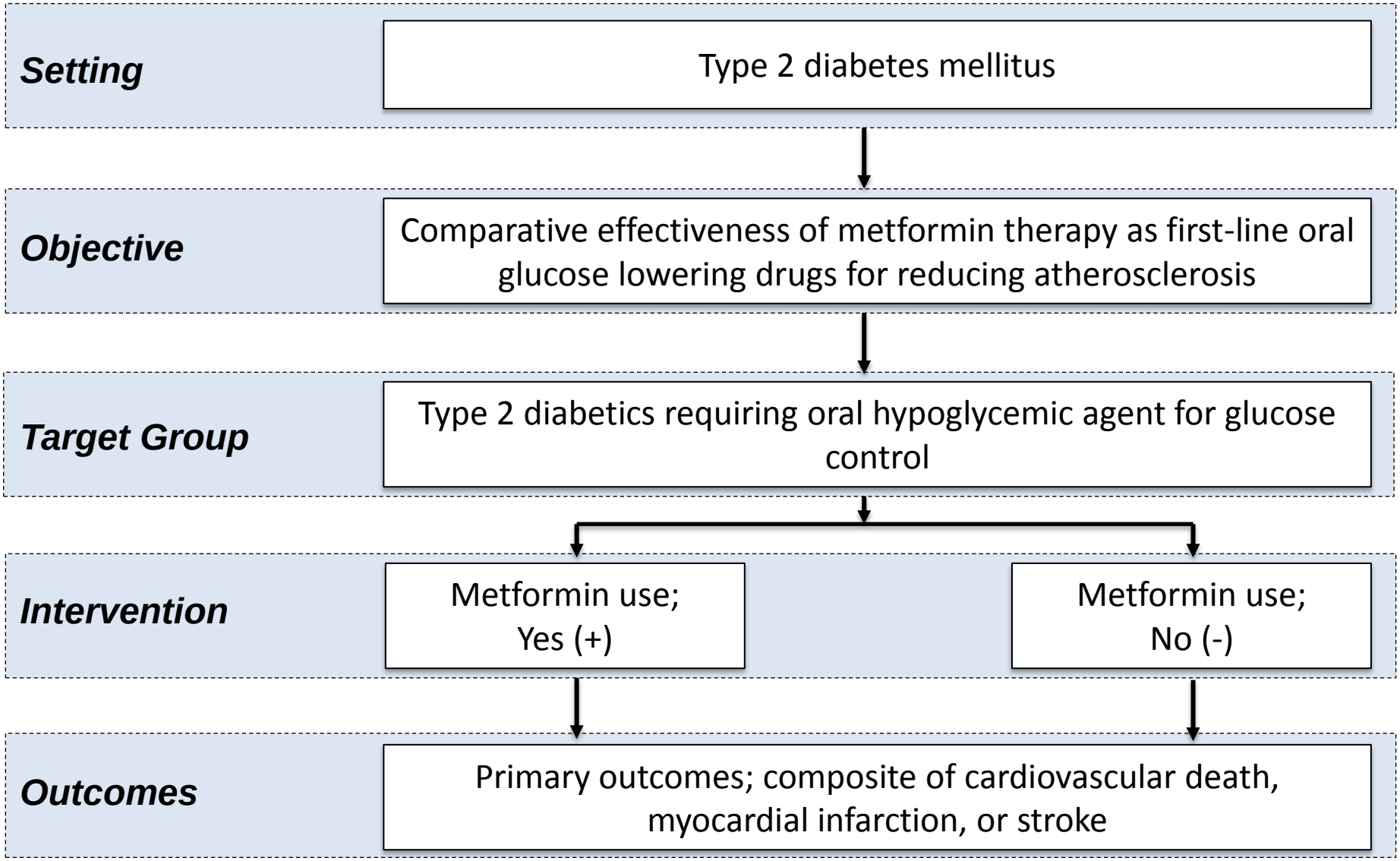
The data characteristics according to the National Health Insurance Service program

Characteristic		Qualification and contribution data	Health insurance claiming data	Health check-up data	Long-term care insurance data
Demographic information	Sex	O	O	O	O
	Age	O	O	O	O
	Region	O			O
	Family information	O			O
	Presence of handicap	O			O
	Death	O			O
	Type of qualification	O			O
	Contribution amount (incomes)	O			O
Medical use	Medical service use		O	O	
	Medical costs		O		
Diseases information	Chronic diseases		O	O	O
	Accident/Poisoning		O		
	Health check-up			O	
	Cognitive function			O	O
Lifestyle and habits	Smoking			O	
	Alcohol			O	
	Obesity			O	
	Exercise			O	
Basic laboratory data				O	

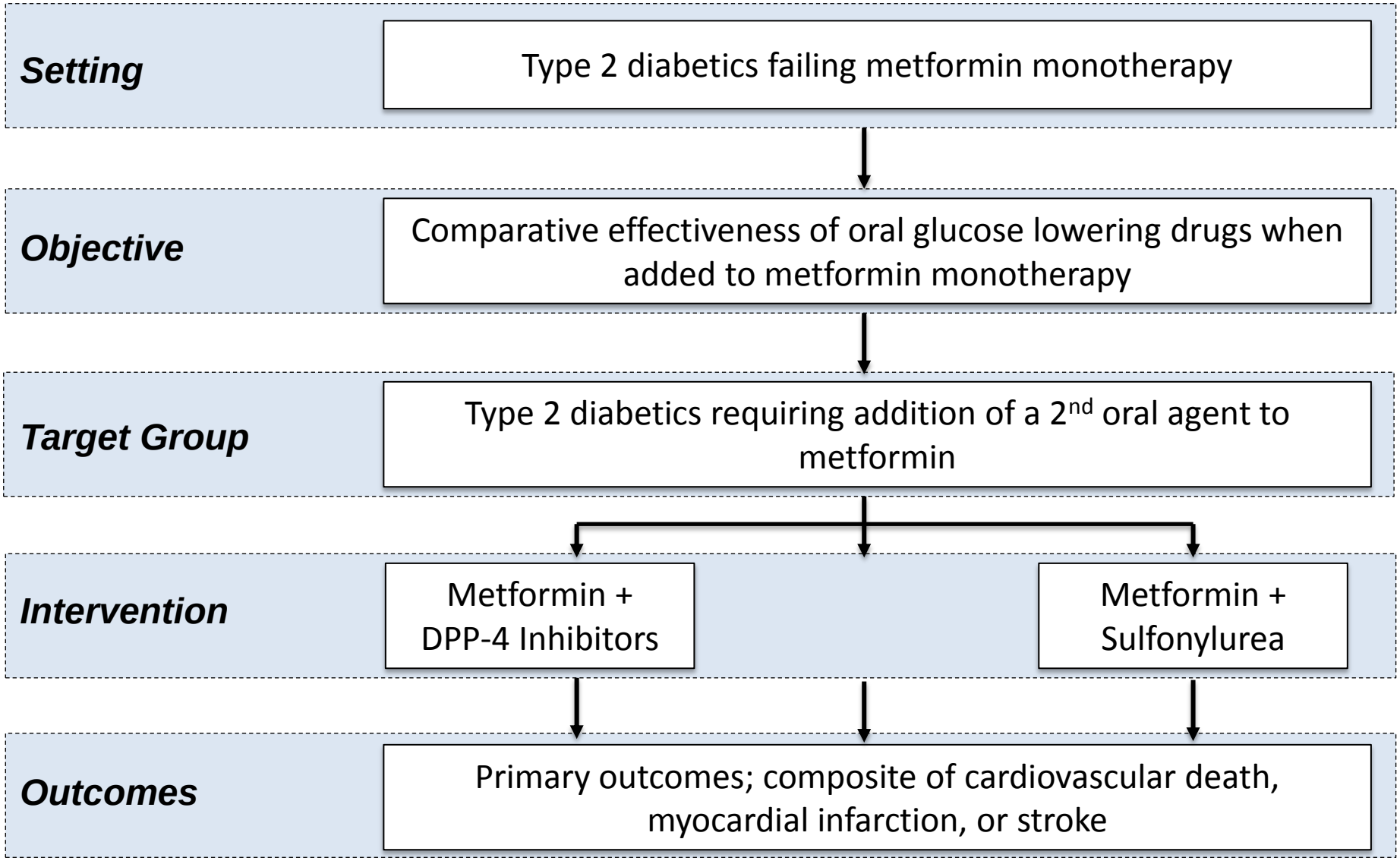
Song SO et al. Diabetes Metab J 2014;38:395-403

Comparative Effectiveness of Metformin Initial Therapy and Add-On Second-Line Drugs on Major Cardiovascular Events Among Patients with Type 2 Diabetes: Observational Study of Administrative Databases

Comparative Effectiveness of First-Line Oral Hypoglycemic Agents



Comparative Effectiveness of Second-Line Oral Hypoglycemic Agents



Project: Oral Antidiabetes Drugs and the Risk of Major Cardiovascular Events: Comparative Effective Research of Metformin-Based Therapy in Adult Patients with Type 2 Diabetes Mellitus From A Nationwide Population-Based Study

Target Journal: NEJM (1st)/ Lancet (2nd) / JAMA (3rd) / BMJ (4th) / AIM (5th) / Circulation (6th) / and then others

Co-Principal Investigator: Duk-Woo Park, MD; Woo-Jae Lee, MD (co-corresponding authors)
Co-Investigator(s): Min-Jung Ko, PhD (first author) and other co-authors

Principal Statistician: Yoon-Jung Kim, MPH
Mentoring Statistician:

Specific Aims:

- Given the common and increasing use of several anti-diabetes drugs and lacking information regarding the relative benefits and disadvantages to cardiovascular health, we investigated the temporal pattern of use over last decade and the risk of major cardiovascular events associated with prescription of different classes of oral antidiabetes drugs focused on metformin-based therapy in routine clinical care.

1. Based on current guideline that metformin is advocated as first line pharmacotherapy for type 2 diabetes, we evaluate the current pattern and clinical impact of metformin use on cardiovascular events.
2. We also evaluate the comparative effectiveness of specific second-line antidiabetic drugs (sulphonylureas [SU], thiazolidinedione [TZD], or dipeptidyl peptidase 4 [DPP-4] inhibitors) added-on metformin therapy.

Population:

- Adult patients aged 18 years or older with type 2 diabetes mellitus who received oral diabetes agents for at least the past 365 days linked to national prescription claim records (i.e., pharmacy data sets for prescription records linked to Health Insurance Review & Assessment Service).
- Cohort entry date for each patient was the date of first prescription, and the exit date (censoring date) was earliest of: a) date of death; b) date of recording of clinical event; c) no contact of at least 365 days on claim data (any data on inpatient, outpatient, or pharmacy use) or d) end of study period (December 31, 2012).
- Data on pharmacy, demographic, clinical covariates, or laboratory value were collected from Korean's National Health Insurance Service and Health Insurance Review & Assessment Service database, between 1/Jan/2005- 31/December/2011.
  - Exclude patients with insulin treatment more than at least 6 months at any time period
  - Exclude patients with malignancy at baseline

Endpoints:

- Primary Endpoint:
  - Major cardiovascular event, defined as composite of death from cardiovascular causes, nonfatal myocardial infarction (MI), or nonfatal stroke.
- Secondary Endpoints:
  - Each component of primary endpoint; death from cardiovascular causes, MI, or stroke
  - All-cause mortality

CardioVascular Research Foundation
University of Ulsan College of Medicine
Asan Medical Center

- Composite of all-cause death, nonfatal MI, or nonfatal stroke
- Congestive heart failure

- Event-assessment:
  - The long-term follow-up was based on merging of national registries of the Korean's National Health Insurance Service; Health Insurance Review & Assessment Service; and the National Population Registry of the Korea National Statistical Office database on the basis of the unique personal identification number of each Korean citizen.
  - We obtained data regarding hospitalization for acute MI (as defined in the International Classification of Diseases, 10th revision, disease codes, I21-I23, I25.2), stroke (disease code, I60-64, I67-68, I69), and congestive heart failure (disease code, I50) from the Health Insurance Review & Assessment Service through December 31, 2012 which ensure at least 1-year of follow-up.
  - Data on vital status, date of death, and cause of death were obtained from the National Population Registry of the Korea National Statistical Office through December31, 2012, from the Korea National Statistical Office with the use of a unique personal identification number.
  - The merging of the national data was performed by the National Evidence-Based Healthcare Collaborating Agency (NECA) and was approved by the institutional review board of the NECA.

Analysis Objectives & Tasks:

1. **Objective:** Summarize temporal pattern of oral antidiabetes drugs prescribed among adult patients with type 2 diabetes.

   Analysis: We identified oral antidiabetes treatments of individual patients from prescription records using the Health Insurance Review & Assessment Service database: monotherapy (metformin, sulfonylureas [SU], thiazolidinedione [TZD], incretin mimetic, and other oral antidiabetes drugs [i.e., acarbose, naeglinide, repaglinide]) and combination therapy (two-, three-, and more than four anti-diabetes drug combinations).
   We summarize the temporal pattern of monotherapy and combination therapy and the temporal change of relative proportion of each anti-diabetes drug from 2002 to 2012.

   - See Appendix for proposed Figure 1: Temporal trend of type and number of oral antidiabetes drugs stratified by number (Figure 1A) and classes (Figure 1B).

**For fair comparison of baseline covariates and outcomes according to specific antidiabetic drugs and reducing non-systematic misclassification errors, drug-group classification is essential. For drug-group classification, we primarily used patient-level analysis for drug treatment categorized by drug class. To further assess the robustness of our findings, we performed a series of additional sensitivity analyses (3 different styles) to confirm whether overall findings regarding comparative effectiveness were consistent regardless of analytic methods.

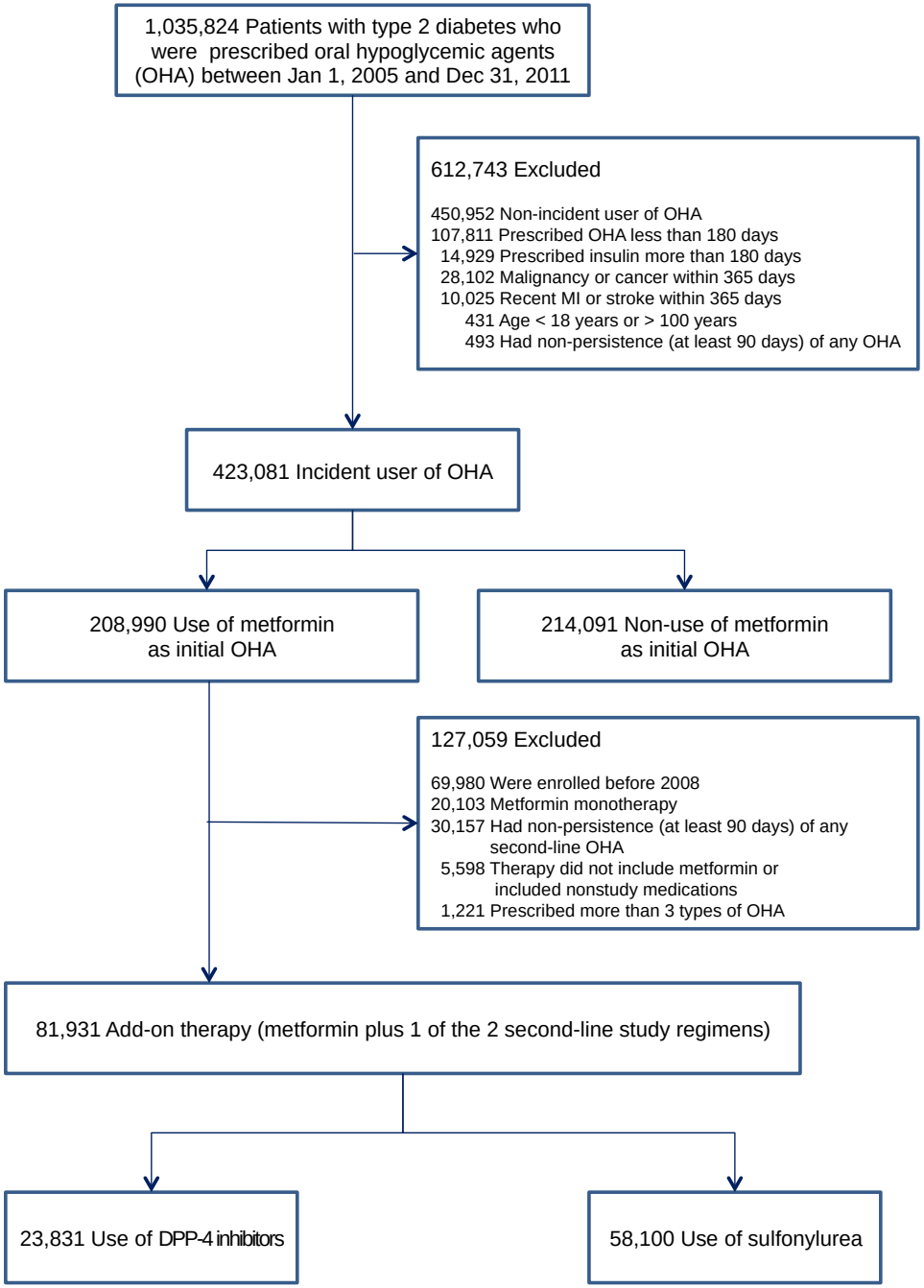
1) Primary analysis [incident user design analysis]: incident (new) users of metformin from January 2005 through December 2011 will be identified. Among metformin initiators, comparison of second-line antidiabetic drugs (SU vs. TZD vs. DPP-4 inhibitors) add-on metformin will be performed. Follow-up will continue through a study outcome, a switch to or addition of another antidiabetic drugs, the 365 days without claim data on antidiabetic medications, or end of the study (December 31, 2012).

CardioVascular Research Fou
TY OF ULSAN
E MEDICINE
ASAN
Medical Center

Table 1. Patient Characteristics According to Use of Metformin

Variable	Metformin user (N=)	Metformin non user (N=)	P value	자료원	변수유형
No. of intervals					
No. of patients					
Demographics					
Age, median (IQR), years				청구자료 - T20	
Female				청구자료 - T20	
BMI, median (IQR)				검진자료	
Duration of diabetes at prescription (years)				청구자료 - T40	ICD-10
Systolic blood pressure (mmHg)				검진자료	
Diastolic blood pressure (mmHg)				검진자료	
Heart rate (rate / min)				검진자료	
Clinical history or risk factors					
Hypertension				청구자료 - T40	ICD-10
Hypercholesterolemia				청구자료 - T40	ICD-10
Current smoker				청구자료 - T40	ICD-10
Family history of CAD				검진자료	
Chronic lung disease (COPD or emphysema or asthma)				청구자료 - T40	ICD-10
Coronary artery disease				청구자료 - T40	ICD-10
Prior myocardial infarction				청구자료 - T40	ICD-10
Prior stroke				청구자료 - T40	ICD-10
Carotid or cerebrovascular disease					
Peripheral vascular disease				청구자료 - T40	ICD-10
Prior coronary-artery bypass grafting				청구자료 - T40	ICD-10
Prior coronary angioplasty				청구자료 - T40	ICD-10
Laboratory data at index prescription					
Total cholesterol				검진자료	
LDL-cholesterol				검진자료	
HDL-cholesterol(2008년 이후 입원)				검진자료	
Triglyceride				검진자료	
Serum creatinine				검진자료	
Fasting serum glucose					
Concomitant other antidiabetes drugs				청구자료 - T30	아프코드
Sulphonylureas				청구자료 - T30	아프코드
Thiazolidinediones				청구자료 - T30	아프코드
DPP IV-inhibitors				청구자료 - T30	아프코드
GLP-1 analogue				청구자료 - T30	아프코드
All others				청구자료 - T30	아프코드
Concomitant cardiovactive medications					
Aspirin				청구자료 - T30	아프코드
Antiplatelet agents				청구자료 - T30	아프코드
Anticoagulants				청구자료 - T30	아프코드
Statin				청구자료 - T30	아프코드
β-blocker				청구자료 - T30	아프코드
Calcium-channel blocker				청구자료 - T30	아프코드
ACE inhibitors or ARB				청구자료 - T30	아프코드
Diuretics				청구자료 - T30	아프코드

CardioVascular Research Fou
ITY OF ULSAN
E MEDICINE
ASAN Medical Center

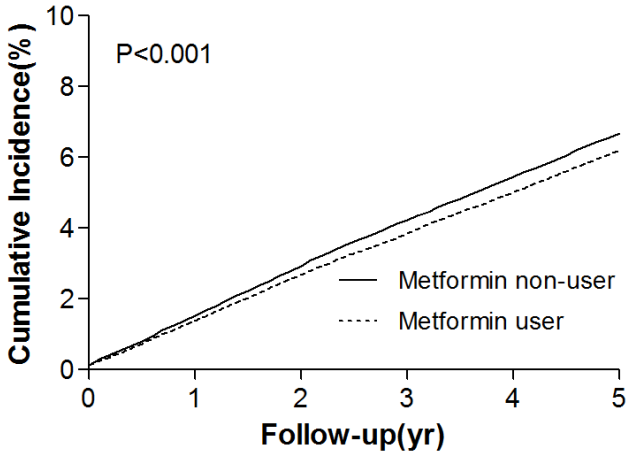

1,035,824 Patients with type 2 diabetes who were prescribed oral hypoglycemic agents (OHA) between Jan 1, 2005 and Dec 31, 2011
612,743 Excluded
450,952 Non-incident user of OHA
107,811 Prescribed OHA less than 180 days
14,929 Prescribed insulin more than 180 days
28,102 Malignancy or cancer within 365 days
10,025 Recent MI or stroke within 365 days
431 Age < 18 years or > 100 years
493 Had non-persistence (at least 90 days) of any OHA
423,081 Incident user of OHA
208,990 Use of metformin as initial OHA
214,091 Non-use of metformin as initial OHA
127,059 Excluded
69,980 Were enrolled before 2008
20,103 Metformin monotherapy
30,157 Had non-persistence (at least 90 days) of any second-line OHA
5,598 Therapy did not include metformin or included nonstudy medications
1,221 Prescribed more than 3 types of OHA
81,931 Add-on therapy (metformin plus 1 of the 2 second-line study regimens)
23,831 Use of DPP-4 inhibitors
58,100 Use of sulfonylurea

Baseline Characteristics of Patients

Characteristic	Before Matching: Metformin User (N = 208,990)	Before Matching: Metformin Non-User (N = 214,091)	Before Matching: Standardized Differences (%)	After Matching: Metformin User (N = 159,509)	After Matching: Metformin Non-User (N = 159,509)	After Matching: Standardized Difference (%)
Demographics						
Age, median (IQR), y	59 (49–69)	62 (52–71)	19.3	61 (50–69)	61 (51–70)	3.2
Female sex	98,125 (47.0)	106,149 (49.6)	4.3	77,932 (48.9)	78,184 (49.0)	0.3
Income quintile[b]						
1	36,259 (17.4)	37,477 (17.5)	0.3	27,964 (17.5)	27,847 (17.5)	0.2
2	30,312 (14.5)	31,161 (14.6)	0.1	22,683 (14.2)	22,507 (14.1)	0.3
3	64,182 (30.7)	66,447 (31.0)	0.6	51,669 (32.4)	51,808 (32.5)	0.2
4	35,232 (16.9)	35,629 (16.6)	0.5	25,819 (16.2)	25,907 (16.2)	0.1
5	43,005 (20.6)	43,377 (20.3)	0.7	31,374 (19.7)	31,440 (19.7)	<0.1
Risk factors and clinical history						

	Before Matching			After Matching		
Characteristic	Metformin User (N = 208,990)	Metformin Non-User (N = 214,091)	Standardized Differences (%)	Metformin User (N = 159,509)	Metformin Non-User (N = 159,509)	Standardized Difference (%)
Hypertension	145,332 (69.5)	154,520 (72.2)	4.7	113,915 (71.4)	114,270 (71.6)	0.4
Hyperlipidemia	129,175 (61.8)	118,101 (55.2)	11.1	92,211 (57.8)	92,300 (57.9)	0.1
Current smoker[c]	44,418 (42.3)	40,993 (41.4)	1.5	30,234 (41.2)	30,167 (41.1)	0.2
Chronic lung disease	17,036 (8.2)	21,971 (10.3)	6.1	13,907 (8.7)	13,950 (8.8)	<0.1
Coronary artery disease	39,959 (19.1)	41,524 (19.4)	0.6	30,146 (18.9)	30,088 (18.9)	<0.1
Carotid or cerebrovascular disease	63,811 (30.5)	67,402 (31.5)	1.7	49,033 (30.7)	48,962 (30.7)	<0.1
Peripheral vascular disease	32,939 (15.8)	35,714 (16.7)	2.1	25,944 (16.3)	25,887 (16.2)	<0.1
Renal disease	41,110 (19.7)	44,482 (20.8)	2.3	32,099 (20.1)	32,181 (20.2)	0.1
Prior PCI	2,132 (1.0)	1,716 (0.8)	1.9	1,334 (0.8)	1,360 (0.9)	<0.1
Prior CABG	189 (0.1)	166 (0.1)	0.3	135 (0.1)	136 (0.1)	0.3
Charlson comorbidity index						
0	353 (0.2)	371 (0.2)	<0.1	325 (0.2)	292 (0.2)	0.4
1	6,352 (3.0)	5,007 (2.3)	3.5	4,117 (2.6)	4,223 (2.7)	0.4
2	13,731 (6.6)	10,925 (5.1)	5.0	8,714 (5.5)	9,166 (5.8)	1.0
≥3	188,554 (90.2)	197,788 (92.4)	6.2	146,353 (91.8)	145,828 (91.4)	1.0

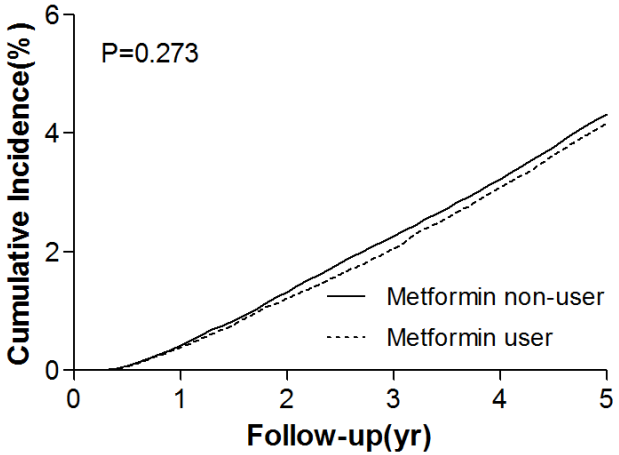
A Primary Outcome



No. at risk

Metformin user	159509	113636	96512	81791	68014	54288
Metformin non-user	159509	110745	92172	77394	63937	50766

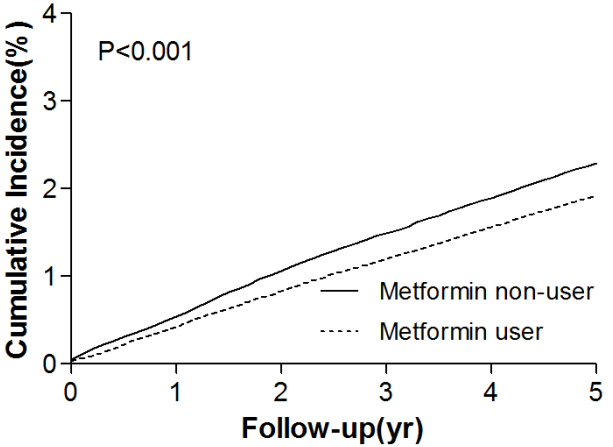
B All-Cause Death



No. at risk

Metformin user	159509	113669	96566	81842	68048	54311
Metformin non-user	159509	110774	92227	77451	63985	50802

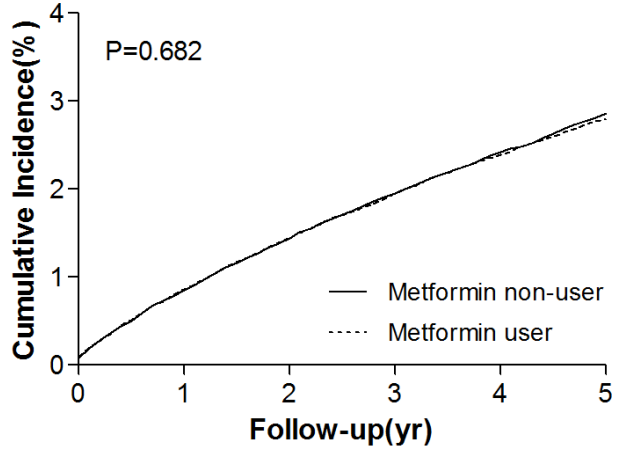
C Myocardial Infarction



No. at risk

Metformin user	159509	113627	96494	81764	67988	54264
Metformin non-user	159509	110732	92165	77378	63922	50750

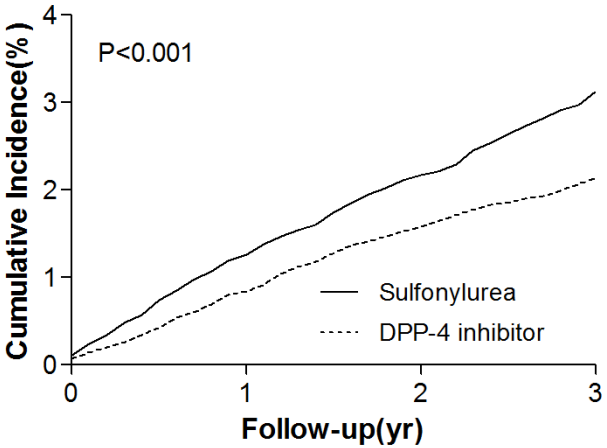
D Ischemic Stroke



No. at risk

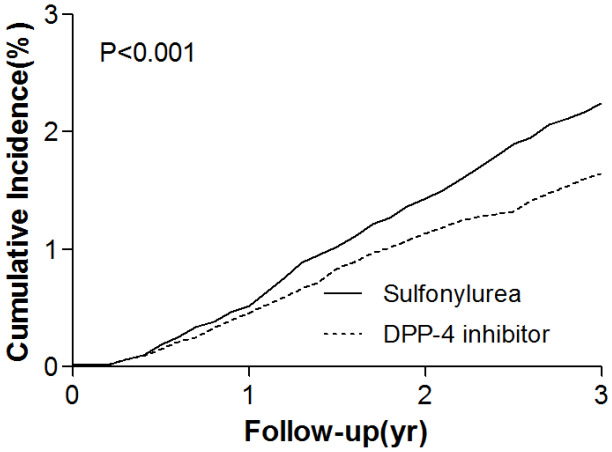
Metformin user	159509	113585	96457	81735	67963	54247
Metformin non-user	159509	110717	92131	77360	63900	50744

A Primary Outcome



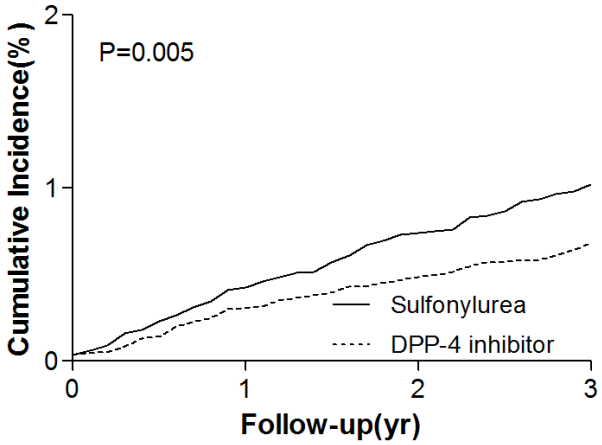
No. at risk				
DPP4 inhibitor	22039	16015	10885	5887
Sulfonylurea	22039	15461	10303	5464

B All-Cause Death



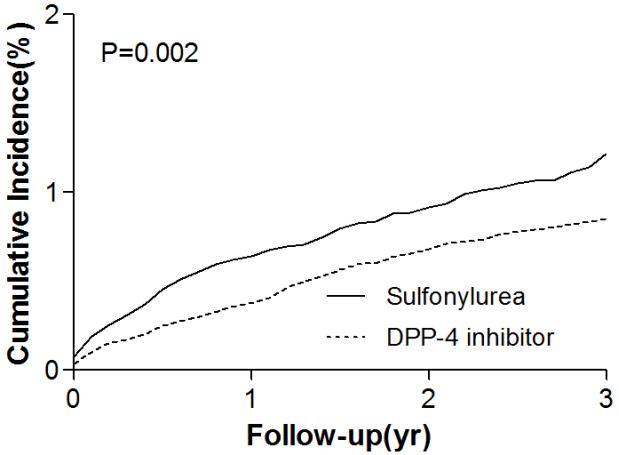
No. at risk				
DPP4 inhibitor	22039	16016	10888	5888
Sulfonylurea	22039	15460	10302	5463

C Myocardial Infarction



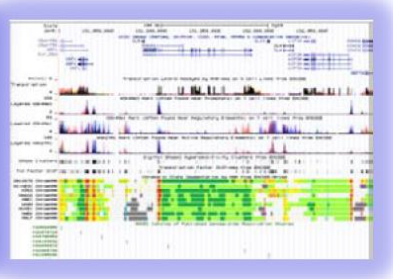
No. at risk				
DPP4 inhibitor	22039	16015	10885	5887
Sulfonylurea	22039	15460	10302	5463

D Ischemic Stroke



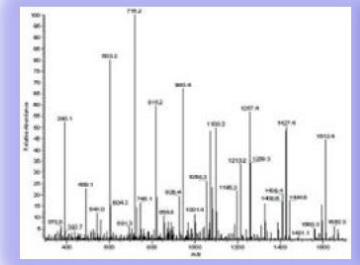
No. at risk				
DPP4 inhibitor	22039	16011	10882	5885
Sulfonylurea	22039	15460	10300	5464

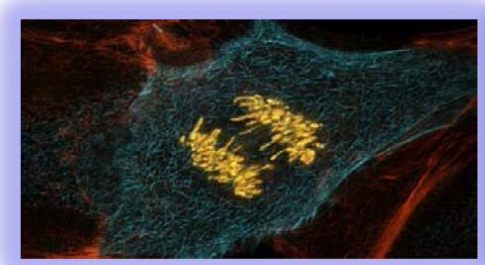
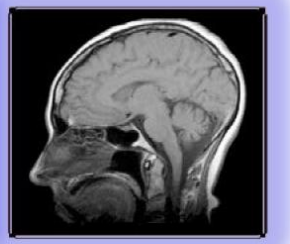
Endless Variety in Big Data

Genomic

Other
'Omics



Imaging

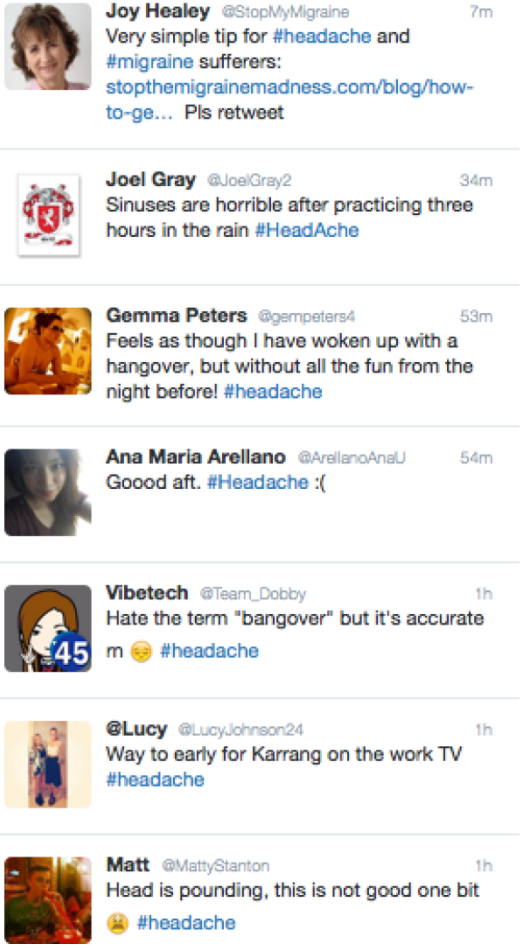
Phenotypic



Exposure

Clinical


Variability and Unstructed in Big Data

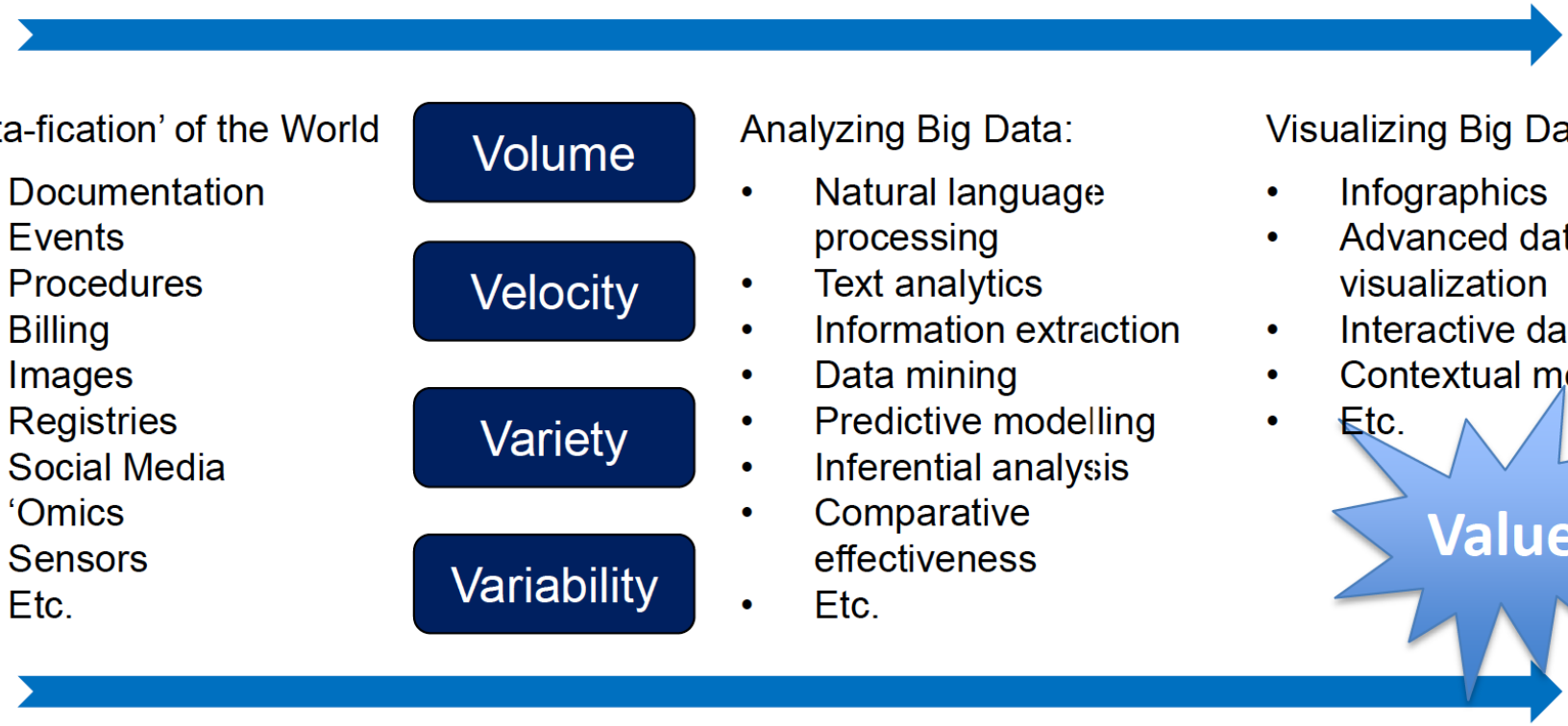


#headache

Huge Variability

- 95% of the world's data is unstructured
  - Text, images, video, voice, etc.
  - Most healthcare data is unstructured
- New data types are emerging
  - Messaging, social media, sensor data

Turning Big Data into Value



In 20 Years…Big Data Era

- All people in developed nations will have —
  - An electronic health record
  - Biological samples
  - Digitized images
- Healthcare will be personalized using an individual's images, samples and clinical data.
- The health of a community will be monitored using aggregate records.