The National Academies of SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Examining the Impact of Real-World Evidence on Medical Product Development: A Workshop Series

Workshop 2:

Practical Approaches

Workshop Briefing Materials

The National Academies of SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Table of Contents

The National Academies of SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Workshop Information

Dates:

March 6-7, 2018

Time:

Day 1: 8:30am – 5:00pm
Day 2: 8:30am – 12:30pm

Location:

National Academy of Sciences Building
Room 120
2101 Constitution Ave NW
Washington, DC 20418

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Forum on
DRUG DISCOVERY, DEVELOPMENT,
and TRANSLATION

Examining the Impact of Real-World Evidence on Medical Product Development: A Three-Part Workshop Series

Washington, D.C.

Background and Objectives

Randomized, controlled clinical trials (RCTs) have traditionally served as the gold standard for evidence generation in support of medical product development and approval. However, it is increasingly recognized that RCTs have inherent limitations, particularly with regard to generalizability, and time and monetary investment. Data from sources supplemental to RCTs, such as safety surveillance, observational studies, registries, claims, or patient-centered outcomes research, would be valuable to support biomedical research, including medical product development and evaluation.

This three-part workshop series will provide a format for examining the practicalities of collection of data from such real-world sources and deriving real-world evidence for the evaluation of medical products, including drugs, biologics, and devices. Each 1.5 day workshop will include presentations and perspectives from thought and knowledge leaders representing a range of disciplines, including but not limited to federal regulatory and funding agencies, clinical and academic medicine and research, medical professional organizations, the regulated biopharmaceutical industry, patients and patient-focused and disease-advocacy organizations, payers, consumer organizations, health systems, and other interested stakeholders that represent the myriad views of those involved in drug, biologic, and device discovery, development, translation, and regulation. The workshop audiences are expected to be similarly diverse, and they will have opportunities to engage in discussion during the workshops. The series will employ case studies to illustrate the current state and to illuminate potential ways forward; staff or invited experts will prepare background papers describing the characteristics of, and gaps in, current data generation efforts. Thought leaders will be invited to react to and build on the papers.

Workshop Topics and Flow

- **Workshop One** (*September 19-20, 2017*) focused on how to align incentives to support collection and use of real-world evidence in health product review, payment, and delivery. Incentives need to address barriers impeding the uptake of real-world evidence, including barriers to transparency.
- **Workshop Two** (*March 6-7, 2018*) will be a "town-hall" style meeting to illuminate what types of data are appropriate for what specific purposes and suggest approaches for data collection and evidence use by developing and working through example use cases.
- **Workshop Three** (*July 17-18, 2018*) will examine and suggest approaches for operationalizing the collection and use of real-world evidence.

Planning Committee

Mark McClellan (Co-Chair), Duke-Margolis Center for Health Policy
Gregory Simon (Co-Chair), Kaiser Permanente Washington Health Research Institute
Jeff Allen, Friends of Cancer Research
Andrew Bindman, UCSF
Adam Haim, NIMH, NIH
Michael Horberg, Kaiser Permanente Mid-Atlantic Medical Group
Petra Kaufmann, NCATS, NIH
Richard Kuntz, Medtronic, Inc.
Elliott Levy, Amgen Inc.
David Madigan, Columbia University
Deven McGraw, Ciitizen
Richard Platt, Harvard Medical School
Patrick Vallance, GlaxoSmithKline
Joanne Waldstreicher, Johnson&Johnson
Marcus Wilson, HealthCore, Inc

The National Academies of Sciences, Engineering, and Medicine are private, nonprofit institutions that provide expert advice on some of the most pressing challenges facing the nation and the world. Our work helps shape sound policies, inform public opinion, and advance the pursuit of science, engineering, and medicine. For more information about these workshops, contact Amanda Wagner Gee (agee@nas.edu).

The National Academies of | SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Examining the Impact of Real-World Evidence on Medical Product Development: A Three-Part Workshop Series

Workshop Two: Practical Approaches

March 6–7, 2018

National Academy of Sciences Building, Room 120
2101 Constitution Ave. NW, Washington, DC 20418

The National Academies of Sciences, Engineering, and Medicine (National Academies) is convening a three-part workshop series examining how real-world evidence development and uptake can enhance medical product development and evaluation. The workshops will advance discussions and common knowledge about complex issues relating to the generation and utilization of real-world evidence, including fostering development and implementation of the science and technology of real-world evidence generation and utilization.

- Workshop One (*September 19-20, 2017*) focused on how to align incentives to support collection and use of real-world evidence in health product review, payment, and delivery. Incentives need to address barriers impeding the uptake of real-world evidence, including barriers to transparency.
- **Workshop Two (*March 6-7, 2018*) will illuminate what types of data are appropriate for what specific purposes and suggest practical approaches for data collection and evidence use by developing and working through example use cases.**
- Workshop Three (*July 17-18, 2018*) will examine and suggest approaches for operationalizing the collection and use of real-world evidence.

DAY 1: March 6, 2018

8:30 a.m. Breakfast Available Outside the Room 120

8:40 a.m. **Welcome and Opening Remarks**

GREG SIMON, *Workshop Series Co-Chair*
Investigator
Kaiser Permanente Washington Health Research Institute

SESSION I: WHEN CAN WE RELY ON REAL-WORLD DATA?

Session discussion questions:

- When can we have confidence in EHR data from real-world practice to accurately assess study eligibility, key prognostic factors, and study outcomes?
- When can we have confidence in data generated outside of clinical settings (e.g., mobile phones, connected glucometers, connected blood pressure monitors)?
- When does adjudication or other post-processing of real-world data add value?

Moderator: Greg Daniel, Duke-Margolis Center for Health Policy

Session Discussants

JESSE BERLIN
Vice President and Global Head, Epidemiology
Johnson & Johnson

ANDY BINDMAN
Professor of Medicine
University of California San Francisco

ADRIAN HERNANDEZ
Vice Dean for Clinical Research
Duke University School of Medicine

9:00 a.m. **Introduction and background to inform the discussion: Novel oral anticoagulants in comparison with warfarin**

9:20 a.m. **Open discussion with audience**

- What questions can characterize the utility of any real-world data source and signal reliability before a study is performed (examples below)?
    - When is accuracy good enough to reasonably and consistently identify the right population?
    - When is accuracy good enough to reasonably and consistently assess the exposure or intervention?
    - When is accuracy good enough to reasonably and consistently assess the right outcome?
    - Are there any big safety issues that would be missed?
    - Are there concerns about data collection or entry, particularly in relation to creating systemic bias?
    - When is expert adjudication necessary to confirm that the recorded data is reliable and/or reasonably complete?
- What information is needed to answer such questions?

10:40 a.m. **BREAK**
(*Workgroup participants gather to synthesize audience feedback*)

11:00 a.m. **Workgroup presents synthesis of audience feedback**

SESSION II: WHEN CAN WE RELY ON REAL-WORLD TREATMENT?

Session discussion questions:

- When conducting research in a real-world setting, are there situations that would require special guidance, knowledge, or experience in order for clinicians to adequately monitor participant safety and respond appropriately to adverse events?
- When does variation between comparison groups (socioeconomic, demographic, etc.); in treatment fidelity; in provider behavior and preferences; or in adherence yield a valid signal about real-world effectiveness, and when is it just noise?

Moderator: Khaled Sarsour, Genetech | A Member of the Roche Group

Session Discussants

MICHAEL HORBERG
Executive Director, Research, Community Benefit, and Medicaid Strategy
Executive Director, Mid-Atlantic Permanente Research Institute
Kaiser Permanente Mid-Atlantic Permanente Medical Group

GREG SIMON
Investigator
Kaiser Permanente Washington Health Research Institute

ROBERT TEMPLE
Deputy Director for Clinical Science
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

11:15 a.m. **Introduction and presentation to inform discussion on participant monitoring: study on Lithium for Suicidal Behavior in Mood Disorders**

IRA KATZ
Psychiatrist
Corporal Michael J. Crescenz VA Medical Center, Philadelphia

11:35 p.m. **Open discussion with audience**

- What conditions make self-monitoring and reporting acceptable?
- Does this vary for treatments at different stages of product development or with different baseline knowledge about use in varied patient types and treatment conditions?
- Can we draw any generalizable lessons about cases in which self-monitoring is acceptable and safe?

12:15 p.m. **Introduction and presentation to inform discussion on signal detection: Novel Oral Anticoagulants in comparison with warfarin**

12:30 p.m. **Open discussion with audience**

- What conditions and training prepare clinical care providers to monitor patient safety outside a tightly controlled environment?
- How does this vary for treatments at different stages of product development or with different baseline knowledge about use in varied patient types and treatment conditions?
- How do you decide which variables require strict adherence to "protocol" and which can be allowed to vary?

1:00 p.m. **BREAK** (Lunch available Outside Room 120)
(*Workgroup participants gather to synthesize audience feedback*)

2:00 p.m. **Workgroup presents synthesis of audience feedback**

SESSION III: WHEN CAN WE LEARN FROM REAL-WORLD TREATMENT ASSIGNMENT?

Session discussion questions:

- When can we have confidence in inference from cluster-randomized or stepped-wedge study designs?
- Under what conditions can we trust inference from observational or naturalistic comparisons?
- How could we judge the validity of observational comparisons in advance, rather than waiting until we've observed the result?

Moderator: Richard Platt, Harvard Medical School

Session Discussants

Rob Califf
Vice Chancellor, Health Data Science, Duke University
Verily Life Sciences

DAVID MADIGAN
Professor of Statistics
Dean, Faculty of Arts and Sciences
Columbia University

DAVID MARTIN
Associate Director for Real-World Evidence Analytics
U.S Food and Drug Administration

2:20 p.m. **Introduction and presentation to inform the discussion: Healthcare Database Analyses of Medical Products for Regulatory Decision Making**

SEBASTIAN SCHNEEWEISS
Professor of Medicine and Epidemiology
Harvard Medical School
Brigham & Women's Hospital

2:50 p.m. **Open discussion with audience**

- Random assignment is always preferable, but when is the cost (in time, money, infrastructure, patient exposure) truly necessary?
- How can we know that the effects from unmeasured confounders are not so large that they would change a decision based on information from an observational study?
- What are some of conditions under which there is more confidence in inference from non-randomized comparisons (*examples of some conditions below*)?
    - Expectation of large effects
    - Outcome proximal to treatment
    - High degree of similarity between comparison groups
    - Pathway from treatment to outcome is relatively clear, and without lots of complexity or reciprocal effects
    - Treatment allocation method is relatively transparent

3:40 p.m. **BREAK**

4:00 p.m. **Open discussion with audience and reflections on the discussion from panelists**

5:00 p.m. **ADJOURN WORKSHOP DAY 1**

DAY 2: MARCH 7, 2018

8:30 a.m. Breakfast Available Outside the Room 120

SESSION IV: SYNTHESIZING THE USE CASES

Session Objective:

- Discuss key considerations presented in each session on Day 1
- Consider components of a potential "checklist" for using real-world evidence

9:00 a.m. **Welcome and recap of Day 1**

GREG SIMON, *Workshop Series Co-Chair*
Investigator
Kaiser Permanente Washington Health Research Institute

MARK MCCLELLAN, *Workshop Series Co-Chair*
Director
Duke-Margolis Center for Health Policy

9:20 a.m. **Open discussion with audience of outputs from Day 1 and potential components to a "checklist" for using RWE**

10:40 a.m. **BREAK**

11:00 a.m. **Open discussion with audience of outputs from Day 1 and potential components to a "checklist" for using RWE**

12:30 p.m. **ADJOURN WORKSHOP DAY 2**

Future Workshop Objectives

WORKSHOP THREE. Examine and suggest approaches for operationalizing the collection and use of real-world evidence. *(July 17-18, 2018, Washington, DC)*

- Applications for using real-world evidence to supplement traditional clinical trials, pragmatic/effectiveness trials, or routine clinical application.
- Mechanisms for determining which discrete types of real-world evidence could support regulatory decisions.
- Operational challenges and barriers for generating and incorporating real-world evidence in the context of a learning health system and how clinicians can best be involved in the collection and utilization of real-world evidence.

The National Academies of SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Examining the Impact of Real-World Evidence on Medical Product Development: A Workshop Series

Workshop 2: Practical Approaches

PLANNING COMMITTEE BIOGRAPHIES

CO-CHAIRS:

MARK MCCLELLAN, M.D., Ph.D., is the Robert J. Margolis Professor of Business, Medicine, and Policy, and Director of the Duke-Margolis Center for Health Policy at Duke University with offices at Duke and in Washington DC. Dr. McClellan is a doctor and an economist, and his work has addressed a wide range of strategies and policy reforms to improve health care, including payment reforms to promote better outcomes and lower costs, methods for development and use of real-world evidence, and approaches for more effective drug and device innovation. Dr. McClellan is a former administrator of the Centers for Medicare & Medicaid Services (CMS) and former commissioner of the U.S. Food and Drug Administration (FDA), where he developed and implemented major reforms in health policy. He was also a Senior Fellow at the Brookings Institution and a professor of economics and medicine at Stanford University.

GREGORY SIMON, M.D., M.P.H., is an investigator at Group Health Research Institute and a psychiatrist in Group Health's Behavioral Health Service. He is also a Research Professor in the Department of Psychiatry and Behavioral Sciences at the University of Washington and chair of the national scientific advisory board of the Depression and Bipolar Support Alliance. Dr. Simon completed residency training in internal medicine at the University of Washington, residency training in psychiatry at the Massachusetts General Hospital, and fellowship training in the Robert Wood Johnson Clinical Scholars program at the University of Washington.
Dr. Simon's research focuses on improving access to and quality of care for mood disorders, both unipolar depression and bipolar disorder. Specific areas of research include improving adherence to medication, increasing the availability of effective psychotherapy, evaluating peer support by and for people with mood disorders, suicide prevention, cost-effectiveness of treatment, and comorbidity of mood disorders with chronic medical conditions. Dr. Simon currently leads the Mental Health Research Network, an NIMH-funded cooperative agreement supporting population-based mental health research across 13 large health systems.

PLANNING COMMITTEE:

JEFF ALLEN, Ph.D., is President and CEO of Friends of Cancer Research (Friends). Friends is an advocacy organization based in Washington, DC that drives collaboration among partners from every healthcare sector to power advances in science, policy, and regulation that speed life-saving treatments to patients. During the past 20 years, Friends has been instrumental in the creation and implementation of policies ensuring patients receive the best treatments in the fastest and safest way possible. For over 10 years, Jeff has been a driving force in the growth and success of the organization. Under his leadership, Friends has evolved into a nimble, forward-thinking policy, public affairs, and research organization. As President and CEO, he leads the strategic development and implementation of Friends' scientific, policy, research, and legislative initiatives, as well as overseeing Board governance and organizational operations.

As a thought leader on many issues related to the Food and Drug Administration, regulatory strategy, and healthcare policy, he is regularly published in prestigious medical journals and policy publications. In addition to participating in major scientific and policy symposiums around the country each year, Jeff has had the honor to be called to testify before Congress on multiple occasions and regularly contributes his expertise to the legislative process. Recent Friends initiatives include the establishment of the Breakthrough Therapies designation and the development of the Lung Cancer Master Protocol (Lung-MAP), a unique partnership that will accelerate and optimize clinical trial conduct for new drugs. Jeff has the privilege to also serve on a

variety of influential committees, boards, and advisory councils including the Alliance for a Stronger FDA (Board Member, Past President), the Medical Evidence Development Consortium (MED-C; Board Chair), Lung-MAP Senior Leadership Team Member, and a participant on working groups convened by the National Academies of Medicine and President's Council of Advisors on Science and Technology (PCAST). Prior to joining Friends, Jeff was an endocrinology researcher at the National Institutes of Health. His background in cancer research focused upon molecular changes associated with cancer formation as well as treatments to prevent cancer progression. Jeff received his Ph.D. in cell and molecular biology from Georgetown University, and holds a Bachelor's of Science in Biology from Bowling Green State University.

ANDREW BINDMAN, M.D., was appointed as Director of AHRQ on May 2, 2016. Prior to his appointment, Dr. Bindman served as Professor of Medicine and Epidemiology & Biostatistics at the University of California, San Francisco (UCSF). He is a primary care physician with Federal and State health policy experience who has practiced, taught, and conducted health services research at San Francisco General Hospital, an urban safety-net hospital, for almost 30 years. During that time, he led the development of a nationally recognized academic division focused on improving the care of vulnerable populations and a State-university partnership with California's Medicaid program that promotes translating research into policy. Dr. Bindman has published more than 150 peer-reviewed scientific articles focused on primary care and on low-income individuals' access to and quality of care. Through his work, Dr. Bindman helped to establish the association between access to care and preventable hospitalizations for ambulatory care-sensitive conditions (what are now called Prevention Quality Indicators [PQIs]). Dr. Bindman has used PQIs to evaluate Medicaid programs and to design interventions to improve quality of care for low-income patients with chronic disease. He has also promoted a participatory research model with policymakers as a way to translate research into evidence-based policy. Dr. Bindman is a Senior Associate Editor of the journal Health Services Research and he was elected to the National Academy of Medicine in 2015.

At UCSF, Dr. Bindman contributed to the training of primary care physicians and the development of health services researchers. He has been the Director of UCSF's Primary Care Research Fellowship, the developer of a course on translating research into policy, and a co-editor of the textbook Medical Management of Vulnerable and Underserved Populations. In 2005, Dr. Bindman received an achievement award from the Health Resources and Services Administration in recognition of his contributions to research training in health care disparities and in improving the diversity of the Nation's health care workforce. He served on AHRQ's Health Care Research Training Study Section from 2005-2009. In 2009-2010, Dr. Bindman was a Robert Wood Johnson Health Policy Fellow who worked as a staff member on the Energy and Commerce Committee in the U.S. House of Representatives. From September 2011 until June 30, 2014, Dr. Bindman served as a senior advisor within the Assistant Secretary for Planning and Evaluation's Office of Health Policy, where he worked to establish new Medicare payment codes for transitional care and chronic care management. From July 2014 until November 2015, Dr. Bindman was a senior advisor to the Centers for Medicare & Medicaid Services, where he helped launch the Innovation Accelerator Program to support care transformation in State Medicaid programs.

ADAM HAIM, Ph.D., is the Chief of the Treatment and Preventive Intervention Research Branch within the Division of Services and Intervention Research at the National Institute of Mental Health (NIMH). Dr. Haim manages a broad portfolio of research focused on evaluating the efficacy and effectiveness of pharmacologic, psychosocial and combination interventions on mental and behavior disorders. He is also a thought leader in the development, evaluation and implementation of technology enhanced mental health interventions. Dr. Haim is a licensed clinical psychologist and earned his doctoral degree in clinical psychology from State University of New York at Albany and completed his research fellowship at the NIMH Intramural Program in the Division of Clinical Neuroendocrinology.

MICHAEL HORBERG, M.D., M.A.S, FACP, is Executive Director Research and Community Benefit of Mid-Atlantic Permanente Medical Group (MAPMG) and the director of the Mid-Atlantic Permanente Research Institute (MAPRI). He is also director of HIV/AIDS for Kaiser Permanente. Dr. Horberg has been appointed to serve on the Presidential Advisory Council on HIV/AIDS (PACHA), and co-chairs the Access to Care and Improved Outcomes Committee of PACHA. Dr. Horberg is a Fellow of the American College of Physicians, and he presently serves as Vice-Chair of the Board of Directors of the HIV Medicine Association of the Infectious Disease Society of America. He has co-chaired the NCQA/AMA/HRSA/IDSA sponsored Expert Panel on HIV-

related provider performance measures. He is Assistant Clinical Professor at Stanford University Medical School. Dr. Horberg is past-president of the national Gay and Lesbian Medical Association. His HIV research interests are health service outcomes for HIV-infected patients (including HIV quality measures and care improvement, and determinants of optimized multidisciplinary care for maximized HIV outcomes), medication adherence issues in these patients, and epidemiology of the disease. He graduated from Boston University's College of Liberal Arts and School of Medicine (with honors of Summa cum Laude and Phi Beta Kappa) and completed his internal medicine residency at Michael Reese Hospital in Chicago (University of Chicago affiliate). He received his Master of Advanced Studies (Clinical Research) from University of California San Francisco.

PETRA KAUFMANN, M.D., M.SC., is the director of both the Office of Rare Diseases Research and the Division of Clinical Innovation. Her work includes overseeing NCATS' Rare Diseases Clinical Research Network, Genetic and Rare Diseases Information Center, and Clinical and Translational Science Awards Program as well as the NIH/NCATS Global Rare Diseases Patient Registry Data Repository/GRDR® program. Kaufmann focuses on engaging a broad range of stakeholders to accelerate translation from discovery to health benefits through use of innovative methods and tools in translational research and training. Before joining NCATS, Kaufmann was the director of the Office of Clinical Research at the National Institute of Neurological Disorders and Stroke (NINDS), where she worked with investigators to plan and execute a large portfolio of clinical research studies and trials in neurological disorders, including many in rare diseases. She established NeuroNEXT, a trial network for Phase II trials using a central institutional review board, streamlined contracting, active patient participation in all project phases, and a scientific and legal framework for partnership with industry. Kaufmann also promoted data sharing, working with multiple stakeholders from the academic, patient organization and industry sectors to develop data standards for more than 10 neurological diseases.

A native of Germany, Kaufmann earned her M.D. from the University of Bonn and her M.Sc. in biostatistics from Columbia University's Mailman School of Public Health. She completed an internship in medicine at St. Luke's/Roosevelt (now part of Mt. Sinai) in New York City, training in neurology and clinical neurophysiology at Columbia University, and a postdoctoral fellowship in the molecular biology of mitochondrial diseases at Columbia's H. Houston Merritt Clinical Research Center for Muscular Dystrophy and Related Diseases. Before joining NINDS, Kaufmann was a tenured associate professor of neurology at Columbia, where she worked as a researcher and clinician in the neuromuscular division, the electromyography laboratories and the pediatric neuromuscular clinic. She has served on scientific advisory committees for many rare disease organizations and is a member of the American Academy of Neurology Science Committee, the International Rare Disease Research Consortium Interdisciplinary Scientific Committee and the Clinical Trial Transformation Initiative Steering Committee. Kaufmann is board certified in neurology, neuromuscular medicine and electrodiagnostic medicine. Kaufmann's research focus is on the clinical investigation of rare diseases, such as spinal muscular atrophy, amyotrophic lateral sclerosis and mitochondrial diseases. She currently sees patients in the Muscular Dystrophy Association Clinic at Children's National Medical Center in Washington, D.C.

RICHARD KUNTZ, M.D., is Senior Vice President and Chief Scientific, Clinical and Regulatory Officer of Medtronic and serves as a member of the Company's Executive Committee. In this role, which he assumed in August 2009, Kuntz oversees the company's global regulatory affairs, health policy and reimbursement, clinical research activities, and corporate technology. Kuntz joined Medtronic in October 2005, as Senior Vice President and President of Medtronic Neuromodulation, which encompasses the company's products and therapies used in the treatment of chronic pain, movement disorders, spasticity, overactive bladder and urinary retention, benign prostatic hyperplasia, and gastroparesis. In this role he was responsible for the research, development, operations and product sales and marketing for each of these therapeutic areas worldwide.Kuntz brings to Medtronic a broad background and expertise in many different areas of healthcare. Prior to Medtronic he was the Founder and Chief Scientific Officer of the Harvard Clinical Research Institute (HCRI), a university-based contract research organization which coordinates National Institutes of Health (NIH) and industry clinical trials with the United States Food and Drug Administration (FDA). Kuntz has directed over 100 multicenter clinical trials and has authored more than 250 original publications. His major interests are traditional and alternative clinical trial design and biostatistics.

Kuntz also served as Associate Professor of Medicine at Harvard Medical School, Chief of the Division of Clinical Biometrics, and an interventional cardiologist in the division of cardiovascular diseases at the

Brigham and Women's Hospital in Boston, MA. Kuntz has served as a member of the Board of Governors of PCORI (Patient Centered Outcomes Research Institute) since it was established in 2010 as part of the Affordable Care Act.Kuntz graduated from Miami University, and received his medical degree from Case Western Reserve University School of Medicine. He completed his residency and chief residency in internal medicine at the University of Texas Southwestern Medical School, and then completed fellowships in cardiovascular diseases and interventional cardiology at the Beth Israel Hospital and Harvard Medical School, Boston. Kuntz received his master's of science in biostatistics from the Harvard School of Public Health.

ELLIOTT LEVY, M.D., is senior vice president, Global Development, at Amgen. He is responsible for the clinical development of Amgen's investigative and marketed products. Before joining Amgen, Dr. Levy spent 17 years at Bristol-Myers Squibb (BMS) in clinical development and pharmacovigilance. He has contributed to the development and approval of numerous new therapies for cardiovascular, metabolic, inflammatory, and malignant diseases, and led large organizations through periods of transformative change. Dr. Levy is a graduate of the Yale School of Medicine, where he also trained in internal medicine and nephrology. Dr. Levy was also a member of the Renal Division at Brigham and Women's Hospital in Boston, Massachusetts, where he was an investigator in federally sponsored outcomes research as well as industry-sponsored clinical trials.

DAVID MADIGAN, Ph.D., received a bachelor's degree in Mathematical Sciences and a Ph.D. in Statistics, both from Trinity College Dublin. He has previously worked for AT&T Inc., Soliloquy Inc., the University of Washington, Rutgers University, and SkillSoft, Inc. He has over 100 publications in such areas as Bayesian statistics, text mining, Monte Carlo methods, pharmacovigilance and probabilistic graphical models. He is an elected Fellow of the American Statistical Association and of the Institute of Mathematical Statistics. He recently completed a term as Editor-in-Chief of Statistical Science.

DEVEN MCGRAW, J.D., M.P.H, currently serves as the Chief Regulatory Officer for Ciitizen, a position she assumed in December 2017. Prior to her current position, Ms. McGraw served as the Deputy Director for Health Information Privacy in the HHS Office of Civil Rights, a position she held from 2015 to 2017. She is a well-respected expert on the HIPAA Rules and has a wealth of experience in both the private sector and the non-profit advocacy world. Prior to her position at OCR, she was a partner in the healthcare practice of Manatt, Phelps & Phillips, LLP. She previously served as the Director of the Health Privacy Project at the Center for Democracy & Technology, which is a leading consumer voice on health privacy and security policy issues, and as the Chief Operating Officer at the National Partnership for Women & Families, where she provided strategic leadership and substantive policy expertise for the Partnership's health policy agenda. Ms. McGraw graduated magna cum laude from the University of Maryland. She earned her J.D., magna cum laude, and her L.L.M. from Georgetown University Law Center and was Executive Editor of the Georgetown Law Journal. She has a Master of Public Health from Johns Hopkins School of Hygiene and Public Health.

RICHARD PLATT, M.D., M.S., is Professor and Chair of the Department of Population Medicine at Harvard Pilgrim Health Care Institute. He has extensive experience in developing systems and capabilities for using routinely collected electronic health information to support public health surveillance, medical product safety assessments, comparative effectiveness and outcomes research, and quality improvement programs. Dr. Platt is Principal Investigator of the FDA Sentinel System. He co-leads the coordinating center of PCORI's National Patient Centered Clinical Research Network (PCORnet). He also co-leads the coordinating center of the NIH Health Care System Research Collaboratory, and he leads a CDC Prevention Epicenter. He is a member of the Institute of Medicine Roundtable on Value and Science Driven Healthcare and the Association of American Medical Colleges Advisory Panel on Research. Dr. Platt is a graduate of Harvard Medical School and the Harvard School of Public Health. He is clinically trained in internal medicine and infectious diseases.

PATRICK VALLANCE, M.D., is President of Research and Development at GlaxoSmithKline (GSK) and a member of the GSK Corporate Executive Team. Prior to joining GSK in 2006, Patrick Vallance was a clinical academic and as Professor of Medicine led the Division of Medicine at University College London and Consultant Physician at UCL. His academic work was in the field of cardiovascular biology and ranged from chemistry through to use of large electronic health records. Patrick Vallance is a Fellow of the Academy of Medical

Sciences. He has been on the Board of the UK Office for Strategic Co-ordination of Health Research since 2009. He is also a director of Genome Research Limited.

JOANNE WALDSTREICHER, M.D., is Chief Medical Officer, Johnson & Johnson. In this role, she has oversight across pharmaceuticals, devices and consumer products for safety, epidemiology, clinical and regulatory operations transformation, internal and external partnerships and collaborations supporting development of the ethical science, technology and R&D policies, including those related to clinical trial transparency and compassionate access. Joanne also chairs the Pharmaceuticals (Janssen) R&D Development Committee and supports the Device and Consumer Development Committees, which review late stage development programs in the pipeline. She also holds an appointment as a Faculty Affiliate of the Division of Medical Ethics, Department of Population Health, New York University School of Medicine.

Among her prior roles in Janssen, the pharmaceutical sector of Johnson & Johnson, Joanne was responsible for late-stage development spanning the areas of neuroscience, cardiovascular and metabolism including INVOKANA®, XARELTO®, INVEGA SUSTENNA®, and INVEGA TRINZA®. Before joining Johnson & Johnson in 2002, Joanne was head of the Endocrinology and Metabolism clinical research group at Merck Research Laboratories, and responsible for overseeing clinical development of MEVACOR®, ZOCOR®, PROSCAR® and PROPECIA®, and for clinical development programs in atherosclerosis, obesity, diabetes, urology and dermatology. During that time, she received numerous awards and distinctions, including the Merck Research Laboratory Key Innovator Award.

Joanne received both the Jonas Salk and Belle Zeller scholarships from the City University of New York and graduated Summa Cum Laude from Brooklyn College. Joanne graduated Cum Laude from Harvard Medical School in 1987, and completed her internship and residency at Beth Israel Hospital, and her endocrinology fellowship at MGH. She has won numerous awards and scholarships, and has authored numerous papers and abstracts. In December, 2016, Joanne was named Healthcare Champion of the Year for Women by the National Association of Female Executives. She has authored numerous papers and abstracts. In December, 2016, Joanne was named Healthcare Champion of the Year for Women by the National Association of Female Executives.

MARCUS WILSON, PHARM.D., is President of HealthCore, Anthem's wholly owned outcomes research subsidiary. HealthCore utilizes Anthem's rich data and extensive provider network to meet the evidence development needs of a broad array of healthcare stakeholders including Anthem, State and Federal agencies and the Life Sciences industry. He has been extensively involved in efforts to utilize electronic healthcare data environments to accelerate healthcare evidence development and to facilitate clinical decision support for more than 20 years. Prior to co-founding HealthCore in 1996, Dr Wilson spent seven years within an integrated delivery system owned by BCBS of Delaware where he oversaw the physician and patient clinical decision support, pharmacy policy and clinical trials programs.

Dr. Wilson is active on a number of boards and national committees including serving as chair of the Innovations in Medical Evidence Development (IMEDS) Steering Committee, Reagan-Udall Foundation for the FDA; chair of the Research Committee for the Academy of Managed Care Pharmacy (AMCP) & the AMCP Foundation; the FDA Mini-Sentinel Program's Project Operations Council; Board of Directors for the Center for Medical Technology Policy (CMTP); and the Dean's Roundtable, College of Science, Virginia Tech. He is a past member of the Board of Directors for the International Society of Pharmacoeconomics and Outcomes Research (ISPOR) and is a reviewer for multiple journals. Dr. Wilson received his Bachelor of Science in Biochemistry from Virginia Tech and his Doctor of Pharmacy degree from the Medical College of Virginia. He completed a residency in Family Medicine at the Medical University of South Carolina prior to joining the faculty at the Philadelphia College of Pharmacy where he taught didactic and experiential courses in therapeutics and clinical decision support. His experiential site within the HMO of Delaware eventually served as the foundational assets for HealthCore. HealthCore was eventually acquired by WellPoint in 2003 just prior to the WellPoint-Anthem merger.

The National Academies of SCIENCES ENGINEERING MEDICINE

Forum on DRUG DISCOVERY, DEVELOPMENT, and TRANSLATION

Examining the Impact of Real-World Evidence on Medical Product Development: A Workshop Series

Workshop 2: Practical Approaches

SPEAKER AND DISCUSSANT BIOGRAPHIES

JESSE BERLIN, SC.D., is Vice President and Global Health of Epidemiology with Johnson & Johnson. Dr. Berlin received his doctorate in biostatistics from the Harvard School of Public Health in 1988. After spending 15 years as a faculty member at the University of Pennsylvania, in the Center for Clinical Epidemiology and Biostatistics, under the direction of Dr. Brian Strom, Jesse left Penn to join Janssen Research & Development as a Senior Director in Biostatistics. After two years, he was promoted to Vice President for Epidemiology. He now serves in his current position across all of Johnson & Johnson, with responsibility for pharmaceuticals, devices and consumer products. He has authored or coauthored over 230 publications in a wide variety of clinical and methodological areas, including papers on the study of meta-analytic methods as applied to both randomized trials and epidemiology. He served on an Institute of Medicine Committee that developed recommendations for the use of systematic reviews in clinical effectiveness research, and served on the Scientific Advisory Committee to the Observational Medical Outcomes Partnership, a public-private partnership aimed at understanding methodology for assessing drug safety in large, administrative databases. He served as a member of working group X for CIOMS (The Council for International Organizations of Medical Sciences), which has published guidelines for meta-analysis of drug safety data in the regulatory context. He was elected as a fellow of the American Statistical Association in 2004. In 2013, Dr. Berlin received the Lagakos Distinguished Alumni Award from the Department of Biostatistics at the Harvard School of Public Health.

ROBERT M. CALIFF, M.D., MACC, is the Vice Chancellor for Health Data Science, the Donald F. Fortin Professor of Cardiology, and Professor of Medicine at the Duke University School of Medicine. He was the Commissioner of Food and Drugs in 2016-2017 and Deputy Commissioner for Medical Products and Tobacco from February 2015 until his appointment as Commissioner in February 2016. Prior to joining the FDA, Dr. Califf was a professor of medicine and vice chancellor for clinical and translational research at Duke University. He also served as director of the Duke Translational Medicine Institute and founding director of the Duke Clinical Research Institute. A nationally and internationally recognized expert in cardiovascular medicine, health outcomes research, healthcare quality, and clinical research, Dr. Califf has led many landmark clinical trials and is one of the most frequently cited authors in biomedical science, with more than 1,200 publications in the peer-reviewed literature. Dr. Califf is a Member of the National Academy of Medicine (formerly known as the Institute of Medicine [IOM]) as of 2016, one of the highest honors in the fields of health and medicine. Dr. Califf has served on numerous IOM committees, and he has served as a member of the FDA Cardiorenal Advisory Panel and FDA Science Board's Subcommittee on Science and Technology. Dr. Califf has also served on the Board of Scientific Counselors for the National Library of Medicine, as well as on advisory committees for the National Cancer Institute, the National Heart, Lung, and Blood Institute, the National Institute of Environmental Health Sciences and the Council of the National Institute on Aging. He has led major initiatives aimed at improving methods and infrastructure for clinical research, including the Clinical Trials Transformation Initiative (CTTI), a public-private partnership co-founded by the FDA and Duke. He also served as the principal investigator for Duke's Clinical and Translational Science Award and the NIH Health Care Systems Research Collaboratory coordinating center and co-PI of the Patient Centered Outcomes Research Institute Network. Dr. Califf is a graduate of Duke University School of Medicine. He completed a residency in internal medicine at the University of California, San Francisco and a fellowship in cardiology at Duke.

GREGORY DANIEL, PHD, MPH is the Deputy Director of the Duke-Robert J. Margolis, MD Center for Health Policy and a Clinical Professor in Duke's Fuqua School of Business. Dr. Daniel directs the DC-based office of the Center and leads the Center's pharmaceutical and medical device policy portfolio, which includes developing policy and data strategies for improving development and access to innovative pharmaceutical and medical device technologies. This includes post-market evidence development to support increased value, improving regulatory science and drug development tools, optimizing biomedical innovation, and supporting drug and device value-based payment reform. Dr. Daniel is also Adjunct Associate Professor in the Division of Pharmaceutical Outcomes and Policy at the UNC Eshelman School of Pharmacy. Previously, he was Managing Director for Evidence Development & Biomedical Innovation in the Center for Health Policy and Fellow in Economic Studies at the Brookings Institution and Vice President, Government and Academic Research at HealthCore (an Anthem, Inc. company). In addition to health and pharmaceutical policy, Dr. Daniel's research expertise includes real world evidence (RWE) development utilizing electronic health data in the areas of health outcomes and pharmacoeconomics, comparative effectiveness, and drug safety and pharmacoepidemiology. Dr. Daniel received a PhD in pharmaceutical economics, policy and outcomes from the University of Arizona, as well as an MPH, MS, and BS in Pharmacy all from The Ohio State University.

ADRIAN HERNANDEZ, M.D., was named Vice Dean for Clinical Research for the Duke University School of Medicine on September 5, 2017. As Vice Dean for Clinical Research, Dr. Hernandez has direct responsibility for advancing the clinical research mission of the School of Medicine. He works with leaders in single and multi-site based human research, patient care delivery, information technology, and health data science within School of Medicine departments, and centers and institutes including the Duke Clinical Research Institute (DCRI), Clinical and Translational Science Institute, and Margolis Center for Health Policy, in order to achieve the vision of advancing health and executing a coordinated strategy in clinical research to evolve the model of care and improve outcomes. Dr. Hernandez oversees the Institutional Review Board (IRB) for Duke Health, the Duke Office of Clinical Research, the Office of Regulatory Affairs & Quality, and the Research Integrity Office. A Professor of Medicine in the Division of Cardiology, Dr. Hernandez previously served as Director of Health Services and Outcomes Research and was a Faculty Associate Director of the DCRI. He has extensive experience in clinical research ranging from clinical trials to outcomes and health services research. He leads research programs focused on understanding population health, generating real-world evidence, and improving patient-centered outcomes through development of new therapies and better care delivery in the national health system. Dr. Hernandez is the coordinating center principal investigator for multiple networks and clinical trials such as the National Heart, Lung, and Blood Institute's Heart Failure Clinical Research Network, PCORI's National Patient-Centered Clinical Research Network (PCORnet) and the NIH's Health System Collaboratory. A central aim of these networks is to transform clinical research by uniting patients, clinicians, health systems, and electronic health data to improve population health and decision making. Dr. Hernandez has over 450 published articles in top peer-reviewed journals including the New England Journal of Medicine, Journal of the American Medical Association, and Lancet. Dr. Hernandez received his medical degree from the University of Texas-Southwestern at Dallas and completed his residency in internal medicine at the University of California San Francisco School of Medicine, before completing a fellowship in cardiology at Duke University. In 2004, Dr. Hernandez joined the Duke faculty as an assistant professor.

IRA KATZ, M.D., Ph.D., is Senior Consultant for Program Evaluation in the Department of Veterans Affairs Office of Mental Health and Suicide Prevention. He worked to develop VA's predictive model using clinical and administrative data to identify which Veterans were at the greatest risk for suicide and related outcomes, and to translate prediction into preventive interventions. He also serves as lead investigator for a large scale VA cooperative study evaluating the effectiveness and safety of lithium for suicide prevention. Before he assumed his current positive, he served as VA's Deputy Chief Patient Care Services Officer for Mental Health, and, before that, as Professor of Psychiatry at the University of Pennsylvania, Director of the Division of Geriatric Psychiatry at Penn and the Philadelphia VA, and as principal investigator of both VA- and NIMH- supported research centers.

DAVID MARTIN, M.D., M.P.H., is the Associate Director for Real-World Evidence Analytics with the Office of Medical Policy, Center for Drug Evaluation and Research, Food and Drug Administration. Previously, he was assigned to FDA/CDER as the FDA Liaison to the Reagan Udall Foundation Innovation in Medical Evidence Development and Surveillance program. IMEDS enables routine private-sector queries related to medical products using a distributed database approach modeled on the FDA's Sentinel system. Dr. Martin is also the principal investigator for a Patient Centered Outcomes Trust Fund project that is incorporating patient data collected through a mobile device application into Sentinel and PCORnet. As a former Branch Chief and Division Director in the FDA Center for Biologics Evaluation and Research, he led analyses of spontaneous reports, formalized risk management planning, and played a key role in the development of the Sentinel system. Before joining the FDA, Dr. Martin served in the U.S. Air Force as a flight and occupational medicine physician. He received his bachelor's degree at the Citadel, his medical degree at the Johns Hopkins University School of Medicine, and his master of public health degree at the Johns Hopkins University Bloomberg School of Public Health.

KHALED SARSOUR, Ph.D., is a principal scientist at the global real world data and personalized health care group at Genentech, Inc. He joined Genentech in 2012 where he leads a team of scientists delivering real-world evidence and insights across the phases of drug development in immunology and infectious diseases. Before joining Genentech, Dr. Sarsour spent 5 years at Eli Lilly and Company in the global patient outcomes and real world evidence group. Dr. Sarsour has coauthored more than 30 peer-reviewed publications and many more conference abstracts in a range of therapeutic areas investigating the safety and real-world effectiveness of treatments. He holds a doctorate in epidemiology from the school of public health at the University of California, Berkeley.

SEBASTIAN SCHNEEWEISS, M.D., Sc.D., is Professor of Medicine and Epidemiology at Harvard Medical School and Vice Chief of the Division of Pharmacoepidemiology and Pharmacoeconomics of the Department of Medicine, Brigham and Women's Hospital. His research is funded by multiple NIH, PCORI, and FDA grants and focuses on the comparative effectiveness and safety of biopharmaceuticals. He has developed analytic methods to improve the validity of epidemiologic studies using complex healthcare databases particularly for newly marketed medical products. Bundling such improved methods into a rapid-cycle analytics framework is the overarching theme of his research. His work is published in more than 300 articles, many of them in high-ranking journals. He is Aetion Inc.'s Science Lead of an evidence platform that enables rapid-cycle analytics for healthcare databases. Dr. Schneeweiss is Director of the Harvard-Brigham Drug Safety Research Center funded by FDA/CDER and Methods Lead of the FDA Sentinel program. He is voting consultant to the FDA Drug Safety and Risk Management Advisory Committee and member of the Methods Committee of the Patient Centered Outcomes Research Institute. He is Past President of the International Society for Pharmacoepidemiology and is Fellow of the American College of Epidemiology, the American College of Clinical Pharmacology, and the International Society for Pharmacoepidemiology. He received his medical training at the University of Munich Medical School and his doctoral degree in Pharmacoepidemiology from Harvard University.

ROBERT TEMPLE, M.D., serves as CDER's Deputy Center Director for Clinical Science and also Acting Deputy Director of the Office of Drug Evaluation I (ODE-I). He has served in this capacity since the office's establishment in 1995. Dr. Temple received his medical degree from the New York University School of Medicine in 1967. In 1972 he joined CDER as a review Medical Officer in the Division of Metabolic and Endocrine Drug Products. He later moved into the position of Director of the Division of Cardio-Renal Drug Products. In his current position, Dr. Temple oversees ODE-1 which is responsible for the regulation of cardio-renal, neuropharmacologic, and psychopharmacologic drug products. Dr. Temple has a long-standing interest in the design and conduct of clinical trials. He has written extensively on this subject, especially on choice of control group in clinical trials, evaluation of active control trials, trials to evaluate dose-response, and trials using "enrichment" designs.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 8, 2011 VOL. 365 NO. 10

Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation

Manesh R. Patel, M.D., Kenneth W. Mahaffey, M.D., Jyotsna Garg, M.S., Guohua Pan, Ph.D., Daniel E. Singer, M.D., Werner Hacke, M.D., Ph.D., Günter Breithardt, M.D., Jonathan L. Halperin, M.D., Graeme J. Hankey, M.D., Jonathan P. Piccini, M.D., Richard C. Becker, M.D., Christopher C. Nessel, M.D., John F. Paolini, M.D., Ph.D., Scott D. Berkowitz, M.D., Keith A.A. Fox, M.B., Ch.B., Robert M. Califf, M.D., and the ROCKET AF Steering Committee, for the ROCKET AF Investigators*

ABSTRACT

BACKGROUND

The use of warfarin reduces the rate of ischemic stroke in patients with atrial fibrillation but requires frequent monitoring and dose adjustment. Rivaroxaban, an oral factor Xa inhibitor, may provide more consistent and predictable anticoagulation than warfarin.

METHODS

In a double-blind trial, we randomly assigned 14,264 patients with nonvalvular atrial fibrillation who were at increased risk for stroke to receive either rivaroxaban (at a daily dose of 20 mg) or dose-adjusted warfarin. The per-protocol, as-treated primary analysis was designed to determine whether rivaroxaban was noninferior to warfarin for the primary end point of stroke or systemic embolism.

RESULTS

In the primary analysis, the primary end point occurred in 188 patients in the rivaroxaban group (1.7% per year) and in 241 in the warfarin group (2.2% per year) (hazard ratio in the rivaroxaban group, 0.79; 95% confidence interval [CI], 0.66 to 0.96; P<0.001 for noninferiority). In the intention-to-treat analysis, the primary end point occurred in 269 patients in the rivaroxaban group (2.1% per year) and in 306 patients in the warfarin group (2.4% per year) (hazard ratio, 0.88; 95% CI, 0.74 to 1.03; P<0.001 for noninferiority; P=0.12 for superiority). Major and nonmajor clinically relevant bleeding occurred in 1475 patients in the rivaroxaban group (14.9% per year) and in 1449 in the warfarin group (14.5% per year) (hazard ratio, 1.03; 95% CI, 0.96 to 1.11; P=0.44), with significant reductions in intracranial hemorrhage (0.5% vs. 0.7%, P=0.02) and fatal bleeding (0.2% vs. 0.5%, P=0.003) in the rivaroxaban group.

CONCLUSIONS

In patients with atrial fibrillation, rivaroxaban was noninferior to warfarin for the prevention of stroke or systemic embolism. There was no significant between-group difference in the risk of major bleeding, although intracranial and fatal bleeding occurred less frequently in the rivaroxaban group. (Funded by Johnson & Johnson and Bayer; ROCKET AF ClinicalTrials.gov number, NCT00403767.)

From the Duke Clinical Research Institute (M.R.P., K.W.M., J.G., J.P.P., R.C.B.) and Duke Translational Medicine Institute (R.M.C.), Duke University Medical Center, Durham, NC; Johnson & Johnson Pharmaceutical Research and Development, Raritan (G.P., C.C.N.), and Bayer HealthCare Pharmaceuticals, Montville (J.F.P., S.D.B.) — both in New Jersey; Massachusetts General Hospital and Harvard Medical School — both in Boston (D.E.S.); Ruprecht-Karls-University, Heidelberg (W.H.), and Hospital of the University of Münster, Münster (G.B.) — both in Germany; the Cardiovascular Institute, Mount Sinai Medical Center, New York (J.L.H.); Royal Perth Hospital, Perth, WA, Australia (G.J.H.); and the University of Edinburgh and Royal Infirmary of Edinburgh — both in Edinburgh (K.A.A.F.). Address reprint requests to Dr. Patel at Duke Clinical Research Institute, Duke University Medical Center, Rm. 0311 Terrace Level, 2400 Pratt St., Durham, NC 27705, or at manesh.patel@duke.edu.

*A complete listing of the steering committee members and trial investigators in the Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF) is provided in the Supplementary Appendix, available at NEJM.org.

This article (10.1056/NEJMoa1009638) was published on August 10, 2011, at NEJM.org.

N Engl J Med 2011;365:883-91.

Atrial fibrillation is associated with an increase in the risk of ischemic stroke by a factor of four to five[1] and accounts for up to 15% of strokes in persons of all ages and 30% in persons over the age of 80 years.[2] The use of vitamin K antagonists is highly effective for stroke prevention in patients with nonvalvular atrial fibrillation and is recommended for persons at increased risk.[3-5] However, food and drug interactions necessitate frequent coagulation monitoring and dose adjustments, requirements that make it difficult for many patients to use such drugs in clinical practice.[6-8]

Rivaroxaban is a direct factor Xa inhibitor that may provide more consistent and predictable anticoagulation than warfarin.[9,10] It has been reported to prevent venous thromboembolism more effectively than enoxaparin in patients undergoing orthopedic surgery[11,12] and was noninferior to enoxaparin followed by warfarin in a study involving patients with established venous thrombosis.[13] This trial was designed to compare once-daily oral rivaroxaban with dose-adjusted warfarin for the prevention of stroke and systemic embolism in patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke.[14]

Methods

Study Design and Oversight

The Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET AF) was a multicenter, randomized, double-blind, double-dummy, event-driven trial that was conducted at 1178 participating sites in 45 countries.[14] The study was supported by Johnson & Johnson Pharmaceutical Research and Development and Bayer HealthCare. The Duke Clinical Research Institute coordinated the trial, managed the database, and performed the primary analyses independently of the sponsors. Pertinent national regulatory authorities and ethics committees at participating centers approved the protocol, which is available with the full text of this article at NEJM.org. The members of an international executive committee designed the trial, were responsible for overseeing the study's conduct, retained the ability to independently analyze and present the data, made the decision to submit the manuscript for publication, and take responsibility for the accuracy and completeness of the data and all analyses. The first academic author wrote the initial draft of the manuscript.

Study Participants

We recruited patients with nonvalvular atrial fibrillation, as documented on electrocardiography, who were at moderate-to-high risk for stroke. Elevated risk was indicated by a history of stroke, transient ischemic attack, or systemic embolism or at least two of the following risk factors: heart failure or a left ventricular ejection fraction of 35% or less, hypertension, an age of 75 years or more, or the presence of diabetes mellitus (i.e., a $CHADS_2$ score of 2 or more, on a scale ranging from 1 to 6, with higher scores indicating a greater risk of stroke). According to the protocol, the proportion of patients who had not had a previous ischemic stroke, transient ischemic attack, or systemic embolism and who had no more than two risk factors was limited to 10% of the cohort for each region; the remainder of patients were required to have had either previous thromboembolism or three or more risk factors. Complete inclusion and exclusion criteria are provided in the Supplementary Appendix, available at NEJM.org. All patients provided written informed consent.

Study Treatment

Patients were randomly assigned to receive fixed-dose rivaroxaban (20 mg daily or 15 mg daily in patients with a creatinine clearance of 30 to 49 ml per minute) or adjusted-dose warfarin (target international normalized ratio [INR], 2.0 to 3.0). Patients in each group also received a placebo tablet in order to maintain blinding. Randomization was performed with the use of a central 24-hour, computerized, automated voice-response system. A point-of-care device was used to generate encrypted values that were sent to an independent study monitor, who provided sites with either real INR values (for patients in the warfarin group in order to adjust the dose) or sham values (for patients in the rivaroxaban group receiving placebo warfarin) during the course of the trial. Sham INR results were generated by means of a validated algorithm reflecting the distribution of values in warfarin-treated patients with characteristics similar to those in the study population.[15]

It was intended that patients would continue to take the assigned therapy throughout the course of the trial, unless discontinuation was considered to be clinically indicated. Follow-up

procedures and restrictions on concomitant medications are summarized in the Supplementary Appendix.

OUTCOMES

The primary efficacy end point was the composite of stroke (ischemic or hemorrhagic) and systemic embolism. Brain imaging was recommended to distinguish hemorrhagic from ischemic stroke. In the presence of atherosclerotic peripheral arterial disease, the diagnosis of embolism required angiographic demonstration of abrupt arterial occlusion.

Secondary efficacy end points included a composite of stroke, systemic embolism, or death from cardiovascular causes; a composite of stroke, systemic embolism, death from cardiovascular causes, or myocardial infarction; and individual components of the composite end points. The principal safety end point was a composite of major and nonmajor clinically relevant bleeding events. Bleeding events involving the central nervous system that met the definition of stroke were adjudicated as hemorrhagic strokes and included in both the primary efficacy and safety end points. Other overt bleeding episodes that did not meet the criteria for major or clinically relevant nonmajor bleeding were classified as minor episodes.

An independent clinical end-point committee applied protocol definitions to adjudicate all suspected cases of stroke, systemic embolism, myocardial infarction, death, and bleeding events that contributed to the prespecified end points. Detailed definitions of the end-point events are provided in the Supplementary Appendix.

STATISTICAL ANALYSIS

The primary hypothesis was that rivaroxaban would be noninferior to warfarin for the prevention of stroke or systemic embolism. The primary analysis was prespecified to be performed in the per-protocol population, which included all patients who received at least one dose of a study drug, did not have a major protocol violation, and were followed for events while receiving a study drug or within 2 days after discontinuation (group A in Fig. 1 in the Supplementary Appendix).[16-19]

For the primary analysis, we determined that a minimum of 363 events would provide a power of 95% to calculate a noninferiority margin of 1.46 with a one-sided alpha level of 0.025. However, 405 events were selected as the prespecified target to ensure a robust statistical result. On the basis of a projected event rate of 2.3% per 100 patient-years in the warfarin group and a projected 14% rate of annual attrition, it was estimated that approximately 14,000 patients would need to be randomly assigned to a study group.

If noninferiority was achieved in the primary analysis, a closed testing procedure was to be conducted for superiority in the safety population during treatment, which included patients who received at least one dose of a study drug and were followed for events, regardless of adherence to the protocol, while they were receiving the assigned study drug or within 2 days after discontinuation (group B in Fig. 1 in the Supplementary Appendix). Key secondary efficacy end points were also tested for superiority in the as-treated safety population.[20] Testing for noninferiority and superiority was also performed in the intention-to-treat population, which included all patients who underwent randomization and were followed for events during treatment or after premature discontinuation (group C in Fig. 1 in the Supplementary Appendix).

In addition, we performed post hoc analyses of events in the intention-to-treat population and events occurring during the end-of-study transition to open-label treatment with conventional anticoagulant agents. In the warfarin group, we used the method of Rosendaal et al.[21] to calculate the overall time that INR values fell within the therapeutic range. Comparative analyses of treatment efficacy were performed according to quartiles of time that INR values fell within the therapeutic range at the participating clinical sites.

Event rates per 100 patient-years are presented as proportions of patients per year. Hazard ratios, confidence intervals, and P values were calculated with the use of Cox proportional-hazards models with treatment as the only covariate. Testing for noninferiority was based on a one-sided significance level of 0.025; testing for superiority was based on a two-sided significance level of 0.05.

RESULTS

RECRUITMENT AND FOLLOW-UP

From December 18, 2006, through June 17, 2009, a total of 14,264 patients underwent randomization (Fig. 1 in the Supplementary Appendix). The study was terminated on May 28, 2010. The proportions of patients who permanently stopped

Table 1. Characteristics of the Intention-to-Treat Population at Baseline.

Characteristic	Rivaroxaban (N=7131)	Warfarin (N=7133)
Age — yr		
Median	73	73
Interquartile range	65–78	65–78
Female sex — no. (%)	2831 (39.7)	2832 (39.7)
Body-mass index*		
Median	28.3	28.1
Interquartile range	25.2–32.1	25.1–31.8
Blood pressure — mm Hg		
Systolic		
Median	130	130
Interquartile range	120–140	120–140
Diastolic		
Median	80	80
Interquartile range	70–85	70–85
Type of atrial fibrillation — no. (%)		
Persistent	5786 (81.1)	5762 (80.8)
Paroxysmal	1245 (17.5)	1269 (17.8)
Newly diagnosed or new onset	100 (1.4)	102 (1.4)
Previous medication use — no. (%)		
Aspirin	2586 (36.3)	2619 (36.7)
Vitamin K antagonist	4443 (62.3)	4461 (62.5)
$CHADS_2$ risk of stroke†		
Mean score (±SD)	3.48±0.94	3.46±0.95
Score — no. (%)		
2	925 (13.0)	934 (13.1)
3	3058 (42.9)	3158 (44.3)
4	2092 (29.3)	1999 (28.0)
5	932 (13.1)	881 (12.4)
6‡	123 (1.7)	159 (2.2)
Coexisting condition — no. (%)		
Previous stroke, systemic embolism, or transient ischemic attack	3916 (54.9)	3895 (54.6)
Congestive heart failure	4467 (62.6)	4441 (62.3)
Hypertension	6436 (90.3)	6474 (90.8)
Diabetes mellitus	2878 (40.4)	2817 (39.5)
Previous myocardial infarction‡	1182 (16.6)	1286 (18.0)
Peripheral vascular disease	401 (5.6)	438 (6.1)
Chronic obstructive pulmonary disease	754 (10.6)	743 (10.4)
Creatinine clearance — ml/min§		
Median	67	67
Interquartile range	52–88	52–86

* The body-mass index is the weight in kilograms divided by the square of the height in meters.
† The $CHADS_2$ score for the risk of stroke ranges from 1 to 6, with higher scores indicating an increased risk. Three patients (one in the rivaroxaban group and two in the warfarin group) had a $CHADS_2$ score of 1.
‡ P<0.05 for the between-group comparison.
§ Creatinine clearance was calculated with the use of the Cockcroft–Gault formula.

their assigned therapy before an end-point event and before the termination date were 23.7% in the rivaroxaban group and 22.2% in the warfarin group. The median duration of treatment exposure was 590 days; the median follow-up period was 707 days. Only 32 patients were lost to follow-up. Because of violations in Good Clinical Practice guidelines at one site that made the data unreliable, 93 patients (50 in the rivaroxaban group and 43 in the warfarin group) were excluded from all efficacy analyses before unblinding. An additional issue with data quality was raised at another trial site, but this issue was resolved without the exclusion of the patients from the analysis (for details, see the Supplementary Appendix).

PATIENT CHARACTERISTICS AND TREATMENTS

Key clinical characteristics of the patients who underwent randomization are shown in Table 1. The median age was 73 years (a quarter of the patients were 78 years of age or older), and 39.7% of the patients were women. The patients had substantial rates of coexisting illnesses: 90.5% had hypertension, 62.5% had heart failure, and 40.0% had diabetes; 54.8% of the patients had had a previous stroke, systemic embolism, or transient ischemic attack. The mean and median $CHADS_2$ scores were 3.5 and 3.0, respectively. Data on medication use at baseline are provided in Table 1 in the Supplementary Appendix. Previous use of vitamin K antagonists was reported by 62.4% of patients. At some time during the study, 34.9% of patients in the rivaroxaban group and 36.2% of those in the warfarin group took aspirin concurrently with the assigned study drug. Among patients in the warfarin group, INR values were within the therapeutic range (2.0 to 3.0) a mean of 55% of the time (median, 58%; interquartile range, 43 to 71).

PRIMARY OUTCOME

In the per-protocol population (the patients included in the primary efficacy analysis), stroke or systemic embolism occurred in 188 patients in the rivaroxaban group (1.7% per year) and in 241 patients in the warfarin group (2.2% per year) (hazard ratio in the rivaroxaban group, 0.79; 95% confidence interval [CI], 0.66 to 0.96; P<0.001 for noninferiority) (Table 2 and Fig. 1A). In the as-treated safety population, primary events occurred in 189 patients in the rivaroxaban group (1.7% per year) and in 243 patients in the warfarin

Table 2. Primary End Point of Stroke or Systemic Embolism.*

Study Population	Rivaroxaban			Warfarin			Hazard Ratio (95% CI)†	P Value	
	No. of Patients	No. of Events	Event Rate	No. of Patients	No. of Events	Event Rate		Noninferiority	Superiority
			no./100 patient-yr			*no./100 patient-yr*			
Per-protocol, as-treated population‡	6958	188	1.7	7004	241	2.2	0.79 (0.66–0.96)	<0.001	
Safety, as-treated population	7061	189	1.7	7082	243	2.2	0.79 (0.65–0.95)		0.02
Intention-to-treat population§	7081	269	2.1	7090	306	2.4	0.88 (0.75–1.03)	<0.001	0.12
During treatment		188	1.7		240	2.2	0.79 (0.66–0.96)		0.02
After discontinuation		81	4.7		66	4.3	1.10 (0.79–1.52)		0.58

* The median follow-up period was 590 days for the per-protocol, as-treated population during treatment; 590 days for the safety, as-treated population during treatment; and 707 days for the intention-to-treat population.
† Hazard ratios are for the rivaroxaban group as compared with the warfarin group.
‡ The primary analysis was performed in the as-treated, per-protocol population during treatment.
§ Follow-up in the intention-to-treat population continued until notification of study termination.

group (2.2% per year) (hazard ratio, 0.79; 95% CI, 0.65 to 0.95; P=0.01 for superiority). Among all randomized patients in the intention-to-treat analysis, primary events occurred in 269 patients in the rivaroxaban group (2.1% per year) and in 306 patients in the warfarin group (2.4% per year) (hazard ratio, 0.88; 95% CI, 0.74 to 1.03; P<0.001 for noninferiority; P=0.12 for superiority) (Fig. 1B).

During treatment in the intention-to-treat population, patients in the rivaroxaban group had a lower rate of stroke or systemic embolism (188 events, 1.7% per year) than those in the warfarin group (240 events, 2.2% per year) (P=0.02) (Table 2 and Fig. 2). Among patients who stopped taking the assigned study drug before the end of the study, during a median of 117 days of follow-up after discontinuation, primary events occurred in 81 patients in the rivaroxaban group (4.7% per year) and in 66 patients in the warfarin group (4.3% per year) (P=0.58). (Details regarding the time to events in patients who completed the study and were switched to standard medical therapy are provided in Fig. 2 in the Supplementary Appendix.)

BLEEDING OUTCOMES

Major and clinically relevant nonmajor bleeding occurred in 1475 patients in the rivaroxaban group and in 1449 patients in the warfarin group (14.9% and 14.5% per year, respectively; hazard ratio in the rivaroxaban group, 1.03; 95% CI, 0.96 to 1.11; P=0.44) (Table 3). Rates of major bleeding were similar in the rivaroxaban and warfarin groups (3.6% and 3.4%, respectively; P=0.58). Decreases in hemoglobin levels of 2 g per deciliter or more and transfusions were more common among patients in the rivaroxaban group, whereas fatal bleeding and bleeding at critical anatomical sites were less frequent. Rates of intracranial hemorrhage were significantly lower in the rivaroxaban group than in the warfarin group (0.5% vs. 0.7% per year; hazard ratio, 0.67; 95% CI, 0.47 to 0.93; P=0.02). Major bleeding from a gastrointestinal site was more common in the rivaroxaban group, with 224 bleeding events (3.2%), as compared with 154 events in the warfarin group (2.2%, P<0.001) (Table 2 in the Supplementary Appendix). (Data on nonhemorrhagic adverse events are provided in Table 3 in the Supplementary Appendix.)

SECONDARY EFFICACY OUTCOMES

The rates of secondary efficacy outcomes in the as-treated safety population are presented in Table 4 in the Supplementary Appendix. During treatment, myocardial infarction occurred in 101 patients in the rivaroxaban group and in 126 patients in the warfarin group (0.9% and 1.1% per year, respectively; hazard ratio in the rivaroxaban group, 0.81; 95% CI, 0.63 to 1.06; P=0.12). In the same analysis population, there were 208 deaths in the rivaroxaban group and 250 deaths in the warfarin group (1.9% and 2.2% per year, respectively; hazard ratio, 0.85; 95% CI, 0.70 to 1.02;

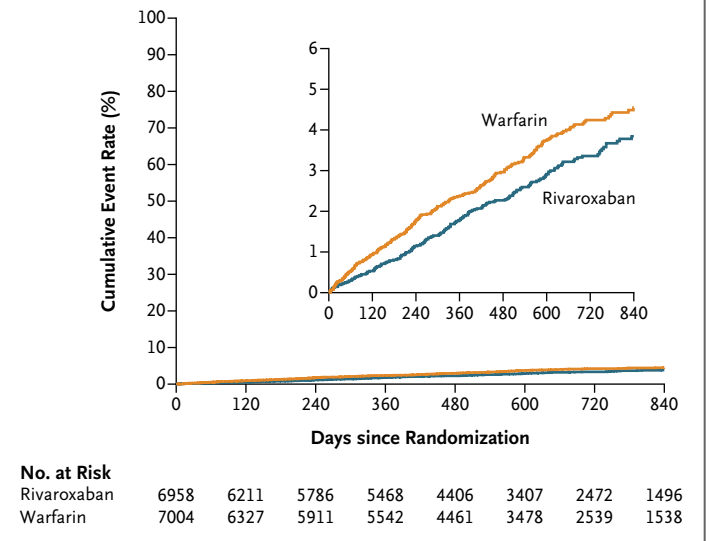


Figure 1. Cumulative Rates of the Primary End Point (Stroke or Systemic Embolism) in the Per-Protocol Population and in the Intention-to-Treat Population.

P=0.07). In addition, in the intention-to-treat analysis throughout the trial, there were 582 deaths in the rivaroxaban group and 632 deaths in the warfarin group (4.5% and 4.9% per year, respectively; hazard ratio, 0.92; 95% CI, 0.82 to 1.03; P=0.15).

SELECTED SUBGROUP ANALYSES

The effect of rivaroxaban, as compared with warfarin, in both efficacy and safety analyses was consistent across all prespecified subgroups (Fig. 3, 4, and 5 in the Supplementary Appendix). Furthermore, the effect of rivaroxaban did not differ across quartiles of the duration of time that INR values were within the therapeutic range according to study center (P=0.74 for interaction) (Table 5 in the Supplementary Appendix). Within the highest quartile according to center, the hazard ratio with rivaroxaban versus warfarin was 0.74 (95% CI, 0.49 to 1.12).

DISCUSSION

In this randomized trial, we compared rivaroxaban with warfarin for the prevention of stroke or systemic embolism among patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke. In both the primary analysis, which included patients in the per-protocol population, and in the intention-to-treat analysis, we found that rivaroxaban was noninferior to warfarin. In the primary safety analysis, there was no significant difference between rivaroxaban and warfarin with respect to rates of major or nonmajor clinically relevant bleeding.

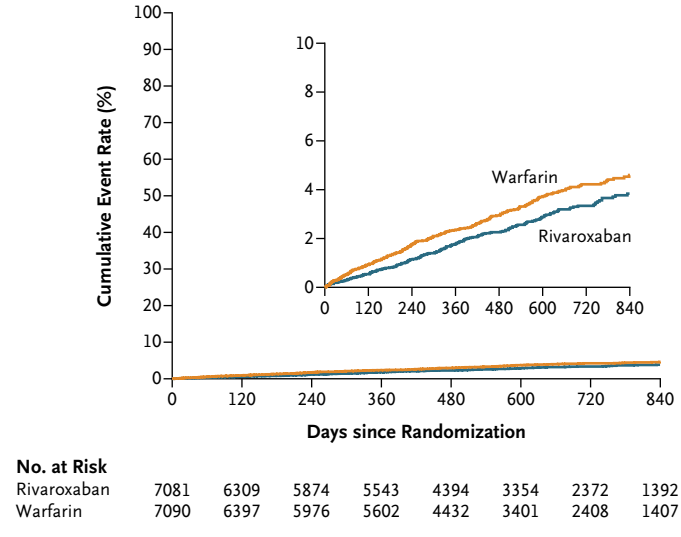
As prespecified in the statistical-analysis plan, we analyzed the trial data in a variety of ways because we anticipated that some patients would discontinue the study treatment and we wished to evaluate both noninferiority and superiority. Although an intention-to-treat analysis is the standard method for assessing superiority in a randomized trial, noninferiority is best established when patients are actually taking the randomized treatment.[16-19] Thus, the primary analysis was performed in the per-protocol population during receipt of the randomly assigned therapy. In the intention-to-treat population, we found no significant between-group difference in a conventional superiority analysis. In contrast, in the analyses of patients receiving at least one dose of a study drug who were followed for events during treatment, we found that rivaroxaban was superior to warfarin. The difference between these results reflects the fact that among patients who discontinued therapy before the conclusion of the trial, no significant difference in outcomes would have been anticipated, and none was seen.

The most worrisome complication of anticoagulation is bleeding. Rates of major and nonmajor clinically relevant bleeding, the main measure of treatment safety, were similar in the rivaroxaban and warfarin groups. Bleeding that proved fatal or involved a critical anatomical site occurred less frequently in the rivaroxaban group, mainly

because of lower rates of hemorrhagic stroke and other intracranial bleeding. In contrast, bleeding from gastrointestinal sites, including upper, lower, and rectal sites, occurred more frequently in the rivaroxaban group, as did bleeding that led to a drop in the hemoglobin level or bleeding that required transfusion. Even though patients in our trial were at increased risk for bleeding events, rates of major bleeding were similar to those in other recent studies involving patients with atrial fibrillation.[4,15,22,23]

Among patients in our study who survived and did not reach the primary end point, the rate of premature, permanent cessation of randomized treatment (14.3% in year 1) was slightly higher than in other studies (average, 11%).[15,23] This may have been a consequence of the trial's double-blind design or the inclusion of patients with more coexisting illnesses. Among patients who permanently discontinued their assigned treatment before the end of the study, only about half were treated thereafter with a vitamin K antagonist. This observation suggests that for at least some of the patients who participated in the trial, the risks of open-label therapy with currently available anticoagulants were ultimately judged to outweigh the risk of stroke or systemic embolism. Event rates were similar at 30 days and 1 year after withdrawal, suggesting that the mechanism of events did not involve hypercoagulability early after withdrawal of rivaroxaban. Events occurring at the end of the study were probably related to increased difficulty in achieving the transition from blinded trial therapy to the open-label use of a vitamin K antagonist when the patient had previously been assigned to the rivaroxaban group, since presumably many patients who had previously been assigned to the warfarin group would have already had a therapeutic INR.

Among patients in the warfarin group, the proportion of time in which the intensity of anticoagulation was in the therapeutic range (mean, 55%), which was calculated from all INR values during the study and for 7 days after warfarin interruptions, was lower than in previous studies of other new anticoagulants in patients with atrial fibrillation (range, 64 to 68%). Among these trials, the only study of blinded treatment was limited to North American sites, which may have facilitated trial compliance.[15] Most earlier trials of warfarin included fewer high-risk patients,[3] and no previous studies addressed patient populations with overall levels of coexisting illnesses and geographic diversity that were similar to those of the patients in our study.[24] Significant variations in the duration of time in the therapeutic range may reflect regional differences and differential skill in managing warfarin.[25] In a recent analysis of anticoagulation management involving more than 120,000 patients in the Veterans Affairs health care system, the mean proportion of time in the therapeutic range was 58%, with significant variation across sites.[24] The efficacy of rivaroxaban, as compared with warfarin, was as favorable in centers with the best INR control as in those with poorer control.



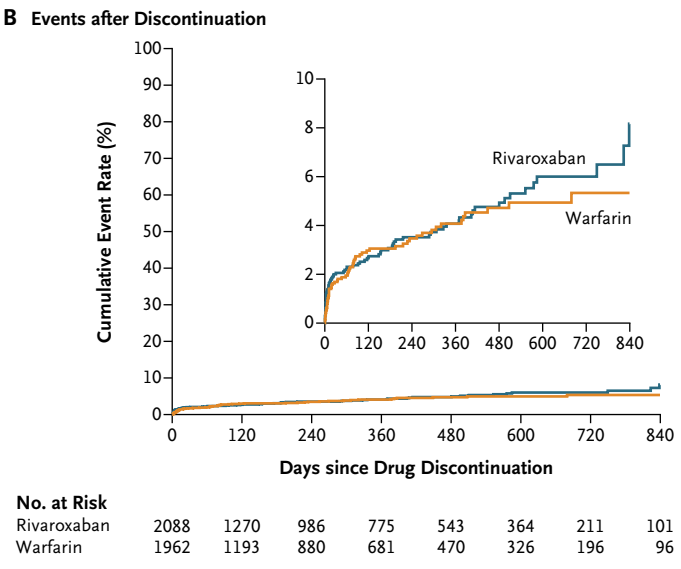


Figure 2. Cumulative Rates of the Primary End Point during Treatment and after Discontinuation in the Intention-to-Treat Population.

Table 3. Rates of Bleeding Events.*

Variable	Rivaroxaban (N=7111)		Warfarin (N=7125)		Hazard Ratio (95% CI)†	P Value‡
	Events	Event Rate	Events	Event Rate		
	no. (%)	*no./100 patient-yr*	*no. (%)*	*no./100 patient-yr*		
Principal safety end point: major and nonmajor clinically relevant bleeding§	1475 (20.7)	14.9	1449 (20.3)	14.5	1.03 (0.96–1.11)	0.44
Major bleeding						
Any	395 (5.6)	3.6	386 (5.4)	3.4	1.04 (0.90–1.20)	0.58
Decrease in hemoglobin ≥2 g/dl	305 (4.3)	2.8	254 (3.6)	2.3	1.22 (1.03–1.44)	0.02
Transfusion	183 (2.6)	1.6	149 (2.1)	1.3	1.25 (1.01–1.55)	0.04
Critical bleeding¶	91 (1.3)	0.8	133 (1.9)	1.2	0.69 (0.53–0.91)	0.007
Fatal bleeding	27 (0.4)	0.2	55 (0.8)	0.5	0.50 (0.31–0.79)	0.003
Intracranial hemorrhage	55 (0.8)	0.5	84 (1.2)	0.7	0.67 (0.47–0.93)	0.02
Nonmajor clinically relevant bleeding	1185 (16.7)	11.8	1151 (16.2)	11.4	1.04 (0.96–1.13)	0.35

* All analyses of rates of bleeding are based on the first event in the safety population during treatment.
† Hazard ratios are for the rivaroxaban group as compared with the warfarin group and were calculated with the use of Cox proportional-hazards models with the study group as a covariate.
‡ Two-sided P values are for superiority in the rivaroxaban group as compared with the warfarin group.
§ Minimal bleeding events were not included in the principal safety end point.
¶ Bleeding events were considered to be critical if they occurred in intracranial, intraspinal, intraocular, pericardial, intraarticular, intramuscular (with compartment syndrome), or retroperitoneal sites.

In conclusion, in this trial comparing a once-daily, fixed dose of rivaroxaban with adjusted-dose warfarin in patients with nonvalvular atrial fibrillation who were at moderate-to-high risk for stroke, rivaroxaban was noninferior to warfarin in the prevention of subsequent stroke or systemic embolism. There were no significant differences in rates of major and clinically relevant nonmajor bleeding between the two study groups, although intracranial and fatal bleeding occurred less frequently in the rivaroxaban group.

Supported by Johnson & Johnson Pharmaceutical Research and Development and Bayer HealthCare.

Dr. Patel reports receiving consulting fees from Ortho McNeil Janssen and Bayer HealthCare and serving on an advisory board for Genzyme; Dr. Mahaffey, receiving consulting fees from Adolor, Alexion, Amgen, Argolyn Bioscience, AstraZeneca, Bayer, Bristol-Myers Squibb, Daiichi Sankyo, Eli Lilly, Elsevier, Forest Labs, Genentech, GlaxoSmithKline, Guidant, Ikarja, Johnson & Johnson, Merck, Novartis, Pfizer, Proctor & Gamble, Sanofi Aventis, Schering Plough, Scios, and WebMD, grant support from Abbott Vascular, Amgen, Amylin, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, CardioKinetix, Cierra, Cordis, Edwards LifeSciences, Eli Lilly, Genentech, GlaxoSmithKline, Guidant, Innocoll Pharmaceuticals, Johnson & Johnson, KCI Medical, Luitpold Pharmaceutical, Medtronic, Merck, Momenta Pharmaceuticals, Novartis, Portola Pharmaceutical, Pozen, Regado Biotechnologies, Sanofi Aventis, Schering Plough, and the Medicines Company, and lecture fees from Johnson & Johnson; Dr. Pan, being an employee of Johnson & Johnson; Dr. Singer, receiving consulting fees from Johnson & Johnson, Bayer HealthCare, Boehringer Ingelheim, Daiichi Sankyo, Merck, Sanofi Aventis, Medtronic, and St. Jude Medical, grant support from Daiichi Sankyo, and lecture fees from Bristol-Myers Squibb and Pfizer; Dr. Hacke, serving as a board member for Sygnis Pharma Germany and receiving consulting fees, grant support, lecture fees, and payments for the development of educational presentations from Boehringer Ingelheim, consulting fees from Photothera USA and Codman USA, and lecture fees from Bayer and Photothera USA; Dr. Breithardt, receiving consulting fees from Bayer HealthCare, Sanofi Aventis, Bristol-Myers Squibb, Boehringer Ingelheim, Boston Scientific, and Otsuka Pharmaceutical, grant support from Sanofi Aventis, St. Jude, and Meda Pharma, and lecture fees from Sanofi Aventis, Boehringer Ingelheim, Bayer HealthCare, and Boston Scientific; Dr. Halperin, receiving consulting fees from Astellas Pharma, Boehringer Ingelheim, Bristol-Myers Squibb, Daiichi Sankyo, Pfizer, and Sanofi Aventis; Dr. Hankey, receiving consulting fees from Sanofi Aventis, Schering Plough, and Boehringer Ingelheim and lecture fees from Sanofi Aventis and Pfizer; Dr. Becker, receiving grant support from Johnson & Johnson, Bayer, Regado Biosciences, and Bristol-Myers Squibb, consulting fees from Boehringer Ingelheim, and lecture fees from Merck and AstraZeneca; Dr. Nessel, being an employee of and having an equity interest in Johnson & Johnson; Drs. Paolini and Berkowitz, being employees of Bayer HealthCare, and Dr. Paolini, having an equity interest in Bayer HealthCare; Dr. Fox, receiving grant support and lecture fees from Eli Lilly and lecture fees from Sanofi Aventis and AstraZeneca; Dr. Califf, receiving consulting fees from Kowa, Nile, Orexigen, Sanofi Aventis, Novartis, and Xoma and grant support from Novartis, Merck, and Amilyn/Lilly and having an equity interest in Nitrox. No other potential conflict of interest relevant to this article was reported. Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Lisa G. Berdan, Karen Hannan, and Kimberly Schwabe for their operational oversight of the trial and Elizabeth Cook for providing editorial support in the preparation of the manuscript.

REFERENCES

1. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983-8.
2. *Idem*. Atrial fibrillation: a major contributor to stroke in the elderly: the Framingham Study. Arch Intern Med 1987;147:1561-4.
3. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med 1999;131:492-501.
4. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007;146:857-67.
5. Singer DE, Albers GW, Dalen JE, Go AS, Halperin JL, Manning WJ. Antithrombotic therapy in atrial fibrillation: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126:Suppl:429S-456S.
6. Albers GW, Yim JM, Belew KM, et al. Status of antithrombotic therapy for patients with atrial fibrillation in university hospitals. Arch Intern Med 1996;156:2311-6.
7. Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV, Singer DE. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. Ann Intern Med 1999;131:927-34.
8. Piccini JP, Hernandez AF, Zhao X, et al. Quality of care for atrial fibrillation among patients hospitalized for heart failure. J Am Coll Cardiol 2009;54:1280-9.
9. Kubitza D, Becka M, Wensing G, Voith B, Zuehlsdorf M. Safety, pharmacodynamics, and pharmacokinetics of BAY 59-7939 — an oral, direct Factor Xa inhibitor — after multiple dosing in healthy male subjects. Eur J Clin Pharmacol 2005;61:873-80.
10. Kubitza D, Becka M, Roth A, Mueck W. Dose-escalation study of the pharmacokinetics and pharmacodynamics of rivaroxaban in healthy elderly subjects. Curr Med Res Opin 2008;24:2757-65.
11. Eriksson BI, Borris LC, Friedman RJ, et al. Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. N Engl J Med 2008;358:2765-75.
12. Lassen MR, Ageno W, Borris LC, et al. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty. N Engl J Med 2008;358:2776-86.
13. The EINSTEIN Investigators. Oral rivaroxaban for symptomatic venous thromboembolism. N Engl J Med 2010;363:2499-510.
14. ROCKET AF Study Investigators. Rivaroxaban — once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and Embolism Trial in Atrial Fibrillation: rationale and design of the ROCKET AF study. Am Heart J 2010;159(3):340.e1-347.e1.
15. Albers GW, Diener HC, Frison L, et al. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: a randomized trial. JAMA 2005;293:690-8.
16. Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: per protocol or intent-to-treat. Stat Med 2006;25:1169-81.
17. Brittain E, Lin D. A comparison of intent-to-treat and per-protocol results in antibiotic non-inferiority trials. Stat Med 2005;24:1-10.
18. Draft: Guidance for Industry: non-inferiority clinical trials. Silver Spring, MD: Food and Drug Administration, 2010. (http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.)
19. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA 2006;295:1152-60. [Erratum, JAMA 2006;296:1842.]
20. Soulakova JN. Comparison of several testing strategies for combination drug efficacy trials based on the closure principle. Stat Med 2009;28:260-73.
21. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993;69:236-9.
22. Connolly S, Pogue J, Hart R, et al. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet 2006;367:1903-12.
23. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009;361:1139-51. [Erratum, N Engl J Med 2010;363:1877.]
24. Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Risk-adjusted percent time in therapeutic range as a quality indicator for outpatient oral anticoagulation: results of the Veterans Affairs Study to Improve Anticoagulation (VARIA). Circ Cardiovasc Qual Outcomes 2011;4:22-9.
25. Wallentin L, Yusuf S, Ezekowitz MD, et al. Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. Lancet 2010;376:975-83.

Received: 28 July 2017 | Revised: 22 October 2017 | Accepted: 15 November 2017

DOI: 10.1002/pds.4375

WILEY

ORIGINAL REPORT

Prospective surveillance pilot of rivaroxaban safety within the US Food and Drug Administration Sentinel System

Elizabeth A. Chrischilles[1] | Joshua J. Gagne[2] | Bruce Fireman[3] | Jennifer Nelson[4] | Sengwee Toh[5] | Azadeh Shoaibi[6] | Marsha E. Reichman[8] | Shirley Wang[2] | Michael Nguyen[8] | Rongmei Zhang[9] | Rima Izem[9] | Margie R. Goulding[8] | Mary Ross Southworth[7] | David J. Graham[8] | Candace Fuller[5] | Hannah Katcoff[5] | Tiffany Woodworth[5] | Catherine Rogers[5] | Ryan Saliga[5] | Nancy D. Lin[10] | Cheryl N. McMahill-Walraven[11] | Vinit P. Nair[12] | Kevin Haynes[13] | Ryan M. Carnahan[1]

[1] Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, IA, USA

[2] Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[3] Kaiser Permanente Northern California, Oakland, CA, USA

[4] Biostatistics Unit, Group Health Research Institute and Department of Biostatistics, University of Washington, Seattle, WA, USA

[5] Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

[6] Center for Biologics Evaluation and Research, US Food and Drug, Rockville, MD, USA

[7] Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

[8] Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

[9] Division of Biometric VII, Office of Biostatistics, Office of Translation Sciences, US Food and Drug Administration, Silver Spring, MD, USA

[10] Optum Epidemiology, Waltham, MA, USA

[11] Aetna, Blue Bell, PA, USA

[12] Humana Inc. Louisville, KY, USA

[13] HealthCore Inc. Alexandria, VA, USA

Correspondence
E. Chrischilles, Department of Epidemiology, College of Public Health, University of Iowa, 145 N. Riverside Dr., Iowa City, IA 52242-2007, USA.
Email: e-chrischilles@uiowa.edu

Funding information
US FDA, Grant/Award Number: HHSF223200910006I

Abstract

Purpose: The US Food and Drug Administration's Sentinel system developed tools for sequential surveillance.

Methods: In patients with non-valvular atrial fibrillation, we sequentially compared outcomes for new users of rivaroxaban versus warfarin, employing propensity score matching and Cox regression. A total of 36 173 rivaroxaban and 79 520 warfarin initiators were variable-ratio matched within 2 monitoring periods.

Results: Statistically significant signals were observed for ischemic stroke (IS) (first period) and intracranial hemorrhage (ICH) (second period) favoring rivaroxaban, and gastrointestinal bleeding (GIB) (second period) favoring warfarin. In follow-up analyses using primary position diagnoses from inpatient encounters for increased definition specificity, the hazard ratios (HR) for rivaroxaban vs warfarin new users were 0.61 (0.47, 0.79) for IS, 1.47 (1.29, 1.67) for GIB, and 0.71 (0.50, 1.01) for ICH. For GIB, the HR varied by age: <66 HR = 0.88 (0.60, 1.30) and 66+ HR = 1.49 (1.30, 1.71).

Conclusions: This study demonstrates the capability of Sentinel to conduct prospective safety monitoring and raises no new concerns about rivaroxaban safety.

KEYWORDS

anticoagulants, outcome assessment, pharmacoepidemiology, product surveillance, post-marketing

1 | INTRODUCTION

Although sequential methods have been commonly applied in randomized trials, their use in observational settings is relatively new.[1] The Vaccine Safety Datalink used these methods to detect potential safety signals more rapidly than would be possible with a single retrospective evaluation, while controlling the overall Type I error rate across the multiple analysis periods.[2] Most subsequent applications, including within Vaccine Safety Datalink and in other settings such as Medicare data, have used either self-controlled or historically controlled designs to address confounding,[3-7] while 1 study implemented exposure matching on individual confounders.[8] Although the general challenges of sequential monitoring in observational settings have been explored,[9] less is known about sequential implementation of propensity score matching (PSM) in a setting like Sentinel. The Sentinel system, which is sponsored by the Food and Drug Administration (FDA), was created to improve medical product safety surveillance. It is a distributed database with more 100 million individuals from 18 Data Partners.

This paper describes the results of a pilot project to test a sequential PSM approach by examining the safety of rivaroxaban (Xarelto®) among patients with atrial fibrillation (AF) in the drug's early uptake period. Atrial fibrillation (AF) affects an estimated 2.9 million people in the United States[10] and is associated with a 4- to 5-fold increase in ischemic stroke risk.[11-13] Anticoagulation therapy with warfarin has long-established efficacy for reducing the risk of thromboembolic events, but this therapy also increases the risk of serious bleeding. Warfarin has other disadvantages including multiple diet and food interactions and a narrow therapeutic window requiring frequent international normalization ratio testing.

Rivaroxaban (Xarelto®) was the second non-vitamin K antagonist oral anticoagulant to receive FDA approval. In the study supporting its approval for stroke prevention in non-valvular AF (ROCKET AF—Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation), rivaroxaban was found to be non-inferior to warfarin therapy for the primary composite endpoint of time to first occurrence of stroke (any type) or non-CNS systemic embolism (HR 0.88; 95% CI 0.74, 1.03).[14] In ROCKET AF, compared with warfarin, rivaroxaban had a similar effect on ischemic stroke (2.9 events per 100 person-years for rivaroxaban vs 2.9 events per 100 person-years for warfarin), decreased the risk of intracranial hemorrhage (including hemorrhagic stroke) (0.5 events per 100 person-years vs 0.7 events per 100 person-years), and increased the risk of major gastrointestinal bleeding (2.0 events per 100 person-years vs 1.2 events per 100 person-years).

KEY POINTS

- In the study supporting rivaroxaban (Xarelto®) approval for stroke prevention in non-valvular atrial fibrillation, compared with warfarin, rivaroxaban had a similar effect on ischemic stroke, but decreased the risk of intracranial hemorrhage and increased the risk of major gastrointestinal bleeding.
- This study used new FDA Sentinel sequential monitoring capabilities to examine the safety of rivaroxaban among patients with atrial fibrillation during the drug's early uptake period in 4 large Data Partners in the FDA Sentinel distributed database with diverse patient populations.
- An indication of a lower risk of ischemic stroke in the rivaroxaban group compared with warfarin was detected early and persisted with additional monitoring and sensitivity analysis.

2 | METHODS

We used a new user cohort design comparing rivaroxaban to warfarin on 3 outcomes—gastrointestinal bleeding, ischemic stroke, and intracranial hemorrhage (including hemorrhagic stroke). Variable ratio PS matching was chosen to make use of the large number of warfarin users. Risk for each outcome was separately evaluated with time-to-event analyses using Cox regression. Sequential testing controlling overall type 1 error was used. A detailed surveillance plan was published, and, prior to conducting the second sequential test, amended to reduce the number of sequential tests from 5 to 2, and to also reflect the refinements made to the sequential monitoring tool.[15]

2.1 | Data source

The 4 largest (Aetna, Humana, Optum, and HealthCore) of Sentinel's 18 Data Partners were selected to participate in this pilot project, with data from November 1, 2011 through April 30, 2015 (Appendix Table A1). Sentinel is a public health surveillance activity that is not under the purview of institutional review boards.[16,17]

2.2 | Study cohort

We employed the Cohort Identification and Descriptive Analysis (CIDA) tool in combination with the PSM tool.[18] We identified new users of either drug who were age 21 years and older on the date of cohort entry and who, in the 183 days before rivaroxaban or warfarin initiation, were continuously enrolled in a participating health plan with medical coverage and pharmacy benefits, did not have a pharmacy dispensing claim for oral anticoagulants (rivaroxaban, warfarin, dabigatran, apixaban, or edoxaban), had a diagnosis of AF or atrial flutter (ICD-9-CM 427.31 or 427.32), and did not have codes for mitral stenosis, mechanical heart valve, joint replacement, renal dialysis, or a history of renal transplant. Codes for atrial flutter were included because a large fraction of these patients have AF along with atrial flutter, and per clinical practice guidelines, they should be treated similarly.[19] Rivaroxaban is not indicated for valvular AF and has not been studied in patients on hemodialysis. Patients with joint replacement or only taking the 10-mg rivaroxaban dosage, which is labeled only for prophylaxis of deep venous thromboembolism following hip or knee replacement surgery, were excluded because this

evaluation focused on the AF indication. We defined the dispensing date of the first eligible prescription of either drug as the index date.

2.3 | Outcomes

Intracranial hemorrhage, gastrointestinal bleeding, and ischemic stroke were identified by ICD-9-CM codes recorded as non-secondary diagnoses associated with inpatient health care claims (Appendix Table A2). Sentinel classifies diagnosis codes on acute inpatient encounters as either primary, secondary, unable to classify, or missing. Definitions that used all non-secondary (defined as primary position, unable to classify, or missing) codes were initially implemented for the outcome definitions. At the time surveillance was initiated, definitions using only primary position diagnosis codes were known to sometimes result in implausible variability in incidence rates across Data Partners. The source of this variability was corrected before the end of surveillance, and thus outcome definitions that included only diagnosis codes in the primary position were used for all end-of-surveillance analyses. The positive predictive values for these primary position-only definitions are known to exceed 85% (see Appendix Table A2 footnote for details).[20-25]

2.4 | Follow-up

Follow-up for each outcome ended at the earliest of any of the following: occurrence of that outcome event, initiation of a different anticoagulant, health plan disenrollment (excluding gaps of less than 45 days), death, discontinuation of the initiated therapy defined as failure to refill 7 days after the end of an exposure episode, or reaching the end of the assessment period (Appendix Table A1). Exposure episodes were defined as beginning on the day after the index date and lasting for the period specified in the days' supply field of the prescription claim. Serial fills of a study drug with gaps of 7 days or less between fills based on days' supply were merged into 1 exposure episode via a stockpiling algorithm.[26] Only the first eligible episode per person was included in the analysis.

2.5 | Covariates

Over 70 covariates were specified a priori, including risk factors for bleeding, risk factors for ischemic stroke, measures of overall health status, and medications (Appendix Tables A3 and A4). In addition, combined Charlson/Elixhauser comorbidity score,[27] as well as 8 measures of health care utilization intensity, were selected from pre-defined algorithms: number of filled prescriptions, unique generic drugs, unique drug classes, inpatient hospital encounters, non-acute institutional encounters, emergency department encounters, ambulatory encounters, and other ambulatory encounters such as telemedicine and email consults.[28] Covariates were based on data from the 183-day baseline period prior to initiation of the anticoagulant.

2.6 | Interim tool changes

The first analysis period used a prototype of the tools, while the second (last) analysis and end-of-surveillance analyses used updated CIDA and PSM tools. There were 2 important changes. First, the CIDA prototype initially misclassified physician service encounters occurring during inpatient stays as secondary diagnosis codes, while the updated CIDA tool correctly classified these diagnoses as position unspecified codes. Second, the updated PSM tool was modified to correctly retain matches throughout surveillance, rather than allow re-matching to occur with each sequential analysis period. The time between the first analysis and the second (last) sequential analysis was 18 months, rather than the planned quarterly intervals due to the time required to update the tools. By the second analysis, the target sample size had been achieved.

2.7 | Statistical analysis

Variable ratio PS matching was used to control for confounding where a new rivaroxaban user was matched to up to 10 new warfarin users from the same Data Partner.[29] PSs were estimated in each Data Partner, using a logistic regression model to estimate patients' probability of initiating rivaroxaban versus warfarin, and included all covariates from the 183-day baseline period in the model. A nearest-neighbor matching algorithm was used with a maximum matching caliper of 0.05 on the PS scale for analysis periods 1 and 2, and 0.01 for the end of surveillance.

We examined the distribution of PS values and checked covariate balance between rivaroxaban and warfarin cohorts within each Data Partner. We compared baseline characteristics between cohorts pooled across Data Partners before and after PS matching using standardized mean differences. A standardized mean difference ≥ 0.10 or ≤ -0.10 was used to indicate potential imbalance.[30] Baseline characteristics of the matched warfarin users were weighted by the inverse of the number of users in a matched set because of variability in the number of matches per set.

2.8 | Sequential analysis and testing

Using the matched data and combining across Data Partners, at each analysis period a separate Cox regression model with time-since-drug-initiation as the time scale, stratified by Data Partner and matched set, was used to estimate the hazard ratio (HR) comparing rivaroxaban and warfarin users for each of the 3 outcomes.

At each analysis period, a 2-sided test based on the standardized Wald statistic from the Cox regression analysis (ie, log(HR)/ stderr (log(HR)) was computed using model-based standard errors. This standardized test statistic was compared with a preset, constant group sequential signaling threshold with a total alpha of 0.05 for all sequential tests.

2.9 | End-of-surveillance analysis

The sequential tests yielded signals for all 3 outcomes. To further investigate, we conducted additional analyses using only diagnosis codes in the primary position for greater specificity. The HR estimates we report in our tables are from these end-of-surveillance analyses.

3 | RESULTS

3.1 | Cohort characteristics

Across the total monitoring time, we identified a total of 41 800 eligible rivaroxaban initiators and 87 907 eligible warfarin initiators (Table 1) with average follow-up of 139 days and 157 days, respectively, in the gastrointestinal bleeding analysis. For simplicity, we only present the descriptive statistics for the gastrointestinal bleeding cohort below. The sample sizes for the cohorts for the other outcome events were similar, with small differences explained by exclusion of patients with the particular outcome event on the index date.

Before PS matching, new users of rivaroxaban were on average 4 years younger and had fewer stroke or bleeding risk factors than warfarin users (Table 1; complete profiles are shown in Appendix Tables A3 and A4). The proportion of patients with a prior recorded ischemic stroke diagnosis was 7.5% for rivaroxaban users and 11.6% for warfarin users. Prior gastrointestinal bleeding had occurred in 3.6% and 5.5%, and intracranial hemorrhage in 0.6%, and 1.3%, of rivaroxaban and warfarin users, respectively. These baseline conditions include recent as well as a more distant history of the condition recorded in inpatient or outpatient settings during the baseline period. The matched cohorts were well balanced on all baseline confounders (Table 1 and Appendix Table A4).

Overall, 36 173 of 41 800 (86.5%) eligible rivaroxaban initiators and 79 520 of 87 907 (90.5%) eligible warfarin initiators were matched (Table 1). After accounting for the varying matching ratios and for loss of an entire matched set after it no longer included both rivaroxaban and warfarin users, the potentially informative mean follow-up was 85 days for rivaroxaban and 71 days for warfarin (Table 2). The number of matched sets more than doubled between the first and final sequential analysis. For example, in the first sequential analysis, 14 550 rivaroxaban users were matched with 46 539 warfarin users in the gastrointestinal bleeding analysis. In the final sequential analysis for the same outcome, 36 173 rivaroxaban users were matched with 79 520 warfarin users.

3.2 | Propensity score-matched sequential analysis results

At the time of the initial analysis that employed the non-secondary outcome definitions, after controlling for confounding using PS matching, the HR for ischemic stroke, 0.64, was significantly less than 1.0 with a test statistic that exceeded the threshold for rejecting the null hypothesis ($P = 0.0036$). The test statistics for gastrointestinal bleeding and intracranial hemorrhage had not exceeded the signaling threshold. At the second and final sequential analysis that employed the non-secondary outcome definitions, the null hypothesis was rejected for both bleeding outcomes with HRs of 1.30 ($P < 0.0001$) for gastrointestinal bleeding and 0.73 ($P = 0.0159$) for intracranial hemorrhage.

3.3 | End-of-surveillance propensity score matched results

Table 2 presents the end-of-surveillance PS matched HRs. Using a more specific outcome definition, the HR for ischemic stroke was 0.61 (0.47, 0.79), for gastrointestinal bleeding was 1.47 (1.29, 1.67), and for intracranial hemorrhage was 0.71 (0.50, 1.01). Histograms of propensity scores for the unmatched and matched cohort for each Data Partner are displayed in Appendix Figures 1 and 2.

In subgroup analyses, we did not find evidence to support that the associations varied significantly in patients with and without prior history of any of the events (Table 3). For gastrointestinal bleeding, the 2 age groups differed significantly ($P = 0.0002$) with an increased HR only observed among those aged 66 years and over (Table 3).

4 | DISCUSSION

This assessment demonstrates the capability of Sentinel to conduct prospective drug safety monitoring using a multi-site distributed database, and to do this with sophisticated re-usable programming tools. This enables highly customized analyses to be done more quickly and in a substantially larger and more heterogeneous patient population[31] than is otherwise possible in a single database system. Test statistics for all 3 outcomes exceeded the signaling threshold during surveillance: ischemic stroke during the first analysis period and both bleeding outcomes in the second period. When in-depth follow-up analyses were conducted, new rivaroxaban users had a 39% decrease in hazard of ischemic stroke, a 47% increase in hazard of gastrointestinal bleeding, and a HR for intracranial hemorrhage (HR = 0.71) that was no longer statistically significant (95% confidence interval: 0.50, 1.01).

The strengths of this assessment are several. The large population enabled several important subgroup analyses. The inclusion of patients who were dispensed anticoagulants in clinical settings across 4 large national health insurers provides real-world evidence to complement clinical trial evidence. There was a broad age range with which to examine effects in younger users. The active comparator new user cohort design with PS matching is a strong design with which to ensure that the study cohorts are as similar as possible except for the drug exposure. Finally, the sequential design enabled analysis of data as information accrued.

Most prior adaptations of sequential analyses in observational settings have involved monitoring of vaccines, which are administered at a single point in time, for outcomes that occur acutely following their receipt (eg, within days or weeks). Conducting sequential surveillance for chronically used drugs with longer-term adverse events follow-up periods is more challenging because "at risk" windows for a given individual are likely to span multiple sequential analysis periods. This necessitates the ability to link individual-level data over time.

When conducting prospective analyses in a dynamic health care data environment, it can be advantageous to incorporate newly updated data over time as these data may represent important corrections or previously missing data. However, doing this poses unique challenges when implementing PS matching as a confounder adjustment strategy in a sequential analysis framework. This test of the Sentinel PSM tool identified that small updates to confounder data in prior analysis periods can alter the estimated PS for individuals. This can then result in different matches being made using updated data than were first made using the originally captured data. Maintaining the same matched sample over time is important for cohort stability and

TABLE 1 Selected baseline patient characteristics, by study drug, unmatched and matched cohorts[a]

	Unmatched			Matched[a]		
Selected Characteristics	N(%) Rivaroxaban	N(%) Warfarin	Standardized Difference	N(%) Rivaroxaban	N(%) Warfarin	Standardized Difference
	41 800	87 907		36 173	79 520	
Gender (F)	16 374 (39.2)	37 017 (42.1)	−0.06	14 669 (40.6)	14 574 (40.3)	0.005
Age-mean (SD)	69.7 (10.7)	73.4 (10.6)	−0.352	71.1 (10.4)	71.1 (10.7)	0
Combined comorbidity score—mean (SD)	2.4 (2.4)	3.2 (2.8)	−0.313	2.5 (2.4)	2.5 (2.4)	−0.007
Atrial fibrillation	36 581 (87.5)	77 568 (88.2)	−0.022	31 630 (87.4)	31 866 (88.1)	−0.02
Atrial flutter	7627 (18.2)	12 454 (14.2)	0.111	5994 (16.6)	6008 (16.6)	−0.001
GI bleed	1507 (3.6)	4841 (5.5)	−0.091	1393 (3.9)	1426 (3.9)	−0.005
Intracranial hemorrhage	231 (0.6)	1152 (1.3)	−0.079	224 (0.6)	239 (0.7)	−0.005
Ischemic stroke	3150 (7.5)	10 207 (11.6)	−0.139	3031 (8.4)	2973 (8.2)	0.006
Hypertension	32 865 (78.6)	71 386 (81.2)	−0.064	28 662 (79.2)	28 683 (79.3)	−0.001
Hyperlipidemia	12 819 (30.7)	25 265 (28.7)	0.042	10 884 (30.1)	11 046 (30.5)	−0.01
Heart failure or cardiomyopathy	15 110 (36.1)	39 359 (44.8)	−0.176	13 781 (38.1)	13 940 (38.5)	−0.009
Peripheral vascular disease	6638 (15.9)	18 645 (21.2)	−0.137	6234 (17.2)	6277 (17.4)	−0.003
Diabetes	12 505 (29.9)	31 905 (36.3)	−0.136	11 417 (31.6)	11 398 (31.5)	0.001
Venous thromboembolism	2525 (6.0)	10 598 (12.1)	−0.211	2456 (6.8)	2340 (6.5)	0.013
Walker use	886 (2.1)	3126 (3.6)	−0.087	844 (2.3)	807 (2.2)	0.007
Home oxygen	2240 (5.4)	7017 (8.0)	−0.105	2123 (5.9)	2078 (5.7)	0.005
Health service utilization intensity: Mean (SD)						
# filled Rx	18.1 (14.8)	18.5 (15)	−0.029	18.1 (14.8)	18.3 (14.8)	−0.013
# inpatient hospital encounters	0.6 (0.8)	0.7 (0.9)	−0.178	0.6 (0.8)	0.6 (0.8)	0.011
# non-acute institutional encounters	0.2 (0.8)	0.4 (1)	−0.151	0.2 (0.9)	0.2 (0.8)	−0.004
# emergency room encounters	0.6 (1.1)	0.6 (1.3)	0.003	0.5 (1)	0.5 (1.2)	0.012
# ambulatory encounters	11.1 (9.6)	13.1 (11.5)	−0.192	11.5 (9.9)	11.6 (9.4)	−0.012

[a]Values for the matched cohort are weighted to incorporate the average values of all the matches in the matched set.

TABLE 2 Propensity score-matched end of surveillance[a] Cox regression analyses comparing rivaroxaban with warfarin, by health outcome

Outcome/ Comparator	New Users	Person-Years at Risk	Events	Adjusted Incidence Rate per 1000 Person-Years[c]	Adjusted Hazard Ratio (95% CI)[b]
Ischemic stroke					
Rivaroxaban	36,512	8,572	82	9.57	0.61 (0.47, 0.79)
Warfarin	80,180	15,672	268	17.10	
Gastrointestinal bleeding					
Rivaroxaban	36,173	8,427	423	50.20	1.47 (1.29, 1.67)
Warfarin	79,520	15,384	651	34.82	
Intracranial hemorrhage					
Rivaroxaban	36,171	8,502	46	5.41	0.71 (0.50, 1.01)
Warfarin	79,529	15,551	143	7.49	

[a]Monitoring period started November 1, 2011 for all Data Partners, but the end date varied among Data Partners: April 30, 2014, December 31, 2014, March 31, 2015, and April 30, 2015. Matching caliper for this analysis was 0.01.

[b]Hazard ratios estimated by stratified Cox regression conditioned on Data Partner and PS matched set. Confidence intervals are nominal 95% intervals for the final hazard ratio estimates.

[c]Incidence rates adjusted for censoring in matched sets and variable ratio matching. See Appendix B for further detail.

TABLE 3 Propensity score-matched[a] end of surveillance Cox regression analysis comparing rivaroxaban with warfarin, by health outcome and subgroup

Outcome/Subgroup	Hazard Ratio (95% CI)[b]
Ischemic stroke	
Age group:	
Patients age 21–65	1.09 (0.61, 1.96)
Patients age 66 and over	0.60 (0.45, 0.79)
Baseline history of outcome event:	
Patients without baseline ischemic stroke	0.68 (0.49, 0.93)
Patients with baseline ischemic stroke	0.61 (0.40, 0.94)
Gastrointestinal bleeding	
Age group:	
Patients age 21–65	0.88 (0.60, 1.30) *
Patients age 66 and over	1.49 (1.30, 1.71) *
Baseline history of outcome event:	
Patients without baseline gastrointestinal bleeding	1.52 (1.32, 1.76)
Patients with baseline gastrointestinal bleeding	1.36 (0.94, 1.95)
Intracranial hemorrhage	
Age group:	
Patients age 21–65	0.61 (0.20, 1.88)
Patients age 66 and over	0.77 (0.54, 1.10)
Baseline history of outcome event:	
Patients without baseline intracranial hemorrhage	0.66 (0.46, 0.94)
Patients with baseline intracranial hemorrhage	6.47 (0.87, 48.19)

[a]Monitoring period started November 1, 2011 for all Data Partners, but the end date varied among Data Partners: April 30, 2014, December 31, 2014, March 31, 2015, and April 30, 2015. Matching caliper for this analysis was 0.01.

[b]Hazard ratio estimated by stratified Cox regression conditioned on Data Partner and PS matched set. Confidence intervals are nominal 95% intervals for the final hazard ratio estimates.

*The null hypothesis that the 2 age subgroups differ by chance alone was rejected (Chi-square [1 degree of freedom] = 13.7, $p = .0002$).

minimizing the sampling variability of results. This was achieved, but additional programming enhancements were needed.

The CIDA + PSM tools attempt to strike a balance between semi-automating decisions and analyses so planning and implementation can be conducted more rapidly, but also retaining many of the design controls that would be implemented under a more traditional customized protocol approach. This assessment was able to extensively control for over 70 confounding variables, implement a PS-matched new user cohort design, and accomplish this with pre-programmed tools.

From a safety perspective, it is often desirable to conduct more frequent tests in order to either identify potential signals as rapidly as possible or provide reassurance that there is no evidence for a major safety concern. However, each time an analysis is conducted, resources (which are not unlimited) must be devoted to oversee and manage the receipt of the data, and to review, troubleshoot, interpret, and act on the results. The Sentinel Data Partners that participated in this surveillance activity refreshed their data on a quarterly basis. Thus, for this evaluation, quarterly testing (5 times) was originally selected as the most frequent rate of testing that would both provide potentially valuable new information at each analysis and also be practically feasible with available resources.[15] Refreshed data can only be used in surveillance analysis after they pass the Sentinel quality assurance processes. Although refreshed quarterly, the included data are from 6 to 9 months prior because the Data Partners prefer to use stable adjudicated data for Sentinel.

In this real-world example including new data as it accumulated over time, we found results that were partly consistent with those of the pivotal trial, the ROCKET-AF, a randomized trial of 14 264 patients with nonvalvular AF. The HRs for gastrointestinal bleeding (favoring warfarin over rivaroxaban) were similar in the 2 studies (Sentinel: HR: 1.47; 95% CI: 1.14, 1.76, and ROCKET-AF: HR: 1.61; 95% CI: 1.30, 1.99). However, for the ischemic stroke outcome, while rivaroxaban use was protective compared with warfarin in Sentinel (HR: 0.61, 95% CI: 0.47, 0.79), there was no difference on that outcome in the ROCKET-AF trial (HR: 0.94; 95% CI: 0.75, 1.17). The ROCKET-AF trial did find rivaroxaban was non-inferior, but not superior, to warfarin for the composite endpoint of stroke (ischemic or hemorrhagic) and non-

central nervous system systemic embolism. We did not examine such a composite outcome in the Sentinel study. In spite of differences in outcome definitions, method of capture, and population eligibility criteria, it is interesting that incidence rates for ischemic stroke and intracranial hemorrhage among warfarin users were quite similar between the 2 studies. Gastrointestinal bleeding rates were higher among Sentinel warfarin users than among warfarin users in the trial.

Our study population was drawn from patients who received anticoagulant therapy in routine ambulatory care settings and extends findings in meaningful ways beyond the randomized clinical trial setting. First, we included patients with a broad range in age and baseline stroke and bleeding risk. This enabled subgroup analyses by age and prior history of the outcome events. The HR estimates for ischemic stroke, for instance, were quite similar in patients with and without ischemic stroke diagnosis codes during the baseline period. This extends evidence beyond the high risk population included in the ROCKET-AF trial in which patients had to have either a history of stroke, transient ischemic attack, or systemic embolism or at least 2 risk factors (congestive heart failure, age 75 years or more, and diabetes). However, our estimate for those with "no prior stroke" includes people with other cerebrovascular conditions such as transient ischemic attacks and therefore may not accurately represent the risk for lowest risk individuals without these prior conditions. We also found that the overall elevated HR for gastrointestinal bleeding with rivaroxaban was not evident among those under age 66 years. The study aimed to evaluate rivaroxaban safety when used for non-valvular AF. Although all patients had AF diagnosis codes and those with codes for other indications were excluded, it is possible that some were taking anticoagulants for other indications.

This Sentinel assessment measures short-term effects (average follow-up duration less than 3 months). Follow-up in this study was short for several reasons. First, rivaroxaban was a newly approved drug with use increasing over the study period. Thus, many patients entered the cohort near the end of the study and were censored at the end of the study. For example, over half of the patients were added in the final analysis period when 10 to 21 months of new data were added. Second, real-world adherence to chronic medications is known to be low, and any on-treatment analysis like ours will have short follow-up time on average.[32,33] Health plan membership churn such as occurs with changes in employment is a third contributing factor. Finally, stratification by matched set in the Cox regression led to censoring of follow-up for the entire matched set when either the rivaroxaban user was censored, or all of the warfarin users in the matched set were censored. Most matched sets had more than 1 warfarin users and so, on average, warfarin follow-up was shortened more than rivaroxaban follow-up by this feature of the analysis. A recent FDA Medicare study found that incidence rates for both stroke and gastrointestinal bleeding were highest in the first 90 days of treatment for both dabigatran and warfarin, and dropped substantially thereafter. The same FDA study showed constant HRs in a time-varying Cox model. (Table 3, online supplement).[34]

Although we adjusted for many variables, there could still be residual confounding, such as would occur if rivaroxaban users were less likely than warfarin users to be smokers or obese, as these are potential risk factors that are incompletely captured in health plan databases. In addition, with a look-back period of 183 days, misclassification of the baseline covariates may be present. While our assessment of the PS-matched cohort suggests excellent balance in measured patient covariates, we cannot assess balance in unobserved covariates.

New use was defined by a minimum of 183 days of non-use of any anticoagulant. It is possible some patients could have had AF diagnosed in previous years and taken warfarin before a long period of non-adherence or non-problematic AF. The decision was made to require 183 days of continuous health plan enrollment rather than a longer period in order to avoid an anticipated non-trivial loss of sample size.

In summary, this first demonstration of the CIDA and PSM tools to enable prospective surveillance has resulted in important changes that improve FDA's ability to observe stable, matched patient-sets over time. The timely relevance of this study's findings illustrates capacity for Sentinel to play an effective role in post-market monitoring of serious cardiovascular outcomes for novel drugs in a way that complements post-market cardiovascular outcomes trials. Many of the important features used in rigorous observational safety studies were supported by the Sentinel CIDA + PSM tools and were able to be applied for this prospective surveillance activity, including diagnostic output to evaluate covariate balance, extensive covariate adjustment, subgroup analysis, and sensitivity analysis of key parameters. An indication of a lower risk of ischemic stroke in the rivaroxaban group compared with warfarin was detected early and persisted with additional monitoring and sensitivity analysis. Over an average of 3 months after initiating treatment, this study does not raise any new safety concerns regarding use of rivaroxaban.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the contributions of the following people and organizations: Judith C. Maro, PhD, Aetna, HealthCore Inc., Humana Inc., and Optum Epidemiology.

CONFLICT OF INTEREST

All authors are coinvestigators of the FDA-funded Sentinel project (primary investigator: Richard Platt).

Consultancies or honoraria:	Shirley Wang: consultant for Aetion, Inc. Joshua Gagne: consultant Aetion, Inc. and to Optum, Inc.
Grants Received:	Joshua Gagne: Principal Investigator of a grant from Novartis Pharmaceuticals Corporation to the Brigham and Women's Hospital
Other conflicts:	Jennifer Nelson: receives industry funding for a FDA post-marketing requirement contract (through inVentiv Health) to assess factors associated with opioid-related adverse events

ROLE OF THE FUNDING SOURCE

The Sentinel System is sponsored by the US Food and Drug Administration (FDA) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance

capabilities. The Sentinel System is one piece of FDA's Sentinel Initiative, a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise.

AUTHOR CONTRIBUTIONS

E.A.C., J.J.G., B.F., J.N., and R.M.C. wrote the manuscript, designed and performed the research, and analyzed the data; S.T. designed and performed the research and analyzed the data; A.S., M.E.R., S.W., M.N., R. Z., R.I., M.R.G., M.R.S., and D.J.G. designed the research; and C.F., H.K., C.R., R.M.S., N.D.L., C.N.M., V.P.N., N.S., and T.W. performed the research and analyzed the data.

ORCID

Elizabeth A. Chrischilles http://orcid.org/0000-0002-1843-1955
Joshua J. Gagne http://orcid.org/0000-0001-5428-9733
Sengwee Toh http://orcid.org/0000-0002-5160-0810
Marsha E. Reichman http://orcid.org/0000-0001-6963-1060
Ryan M. Carnahan http://orcid.org/0000-0002-7478-4739

REFERENCES

1. Nelson JC, Cook AJ, Yu O, Zhao S, Jackson LA, Psaty BM. Methods for observational post-licensure medical product safety surveillance. *Stat Methods Med Res*. 2015;24(2):177-193.
2. Lieu TA, Kulldorff M, Davis RL, et al. Real-time vaccine safety surveillance for the early detection of adverse events. *Med Care*. 2007;45(Suppl 2):S89-S95.
3. Nelson JC, Yu O, Dominguez-Islas CP, et al. Adapting group sequential methods to observational postlicensure vaccine safety surveillance: results of a pentavalent combination DTaP-IPV-Hib vaccine safety study. *Am J Epidemiol*. 2013;177(2):131-141.
4. Daley MF, Yih WK, Glanz JM, et al. Safety of diphtheria, tetanus, acellular pertussis and inactivated poliovirus (DTaP-IPV) vaccine. *Vaccine*. 2014;32(25):3019-3024.
5. Tseng HF, Sy LS, Liu IL, et al. Postlicensure surveillance for pre-specified adverse events following the 13-valent pneumococcal conjugate vaccine in children. *Vaccine*. 2013;31(22):2578-2583.
6. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM, Group VSDRCAIW. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink Project, 2010-2011. *Vaccine*. 2012;30(11):2024-2031.
7. Lee GM, Greene SK, Weintraub ES, et al. H1N1 and seasonal influenza vaccine safety in the vaccine safety datalink project. *Am J Prev Med*. 2011;41(2):121-128.
8. Gee J, Naleway A, Shui I, et al. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the vaccine safety datalink. *Vaccine*. 2011;29(46):8279-8284.
9. Nelson JC, Cook AJ, Yu O, et al. Challenges in the design and analysis of sequentially monitored postmarket safety surveillance evaluations using electronic observational health care data. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):62-71.
10. Go AS, Hylek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and risk factors in atrial fibrillation (ATRIA) study. *JAMA*. 2001;285(18):2370-2375.
11. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. *Stroke*. 1991;22(8):983-988.
12. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation*. 1998;98(10):946-952.
13. Ott A, Breteler MM, de Bruyne MC, van Harskamp F, Grobbee DE, Hofman A. Atrial fibrillation and dementia in a population-based study. *The Rotterdam Study Stroke*. 1997;28(2):316-321.
14. Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med*. 2011;365(10):883-891.
15. Mini-sentinel prospective routine observational monitoring program tools (PROMPT): rivaroxaban surveillance plan v3.0. 2015. (Accessed 09/13, 2016, at https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_PROMPT_Rivaroxaban-Surveillance-Plan_0_0.pdf.)
16. Forrow S, Campion DM, Herrinton LJ, et al. The organizational structure and governing principles of the Food and Drug Administration's mini-sentinel pilot program. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):12-17.
17. McGraw D, Rosati K, Evans B. A policy framework for public health uses of electronic health data. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):18-22.
18. Routine querying system. 2016. at https://www.sentinelinitiative.org/sites/default/files/SurveillanceTools/RoutineQuerying/Sentinel-Routine_Querying_System-Documentation_3_3_0_0.pdf.)
19. You JJ, Singer DE, Howard PA, et al. Antithrombotic therapy for atrial fibrillation: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141(2):e531S-e575S.
20. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.
21. Wahl PM, Rodgers K, Schneeweiss S, et al. Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population. *Pharmacoepidemiol Drug Saf*. 2010;19(6):596-603.
22. A protocol for assessment of dabigatran and selected safety outcomes v3.0. Mini-sentinel, 2015. at https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Protocol-for-Assessment-of-Dabigatran_0.pdf.)
23. Andrade SE, Harrold LR, Tjia J, et al. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):100-128.
24. Schelleman H, Bilker WB, Brensinger CM, Wan F, Yang YX, Hennessy S. Fibrate/statin initiation in warfarin users and gastrointestinal bleeding risk. *Am J Med*. 2010;123(2):151-157.
25. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiol Drug Saf*. 2011;20(6):560-566.
26. Modular program 3: frequency of select events during exposure to a drug/procedure group of interest (version 7.1). 2014. (Accessed September 24, 2015, at https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel-Modular_Program_3-Documentation.pdf.)
27. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011;64(7):749-759.
28. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol*. 2001;154(9):854-864.
29. Wang SV, Schneeweiss S, Rassen JA. Optimal matching ratios in drug safety surveillance. *Epidimiology*. 2014;25(5):772-773.
30. Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial

infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol*. 2001;54(4):387-398.

31. Psaty BM, Breckenridge AM. Mini-sentinel and regulatory science—big data rendered fit and functional. *N Engl J Med*. 2014;370(23):2165-2167.
32. De Geest S, Sabate E. Adherence to long-term therapies: evidence for action. *European journal of cardiovascular nursing: journal of the Working Group on Cardiovascular Nursing of the European Society of Cardiology*. 2003;2(4):323.
33. Haynes RB, McDonald H, Garg AX, Montague P. Interventions for helping patients to follow prescriptions for medications. The Cochrane database of systematic reviews 2002:Cd000011.
34. Graham DJ, Reichman ME, Wernecke M, et al. Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. *Circulation*. 2015;131(2):157-164.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Chrischilles EA, Gagne JJ, Fireman B, et al. Prospective surveillance pilot of rivaroxaban safety within the US Food and Drug Administration Sentinel System. *Pharmacoepidemiol Drug Saf*. 2018;1–9. https://doi.org/10.1002/pds.4375

Annals of Internal Medicine ORIGINAL RESEARCH

Outcomes of Dabigatran and Warfarin for Atrial Fibrillation in Contemporary Practice

A Retrospective Cohort Study

Alan S. Go, MD; Daniel E. Singer, MD; Sengwee Toh, ScD; T. Craig Cheetham, PharmD, MS; Marsha E. Reichman, PhD; David J. Graham, MD, MPH; Mary Ross Southworth, PharmD; Rongmei Zhang, PhD; Rima Izem, PhD; Margie R. Goulding, PhD; Monika Houstoun, PharmD; Katrina Mott, MS; Sue Hee Sung, MPH; and Joshua J. Gagne, PharmD, ScD

Background: Dabigatran (150 mg twice daily) has been associated with lower rates of stroke than warfarin in trials of atrial fibrillation, but large-scale evaluations in clinical practice are limited.

Objective: To compare incidence of stroke, bleeding, and myocardial infarction in patients receiving dabigatran versus warfarin in practice.

Design: Retrospective cohort.

Setting: National U.S. Food and Drug Administration Sentinel network.

Patients: Adults with atrial fibrillation initiating dabigatran or warfarin therapy between November 2010 and May 2014.

Measurements: Ischemic stroke, intracranial hemorrhage, extracranial bleeding, and myocardial infarction identified from hospital claims among propensity score–matched patients starting treatment with dabigatran or warfarin.

Results: Among 25 289 patients starting dabigatran therapy and 25 289 propensity score–matched patients starting warfarin therapy, those receiving dabigatran did not have significantly different rates of ischemic stroke (0.80 vs. 0.94 events per 100 person-years; hazard ratio [HR], 0.92 [95% CI, 0.65 to 1.28]) or extracranial hemorrhage (2.12 vs. 2.63 events per 100 person-years; HR, 0.89 [CI, 0.72 to 1.09]) but were less likely to have intracranial bleeding (0.39 vs. 0.77 events per 100 person-years; HR, 0.51 [CI, 0.33 to 0.79]) and more likely to have myocardial infarction (0.77 vs. 0.43 events per 100 person-years; HR, 1.88 [CI, 1.22 to 2.90]). However, the strength and significance of the association between dabigatran use and myocardial infarction varied in sensitivity analyses and by exposure definition (HR range, 1.13 [CI, 0.78 to 1.64] to 1.43 [CI, 0.99 to 2.08]). Older patients and those with kidney disease had higher gastrointestinal bleeding rates with dabigatran.

Limitation: Inability to examine outcomes by dabigatran dose (unacceptable covariate balance between matched patients) or quality of warfarin anticoagulation (few patients receiving warfarin had available international normalized ratio values).

Conclusion: In matched adults with atrial fibrillation treated in practice, the incidences of stroke and bleeding with dabigatran versus warfarin were consistent with those seen in trials. The possible relationship between dabigatran and myocardial infarction warrants further investigation.

Primary Funding Source: U.S. Food and Drug Administration.

Ann Intern Med. 2017;167:845-854. doi:10.7326/M16-1157 **Annals.org**
For author affiliations, see end of text.
This article was published at Annals.org on 14 November 2017.

Atrial fibrillation increases ischemic stroke risk by 4- to 5-fold and is the most common significant arrhythmia in adults (1, 2). The burden of atrial fibrillation will continue to increase in the United States, with an estimated prevalence of 6 million to 12 million cases by 2050 (1, 2). The evidence-based cornerstone of stroke prevention remains anticoagulant use. The vitamin K antagonist warfarin reduces ischemic stroke by a relative 68% but can cause intracranial and major extracranial bleeding (3). Furthermore, efficacy and safety of warfarin depend on achieving an international normalized ratio of 2.0 to 3.0 through careful monitoring (4). Warfarin is very effective in settings where high-quality anticoagulation is achieved (5).

Dabigatran, an oral direct thrombin inhibitor, was approved by the U.S. Food and Drug Administration (FDA) in 2010 for patients with nonvalvular atrial fibrillation (6). This approval was based on the RE-LY (Randomized Evaluation of Long-Term Anticoagulation Therapy) trial, which showed that dabigatran (150 mg twice daily) was superior to warfarin for reducing the combined rate of all stroke and systemic embolism (7, 8). Major bleeding was similar with 150 mg of dabigatran twice daily and adjusted-dose warfarin, but more patients had intracranial bleeding and fewer had gastrointestinal bleeding with dabigatran. In addition, the rate of acute myocardial infarction was significantly higher with 150 mg of dabigatran twice daily (7), but this difference was no longer significant after additional events were identified (8). However, meta-analyses of randomized trials involving dabigatran suggested increased myocardial infarction or acute coronary syndromes (9) and gastrointestinal bleeding (10)—findings largely driven by the results of RE-LY and data showing higher extracranial bleeding rates for patients aged 80 years or older (11).

After initial use of dabigatran in practice, published articles and reports to the FDA (9, 12) suggested major bleeding associated with dabigatran. In response, the FDA did preliminary analyses of bleeding risk using data from its Sentinel network (13, 14), where no increased bleeding rates were seen with dabigatran versus warfarin, but adjustment for confounders was limited (15). Given conflicting observational data about the balance of thromboembolic and safety risks with dabigatran versus warfarin (9, 12, 16–26), we examined the incidence of thromboembolism, bleeding, and myocardial infarction associated with initiation of dab-

igatran or warfarin treatment in a large, real-world population with atrial fibrillation in the FDA's Sentinel program.

METHODS

Source Population

The Sentinel program is a national surveillance system sponsored by the FDA for medical products. It includes a central coordinating center and 17 collaborating institutions and health care delivery systems contributing data from administrative, clinical, and pharmacy dispensing databases to the Sentinel Distributed Database (13, 14). Most patients in the database are privately insured. As a public health surveillance activity, Sentinel is not under the purview of institutional review boards (27).

Design and Analytic Sample

The detailed protocol for this analysis is available at www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Protocol-for-Assessment-of-Dabigatran_0.pdf. In brief, we did a retrospective analysis of the Sentinel Distributed Database with a "new user" design (28). The sample consisted of adults aged 21 years or older with atrial fibrillation initiating dabigatran or warfarin therapy between 1 November 2010 and 31 May 2014 (13, 29). Because of the data refresh schedule, the end date varied across sites, but most sites contributed data through 2013. Atrial fibrillation was defined as at least 1 diagnosis of atrial fibrillation or atrial flutter based on International Classification of Diseases, Ninth Revision, Clinical Modification, codes 427.31 and 427.32 from any setting in the 12 months before the date when dabigatran or warfarin was first dispensed (index date). We excluded patients with fewer than 365 days of continuous prescription and medical coverage immediately preceding the index date; any prior dispensing for oral anticoagulants (that is, warfarin, dabigatran, rivaroxaban, or apixaban) during the 365 days before the index date; known mechanical heart valve or mitral stenosis, prior kidney transplant, or long-term dialysis before the index date (based on diagnosis or procedure codes) (30); or residence in a skilled-nursing facility or nursing home at the index date.

Anticoagulant Exposure

We used outpatient pharmacy dispensing data to characterize initiation and longitudinal exposure to dabigatran or warfarin in an "on-treatment" approach to understand outcomes associated with active treatment. We allowed all possible doses and dosing regimens in the analysis for both dabigatran and warfarin. Follow-up started on the index date, and person-time of continuous exposure was based on prescriptions dispensed for the index treatment. In primary analyses, we allowed a grace period of up to 7 days between the estimated end date of any prescription and the start date of the next prescription, based on the days' supply information from each, to consider a patient continuously exposed to the drug of interest.

We separately addressed early drug refills using an approach that attempts to balance possible stockpiling with other situations in which the patient has used up the earlier prescription. Toward that end, we used a 7-day limit for early refills for both dabigatran and warfarin, such that for any refill that occurred within 7 days before the predicted end of a first prescription, the additional days were added to the end of the second prescription for consecutive prescriptions.

Outcomes

Outcomes were ischemic stroke, intracranial hemorrhage, all strokes, and major extracranial bleeding (see **Appendix Table 1**, available at Annals.org, for codes). We followed previously described algorithms (5, 31) using hospital discharge diagnoses in which ischemic stroke was identified by primary discharge diagnoses, intracranial hemorrhage by primary and secondary discharge diagnoses with subclassification of major trauma, and major extracranial hemorrhage by primary discharge diagnoses of extracranial hemorrhage with subclassification of gastrointestinal bleeding. On the basis of primary hospital discharge diagnoses used in previous FDA Sentinel protocols (32), we also identified patients hospitalized for myocardial infarction. Patients were followed through the end of available data from each site or until they were censored because of treatment discontinuation, initiation of the comparator treatment (that is, warfarin or dabigatran), initiation of another anticoagulant treatment, nursing home or skilled-nursing facility admission, health system disenrollment, or death.

Covariates

Using demographic information as well as diagnostic and procedure codes, we identified risk factors for bleeding and those for thromboembolism or myocardial infarction (**Appendix Table 2**, available at Annals.org; codes available on request). We also used relevant diagnostic and procedure codes, records on dispensed prescription medications, and resource use data for proxy measures of overall health status and frailty (**Appendix Table 2**; codes available on request). Finally, we identified prior receipt of angiotensin-converting enzyme inhibitors, angiotensin-receptor blockers, antiangina vasodilators, antiarrhythmics, antiplatelet agents, aspirin, β-blockers, calcium-channel blockers, other antihypertensive agents, antidiabetic drugs, diuretics, estrogens, progestins, heparin and low-molecular-weight heparins, nonsteroidal anti-inflammatory drugs, statins, nonstatin lipid-lowering drugs, and proton-pump inhibitors.

Statistical Analysis

Analyses were done using SAS, version 9.43 (SAS Institute). To construct the matched cohort, we estimated a propensity score for initiating dabigatran therapy using logistic regression (PROC LOGISTIC in SAS) among all eligible patients starting treatment with dabigatran or warfarin within each participating data partner including all covariates described above (33). The

data partner–specific propensity scores were then used to match patients receiving dabigatran with those receiving warfarin in a 1:1 ratio using a nearest-neighbor–matching algorithm with a maximum matching caliper of 0.05 within each data partner. We used the nearest-neighbor–matching macro in the Pharmacoepidemiology Toolbox with the following parameter settings: caliper, 0.05; ratio, 1; fixed_ratio, 1; and balanced, 0 (34, 35).

Given the large sample size, we compared characteristics among those receiving dabigatran or warfarin using standardized differences, which were calculated as the difference in means or proportions of a variable divided by a pooled estimate of the SD of the variable, with a value greater than 0.1 considered to be significant (36, 37). To describe the incidence of stroke, bleeding events, and myocardial infarction, we calculated cohort estimates of event rates per 100 person-years along with 95% CIs and plotted cumulative incidence curves for each outcome. To compare the incidence of these outcomes in patients receiving dabigatran versus warfarin, the prespecified primary analyses used Cox proportional hazards regression (PROC PHREG) in the matched cohort, stratified by data partner. The Cox model included exposure as the only independent variable because the 1:1 matching adjusts for covariates. We confirmed that the proportional hazards assumption had not been violated by examining an exposure-by-time interaction term and by visual inspection of Kaplan–Meier plots. We also estimated incidence rate differences, accounting for stratification by data partner by using inverse variance weights.

We also did a series of sensitivity analyses. First, we assembled a separate variable-ratio–matched cohort, which allowed more than 1 patient receiving warfarin to be matched to each patient receiving dabigatran, and we did Cox regression stratified by data partner and matched set. Second, we did conditional Cox regression only for the outcome of myocardial infarction to understand the difference between 1:1 and variable-ratio–matched results where, in addition to stratifying by data partner, we stratified by matched pair such that both members of the pair were censored at the time either member was censored. We also examined the potential influence of methods for defining drug exposure (that is, a 14-day grace period between prescriptions to define continuous exposure to each drug or a combination of prescription data and outpatient completion of international normalized ratio tests to characterize warfarin use [5]). Finally, we rematched and evaluated whether differential associations existed between treatment groups and outcomes in prespecified subgroups by age (<65 years, 65 to 74 years, 75 to 84 years, and ≥85 years), sex, and reduced kidney function (defined using relevant International Classification of Diseases, Ninth Revision, Clinical Modification, diagnostic codes). Additional details about the matching process can be found in the **Appendix** (available at Annals.org).

Role of the Funding Source

The FDA was involved in the design, conduct, and reporting of the study.

RESULTS

Cohort Assembly and Follow-up

Using data from the 8 participating data partners, we identified and propensity score–matched 25 289 eligible patients newly receiving dabigatran (95.8% of all 26 390 patients starting treatment with dabigatran) and 25 289 patients newly receiving warfarin (30.4% of all 83 084 patients starting treatment with warfarin) from November 2010 through May 2014 (**Appendix Figure** and **Appendix Table 3**, available at Annals.org). The numbers of initially identified patients per data partner are provided in **Appendix Table 4** (available at Annals.org). Among matched users, the mean age was 68.4 years, approximately one third were women, and the comorbidity burden was high (**Table 1**). However, on the basis of standardized differences, no material imbalances existed across characteristics between matched groups in any site or in the overall cohort (**Table 1**). Mean continuous follow-up was 123 days (SD, 149) for patients receiving dabigatran and 102 days (SD, 119) for matched patients receiving warfarin. Median continuous exposure was 66 days (interquartile range, 36 to 151 days) for dabigatran and 66 days (interquartile range, 36 to 123 days) for warfarin. During follow-up, 73.3% of patients receiving dabigatran and 70.8% of those receiving warfarin were censored because of discontinuing or having a significant gap in their index anticoagulant therapy, whereas 6.3% of patients receiving dabigatran and 4.0% of those receiving warfarin switched to another anticoagulant. Data were censored for administrative reasons in 20.5% of the dabigatran group and 25.2% of the warfarin group (**Appendix Table 5**, available at Annals.org).

Ischemic Stroke and Intracranial Hemorrhage

During follow-up, the rate of ischemic stroke in patients receiving dabigatran was 0.80 events per 100 person-years, compared with 0.94 events per 100 person-years in matched patients receiving warfarin (**Table 2** and **Figure 1**). No statistically significant difference existed between dabigatran and warfarin in the incidence of ischemic stroke (hazard ratio [HR], 0.92 [95% CI, 0.65 to 1.28]) (**Figure 2**). Results of sensitivity and subgroup analyses were similar (**Appendix Tables 6** to **14**, available at Annals.org).

The rate of intracranial hemorrhage in patients receiving dabigatran was 0.39 events per 100 person-years, compared with 0.77 events per 100 person-years among matched patients receiving warfarin (**Table 2** and **Figure 1**). The rate of intracranial hemorrhage was significantly lower in the dabigatran group than the warfarin group (HR, 0.51 [CI, 0.33 to 0.79]) (**Figure 2**). Results were similar after excluding traumatic intracranial hemorrhages (**Figure 2**). In additional sensitivity and subgroup analyses, results were similar to those

Table 1. Characteristics of Propensity Score–Matched Patients With Atrial Fibrillation Starting Dabigatran or Warfarin Therapy

Characteristic	Dabigatran (*n* = 25 289)	Warfarin (*n* = 25 289)	Standardized Difference
Demographic characteristics			
Mean age (SD), *y*	68.48 (10.91)	68.34 (11.11)	0.01
Female, *n (%)*	9128 (36.1)	9033 (35.7)	0.01
Health service use			
Mean combined comorbidity score (SD)	2.47 (2.22)	2.44 (2.19)	0.01
Mean prior hospitalizations (SD), *n*	0.66 (0.84)	0.66 (0.85)	0.01
Mean physician visits (SD), *n*	13.97 (8.27)	13.96 (8.28)	0.00
Mean unique National Drug Code numbers (SD), *n*	10.30 (7.27)	10.36 (7.60)	0.01
Medical history, *n (%)*			
Advanced kidney dysfunction	2932 (11.6)	2931 (11.6)	0.00
Advanced liver disease	71 (0.3)	79 (0.3)	0.01
Alcoholism	157 (0.6)	153 (0.6)	0.00
Anemia	1492 (5.9)	1526 (6.0)	0.01
Atrial fibrillation	24 584 (97.2)	24 555 (97.1)	0.01
Atrial flutter	5288 (20.9)	5376 (21.3)	0.01
Chronic heart failure	9766 (38.6)	9596 (37.9)	0.01
Coagulation defects	375 (1.5)	402 (1.6)	0.01
Diabetes mellitus	7622 (30.1)	7473 (29.6)	0.01
Hospitalized gastrointestinal bleeding	272 (1.1)	287 (1.1)	0.01
Hospitalized intracranial bleeding	98 (0.4)	93 (0.4)	0.00
Hyperlipidemia	9887 (39.1)	9947 (39.3)	0.00
Hypertension	20 633 (81.6)	20 603 (81.5)	0.00
Ischemic stroke	2053 (8.1)	2038 (8.1)	0.00
Metastatic cancer	311 (1.2)	304 (1.2)	0.00
Myocardial infarction	1235 (4.9)	1221 (4.8)	0.00
Nonspecific cerebrovascular symptoms	416 (1.6)	426 (1.7)	0.00
Other arterial embolism	216 (0.9)	233 (0.9)	0.01
Other gastrointestinal ulcer disease	273 (1.1)	298 (1.2)	0.01
Other hospitalized bleeding	250 (1.0)	252 (1.0)	0.00
Other ischemic cerebrovascular disease	4435 (17.5)	4505 (17.8)	0.01
Other ischemic heart disease	1258 (5.0)	1249 (4.9)	0.00
Peripheral vascular disease	4401 (17.4)	4397 (17.4)	0.00
Smoking and tobacco use	3797 (15.0)	3810 (15.1)	0.00
Trauma with likely immobilization	1277 (5.0)	1261 (5.0)	0.00
Recent procedures, *n (%)*			
Coronary artery bypass graft surgery	1810 (7.2)	1852 (7.3)	0.01
Other major surgery	1255 (5.0)	1307 (5.2)	0.01
Percutaneous coronary intervention	2886 (11.4)	2879 (11.4)	0.00
Frailty indicators, *n (%)*			
Cane use	94 (0.4)	92 (0.4)	0.00
Commode chair use	277 (1.1)	301 (1.2)	0.01
Home oxygen use	1225 (4.8)	1211 (4.8)	0.00
Osteoporotic fracture	552 (2.2)	530 (2.1)	0.01
Recent fall	824 (3.3)	840 (3.3)	0.00
Walker use	660 (2.6)	678 (2.7)	0.00
Wheelchair use	258 (1.0)	258 (1.0)	0.00
Medication use, *n (%)*			
Angiotensin-converting enzyme inhibitors	9931 (39.3)	10 008 (39.6)	0.01
Angiotensin-receptor blockers	5266 (20.8)	5263 (20.8)	0.00
Antiangina vasodilators	2327 (9.2)	2362 (9.3)	0.00
Antiarrhythmics	8733 (34.5)	8917 (35.3)	0.02
Antiplatelets	3335 (13.2)	3401 (13.4)	0.01
Aspirin	237 (0.9)	216 (0.9)	0.01
β-Blockers	18 087 (71.5)	18 126 (71.7)	0.00
Calcium-channel blockers	10 348 (40.9)	10 264 (40.6)	0.01
Diuretics	11 770 (46.5)	11 598 (45.9)	0.01
Other antihypertensives	2225 (8.8)	2222 (8.8)	0.00
Diabetes drugs	7622 (30.1)	7473 (29.6)	0.01
Estrogens	884 (3.5)	892 (3.5)	0.00
Progestins	301 (1.2)	295 (1.2)	0.00
Heparin and low-molecular-weight heparins	372 (1.5)	465 (1.8)	0.03
Nonsteroidal anti-inflammatory drugs	5336 (21.1)	5421 (21.4)	0.01
Statins	13 458 (53.2)	13 475 (53.3)	0.00
Nonstatin lipid-lowering drugs	3238 (12.8)	3299 (13.0)	0.01
Proton-pump inhibitors	6365 (25.2)	6453 (25.5)	0.01

from the main analysis, although precision was limited by the low number of events (Appendix Tables 6 to 14).

The incidence of combined stroke was lower in patients receiving dabigatran (1.18 events per 100 person-years) than in matched patients receiving warfarin (1.68 events per 100 person-years) (Table 2). The HR was 0.74 (CI, 0.57 to 0.97) with inclusion of trauma-related events and 0.75 (CI, 0.56 to 1.00) with their exclusion (Figure 2). Results were quantitatively consistent in sensitivity analyses, although they were not statistically significant in the variable-ratio–matched analysis (Appendix Table 6) or in subgroup analyses (Appendix Tables 7 to 14).

Major Extracranial Bleeding

Patients receiving dabigatran had a rate of 2.12 events per 100 person-years for major extracranial bleeding (primarily gastrointestinal), and matched patients receiving warfarin had a rate of 2.63 events per 100 person-years (Table 2 and Figure 1), with no significant association between dabigatran use and major extracranial bleeding compared with warfarin (HR, 0.89 [CI, 0.72 to 1.09]) (Figure 2). In sensitivity analyses, results were similar to those from the main analysis (Appendix Table 6); however, in subgroup analyses, compared with warfarin, dabigatran use was associated with a lower rate of major extracranial bleeding in persons aged 64 years or younger (HR, 0.51 [CI, 0.30 to 0.87]) and in women (HR, 0.73 [CI, 0.54 to 0.99]) (Appendix Tables 7 to 13).

We did not see a significant increase in gastrointestinal bleeding associated with dabigatran compared with warfarin in the primary analysis (HR, 1.04 [CI, 0.83 to 1.30]) (Figure 2), in sensitivity analyses related to matching and characterizing warfarin exposure (Appendix Table 6), or in patients aged younger than 75 years (Appendix Tables 7 to 9). However, rates of gastrointestinal bleeding were higher with dabigatran than warfarin in patients aged 75 to 84 years (HR, 1.47 [CI, 1.01 to 2.14]), those aged 85 years or older (HR, 1.84 [CI, 1.05 to 3.20]), and those classified as having reduced kidney function (HR, 1.91 [CI, 1.04 to 3.51]) (Appendix Tables 7 and 14).

Acute Myocardial Infarction

In the primary analysis, the rate of acute myocardial infarction in patients receiving dabigatran was 0.77 events per 100 person-years, compared with 0.43 events per 100 person-years in matched patients receiving warfarin (Table 2 and Figure 1), with an HR of 1.88 (CI, 1.22 to 2.90) (Figure 2). However, in sensitivity analyses using a conditional analytic approach (HR, 1.41 [CI, 0.82 to 2.43]) or a variable-ratio–matching method (HR, 1.13 [CI, 0.78 to 1.64]), the association of dabigatran use with myocardial infarction was smaller and not statistically significant compared with warfarin use. In additional sensitivity analyses using different methods for classifying drug exposure, the association between dabigatran use and myocardial infarction was attenuated (using a 14-day grace period: HR, 1.43 [CI, 0.99 to 2.08]; using an expanded warfarin exposure algorithm: HR, 1.38 [CI, 1.00 to 1.92]) and of borderline statistical significance (Appendix Table 6). Finally, in subgroup analyses, we saw a significant association in men (HR, 2.09 [CI, 1.17 to 3.64]) but not women. We saw notably stronger associations in patients aged 75 to 84 years (HR, 4.09 [CI, 1.39 to 12.03]) and those aged 85 years or older (HR, 5.25 [CI, 1.17 to 23.60]), but CIs were very wide (Appendix Tables 7 to 13).

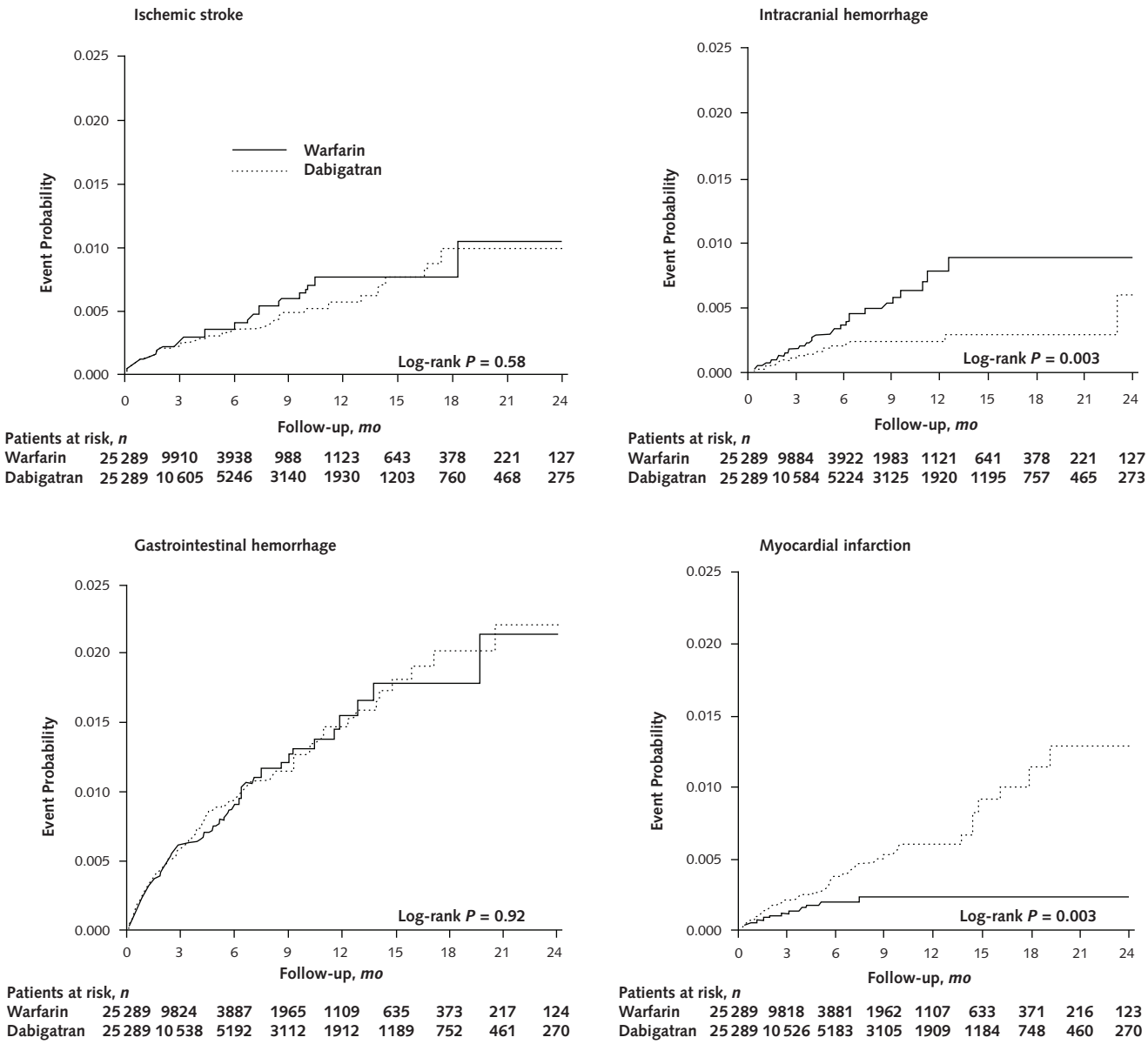
DISCUSSION

In a large cohort of carefully matched patients starting dabigatran or warfarin therapy for atrial fibrillation, we found that dabigatran use was associated with a lower rate of intracranial hemorrhage, similar rates of ischemic stroke and extracranial hemorrhage, and a potentially higher rate of myocardial infarction. Results of sensitivity analyses using different analytic approaches and drug exposure definitions were similar for ischemic stroke and bleeding outcomes. However, the association between dabigatran use and myocardial infarction was smaller and not statistically significant in sensitivity analyses, including those using a conditional modeling approach, variable-ratio matching, or a 14-day grace period between serial prescriptions to

Table 2. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Propensity Score–Matched Patients With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 25 289)		Warfarin (n = 25 289)		Incidence Rate Difference per 100 Person-Years (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	68	0.80	67	0.94	−0.15 (−0.44 to 0.15)
Intracranial hemorrhage	33	0.39	55	0.77	−0.39 (−0.63 to −0.15)
Excluding trauma	18	0.21	38	0.54	−0.32 (−0.52 to −0.13)
Combined stroke	100	1.18	119	1.68	−0.51 (−0.88 to 0.13)
Excluding trauma	85	1.00	102	1.44	−0.44 (−0.79 to −0.09)
Major extracranial bleeding	181	2.12	186	2.63	−0.50 (−0.99 to −0.01)
Gastrointestinal	165	1.93	145	2.05	−0.11 (−0.55 to 0.33)
Nongastrointestinal	16	0.19	41	0.58	−0.39 (−0.59 to −0.19)
Myocardial infarction	66	0.77	30	0.43	0.35 (0.11 to 0.59)

Figure 1. Cumulative incidence curves for clinical outcomes in matched cohorts of patients with atrial fibrillation newly receiving dabigatran and warfarin.



define continuous drug exposure. In subgroup analyses, gastrointestinal bleeding was higher with dabigatran than warfarin in older patients and in those classified as having reduced kidney function. A higher rate of myocardial infarction with dabigatran was also seen in men and those aged 75 years or older.

In RE-LY, 150 mg or 110 mg of dabigatran twice daily was tested versus adjusted-dose warfarin in 18 113 adults with atrial fibrillation (7). Both dabigatran doses resulted in lower intracranial hemorrhage rates; in older participants, the 150-mg dose resulted in fewer ischemic strokes and systemic emboli but more gastrointestinal bleeding (38). Myocardial infarction rates were higher with 150 mg of dabigatran (relative risk, 1.38 [CI, 1.00 to 1.91]; $P = 0.048$) than with warfarin. However, inclusion of additional myocardial infarction events identified after the RE-LY trial database was locked resulted in a slightly lower estimate that was no longer statistically significant for the group receiving dabigatran, 150 mg (relative risk, 1.27 [CI, 0.94 to 1.71]; $P = 0.120$) (8). Results with combined dabigatran doses were similar to those with the 150-mg dose for myocardial infarctions (HR, 1.28 [CI, 0.98 to 1.67]; $P = 0.070$) (39).

Relatively few studies have rigorously evaluated outcomes associated with dabigatran versus warfarin in populations more generalizable to clinical practice, and our analysis materially expands on previous studies.

The higher gastrointestinal bleeding rates we saw in patients aged 75 years or older (38) are consistent with results from RE-LY and a propensity score–matched cohort of patients with atrial fibrillation insured privately and through Medicare Advantage (16). However, in a retrospective Canadian cohort study (26) of matched patients receiving dabigatran and warfarin, bleeding rates with dabigatran, 150 mg, were lower in men (HR, 0.73 [CI, 0.64 to 0.84]) and borderline lower in women (HR, 0.85 [CI, 0.71 to 1.01]) (25). Rates of myocardial infarction with dabigatran in this Canadian cohort were higher but not significant in men (HR, 1.27 [CI, 0.94 to 1.71]) and not significantly different in women (25). Notably, many observational studies either did not have a "new user" design with rigorous individual-level matching or accounted for only a limited number of confounders.

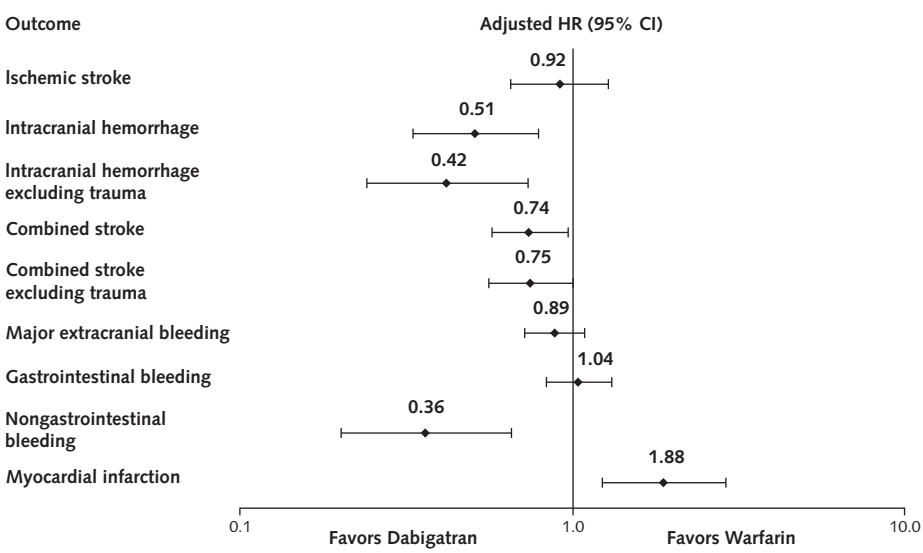
In 67 207 propensity score–matched pairs of Medicare beneficiaries with atrial fibrillation, patients starting dabigatran therapy (24) had lower adjusted rates of ischemic stroke (HR, 0.80 [CI, 0.67 to 0.96]) and intracranial hemorrhage (HR, 0.34 [CI, 0.26 to 0.46]) than those starting warfarin therapy. However, patients receiving dabigatran had excess major gastrointestinal bleeding—particularly women aged 75 years or older and men aged 85 years or older—but had no difference in myocardial infarction (HR, 0.92 [CI, 0.78 to 1.08]) (24). This study used a similar design to ours and saw more events, but it included only persons aged 65 years or older and used different analytic approaches, including stratification by dabigatran dose (24).

Finally, outside of our study, separate analyses of dabigatran versus warfarin were done using Sentinel data for the same periods. These analyses implemented Sentinel's new modular programs, which used propensity score matching based on covariates similar to those in our study. Results were similar for all outcomes except myocardial infarction, where a nonsignificant higher rate was seen with dabigatran (HR, 1.24 [CI, 0.85 to 1.83]). This finding is similar to results of sensitivity analyses in our study.

Strengths of our study include the large sample of highly matched patients newly receiving dabigatran or warfarin, which minimizes certain types of biases (28). Our analysis of data from the unique Sentinel network involved a broad spectrum of patients and practice settings that complement previous analyses in fee-for-service Medicare patients (24). Using extensive matching methods, we accounted for many potential confounders of the associations between anticoagulant choice and outcomes. We also did several sensitivity and subgroup analyses that produced results largely consistent with those of the main analyses, except for the outcomes of myocardial infarction and gastrointestinal bleeding.

Our study had several limitations. Information on outpatient international normalized ratios was not available for most warfarin-treated patients, which affected our ability to characterize longitudinal exposure more accurately and to characterize quality of anticoagulation. We also could not examine outcomes by dabigatran dose because we could not achieve acceptable covariate balance between matched users by dabigatran dose. We studied commercially insured patients, so our results may not be generalizable to uninsured patients or to all practice settings. As in previous studies (24), the duration of continuous exposure to dabigatran or warfarin was relatively short, which limited precision for some outcomes, and we could not measure drug adherence directly. Outpatient serum creatinine data were not uniformly available, which precluded reliable estimates for outcomes by level of kidney func-

Figure 2. Adjusted HRs for thromboembolism, intracranial hemorrhage, extracranial hemorrhage, and acute myocardial infarction among propensity score–matched patients with atrial fibrillation receiving dabigatran and warfarin.



HR = hazard ratio.

tion. Patients receiving long-term dialysis, who may be at higher risk for adverse outcomes with the newer anticoagulants, were excluded (40). Death could not be systematically ascertained, and events that may have occurred in nursing homes or skilled-nursing facilities were not available. We also could not address outcomes that occurred after withdrawal of dabigatran or warfarin. Conditions may have been misclassified on the basis of diagnostic or procedure codes, although this probably would not differ between treatment groups. Finally, despite highly matched cohorts across a wide range of characteristics, we cannot rule out residual confounding.

In conclusion, among insured adults with atrial fibrillation treated in usual care settings, compared with warfarin dabigatran was independently associated with a lower rate of intracranial hemorrhage, no significant differences in the rates of hospitalized ischemic stroke or extracranial hemorrhage, and possibly an increased risk for myocardial infarction. Gastrointestinal bleeding and myocardial infarction were notably higher in patients aged 75 years or older, and gastrointestinal bleeding risk was higher in those with diagnosed kidney dysfunction. Collectively, these results provide reassurance about overall bleeding risks—particularly intracranial hemorrhage—associated with dabigatran use and give insights to potentially assist in decision making about stroke prevention strategies for certain patients with atrial fibrillation. However, given the variability of findings for the outcome of myocardial infarction based on the analytic approach we used and results from other studies, the association between dabigatran and myocardial infarction remains uncertain.

From Kaiser Permanente Northern California, Oakland, California, University of California, San Francisco, San Francisco, California, and Stanford University School of Medicine, Stanford, California; Massachusetts General Hospital, Boston, Massachusetts; Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts; Western University of Health Sciences College of Pharmacy, Pomona, California; U.S. Food and Drug Administration, Silver Spring, Maryland; Kaiser Permanente Northern California, Oakland, California; and Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Note: The Sentinel Initiative is a multifaceted effort by the FDA to develop a national electronic surveillance system that will complement existing methods for monitoring the safety of FDA-regulated medical products. Sentinel collaborators include data and academic partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise.

Disclaimer: The opinions or assertions presented herein are the private views and opinions of the authors and are not to be construed as conveying either an official endorsement or criticism by the U.S. Department of Health and Human Services, the U.S. Public Health Service, or the FDA.

Acknowledgment: The authors thank Alda I. Inveiss for her expert technical assistance. The authors acknowledge the helpful input and contributions of all members (comprising Sentinel investigators and FDA staff) of the working groups for this project, as well as the constructive comments and suggestions from many additional FDA staff, Sentinel collaborators, and Sentinel Operations Center staff during the planning stages. The authors thank programmers and research staff from the Sentinel collaborators that contributed data to the analyses (Aetna, Harvard Pilgrim Health Care Institute, HealthCore, HealthPartners Institute, Humana Comprehensive Health Insights, Kaiser Permanente Northern California, Kaiser Permanente Washington Health Research Institute, and Optum Epidemiology) and the Sentinel collaborators that did the initial workplans but did not contribute data to the included analyses (Kaiser Permanente Colorado, Kaiser Permanente Georgia, Kaiser Permanente Hawaii, Kaiser Permanente Mid-Atlantic, Kaiser Permanente Northwest, Lovelace Clinic Foundation, Marshfield Clinic Research Foundation, Meyers Primary Care Institute, and Vanderbilt University Medical Center).

Financial Support: The Sentinel Coordinating Center is funded by the FDA through contract HHSF223200910006I from the U.S. Department of Health and Human Services and task order HHSF22301007T from the Center for Drug Evaluation and Research.

Disclosures: Dr. Go reports grants from the FDA during the conduct of the study. Dr. Singer reports grants and personal fees from Boehringer Ingelheim outside the submitted work; personal fees from Johnson & Johnson, Merck, Pfizer, Bristol-Myers Squibb, and CVS Health outside the submitted work; and other support from the Department of Population Medicine of Harvard Medical School during the conduct of the study. Dr. Toh reports grants from the FDA during the conduct of the study. Dr. Cheetham reports grants from Bristol-Myers Squibb outside the submitted work. Dr. Gagne reports grants from the FDA during the conduct of the study, personal fees from Optum and Aetion outside the submitted work, and grants from Novartis outside the submitted work. Authors not named here have disclosed no conflicts of interest. Disclosures can also be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M16-1157.

Reproducible Research Statement: *Study protocol:* Available at www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Protocol-for-Assessment-of-Dabigatran_0.pdf. *Statistical code:* Available from Dr. Toh (e-mail, Darren_Toh@harvardpilgrim.org). *Data set:* Not available.

Requests for Single Reprints: Alan S. Go, MD, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612; e-mail, Alan.S.Go@kp.org.

Current author addresses and author contributions are available at Annals.org.

References

1. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. JAMA. 2001;285:2370-5. [PMID: 11343485]

2. **Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, et al.** Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. Circulation. 2006;114:119-25. [PMID: 16818816]
3. **Hart RG, Benavente O, McBride R, Pearce LA.** Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med. 1999;131:492-501. [PMID: 10507957]
4. **You JJ, Singer DE, Howard PA, Lane DA, Eckman MH, Fang MC, et al.** Antithrombotic therapy for atrial fibrillation: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012;141:e531S-e575S. [PMID: 22315271] doi:10.1378/chest.11-2304
5. **Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, et al.** Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? JAMA. 2003;290:2685-92. [PMID: 14645310]
6. **U.S. Food and Drug Administration.** Drugs@FDA: FDA Approved Drug Products—Pradaxa (dabigatran etexilate mesylate) capsules. 2010. Accessed at www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&applno=022512 on 6 October 2017.
7. **Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al; RE-LY Steering Committee and Investigators.** Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med. 2009;361:1139-51. [PMID: 19717844] doi:10.1056/NEJMoa0905561
8. **Connolly SJ, Ezekowitz MD, Yusuf S, Reilly PA, Wallentin L; Randomized Evaluation of Long-Term Anticoagulation Therapy Investigators.** Newly identified events in the RE-LY trial [Letter]. N Engl J Med. 2010;363:1875-6. [PMID: 21047252] doi:10.1056/NEJMc1007378
9. **Uchino K, Hernandez AV.** Dabigatran association with higher risk of acute coronary events: meta-analysis of noninferiority randomized controlled trials. Arch Intern Med. 2012;172:397-402. [PMID: 22231617] doi:10.1001/archinternmed.2011.1666
10. **Sharma M, Cornelius VR, Patel JP, Davies JG, Molokhia M.** Efficacy and harms of direct oral anticoagulants in the elderly for stroke prevention in atrial fibrillation and secondary prevention of venous thromboembolism: systematic review and meta-analysis. Circulation. 2015;132:194-204. [PMID: 25995317] doi:10.1161/CIRCULATIONAHA.114.013267
11. **Lauw MN, Eikelboom JW, Coppens M, Wallentin L, Yusuf S, Ezekowitz M, et al.** Effects of dabigatran according to age in atrial fibrillation. Heart. 2017;103:1015-1023. [PMID: 28213368] doi:10.1136/heartjnl-2016-310358
12. **Legrand M, Mateo J, Aribaud A, Ginisty S, Eftekhari P, Huy PT, et al.** The use of dabigatran in elderly patients. Arch Intern Med. 2011;171:1285-6. [PMID: 21788545] doi:10.1001/archinternmed.2011.314
13. **Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, et al.** Design considerations, architecture, and use of the Mini-Sentinel distributed data system. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:23-31. [PMID: 22262590] doi:10.1002/pds.2336
14. **Forrow S, Campion DM, Herrinton LJ, Nair VP, Robb MA, Wilson M, et al.** The organizational structure and governing principles of the Food and Drug Administration's Mini-Sentinel pilot program. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:12-7. [PMID: 22262588] doi:10.1002/pds.2242
15. **Southworth MR, Reichman ME, Unger EF.** Dabigatran and postmarketing reports of bleeding. N Engl J Med. 2013;368:1272-4. [PMID: 23484796] doi:10.1056/NEJMp1302834
16. **Abraham NS, Singh S, Alexander GC, Heien H, Haas LR, Crown W, et al.** Comparative risk of gastrointestinal bleeding with dabigatran, rivaroxaban, and warfarin: population based cohort study. BMJ. 2015;350:h1857. [PMID: 25910928] doi:10.1136/bmj.h1857
17. **Hernandez I, Baik SH, Piñera A, Zhang Y.** Risk of bleeding with dabigatran in atrial fibrillation. JAMA Intern Med. 2015;175:18-24. [PMID: 25365537] doi:10.1001/jamainternmed.2014.5398
18. **Larsen TB, Gorst-Rasmussen A, Rasmussen LH, Skjøth F, Rosenzweig M, Lip GY.** Bleeding events among new starters and switchers to dabigatran compared with warfarin in atrial fibrillation. Am J Med. 2014;127:650-656.e5.[PMID:24530792]doi:10.1016/j.amjmed.2014.01.031
19. **Larsen TB, Rasmussen LH, Gorst-Rasmussen A, Skjøth F, Lane DA, Lip GY.** Dabigatran and warfarin for secondary prevention of stroke in atrial fibrillation patients: a nationwide cohort study. Am J Med. 2014;127:1172-8.e5. [PMID: 25193361] doi:10.1016/j.amjmed.2014.07.023
20. **Larsen TB, Rasmussen LH, Gorst-Rasmussen A, Skjøth F, Rosenzweig M, Lane DA, et al.** Myocardial ischemic events in 'real world' patients with atrial fibrillation treated with dabigatran or warfarin. Am J Med. 2014;127:329-336.e4. [PMID: 24361757] doi:10.1016/j.amjmed.2013.12.005
21. **Lauffenburger JC, Farley JF, Gehi AK, Rhoney DH, Brookhart MA, Fang G.** Effectiveness and safety of dabigatran and warfarin in real-world US patients with non-valvular atrial fibrillation: a retrospective cohort study. J Am Heart Assoc. 2015;4. [PMID: 25862791] doi:10.1161/JAHA.115.001798
22. **Smythe MA, Forman MJ, Bertran EA, Hoffman JL, Priziola JL, Koerber JM.** Dabigatran versus warfarin major bleeding in practice: an observational comparison of patient characteristics, management and outcomes in atrial fibrillation patients. J Thromb Thrombolysis. 2015;40:280-7. [PMID: 25851800] doi:10.1007/s11239-015-1213-7
23. **Yap LB, Eng DT, Sivalingam L, Rusani BI, Umadevan D, Muhammad Z, et al.** A comparison of dabigatran with warfarin for stroke prevention in atrial fibrillation in an Asian population. Clin Appl Thromb Hemost. 2016;22:792-797. [PMID: 25962393]
24. **Graham DJ, Reichman ME, Wernecke M, Zhang R, Southworth MR, Levenson M, et al.** Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. Circulation. 2015;131:157-64. [PMID: 25359164] doi:10.1161/CIRCULATIONAHA.114.012061
25. **Avgil Tsadok M, Jackevicius CA, Rahme E, Humphries KH, Pilote L.** Sex differences in dabigatran use, safety, and effectiveness in a population-based cohort of patients with atrial fibrillation. Circ Cardiovasc Qual Outcomes. 2015;8:593-9. [PMID: 26508666] doi:10.1161/CIRCOUTCOMES.114.001398
26. **Avgil-Tsadok M, Jackevicius CA, Essebag V, Eisenberg MJ, Rahme E, Behlouli H, et al.** Dabigatran use in elderly patients with atrial fibrillation. Thromb Haemost. 2016;115:152-60. [PMID: 26354766] doi:10.1160/TH15-03-0247
27. **McGraw D, Rosati K, Evans B.** A policy framework for public health uses of electronic health data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:18-22. [PMID: 22262589] doi:10.1002/pds.2319
28. **Ray WA.** Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol. 2003;158:915-20. [PMID: 14585769]
29. **Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, et al.** The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:1-8. [PMID: 22262586] doi:10.1002/pds.2343
30. **Go AS, Fang MC, Udaltsova N, Chang Y, Pomernacki NK, Borowsky L, et al; ATRIA Study Investigators.** Impact of proteinuria and glomerular filtration rate on risk of thromboembolism in atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. Circulation. 2009;119:1363-9. [PMID: 19255343] doi:10.1161/CIRCULATIONAHA.108.816082
31. **Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, et al.** A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. J Am Heart Assoc. 2013;2:e000250. [PMID: 23782923] doi:10.1161/JAHA.113.000250
32. **Cutrona SL, Toh S, Iyer A, Foy S, Daniel GW, Nair VP, et al.** Validation of acute myocardial infarction in the Food and Drug Administration's Mini-Sentinel program. Pharmacoepidemiol Drug Saf. 2013;22:40-54. [PMID: 22745038] doi:10.1002/pds.3310
33. **Gagne JJ, Glynn RJ, Rassen JA, Walker AM, Daniel GW, Sridhar G, et al.** Active safety monitoring of newly marketed medications in a

distributed data network: application of a semi-automated monitoring system. Clin Pharmacol Ther. 2012;92:80-6. [PMID: 22588606] doi:10.1038/clpt.2011.369
34. **Rassen JA, Doherty M, Huang W, Schneeweiss S.** Pharmacoepidemiology Toolbox. 2012. Accessed at www.drugepi.org/dope-downloads/#Pharmacoepidemiology Toolbox on 6 October 2017.
35. **Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S.** One-to-many propensity score matching in cohort studies. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 2:69-80. [PMID: 22552982] doi:10.1002/pds.3263
36. **Austin PC.** Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. Commun Stat Simul Comput. 2009;38:1228-34.
37. **Austin PC.** An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011;46:399-424. [PMID: 21818162]
38. **Eikelboom JW, Wallentin L, Connolly SJ, Ezekowitz M, Healey JS, Oldgren J, et al.** Risk of bleeding with 2 doses of dabigatran compared with warfarin in older and younger patients with atrial fibrillation: an analysis of the randomized evaluation of long-term anticoagulant therapy (RE-LY) trial. Circulation. 2011;123:2363-72. [PMID: 21576658] doi:10.1161/CIRCULATIONAHA.110.004747
39. **Hohnloser SH, Oldgren J, Yang S, Wallentin L, Ezekowitz M, Reilly P, et al.** Myocardial ischemic events in patients with atrial fibrillation treated with dabigatran or warfarin in the RE-LY (Randomized Evaluation of Long-Term Anticoagulation Therapy) trial. Circulation. 2012;125:669-76. [PMID: 22215856] doi:10.1161/CIRCULATIONAHA.111.055970
40. **Chan KE, Edelman ER, Wenger JB, Thadhani RI, Maddux FW.** Dabigatran and rivaroxaban use in atrial fibrillation patients on hemodialysis. Circulation. 2015;131:972-9. [PMID: 25595139] doi:10.1161/CIRCULATIONAHA.114.014113

IN THE CLINIC

In the Clinic is a monthly feature in *Annals* that focuses on practical management of patients with common clinical conditions. It offers evidence-based answers to frequently asked questions about screening, prevention, diagnosis, therapy, and patient education and provides physicians with tools to improve the quality of care. In the Clinic includes CME quizzes offering category 1 CME credit.

For more information on In the Clinic and to read the latest issue, visit www.annals.org/intheclinic.

Current Author Addresses: Dr. Go and Ms. Sung: Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612.
Dr. Singer: Division of General Internal Medicine, Massachusetts General Hospital, 50 Staniford Street, 9th Floor, Boston, MA 02114.
Dr. Toh: Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Landmark Center, 401 Park Drive, Suite 401 East, Boston, MA 02215.
Dr. Cheetham: Western University College of Pharmacy, 309 East Second Street, Pomona, CA 91766.
Drs. Reichman, Graham, and Goulding and Ms. Mott: Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Avenue, Building 22, Silver Spring, MD 20993.
Drs. Southworth and Houstoun: Office of New Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Avenue, Building 22, Suite 4177, Silver Spring, MD 20993.
Drs. Zhang and Izem: Office of Biostatistics, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Avenue, Building 21, Silver Spring, MD 20993.
Dr. Gagne: Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA 02120.

Author Contributions: Conception and design: A.S. Go, S. Toh, T.C. Cheetham, M.E. Reichman, D.J. Graham, M.R. Southworth, R. Zhang, M.R. Goulding, M. Houstoun, K. Mott, J.J. Gagne.
Analysis and interpretation of the data: A.S. Go, D.E. Singer, S. Toh, T.C. Cheetham, M.E. Reichman, D.J. Graham, M.R. Southworth, R. Zhang, R. Izem, M.R. Goulding, M. Houstoun, K. Mott, J.J. Gagne.
Drafting of the article: A.S. Go, R. Izem.
Critical revision of the article for important intellectual content: A.S. Go, D.E. Singer, S. Toh, T.C. Cheetham, M.E. Reichman, D.J. Graham, M.R. Southworth, R. Zhang, R. Izem, M.R. Goulding, M. Houstoun, K. Mott, J.J. Gagne.
Final approval of the article: A.S. Go, D.E. Singer, S. Toh, T.C. Cheetham, M.E. Reichman, D.J. Graham, M.R. Southworth, R. Zhang, R. Izem, M.R. Goulding, M. Houstoun, K. Mott, S.H. Sung, J.J. Gagne.
Provision of study materials or patients: A.S. Go.
Statistical expertise: A.S. Go, S. Toh, M.E. Reichman, R. Zhang, R. Izem, J.J. Gagne.
Obtaining of funding: A.S. Go, S. Toh, M.E. Reichman, J.J. Gagne.
Administrative, technical, or logistic support: A.S. Go, S. Toh, M.E. Reichman, M. Houstoun, K. Mott, S.H. Sung.
Collection and assembly of data: A.S. Go, S. Toh, M.R. Southworth, M. Houstoun, J.J. Gagne.

Appendix: Additional Details About the Propensity Score–Matching Process

Propensity score estimation and 1:1 matching of patients newly receiving dabigatran and warfarin were done separately within each data partner. To test our process, we examined the distribution of propensity scores between treatment groups within each data partner, and that assessment did not reveal any concerns. We also inspected data partner–specific covariate distributions after matching, which showed good balance within each data partner (data not shown).

For our exploratory subgroup analyses, we used the data partner–specific propensity scores to rematch patients rather than fitting separate propensity score models within each subgroup, which would be limited by the small numbers of patients in some subgroups and data partners. In some cases, the rematching resulted in fewer patients being included in the subgroup analyses than in the overall analysis. For example, 20 068 patients in each treatment group were included across the age subgroup analyses, compared with 25 289 in the overall analysis. In other cases, rematching resulted in more patients in a particular subgroup analysis—for example, the female subgroup analysis included 9143 matched pairs, whereas 9128 and 9033 women were in the dabigatran and warfarin groups, respectively, in the overall analysis. We inspected data partner–specific covariate distributions after rematching within each a priori–planned subgroup to evaluate for potential covariate imbalances. Because few patients started the 75-mg dose of dabigatran, we did not see acceptable balance in baseline characteristics in this subgroup with corresponding matched patients receiving warfarin. For example, even after individual propensity score matching, the mean ages of patients in the dabigatran and warfarin groups differed by 5 years (79 vs. 74 years, respectively) and other important confounders differed between groups. For any subgroups in which we did not achieve adequate covariate balance between matched patients, we did not present results.

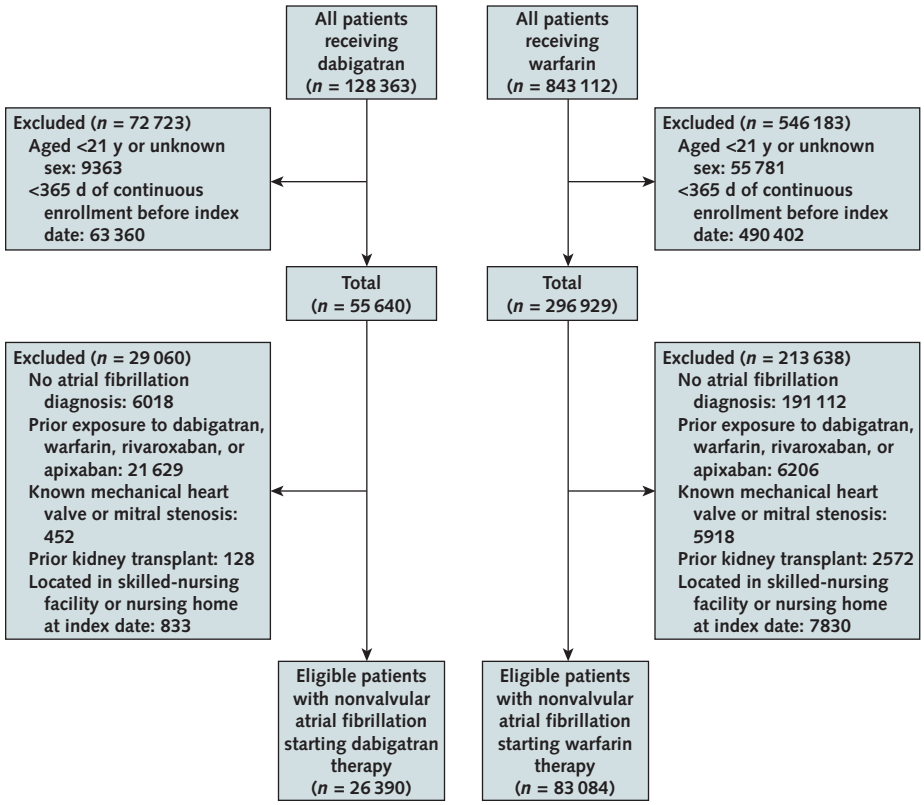
Appendix Table 1. International Classification of Diseases, Ninth Revision, Clinical Modification, Codes Used for Defining Clinical Outcomes

Outcome	Codes
Ischemic stroke	433.x1, 434.x1, 436.xx
Intracranial hemorrhage	430, 431, 432.0, 432.1, 432.9, 852.0x, 852.2x, 852.4x, 853.0
Traumatic	852.0x, 852.2x, 852.4x, 853.0
Major extracranial bleeding	423.0, 455.2, 455.5, 455.8, 456.0, 456.20, 459.0, 530.7, 530.82, 531.0-531.6, 532.0-532.6, 533.0-533.6, 534.0-534.6, 535.01-535.61, 537.83, 562.02, 562.03, 562.12, 562.13, 568.81, 569.3, 569.85, 578.0, 578.1, 578.9, 599.7, 719.11, 784.7, 784.8, and 786.3
Gastrointestinal bleeding	455.2, 455.5, 455.8, 456.0, 456.20, 530.7, 530.82, 531.0-531.6, 532.0-532.6, 533.0-533.6, 534.0-534.6, 535.01-535.61, 537.83, 562.02, 562.03, 562.12, 562.13, 568.81, 569.3, 569.85, 578.0, 578.1, and 578.9
Acute myocardial infarction	410.x0 or 410.x1

Appendix Table 2. Variables Included in Covariate Categories

Category	Included Covariates
Risk factors for bleeding	Prior hospitalized bleeding, other gastrointestinal ulcer disorder, and diagnosed coagulation defects
Risk factors for thromboembolism or myocardial infarction	Age, sex, prior ischemic stroke, transient ischemic attack, other ischemic cerebrovascular events, acute coronary syndrome, percutaneous coronary intervention, coronary artery bypass surgery, other arterial embolism, venous thromboembolism or phlebitis, venous thromboembolism risk indicators, central venous thrombosis, major trauma potentially causing prolonged immobilization, major surgery, chronic heart failure, peripheral arterial disease, hypertension, hyperlipidemia, diabetes mellitus, advanced kidney dysfunction, advanced liver disease, metastatic cancer, alcoholism, smoking, and anemia
Proxy measures of health status and frailty	Number of distinct dispensed medications, number of prior hospitalizations, number of prior physician visits, combined comorbidity score (43), use of home oxygen, wheelchair use, walker use, cane use, commode chair use, prior osteoporotic fracture, and prior mechanical fall
Prescribed medications	Clopidogrel, prasugrel, ticagrelor, ticlopidine, nonsteroidal anti-inflammatory drugs, statins, nonstatin lipid-lowering agents, angiotensin-converting enzyme inhibitors, angiotensin II–receptor blocker, aldosterone receptor antagonists, β-blockers, calcium-channel blockers, prescription H_2-blocker or proton-pump inhibitors, prescription aspirin, antidiabetic drugs, antiarrhythmic drugs, diuretics, other antihypertensives, antiangina vasodilators, estrogens, progestins, selective serotonin receptor inhibitors, selective serotonin reuptake inhibitors and serotonin-norepinephrine reuptake inhibitors, heparin, and low-molecular-weight heparins

Appendix Figure. Study flow diagram.



Appendix Table 3. Baseline Characteristics of Unmatched Cohort

Characteristic	Dabigatran (*n* = 26 390)	Warfarin (*n* = 83 084)	Standardized Difference
Demographic characteristics			
Mean age (SD), *y*	67.9 (11.1)	72.4 (10.8)	0.42
Female, *n (%)*	9271 (35.1)	35 086 (42.2)	0.15
Health service use			
Mean combined comorbidity score (SD)	2.4 (2.2)	3.5 (2.8)	0.42
Mean prior hospitalizations (SD), *n*	0.7 (0.8)	0.9 (1.1)	0.28
Mean physician visits (SD), *n*	10.2 (7.2)	11.0 (8.3)	0.10
Mean unique National Drug Code numbers (SD), *n*	13.9 (8.2)	15.4 (8.9)	0.18
Medical history, *n (%)*			
Advanced kidney dysfunction	2935 (11.1)	18 332 (22.1)	0.30
Advanced liver disease	71 (0.3)	335 (0.4)	0.02
Alcoholism	162 (0.6)	684 (0.8)	0.02
Anemia	1514 (5.7)	7249 (8.7)	0.12
Atrial fibrillation	25 644 (97.2)	80 675 (97.1)	0.00
Atrial flutter	5741 (21.8)	15 282 (18.4)	0.08
Chronic heart failure	9932 (37.6)	38 842 (46.8)	0.19
Coagulation defects	376 (1.4)	3075 (3.7)	0.14
Diabetes mellitus	7770 (29.4)	29 386 (35.4)	0.13
Hospitalized gastrointestinal bleeding	274 (1.0)	2248 (2.7)	0.12
Hospitalized intracranial bleeding	98 (0.3)	664 (0.8)	0.06
Hyperlipidemia	10 417 (39.5)	28 459 (34.3)	0.11
Hypertension	21 374 (81.0)	71 045 (85.5)	0.12
Ischemic stroke	2069 (7.8)	9807 (11.8)	0.13
Metastatic cancer	311 (1.2)	2157 (2.6)	0.10
Myocardial infarction	1242 (4.7)	6556 (7.9)	0.13
Nonspecific cerebrovascular symptoms	419 (1.6)	2283 (2.7)	0.08
Other arterial embolism	217 (0.8)	1710 (2.1)	0.10
Other gastrointestinal ulcer disease	280 (1.1)	1467 (1.8)	0.06
Other hospitalized bleeding	251 (1.0)	1894 (2.3)	0.11
Other ischemic cerebrovascular disease	4532 (17.2)	18 199 (21.9)	0.12
Other ischemic heart disease	1280 (4.9)	5357 (6.4)	0.07
Peripheral vascular disease	4443 (16.8)	20 637 (24.8)	0.20
Smoking and tobacco use	3982 (15.1)	15 862 (19.1)	0.11
Trauma with likely immobilization	1303 (4.9)	6043 (7.3)	0.10
Recent procedures, *n (%)*			
Coronary artery bypass surgery	1821 (7.0)	8572 (10.3)	0.12
Other major surgery	1257 (4.8)	7967 (9.6)	0.19
Percutaneous coronary intervention	2948 (11.2)	11 644 (14.0)	0.09
Frailty indicators, *n (%)*			
Cane use	94 (0.4)	533 (0.6)	0.04
Commode chair use	277 (1.0)	1714 (2.1)	0.08
Home oxygen use	1235 (4.7)	5888 (7.1)	0.10
Osteoporotic fracture	556 (2.1)	3560 (4.3)	0.12
Recent fall	831 (3.1)	4940 (5.9)	0.13
Walker use	664 (2.5)	4126 (5.0)	0.13
Wheelchair use	259 (1.0)	1574 (1.9)	0.08
Medication use, *n (%)*			
Angiotensin-converting enzyme inhibitors	10 157 (38.4)	35 468 (42.7)	0.09
Angiotensin-receptor blockers	5559 (21.1)	15 704 (18.9)	0.05
Antiangina vasodilators	2374 (9.0)	10 749 (12.9)	0.13
Antiarrhythmics	9221 (34.9)	26 488 (31.9)	0.06
Antiplatelets	3472 (13.2)	11 452 (13.8)	0.02
Aspirin	239 (1.0)	2012 (2.4)	0.12
β-blockers	18 910 (71.7)	59 719 (71.9)	0.00
Calcium channel blockers	10 790 (40.9)	33 933 (40.8)	0.00
Diabetes drugs	5357 (20.3)	20 046 (24.1)	0.09
Diuretics	11 975 (45.4)	44 998 (54.2)	0.18
Estrogens	930 (3.5)	2631 (3.2)	0.02
Heparin and low-molecular-weight heparins	372 (1.4)	9999 (12.0)	0.43
Nonsteroidal anti-inflammatory drugs	5667 (21.5)	16 386 (19.7)	0.04
Nonstatin lipid-lowering drugs	3429 (13.0)	8929 (10.7)	0.07
Other antihypertensives	2267 (8.6)	9941 (12.0)	0.11
Proton-pump inhibitors	6591 (25.0)	25 797 (31.0)	0.14
Progestins	328 (1.2)	1235 (1.5)	0.02
Selective serotonin reuptake inhibitors and selective norepinephrine reuptake inhibitors	3854 (14.6)	13 544 (16.3)	0.05
Statins	13 902 (52.7)	46 871 (56.4)	0.08

Appendix Table 4. Number of Patients and Associated Data Extraction Period From Each Participating Data Partner

Data Partner	Start Date	End Date	Years, *n*	Matched Pairs, *n*	Incident Dabigatran Users, *n*	Incident Warfarin Users, *n*
1	11/1/2010	8/30/2013	2.83	5127	5651	9931
2	11/1/2010	12/31/2013	3.17	6193	6363	20 440
3	11/1/2010	9/30/2013	2.92	529	529	10 363
4	11/1/2010	10/31/2013	3.00	5421	5785	12 178
5	11/1/2010	3/31/2014	3.42	6948	6962	23 892
6	11/1/2010	12/30/2013	3.17	314	314	3238
7	11/1/2010	5/31/2014	3.58	368	393	1111
8	11/1/2010	12/31/2012	2.17	389	393	1931

Appendix Table 5. Distribution of Reasons for Censoring, by Treatment Group*

Reason for Censoring	Incident Dabigatran Users (*n* = 25 289)	Incident Warfarin Users (*n* = 25 289)
Clinical outcome	68 (0.3)	67 (0.3)
Discontinuation of index anticoagulant medication	18 537 (73.3)	17 911 (70.8)
Initiation of another anticoagulant	1584 (6.3)	1014 (4.0)
Death and administrative censoring (disenrollment, end of data, admission to nursing home or skilled-nursing facility)	5100 (20.2)	6297 (24.9)

* Values are numbers (percentages).

Appendix Table 6. Association of Dabigatran Versus Warfarin Exposure With Outcomes in Sensitivity Analyses

Outcome	Hazard Ratio (95% CI)		
	Variable-Ratio Matched*	14-Day Grace Period for Classifying Continuous Drug Use (*n* = 25 289 pairs)	Warfarin Use Based on Prescriptions and International Normalized Ratio Testing (*n* = 25 289 pairs)†
Ischemic stroke	0.94 (0.67–1.31)	0.87 (0.64–1.17)	0.83 (0.62–1.12)
Intracranial hemorrhage	0.52 (0.33–0.82)	0.55 (0.38–0.80)	0.53 (0.36–0.78)
Excluding trauma	0.42 (0.24–0.73)	0.49 (0.30–0.79)	0.41 (0.24–0.68)
Combined stroke	0.78 (0.59–1.02)	0.74 (0.58–0.93)	0.71 (0.56–0.90)
Excluding trauma	0.78 (0.58–1.04)	0.75 (0.58–0.97)	0.80 (0.67–0.95)
Major extracranial bleeding	0.91 (0.74–1.10)	0.87 (0.72–1.06)	0.86 (0.72–1.04)
Gastrointestinal	1.04 (0.84–1.28)	1.06 (0.86–1.31)	1.03 (0.84–1.26)
Nongastrointestinal	0.28 (0.16–0.46)	0.32 (0.19–0.54)	0.31 (0.18–0.53)
Myocardial infarction	1.13 (0.78–1.64)	1.43 (0.99–2.08)	1.38 (1.00–1.92)

* Dabigatran, *n* = 25 289. Warfarin, *n* = 83 084.
† Reference 5.

Appendix Table 7. Association of Dabigatran Versus Warfarin Use and Outcomes in Subgroup Analyses

Outcome	Hazard Ratio (95% CI)						
	Age Group				Sex		Reduced Kidney Function (n = 1815 pairs)
	<65 y (n = 9438 pairs)	65-74 y (n = 7334 pairs)	75-84 y (n = 1287 pairs)	≥85 y (n = 2009 pairs)	Male (n = 16 113 pairs)	Female (n = 9143 pairs)	
Ischemic stroke	1.09 (0.55–2.17)	1.10 (0.53–2.30)	0.87 (0.49–01.55)	1.00 (0.41–2.41)	0.86 (0.52–1.40)	1.00 (0.62–1.62)	0.27 (0.06–1.29)
Intracranial hemorrhage	0.39 (0.14–1.11)	0.30 (0.12–0.74)	0.68 (0.34–1.34)	0.67 (0.24–1.83)	0.54 (0.32–0.94)	0.49 (0.24–0.99)	0.72 (0.20–2.54)
Excluding trauma	0.53 (0.18–1.59)	0.19 (0.05–0.65)	0.58 (0.21–1.64)	0.65 (0.17–2.56)	0.51 (0.25–1.02)	0.32 (0.13–0.83)	–
Combined stroke	0.77 (0.44–1.37)	0.64 (0.37–1.12)	0.81 (0.52–1.26)	0.84 (0.43–1.62)	0.71 (0.49–1.03)	0.83 (0.56–1.23)	0.47 (0.18–1.21)
Excluding trauma	0.88 (0.49–1.58)	0.64 (0.35–1.15)	0.82 (0.50–1.37)	0.88 (0.42–1.84)	0.74 (0.49–1.11)	0.81 (0.53–1.23)	0.20 (0.05–0.91)
Major extracranial bleeding	0.51 (0.30–0.87)	0.69 (0.46–1.04)	1.20 (0.86–1.68)	1.60 (0.96–2.69)	1.01 (0.76–1.34)	0.73 (0.54–0.99)	1.59 (0.93–2.72)
Gastrointestinal	0.59 (0.32–1.07)	0.81 (0.52–1.24)	1.47 (1.01–2.14)	1.84 (1.05–3.20)	1.26 (0.92–1.73)	0.78 (0.57–1.07)	1.91 (1.04–3.51)
Nongastrointestinal	0.11 (0.03–0.36)	0.12 (0.03–0.50)	0.29 (0.14–0.61)	0.33 (0.07–1.63)	0.20 (0.10–0.39)	0.22 (0.08–0.58)	0.52 (0.17–1.56)
Myocardial infarction	2.13 (0.98–4.66)	0.97 (0.06–15.56)	4.09 (1.39–12.03)	5.25 (1.17–23.60)	2.06 (1.17–3.64)	1.69 (0.84–3.38)	2.18 (0.20–24.18)

Appendix Table 8. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Patients Younger Than 65 Years With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 9438)		Warfarin (n = 9438)		Hazard Ratio (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	18	0.62	15	0.62	1.09 (0.55–2.17)
Intracranial hemorrhage	5	0.17	12	0.50	0.39 (0.14–1.11)
Excluding trauma	5	0.17	9	0.37	0.53 (0.18–1.59)
Combined stroke	22	0.76	26	1.08	0.77 (0.44–1.37)
Excluding trauma	22	0.76	23	0.95	0.88 (0.49–1.58)
Major extracranial bleeding	21	0.72	37	1.54	0.51 (0.30–0.87)
Gastrointestinal	18	0.62	27	1.12	0.59 (0.32–1.07)
Nongastrointestinal	1	0.04	9	0.37	0.11 (0.03–0.36)
Myocardial infarction	22	0.76	9	0.38	2.13 (0.98–4.66)

Appendix Table 9. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Patients Aged 65–74 Years With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 7334)		Warfarin (n = 7334)		Hazard Ratio (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	16	0.62	13	0.59	1.10 (0.53–2.30)
Intracranial hemorrhage	6	0.23	19	0.87	0.30 (0.12–0.74)
Excluding trauma	3	0.12	5	0.23	0.19 (0.05–0.65)
Combined stroke	22	0.86	31	1.42	0.64 (0.37–1.12)
Excluding trauma	19	0.74	27	1.24	0.64 (0.35–1.15)
Major extracranial bleeding	41	1.60	54	2.48	0.69 (0.46–1.04)
Gastrointestinal	39	1.52	44	2.02	0.81 (0.52–1.24)
Nongastrointestinal	1	0.04	10	0.46	0.12 (0.03–0.50)
Myocardial infarction	11	0.43	19	0.88	0.54 (0.26–1.13)

Appendix Table 10. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Patients Aged 75–84 Years With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 6411)		Warfarin (n = 6411)		Hazard Ratio (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	23	0.98	24	1.25	0.87 (0.49–01.55)
Intracranial hemorrhage	15	0.64	19	0.99	0.68 (0.34–1.34)
Excluding trauma	6	0.26	9	0.47	0.58 (0.21–1.64)
Combined stroke	38	1.63	42	2.20	0.81 (0.52–1.26)
Excluding trauma	29	1.24	32	1.67	0.82 (0.50–1.37)
Major extracranial bleeding	80	3.44	61	3.20	1.20 (0.86–1.68)
Gastrointestinal	71	3.06	44	2.32	1.47 (1.01–2.14)
Nongastrointestinal	6	0.26	16	0.84	0.29 (0.14–0.61)
Myocardial infarction	20	0.86	4	0.21	4.09 (1.39–12.03)

Appendix Table 11. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Patients Aged 85 Years and Older With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 2009)		Warfarin (n = 2009)		Hazard Ratio (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	10	0.47	10	0.55	1.00 (0.41–2.41)
Intracranial hemorrhage	7	0.33	9	0.50	0.67 (0.24–1.83)
Excluding trauma	4	0.19	5	0.28	0.65 (0.17–2.56)
Combined stroke	17	0.81	19	1.05	0.84 (0.43–1.62)
Excluding trauma	14	0.67	15	0.83	0.88 (0.42–1.84)
Major extracranial bleeding	39	1.87	23	1.28	1.60 (0.96–2.69)
Gastrointestinal	37	1.78	19	1.06	1.84 (1.05–3.20)
Nongastrointestinal	1	0.05	2	0.11	0.33 (0.07–1.63)
Myocardial infarction	13	0.63	2	0.11	5.25 (1.17–23.60)

Appendix Table 12. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Adult Men With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome	Dabigatran (n = 16 113)		Warfarin (n = 16 113)		Hazard Ratio (95% CI)
	Patients With Events, n	Incidence Rate per 100 Person-Years	Patients With Events, n	Incidence Rate per 100 Person-Years	
Ischemic stroke	31	0.58	33	0.73	0.86 (0.52–1.40)
Intracranial hemorrhage	21	0.39	35	0.73	0.54 (0.32–0.94)
Excluding trauma	12	0.23	22	0.49	0.51 (0.25–1.02)
Combined stroke	51	0.96	65	1.44	0.71 (0.49–1.03)
Excluding trauma	42	0.79	52	1.15	0.74 (0.49–1.11)
Major extracranial bleeding	101	1.90	94	2.09	1.01 (0.76–1.34)
Gastrointestinal	90	1.70	67	1.49	1.26 (0.92–1.73)
Nongastrointestinal	8	0.15	26	0.58	0.20 (0.10–0.39)
Myocardial infarction	42	0.79	17	0.38	2.06 (1.17–3.64)

Appendix Table 13. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Adult Women With Atrial Fibrillation Receiving Dabigatran or Warfarin

Outcome Type	Dabigatran (n = 9143)		Warfarin (n = 9143)		Hazard Ratio (95% CI)
	Patients With Events, *n*	Incidence Rate per 100 Person-Years	Patients With Events, *n*	Incidence Rate per 100 Person-Years	
Ischemic stroke	36	1.13	32	1.26	1.00 (0.62–1.62)
Intracranial hemorrhage	12	0.38	21	0.83	0.49 (0.24–0.99)
Excluding trauma	6	0.19	16	0.63	0.32 (0.13–0.83)
Combined stroke	48	1.51	51	2.01	0.83 (0.56–1.23)
Excluding trauma	42	1.33	46	1.81	0.81 (0.53–1.23)
Major extracranial bleeding	80	2.53	96	3.79	0.73 (0.54–0.99)
Gastrointestinal	75	2.38	84	3.32	0.78 (0.57–1.07)
Nongastrointestinal	1	0.03	10	0.40	0.22 (0.08–0.58)
Myocardial infarction	24	0.76	12	0.48	1.69 (0.84–3.38)

Appendix Table 14. Frequency and Rates of Thromboembolism, Intra- and Extracranial Hemorrhage, and Acute Myocardial Infarction: Patients With Atrial Fibrillation and Reduced Kidney Function Receiving Dabigatran or Warfarin

Outcome Type	Dabigatran (n = 1815)		Warfarin (n = 1815)		Hazard Ratio (95% CI)
	Patients With Events, *n*	Incidence Rate per 100 Person-Years	Patients With Events, *n*	Incidence Rate per 100 Person-Years	
Ischemic stroke	2	0.37	8	1.55	0.27 (0.06–1.29)
Intracranial hemorrhage	4	0.73	6	1.7	0.72 (0.20–2.54)
Excluding trauma	0	0.00	3	0.58	–
Combined stroke	6	1.10	14	2.72	0.47 (0.18–1.21)
Excluding trauma	2	0.37	11	2.14	0.20 (0.05–0.91)
Major extracranial bleeding	35	6.50	22	4.32	1.59 (0.93–2.72)
Gastrointestinal	30	5.58	16	3.15	1.91 (1.04–3.51)
Nongastrointestinal	2	0.37	5	0.98	0.52 (0.17–1.56)
Myocardial infarction	2	0.37	1	0.20	2.18 (0.20–24.18)

Lithium for Suicidal Behavior in Mood Disorders (Li+)

ClinicalTrials.gov Identifier: NCT01928446

Recruitment Status : Recruiting
First Posted : August 26, 2013
Last Update Posted : February 15, 2018

See **Contacts and Locations**

Sponsor:
VA Office of Research and Development

Information provided by (Responsible Party):
VA Office of Research and Development

Study Details | **Tabular View** | **No Results Posted** | Disclaimer | How to Read a Study Record

Study Description

Go to

Brief Summary:

Observational evidence and findings from clinical trials conducted for other reasons suggest that **lithium**, a drug used for the treatment of bipolar disorder, and, to a lesser extent, depression, may

reduce rates of suicides and suicide attempts. However, this hypothesis has not yet been adequately examined in a randomized clinical trial conducted specifically to test **lithium's** efficacy in preventing suicides. This clinical trial fills this gap.

This study is feasible within the Department of Veterans Affairs (VA) because it is a large, integrated health system with existing programs for identifying patients at risk for suicide and delivering enhanced services. In VA, approximately 12,000 patients with depression or bipolar disorder survive a suicide attempt or related behavior each year, and 15% of them repeat within one year. Experimental treatment in this study will supplement usual care for major depression or bipolar disorder, as well as VA's standard, enhanced management for patients at high risk.

The investigators will recruit 1862 study participants, from approximately 30 VA Hospitals. Participants will be patients with bipolar disorder or depression who have survived a recent episode of suicidal self-directed violence or were hospitalized specifically to prevent suicide. Randomly, half will receive **lithium**, and half will receive placebo. Neither the patients nor their doctors will know whether a particular person has received **lithium** or placebo. The treatment will be administered and the patients will be followed for one year, after which patients will go back to usual care. Recruitment will occur over 3 years.

The investigators are primarily interested in whether **lithium** leads to increases in the time to the first repeated episode of suicidal behavior, including suicide attempts, interrupted attempts, hospitalizations specifically to prevent suicide, and deaths from suicide. In addition, this study will allow us to explore whether **lithium** decreases the total number of suicidal behaviors, and whether it has comparable effects on impulsive and non-impulsive behaviors. If there is an effect of **lithium**, the investigators will be interested in whether or not it could be attributed to improved control of the underlying mental health condition, or, alternatively, whether it represents a direct effect of suicide-related behavior.

Condition or disease	Intervention/treatment	Phase
Depressive Disorder Bipolar Disorder **Suicide** **Suicide**, Attempted	Drug: **Lithium** Drug: Placebo	Phase 2 Phase 3

Hide Detailed Description

Detailed Description:

Objective: To test the hypothesis that lithium augmentation of enhanced usual care will reduce the rate of repeated episodes of suicidal self-directed violence (repeated suicide attempts, interrupted attempts, hospitalizations specifically to prevent suicide, and deaths from suicide) in participants with bipolar disorder or depression who have survived a recent event.

Background: The hypothesis that lithium can prevent suicide in patients with bipolar disorder and depression is based on data from observational studies and randomized clinical trials conducted to evaluate other outcomes. The question about the effectiveness of lithium for suicide prevention is one of major scientific, clinical, and public health significance. There have been no adequately powered clinical trials conducted specifically to evaluate suicide behaviors as an outcome. Two recent randomized clinical trials failed to recruit adequate numbers of subjects to be conclusive.

The VHA, as a large national healthcare system with an established program for identifying new suicide attempts, evaluating patients for underlying mental health and medical conditions, providing needed services, connecting Veterans to state-of-the-art suicide risk management, and monitoring outcomes is uniquely able to conduct a large scale clinical trial of lithium for suicide prevention.

The rationale for the study is based on the following:

- Data from observational studies and double-blind randomized clinical trials suggest that lithium can prevent suicide-related behaviors in patients with bipolar disorder and major depression.
- The high risk of suicide in Veterans receiving health care services from VHA has persisted despite extensive improvements in mental health services and in programs for suicide prevention.
- Each month, there are over 1,100 unique VHA patients with bipolar disorder or depression who attempt suicide and survive.
- Surviving a suicide attempt is the most powerful known risk factor for death from suicide in VA and elsewhere.
- Approximately 15% of VA survivors reattempt or die from suicide within one year.
- Evaluating rates of reattempts in those who have survived attempts is an established and effective method for testing interventions that may prevent suicide.
- Experimental treatment in CSP-590 would supplement usual care for major depression or bipolar disorder.
- Study procedures for the management of suicide risk would meet or exceed VA standards and requirements.
- Study procedures optimize the safety of lithium, including the potential risk of overdoses, and meet or exceed all published practice standards. The trial will utilize multiple strategies to minimize risks including frequent monitoring and assessment, determination of lithium levels during titration and at steady state, and dispensing medications in limited quantities in blister packs.
- The investigator's survey of VA psychiatrists indicates that the question is clinically important and compelling and that a clinical trial that demonstrated the hypothesized effect would transform the clinical management of suicidality.

Design: Randomized, double-blind, placebo-controlled clinical trial of lithium versus placebo augmentation of enhanced usual care.

Patient population: VHA patients with bipolar disorder or depression who have survived a recent episode of suicidal self-directed violence.

Primary outcome: Time to the first repeated episode of suicidal self-directed violence, including suicide attempts, interrupted attempts, hospitalizations specifically to prevent suicide, and deaths from suicide

Duration: Total study duration will be 4.5 years. Recruitment will occur over 3 years. Participants will be followed for one year.

Sample size calculations and number of sites required: The design of the study is based on testing for a 37% reduction in the rate of repeated suicidal self-directed violence, a figure based on an effect size of approximately 43% observed in recent studies and then allowing for attenuation due to non-adherence. Adjusting for potential data loss due to attrition, 90% statistical power to detect a significant 37% reduction in reattempt rates at 5% overall type I error would require 1862 subjects. With recruitment of 20% of eligible subjects over a three year period, this would require approximately 9310 potentially eligible subjects. Based on current suicide surveillance data, this could be achieved with 29 sites.

Study Design

Go to

Study Type :	Interventional (Clinical Trial)
Estimated Enrollment :	1862 participants
Allocation:	Randomized
Intervention Model:	Parallel Assignment
Masking:	Quadruple (Participant, Care Provider, Investigator, Outcomes Assessor)
Primary Purpose:	Prevention
Official Title:	CSP #590 - **Lithium** for Suicidal Behavior in Mood Disorders
Actual Study Start Date :	July 8, 2015
Estimated Primary Completion Date :	May 31, 2019
Estimated Study Completion Date :	August 30, 2019

Resource links provided by the National Library of Medicine

Genetics Home Reference related topics: Bipolar disorder

MedlinePlus related topics: Mood Disorders Suicide

Drug Information available for: Lithium carbonate Lithium citrate

U.S. FDA Resources

Arms and Interventions

Go to

Arm	Intervention/treatment
Experimental: **Lithium** **Lithium** in the form of extended release **lithium** carbonate. Subjects will be started on 600 mg/day (300mg bid) until steady state at target plasma levels between 0.6 and 0.8 meq/liter is achieved. The lowest dose will be 300 mg/day. **Lithium** will be prescribed for the duration of follow-up (1 year).	Drug: **Lithium** **Lithium** in the form of extended release **lithium** carbonate. Subjects will be started on 600 mg/day (300mg bid) until steady state at target plasma levels between 0.6 and 0.8 meq/liter is achieved. The lowest dose will be 300 mg/day. **Lithium** will be prescribed for the duration of follow-up (1 year).
Placebo Comparator: Placebo Placebo tablets will be given to the patients for the duration of follow-up (1 year). Dose adjustments will mimic the intervention arm of the study	Drug: Placebo Oral placebo tablets will be administered for the duration of follow-up (1 year).

Outcome Measures

Go to

Primary Outcome Measures :

1. Time to the first repeated episode of suicidal self-directed violence, including suicide attempts, interrupted attempts and hospitalizations for prevention of attempts. [Time Frame: Primary outcome is assessed from randomization up to 12 months.]

   The primary hypothesis tested is that lithium augmentation of enhanced usual care is superior to enhanced usual care plus placebo for the prevention of repeated episodes of suicidal self-directed violence over time. The investigators posit a one-year repeat rate of 15% in the placebo group and a 37% reduction of events in the intervention group.

   Suicidal self-directed violence includes non-fatal suicide attempts, interrupted attempts (attempts interrupted by patient or by others), hospitalization to prevent suicide and deaths from suicide.

Eligibility Criteria

Go to

Information from the National Library of Medicine

Choosing to participate in a study is an important personal decision. Talk with your doctor and family members or friends about deciding to join a study. To learn more about this study, you or your doctor may contact the study research staff using the contacts provided below. For general information, Learn About Clinical Studies.

Ages Eligible for Study:	Child, Adult, Senior
Sexes Eligible for Study:	All
Accepts Healthy Volunteers:	No

Criteria

Inclusion Criteria:

- Must be a Veteran of the United States Armed Forces
- Survived an episode of suicidal self-directed violence (including suicide attempts and interrupted attempts) that occurred within six months of admission to the study, or they were admitted within the past six months to a mental health inpatient unit specifically to prevent suicide
- Have a diagnosis of an affective disorder meeting DSM-IV-TR (2000) criteria for Bipolar I Disorder, Bipolar II Disorder, or current or recurrent Major Depressive Disorder
- Are able and willing to identify one or more family members, friends, or other contacts and give permission for both clinical providers and the Research Team to contact them if the patient cannot be reached
- Are able to provide informed consent
- There is concurrence from the patient's mental health provider about inclusion/exclusion criteria and confirmation of the providers' willingness to work with the research team in managing the patient during the course of the study. The provider responsible for the patient's general medical care has been made aware of the participation
- Must be registered at a VA Medical Center

Exclusion Criteria:

- Schizophrenia or schizoaffective disorder
- Cognitive impairment defined as a Brief Orientation Memory and Concentration Test score > 10
- Lack of decision-making capacity to evaluate the risks versus the benefits of participation as determined by Jeste's brief instrument for assessing decisional capacity, or adjudication of incompetence and the appointment of a guardian or conservator
- Six or more previous lifetime suicide attempts as ascertained through SPAN, reports from family, or patient self-report
- Current or recent (within six months) use of lithium

- History of significant adverse effects of lithium as ascertained through the medical record or self-report
- Unstable medical conditions or specific medical comorbidity:
  - Congestive heart failure by Framingham criteria
  - QTc greater than or equal to 450 ms for men and greater than or equal to 460 ms for women
  - Chronic renal failure defined by national Kidney Foundation Disease Outcome Quality Initiative (KDOQI) criteria
- Any possibility of being pregnant or not on appropriate birth control
- Lactation and breastfeeding
- Concurrent medications:
  - All diuretics except amiloride
  - Haloperidol
  - Clozapine
- Active substance abuse:
  - Active alcohol or opiate dependence requiring medically supervised withdrawal and stabilization
  - Active cocaine, methamphetamine, other stimulant, hallucinogen, or cannabis abuse requiring stabilization
- Enrollment in another randomized interventional clinical trial

Contacts and Locations

Go to

Information from the National Library of Medicine

To learn more about this study, you or your doctor may contact the study research staff using the contact information provided by the sponsor.

Please refer to this study by its ClinicalTrials.gov identifier (NCT number): ***NCT01928446***

Contacts

Contact: Matthew H Liang, MD	(857) 364-6116	Matthew.Liang@va.gov
Contact: Melynn Nuite, RN BS CCRC	(617) 232-9500	Melynn.Nuite@va.gov

Show 28 Study Locations

Sponsors and Collaborators

VA Office of Research and Development

Investigators

Study Chair: Ira R Katz, MD PhD Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA

More Information

Go to

Responsible Party:	VA Office of Research and Development
ClinicalTrials.gov Identifier:	NCT01928446 History of Changes
Other Study ID Numbers:	590
First Posted:	August 26, 2013 Key Record Dates
Last Update Posted:	February 15, 2018
Last Verified:	February 2018

Individual Participant Data (IPD) Sharing Statement:

Plan to Share IPD:	Yes
Plan Description:	Individual Participant Data will be made available after study closure only to research credentialed Veterans Affairs researchers who submit a valid study question to their IRB of record. A Data Use Agreement will be in effect between the researcher and the coordinating center
URL:	http://

Studies a U.S. FDA-regulated Drug Product:	Yes
Studies a U.S. FDA-regulated Device Product:	No
Product Manufactured in and Exported from the U.S.:	Yes

Keywords provided by VA Office of Research and Development:

Lithium
Placebo
Double-blind methods
Clinical Trials, Randomized
Veterans

Additional relevant MeSH terms:

Suicide
Suicide, Attempted
Lithium Carbonate
Disease
Depressive Disorder
Depression
Bipolar Disorder
Mood Disorders
Pathologic Processes
Mental Disorders
Behavioral Symptoms
Bipolar and Related Disorders
Self-Injurious Behavior
Antidepressive Agents
Psychotropic Drugs
Enzyme Inhibitors
Molecular Mechanisms of Pharmacological Action
Antimanic Agents
Tranquilizing Agents
Central Nervous System Depressants
Physiological Effects of Drugs

Office of Research & Development

VA RESEARCH CURRENTS
Research News from the U.S. Department of Veterans Affairs

Go Back

New studies

Can lithium help stem suicide rate? VA study aims to find out

September 17, 2014

As a drug, lithium has been around since the 1800s. Made from a whitish mineral found in rocks, the drug is widely used today as a mood stabilizer, especially for those with bipolar disorder.

Some studies suggest it may also be useful for preventing suicide. But the theory needs further testing.

Enter VA's Cooperative Studies Program (CSP). The program is gearing up to launch a major trial involving more than 1,800 Veterans from 28 VA medical centers. The study will include only those with bipolar disorder or depression who recently survived a suicide attempt, or were hospitalized to prevent one.

A counselor in action at VA's Center of Excellence for Suicide Prevention, which houses the Veterans Crisis Line and other services. The center will play a role in a VA clinical trial of the drug lithium for suicide prevention. *(Photo by Robert Turtil)*

Some 12,000 VA patients with bipolar disorder or depression survive a suicide attempt every year. Experts say such patients remain at increased risk of suicide for the rest of their lives.

The new study will enroll Veterans for three years and follow each patient one year. The study team will look at outcomes such as repeat suicide attempts and hospitalizations to prevent suicide, as well as deaths from suicide. Half the study volunteers in the randomized, double-blinded trial will get a form of the drug known as lithium carbonate, in an extended-release tablet to minimize side effects. The other half will get a placebo.

All will get VA's standard mental health care, plus extra care coordination: The study team will follow up with each patient throughout the study, and give regular updates to other care providers.

VA Research Currents spoke with three members of the study team to learn more about the research effort, which was announced by President Obama at the American Legion national convention in August 2014.

Study chair **Dr. Ira Katz,** a psychiatrist based in Philadelphia, is a senior consultant for VA's Office of Mental Health Operations.

Study director **Dr. Matt Liang** is a "trialist" at the Boston CSP Coordinating Center who has led close to 30 clinical trials in his career, some with VA but most through the National Institutes of Health and the Centers for Disease Control and Prevention. He is also an internist with Brigham and Women's Hospital and a professor at Harvard Medical School.

Study project manager **Natalie Morgenstern** is a health science specialist for the Boston CSP Coordinating Center.

RELATED ARTICLES

- Research rocks for homeless Vets at the Louisville VA
- PTSD and accelerated aging: How advanced is the science?
- Drug that acts like ketamine—but without the potential for abuse or psychotic effects—eases depression in lab tests
- Program improves cognition, psychiatric symptoms but falls short on

employment front

- Down on the farm

Why study lithium?

Katz: The observation that patients on lithium were less likely to die from suicide has been around at least since the 1980s or early 1990s. It's been seen mainly in studies in which lithium was being used to treat bipolar disorder or depression. There have also been epidemiologic studies in different parts of the world that have found a correlation between lower suicide rates and higher levels of lithium in the drinking water.

This has all helped lay the groundwork for our study. We're looking directly at whether pharmacologic doses of lithium do indeed have an anti-suicide effect.

This has been a confusing and difficult issue from a research standpoint because lithium has a substantial number of side effects, and it is dangerous in over-dosage. In fact, it can be fatal. So the question has always been this: When we see that patients on lithium are less likely to commit suicide, is it because lithium really does have anti-suicide effects, or is it because doctors don't dare give lithium to anyone who is at risk for suicide, for fear they might intentionally overdose?

That sort of puzzle, in which it *looks* like lithium has anti-suicide effects, but it may be because the people who are prescribed the drug are at less risk to begin with, is called an indication bias. That's been the major problem in interpreting the observational studies.

To the extent researchers have been able to tease out the answer from database studies, they've argued that the drug probably does have anti-suicide effects. But the gold standard to know about causality and to establish the effectiveness of a treatment is a randomized clinical trial, and that's what we're doing.

What is the study team doing to ensure lithium is used safely?

Liang: Lithium has a narrow toxicity-efficacy ratio. That means there's a relatively small difference between the effective therapeutic dose and the higher dose that would be toxic. So we've tried very hard to minimize the possibility of harm. We're being extremely careful about who we include, and how we monitor the lithium levels and potential side effects. We have two central consultants, Dr. Malcolm Rogers, a psychiatrist, and Dr. Chester Conrad, a cardiologist, who will be available 24/7 to assist the sites. The protocol meets or exceeds any published guideline on safe usage, and is probably over and above what is done in normal clinical practice, in VA or the general psychiatry community.

Morgenstern: We're also using an extended-release form of the drug. The coating tends to make it more tolerable and decrease the severity and frequency of side effects.

Do scientists understand how lithium might work in the brain to prevent suicide?

Katz: We have some insight on this from looking at what happens in clinical populations receiving lithium. Suicide rates seem to be lower both in patients for whom lithium has worked well to treat the underlying psychiatric symptoms, and in those for whom lithium has worked less well. And that's led to the notion that the drug may have anti-suicide effects that are independent of its effect on depression or bipolar disease.

So while the primary goal of the study is to see if lithium prevents repeat suicide attempts, one of the secondary goals is to determine, if we do have a lithium effect, whether it is due to better control of the symptoms of depression or bipolar disorder. That's an important secondary analysis.

But we really don't know in-depth the cellular or molecular mechanisms of lithium for treatment of bipolar disorder or depression, and we know even less about what could explain its possible anti-suicide effects. We suspect if this study is positive, and we have definite evidence that a drug can prevent suicidal behavior, it will stimulate a good deal of pharmacologic research trying to look for other medications that may have a comparable effect.

How is the study going to recruit participants?

Morgenstern: One source will be referrals from clinicians. There'll also be some targeted outreach, and limited advertising in the form of flyers around the VA medical centers that are taking part. We chose those centers that had higher numbers of patients with documented past suicide attempts. We determined that by using the SPAN database. [SPAN is VA's Suicide Prevention and Application Network, coordinated out of the Center of Excellence for Suicide Prevention at the VA in Canandaigua, New York.] We'll also have access to the screening logs that the sites use, and then we can supplement them with people that they might be missing, so they can try and target those people as well. SPAN will be an invaluable asset. We couldn't do the study without it.

All VA medical centers have suicide prevention coordinators. What role will they play?

Liang: We're collaborating with the suicide prevention coordinators, but we're trying to do it in the most ethical manner possible. They will let potential participants know about the study but won't explain it in any great detail, or actually enroll them. We didn't want any care providers to be conflicted, or to potentially be coercive to people who are potentially vulnerable after a suicide attempt.

If lithium proves effective for suicide prevention, will patients be able to stay on the drug long- term?

Katz: The study itself is one year. In terms of how lithium might be used if we demonstrate an effect, it's important to note that many people with bipolar disorder have been on lithium for or 10, 20, 30 years and managing quite well. We also know that people who have survived a suicide attempt can be at increased risk for suicide for the rest of their life. On the other hand, there are concerns that long-term use of lithium may lead to decreased kidney function.

So the first question will be whether lithium is effective over the time period of the year. A downstream question will be what the risks versus the benefits are of its use over the long term.

Will the patients in the study be followed longer than one year?

Katz: Because they are in the VA system, we'll be able to keep an eye on these patients over the longer term. VA already has an infrastructure and a system for tracking suicide-related behaviors, mainly through the suicide prevention coordinators, who are funded separately from the study. We have that system of care in place, and it is one of the unique benefits of VA.

What might be some next steps after the study ends?

*Katz:*This study is looking at the effects of lithium in doses that are used pharmacologically. If this is positive, our next question might be whether you need such high doses, or whether far lower doses might also be effective. However, the most important question is about how we would translate findings from the study into improved care. For this, we would make sure that mental health staff and other care providers in VA are aware of the results, and that all VA psychiatrists know how to use lithium to prevent suicide.

Liang: This is the first real test of lithium for suicidality. If the results are positive, it will open up a number of opportunities for understanding how the finding might be applied to a broader population, both in and beyond VA.

To learn more about the trial, expected to launch this fall, click here or go to clinicaltrials.gov and enter the search term CSP 590.

Questions about the R&D website? Email the Web Team.

JAMA Internal Medicine | Original Investigation

Use of Health Care Databases to Support Supplemental Indications of Approved Medications

Michael Fralick, MD; Aaron S. Kesselheim, MD, JD, MPH; Jerry Avorn, MD; Sebastian Schneeweiss, MD, ScD

Invited Commentary page 63

Supplemental content

IMPORTANCE Manufacturers of US Food and Drug Administration–approved prescription drugs often apply for additional indications based on randomized clinical trials. Real-world database analyses on a medication's use and outcomes in routine settings of care might help to inform decision making regarding such supplemental indications.

OBJECTIVE To examine whether longitudinal data from a health care database can support the results of a randomized clinical trial that led to a supplemental indication for telmisartan.

DESIGN, SETTING, AND PARTICIPANTS This cohort study of patients newly prescribed telmisartan or ramipril used insurance claims data from a nationwide health care database from January 1, 2003, through September 30, 2009, to compare patient outcomes. This study replicated the inclusion and exclusion criteria used in the Ongoing Telmisartan Alone and in Combination with Ramipril Global End-point Trial (ONTARGET) and used propensity score matching to balance 74 patient characteristics. Data analysis was performed from February 15, 2017, to May 24, 2017.

EXPOSURES Telmisartan use vs ramipril use.

MAIN OUTCOMES AND MEASURES The primary outcome was a composite of myocardial infarction, stroke, or hospitalization for congestive heart failure.

RESULTS Of the 640 951 patients included in the study, 48 053 were newly prescribed ramipril (mean [SD] age, 68.29 [9.52] years; 31 940 male [66.5%]) and 4665 were newly prescribed telmisartan (mean [SD] age, 69.43 [9.60] years; 2413 male [51.7%]). After propensity score matching, a total of 4665 patients were newly prescribed telmisartan (mean [SD] age, 69.43 [9.60] years; 2413 [51.7%]), and 4665 patients were newly prescribed ramipril (mean [SD] age, 69.36 [9.67] years; 2343 male [50.2%]). As seen in ONTARGET, the composite risk of stroke, myocardial infarction, and hospitalization for congestive heart failure was similar for the 2 medications (hazard ratio, 1.0; 95% CI, 0.9-1.1). In addition, the study found that telmisartan was associated with a substantially decreased risk of angioedema (hazard ratio, 0.1; 95% CI, 0.03-0.56) compared with ramipril.

CONCLUSIONS AND RELEVANCE Real-world data analyses of patients receiving routine care provided findings similar to those found in the randomized clinical trial that established telmisartan's supplemental indication. In certain situations, database studies may support supplemental applications for effectiveness for already approved medications.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Michael Fralick, MD, The Program on Regulation, Therapeutics, and Law, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont St, Ste 3030, Boston, MA 02120 (mif823@mail.harvard.edu).

JAMA Intern Med. 2018;178(1):55-63. doi:10.1001/jamainternmed.2017.3919
Published online November 20, 2017.

In December 2016, the 21st Century Cures Act was signed into law in the United States.[1] It contained a provision intended to promote real-world data studies of medication use and outcomes in routine clinical settings in US Food and Drug Administration (FDA) authorization of additional indications for already approved prescription drugs.[1] Such data, with or without randomization, are drawn from health care use data, insurance claims, registry studies, and/or electronic health record systems in typical clinical settings of care.[2-4] Although the FDA has long used such data to clarify the safety of medications, the data can seldom establish a drug's effectiveness. Well-designed randomized clinical trials are the criterion standard for assessing whether a drug is efficacious because random treatment assignment and a controlled research environment can more readily support causal inferences.

In recent years, new methodologic approaches have improved the validity and reproducibility of nonrandomized data, including new-user designs,[5] active comparators, propensity score (PS) matching, and controlling for disease risk scores.[6,7] Other important aspects include assessing covariates before cohort entry (to avoid adjusting for intermediate variables) and defining cohort entry as the time when the patient first receives the exposure of interest (to decrease the possibility of immortal time bias).[6-8]

Can such analytic techniques confirm supplemental indications for already approved drugs? Approximately half of all drugs approved in the United States are later approved for supplemental indications, modifications to the initial indication, or expanded populations.[9,10] Supplemental indications are typically identified on the basis of prospective clinical trials. To determine whether real-world data analyses can confirm a supplemental indication, we identified a supplemental approval amenable to study and applied the same inclusion and exclusion criteria and outcomes measurements that were used in the pivotal randomized clinical trial.

Key Points

Question Can health care databases be used to confirm a supplemental indication that has been demonstrated in a randomized clinical trial for an approved medication?

Findings This cohort study replicated the results of a randomized clinical trial that established the supplemental indication for telmisartan by using data from a US health care database (insurance claims data) available at the time that the supplemental indication was approved. Similar to the randomized clinical trial, our study revealed a decreased risk of angioedema with telmisartan compared with ramipril.

Meaning In certain clinical scenarios, database studies may support supplemental effectiveness applications for already approved medications.

Methods

Our cohort study was conducted in commercially insured patients using the MarketScan health care database provided by Truven (January 1, 2003, through September 30, 2009). This nationwide database captures anonymized longitudinal, individual-level data on health care use, patient demographics, inpatient and outpatient diagnostic and procedural codes, and pharmacy dispensing of prescription drugs for more than 60 million commercially insured people in the United States. The study was approved by the institutional review board at Brigham and Women's Hospital, including a waiver for informed consent, and a valid data licensing agreement was in place. All data were anonymized and deidentified.

Data Sources

To identify an experimental setting, we reviewed all supplemental applications to the FDA from 2005 to 2014 and their accompanying clinical trials.[9] The supplemental indications were classified into 3 mutually exclusive categories: new indication (n = 138), modification (n = 86), and expansion (n = 66) (eAppendix in the Supplement).[9] Of the 138 new indications, 108 (78.3%) of the pivotal clinical trials had a primary outcome that was not identifiable in US longitudinal health care databases (eg, pathology results, change in clinical scores, and radiologic tumor response), 12 (8.7%) did not have an active comparator, 4 (2.9%) were based on in-hospital medication administration (eg, postoperative nausea medication, anesthetic medications), and 14 (10.1%) were potentially replicable with the claims data available to us. Of the 14, we selected telmisartan a priori and did not analyze data for the other 13 (eAppendix in the Supplement).

The angiotensin receptor blocker (ARB) telmisartan (Micardis) was approved as an antihypertensive in 1998. In October 2009, it was approved supplementarily for cardiovascular risk reduction in patients 55 years or older who are unable to take angiotensin-converting enzyme inhibitors (ACE-Is) and have a high risk of major cardiovascular events. Telmisartan was an optimal case study for 3 reasons. First, the primary outcome in the pivotal supplemental indication trial could be accurately identified in health care use data. Second, the randomized clinical trial used an active comparator, the ACE-I ramipril (Altace), which would minimize confounding in cohort studies.[3,7] Third, the inclusion criteria, exclusion criteria, and baseline patient characteristics were identifiable in claims data.

The trial that identified the supplemental indication for telmisartan for cardiovascular risk reduction, Ongoing Telmisartan Alone and in Combination with Ramipril Global Endpoint Trial (ONTARGET), was published in April 2008.[11] ONTARGET's primary objectives were to determine whether telmisartan was at least as effective as ramipril at reducing cardiovascular risk and to assess whether the combination of telmisartan and ramipril was more effective than ramipril alone. The trial was conducted across 733 centers in 40 countries between 2001 and 2008.[11]

Study Cohort

Potentially eligible patients must have had at least 6 months of continuous enrollment in a participating health plan before the date of cohort entry. Our inclusion and exclusion cri-

teria mirrored those of ONTARGET.[11] We included patients 55 years or older who filled a new prescription for telmisartan or ramipril (no fills for either drug or any other ACE-I or ARB during the prior 180 days). Cohort entry date was the first day of a prescription fill. As in ONTARGET, we included patients with a diagnosis of coronary artery disease, peripheral artery disease, cerebrovascular disease, or diabetes mellitus during the 180 days before cohort entry.

As in ONTARGET, we excluded patients with a limited life expectancy (ie, living in a hospice, palliative care facility, or a nursing home and those with cancer), liver disease, syncope or a recent myocardial infarction (within 2 days of cohort entry), transient ischemic attack (within 7 days of cohort entry), percutaneous transluminal coronary angiography (within 30 days of cohort entry), or hospitalization for congestive heart failure during the 180 days before cohort entry. Other exclusion criteria used in ONTARGET were not applied because they were not readily identifiable (known allergy to study medication, unable to tolerate study medication, hemodynamically significant primary valvular or outflow tract obstruction, uncorrected volume or sodium depletion, planned cardiac procedure, blood pressure >160/100 mm Hg despite treatment, significant renal artery stenosis, and angina in the absence of multivessel coronary artery disease) or rare (hereditary fructose intolerance, complex congenital heart disease, primary hyperaldosteronism, and heart transplant). We also excluded patients who previously received any ACE-I or ARB.

Study End Point

Our primary outcome was a composite of myocardial infarction, stroke, or hospitalization for congestive heart failure using the primary discharge diagnosis code for an inpatient visit (see eTable 1 in the Supplement for *International Classification of Diseases, Ninth Revision* codes). These definitions have satisfactory measurement characteristics; the positive predictive value for myocardial infarction was 93% or higher; stroke, 81% or higher; and congestive heart failure, 87% or higher.[12-14] Cardiovascular deaths were included in the composite outcome if they occurred during a hospitalization for myocardial infarction, stroke, or heart failure but not outside the hospital.

Statistical Analysis

Our primary analysis compared the rates of the composite end point among patients initiating treatment with telmisartan vs ramipril. Data were censored for patients when they discontinued use of their initial medication, switched to the comparator medication, experienced a study outcome, disenrolled from their health plan, or died, or on September 30, 2009.[15] To address confounding, we adjusted for 74 patient characteristics, including demographics, comorbid conditions, concurrent medications, and health care use measures, using PS methods (**Table 1** and eTable 2 in the Supplement). To balance patient characteristics, we used 1:1 PS matching with a caliper of 0.05 and did not perform further variable selection. We compared standardized differences to evaluate the level of balance achieved in patient characteristics after PS matching[16] and used unstratified Cox proportional hazards regression to compute hazard ratios (HRs) and 95% CIs. We then performed a predefined secondary analysis that carried forward the exposure to the first-used medication for 365 days.[6]

To assess the robustness of our results, we also sought to confirm the well-established increased risk of angioedema for ramipril, expecting that rates of angioedema would be lower for telmisartan, as also demonstrated in ONTARGET. To further assess robustness, we replicated all study end points using a larger cohort derived from less stringent exclusion criteria by creating a cohort that allowed for past ACE-I or ARB use other than telmisartan or ramipril in the preceding 180 days. All analyses were conducted using the Aetion platform and R, version 3.1.2.5 (R Foundation for Statistical Computing), which has been previously validated for a range of studies[17,18] and for predicting clinical trial findings.[19]

Results

We identified 640 951 patients who filled a prescription for ramipril or telmisartan from January 1, 2003, through September 30, 2009, and had a sufficient baseline enrollment period of at least 180 days. After applying study inclusion and exclusion criteria, 52 739 patients were included (**Figure**), of whom 48 053 were newly prescribed ramipril (mean [SD] age, 68.29 [9.52] years; 31 940 male [66.5%]) and 4665 were newly prescribed telmisartan (mean [SD] age, 69.43 [9.60] years; 2413 male [51.7%]) (a total of 21 patients did not begin follow-up). Patients prescribed ramipril were more likely to be male and have cardiac disease, whereas patients prescribed telmisartan were more likely to have hypertension, kidney disease, and previous transient ischemic attack or stroke and be prescribed a calcium channel blocker (Table 1). After PS matching 4665 telmisartan users (mean [SD] age, 69.43 [9.60] years; 2413 [51.7%]) to 4665 ramipril users (mean [SD] age, 69.36 [9.67] years; 2343 male [50.2%]), these differences were well balanced with standardized differences less than 0.1 (Table 1). Most frequencies of baseline characteristics were consistent with ONTARGET (eg, similar age, rates of hypertension, coronary artery disease, diabetes, and stroke), whereas some were not (ie, lower rates of angina, lower rates of smoking, less documented antiplatelet use, and more women included in our study) (eTable 2 in the Supplement). In the unmatched cohort, mean follow-up time was 232 days (interquartile range, 113-454 days) for the ramipril group and 188 days (interquartile range, 108-427 days) for the telmisartan group. The most common reason for censoring was treatment discontinuation, in 32 135 ramipril users (66.9%) and 3483 telmisartan users (74.7%).

In ONTARGET, the relative risk of the composite outcome of death from cardiovascular causes, myocardial infarction, stroke, or hospitalization for congestive heart failure was 1.01 (95% CI, 0.94-1.09), indicating no significant difference between telmisartan and ramipril. In our study, the PS-matched relative risk of the composite of myocardial infarc-

Table 1. Baseline Patient Characteristics Before Receiving Telmisartan or Ramipril[a,b]

Characteristic	Unmatched Population			PS-Matched Population		
	Ramipril (n = 48 053)	Telmisartan (n = 4665)	Standardized Difference	Ramipril (n = 4665)	Telmisartan (n = 4665)	Standardized Difference
Age, mean (SD), y	68.29 (9.52)	69.43 (9.60)	0.119	69.36 (9.67)	69.43 (9.60)	0.007
Age category, y						
55-59	9747 (20.3)	802 (17.2)	0.149	839 (18.0)	802 (17.2)	0.031
60-64	11 539 (24.0)	985 (21.1)		947 (20.3)	985 (21.1)	
65-69	6262 (13.0)	626 (13.4)		655 (14.0)	620 (13.4)	
70-74	6468 (13.5)	681 (14.6)		666 (14.3)	681 (14.6)	
≥75	14 037 (29.2)	1571 (33.7)		1558 (33.4)	1571 (33.7)	
Male	31 940 (66.5)	2413 (51.7)	0.303	2343 (50.2)	2413 (51.7)	0.030
Date of cohort entry						
First quarter	13 667 (28.4)	1198 (25.7)	0.046	1149 (24.6)	1198 (25.7)	0.053
Second quarter	10 080 (21.0)	1038 (22.3)		1005 (21.5)	1038 (22.3)	
Third quarter	12 730 (26.5)	1310 (28.1)		1395 (29.9)	1310 (28.1)	
Fourth quarter	11 576 (24.1)	1119 (24.0)		1116 (23.9)	1119 (24.0)	
Comorbid conditions						
Hypertension	21 361 (44.5)	2835 (60.8)	0.331	2832 (60.7)	2835 (60.8)	0.001
Coronary artery disease	37 591 (78.2)	3105 (66.6)	0.263	3053 (65.4)	3105 (66.6)	0.024
Diabetes mellitus	14 375 (29.9)	1524 (32.7)	0.059	1514 (32.5)	1524 (32.7)	0.005
PAD	2651 (5.5)	362 (7.8)	0.090	355 (7.6)	362 (7.8)	0.006
Stroke or TIA	5727 (11.9)	730 (15.6)	0.108	783 (16.8)	730 (15.6)	0.031
Angina	11 272 (23.5)	815 (17.5)	0.149	817 (17.5)	815 (17.5)	0.001
Heart failure	7205 (15.0)	510 (10.9)	0.121	526 (11.3)	510 (10.9)	0.011
Renal disease	3549 (7.4)	545 (11.7)	0.147	515 (11.0)	545 (11.7)	0.020
Smoking	1734 (3.6)	115 (2.5)	0.067	128 (2.7)	115 (2.5)	0.017
Previous CABG or PCI	5454 (11.3)	124 (2.7)	0.346	111 (2.4)	124 (2.7)	0.018
Medications						
Statin	22 441 (46.7)	2104 (45.1)	0.032	2073 (44.4)	2104 (45.1)	0.013
β-Blocker	20 957 (43.6)	1926 (41.3)	0.047	1913 (41.0)	1926 (41.3)	0.006
Antiplatelet agent	11 031 (23.0)	1127 (24.2)	0.028	1148 (24.6)	1127 (24.2)	0.010
Calcium channel blocker	5386 (11.2)	833 (17.9)	0.189	825 (17.7)	833 (17.9)	0.004
Diuretic	11 396 (23.7)	1342 (28.8)	0.115	1325 (28.4)	1342 (28.8)	0.008
ACE-I or ARB	0	0	0	0	0	0
Health care use						
Cardiology visit	29 928 (62.3)	2526 (54.1)	0.165	2585 (55.4)	2526 (54.1)	0.025
General practitioner visit	35 314 (73.5)	3571 (76.5)	0.071	3573 (76.6)	3571 (76.5)	0.001
Emergency department visit	9946 (20.7)	907 (19.4)	0.031	911 (19.5)	907 (19.4)	0.002
Influenza vaccination	4141 (8.6)	401 (8.6)	0.001	392 (8.4)	401 (8.6)	0.007
Transthoracic echocardiogram	19 496 (40.6)	1589 (34.1)	0.135	1638 (35.1)	1589 (34.1)	0.022

Abbreviations: ACE-I, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; CABG, coronary artery bypass graft; PAD, peripheral arterial disease; PCI, percutaneous coronary intervention; PS, propensity score; TIA, transient ischemic attack.

[a] Data are presented as number (percentage) of patients unless otherwise indicated.

[b] In the preceding 180 days.

tion, stroke, or hospitalization for congestive heart failure was almost identical (HR, 0.99; 95% CI, 0.85-1.14) (**Table 2**).

Sensitivity Analysis

A sensitivity analysis using the last exposure to the first-used medication for 365 days without considering treatment discontinuation found that the primary end point occurred in 402 ramipril users (86 events per 1000 patients) and 363 telmisartan users (78 events per 1000 patients). This resulted in no significant difference in risk after PS matching (HR, 0.90; 95% CI, 0.77-1.04).

Validation Against a Known Causal Association

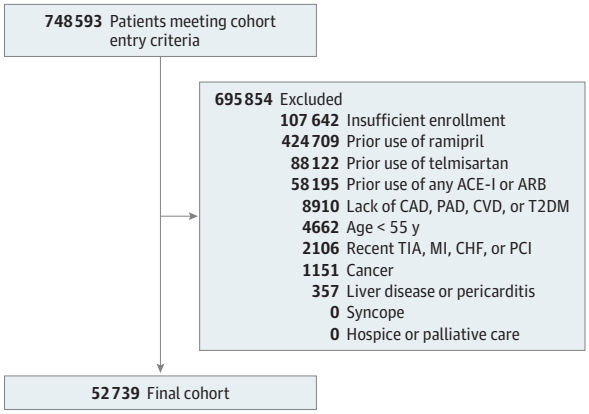
Among PS-matched individuals, there were 18 angioedema events in new users of ramipril (3.1 events per 1000 person-years) and 2 events in new users of telmisartan (0.4 events

per 1000 person-years). A decreased risk (HR, 0.13; 95% CI, 0.03-0.56) of angioedema with telmisartan was also observed in ONTARGET (HR, 0.40; *P* = .01).

Robustness of Findings

In the cohort with less stringent exclusion criteria to allow for past ACE-I or ARB use apart from ramipril or telmisartan, we identified 8656 PS-matched new users of telmisartan and 8656 PS-matched new users of ramipril. In this cohort, there was a similar PS-matched relative risk of the composite of myocardial infarction, stroke, or hospitalization for congestive heart failure (HR, 0.97; 95% CI, 0.88-1.08) (**Table 3**). A decreased PS-matched risk of angioedema with telmisartan compared with ramipril (HR, 0.35; 95% CI, 0.17-0.71) was also revealed.

Figure. Flowchart of Patients in the Study



Cohort criteria included receiving a prescription for ramipril or telmisartan between January 2003 (start of available data) and September 2009 (prior to the Food and Drug Administration approval of the supplemental indication). ACE-I indicates angiotensin converting enzyme inhibitor; ARB, angiotensin receptor blocker; CAD, coronary artery disease; CHF, congestive heart failure; CVD, cerebrovascular disease; MI, myocardial infarction; PAD, peripheral arterial disease; PCI, percutaneous coronary intervention; T2DM, type 2 diabetes mellitus; and TIA, transient ischemic attack.

Discussion

Among patients newly prescribed telmisartan and ramipril before the FDA's decision to approve a supplemental indication for telmisartan, we found results that were almost identical to those of the randomized clinical trial that led to telmisartan's supplemental indication. We further identified and quantified the known causal association between ramipril and angioedema. This finding suggests that our data and analysis plan were sufficiently valid to detect known causal associations first identified in a prospective trial.[20]

This study is one of the largest to analyze real-world data to mirror a large randomized clinical trial that had established the clinical basis for a supplemental indication for a medication. In contrast to ONTARGET, which took approximately 7 years to complete and cost tens of millions of dollars, our study took approximately 12 weeks to implement for less than a hundredth of the cost. The fact that our case study bolstered the conclusions of a trial designed to identify a supplemental indication for a marketed medication and was done relatively efficiently using available data sets, rigorous epidemiologic methods, and modern software platforms supports the concept of conducting similar database analyses as part of routine practice for manufacturers submitting applications for supplemental indications to the FDA.[21]

Results concordant with the pivotal clinical trial can provide regulators with greater confidence in approving the indication, whereas discordant results could warrant deeper reexamination of the clinical trial or nonrandomized data. When results are discordant with the pivotal trial, an in-depth analysis of the trial and the nonrandomized study will be necessary to identify reasons for this discordance. These reasons can include issues related to study design, statistical analysis, and patient population. Additional research will be necessary to help navigate this scenario.[22] Eventually, the FDA can develop empirically based guidance on when database analyses are useful in this context and when they are less reliable as a confirmatory source.

Validity of Nonrandomized Real-world Data Analyses

There have been examples of real-world data providing results before the randomized clinical trial was completed[23-25] and nonrandomized real-world studies that changed prescribing practices for which there will likely never be randomized clinical trial findings.[26-29] A common signal of quality among these studies and our current study was the use of a new-user, active-comparator design. This approach compares 2 groups of patients who newly start taking a medication and avoids comparing 2 groups with intrinsically discrepant risk profiles as would be found using a nonuser comparator or comparing new users with ongoing users. The new-user design with an active comparator allows a more homogeneous baseline population and was one of the main reasons why the observed baseline characteristics for our patients were similar even before matching. By design, approximately 80% of the 74 baseline characteristics were well balanced before PS matching, suggesting that unmeasured factors may be equally balanced. Similar results were observed in the recent new-user, active-comparator study by Graham et al[26] that compared the safety and effectiveness of rivaroxaban with those of dabigatran.

By contrast, some nonrandomized real-world studies[30,31] found results that differed from those in subsequent randomized clinical trials.[32-34] This difference can occur for many reasons, including incorrect study design implementation, reverse causation,[35] immortal time bias,[8] depletion of susceptibles,[36] failure to identify important unmeasured confounding factors, or the inclusion of a different study population than was used in the clinical trial. In particular, comparators that use patients defined as those who did not fill a prescription (nonusers) may introduce treatment selection bias that may not be controllable with any statistical method.[37,38]

Studies such as ours require that inclusion and exclusion criteria and end points be adequately defined in a randomized clinical trial report and subsequently identifiable in the health care data set being studied. Many trials include study

Table 2. Incidence of the Composite End Point, Its Components, and the Risk of Angioedema

Variable	Observational Cohort Study		ONTARGET Clinical Trial	
	Ramipril (n = 4665)	Telmisartan (n = 4665)	Ramipril (n = 8576)	Telmisartan (n = 8542)
Composite End Point				
No. of person-years	5579	4570	NA	NA
No. of events	403	343	1412	1423
Incidence rate per 1000 person-years	72.23	75.05	NA	NA
Rate difference per 1000 person-years	1 [Reference]	2.82 (−7.80 to 13.44)	NA	NA
Unadjusted relative risk	1 [Reference]	0.99 (0.89 to 1.11)	NA	NA
Relative risk	1 [Reference]	0.99 (0.85 to 1.14)[a]	1 [Reference]	1.01 (0.94 to 1.09)
Stroke				
No. of person-years	5808	4718	NA	NA
No. of events	107	86	405	369
Incidence rate per 1000 person-years	18.42	18.23	NA	NA
Rate difference per 1000 person-years	1 [Reference]	−0.20 (−5.40 to 5.00)	NA	NA
Unadjusted relative risk	1 [Reference]	1.08 (0.87 to 1.35)	NA	NA
Relative risk	1 [Reference]	0.95 (0.71 to 1.26)[a]	1 [Reference]	0.91 (0.70 to 1.05)
Myocardial Infarction				
No. of person-years	5824	4726	NA	NA
No. of events	84	68	413	440
Incidence rate per 1000 person-years	14.42	14.39	NA	NA
Rate difference per 1000 person-years	1 [Reference]	−0.03 (−4.64 to 4.57)	NA	NA
Unadjusted relative risk	1 [Reference]	0.97 (0.76 to 1.24)	NA	NA
Relative risk	1 [Reference]	0.92 (0.67 to 1.27)[a]	1 [Reference]	1.07 (0.94 to 1.22)
Hospitalization for Heart Failure				
No. of person-years	5684	4656	NA	NA
No. of events	284	231	354	394
Incidence rate per 1000 person-years	49.97	49.61	NA	NA
Rate difference per 1000 person-years	1 [Reference]	−0.35 (−9.00 to 8.29)	NA	NA
Unadjusted relative risk	1 [Reference]	0.94 (0.82 to 1.08)	NA	NA
Relative risk	1 [Reference]	0.95 (0.79 to 1.13)[a]	1 [Reference]	1.12 (0.97 to 1.29)
Angioedema				
No. of person-years	5885	4772	NA	NA
No. of events	18	2	25	10
Incidence rate per 1000 person-years	3.06	0.42	NA	NA
Rate difference per 1000 person-years	1 [Reference]	−2.64 (−4.17 to −1.11)	NA	NA
Unadjusted relative risk	1 [Reference]	0.18 (0.04 to 0.70)	NA	NA
Relative risk	1 [Reference]	0.13 (0.03 to 0.56)[a]	1 [Reference]	0.4 (P = .01)[b]

Abbreviations: NA, not applicable; ONTARGET, Ongoing Telmisartan Alone and in Combination with Ramipril Global End-point Trial.

[a] Relative risk using 1:1 propensity score matching.

[b] The CIs are not provided in the ONTARGET article.

end points that are not recorded in claims data or electronic health care records (eg, rating scales used in trials of psychiatric medications) or not easily identifiable (eg, progression-free survival used in oncology trials) without requiring challenging natural language processing of free-text information. It would also be difficult to replicate results from randomized clinical trials that include different treatment modalities with substantially different risk-benefit profiles (eg, implantable cardioverter defibrillators compared with medical therapy) because of fundamental differences in risk profiles between the 2 populations.[7,38]

Pharmacoepidemiology analysis of data from nonrandomized, real-world health care databases can be used to support supplemental indications established in prospective randomized clinical trials of marketed medications. This is powerful because they represent outcomes in settings of typical care, rather than the highly controlled research environments of RCTs, and can be accomplished quickly and inexpensively. The

analyses can also include subgroups of patients who are underrepresented in clinical trials, including elderly individuals, patients with many comorbidities, pregnant women, and other at-risk groups. In our study, for example, 50% of patients were women compared with approximately 26% in ONTARGET. Finally, such studies can evaluate a larger population of patients and can assess end points that trials are often underpowered to detect, such as rare adverse events.

Limitations

Our observed null finding might reflect limitations within our data set (eg, lack of out-of-hospital death data), duration of follow-up, or study design rather than a true observation. It is well established that noninferiority can appear to be present because of inadequate rigor or scale in any study, whether a randomized clinical trial or an observational analysis.[39,40] However, this does not explain the increased risk of angioedema that we observed with ramipril but not telmisartan. Some authors[41,42] have questioned the value of PS matching over traditional risk-adjusted regression analysis, neither of which guarantee full account for unmeasured confounding. However, our unadjusted primary, secondary, and sensitivity analyses did not change meaningfully after PS matching. Another limitation of our study was an inability to assess medication adherence beyond prescription filling, although this is generally seen as a valid measure of actual use.[15]

Conclusions

The FDA is currently considering how it will use nonrandomized, real-world data as part of supplemental indication applications.[42,43] In the absence of large-scale empirical comparative analyses that identify the reasons for failure and success to replicate randomized controlled findings with real-world data analyses, we performed a case study that highlights some important considerations. Many context-specific questions about study design, confounding control, data quality, and outcome validity will need to be considered.[4,6] Preregistering study designs and analysis plans and providing a publicly available summary of the results when available, similar to the current practice of randomized clinical trials, promotes ethical conduct of these studies.

Even well-designed analyses sometimes result in incorrect conclusions, and some randomized clinical trials may be inaccurate.[44] Retrospective reviews of the literature[34,45-48] provide single summarizations of the differences between these 2 approaches but provide few insights on the validity of individual real-world data analyses. To establish a meaningful baseline, the FDA will need many sets of randomized clinical trials with prospectively designed, nonrandomized analyses to match the populations included in randomized clinical trials across a range of clinical questions, each investigated with a set of designs and methods following rigorous epidemiologic principles.

Regulators have a difficult task in providing specific rules for decision making in this maturing yet still developing and highly context-specific field. However, if done selectively and with principled methods, it might be feasible to use nonrandomized, real-world data to provide supportive evidence in establishing supplemental drug indications.

Table 3. Results From Secondary Analyses That Expanded the Population by Including Patients Who Had Used Angiotensin-Converting Enzyme Inhibitors or Angiotensin Receptor Blockers in the Past 180 Days

Variable	Ramipril (n = 8656)	Telmisartan (n = 8656)
Composite End Point		
No. of person-years	10 227	8749
No. of events	799	695
Incidence rate per 1000 person-years	78.13	79.44
Unadjusted	1 [Reference]	0.96 (0.89-1.04)
PS matched	1 [Reference]	0.97 (0.88-1.08)
Stroke		
No. of person-years	10 760	9098
No. of events	184	173
Incidence rate per 1000 person-years	17.1	19.01
Unadjusted	1 [Reference]	1.08 (0.92-1.27)
PS matched	1 [Reference]	1.07 (0.87-1.32)
Myocardial Infarction		
No. of person-years	10 755	9132
No. of events	156	126
Incidence rate per 1000 person-years	14.5	13.8
Unadjusted	1 [Reference]	0.87 (0.72-1.04)
PS matched	1 [Reference]	0.91 (0.72-1.15)
Hospitalization for Heart Failure		
No. of person-years	10 433	8928
No. of events	589	481
Incidence rate per 1000 person-years	56.46	53.88
Unadjusted	1 [Reference]	0.93 (0.84-1.02)
PS matched	1 [Reference]	0.91 (0.81-1.03)
Angioedema		
No. of person-years	10 885	9220
No. of events	32	10
Incidence rate per 1000 person-years	2.94	1.08
Unadjusted	1 [Reference]	0.43 (0.23-0.82)
PS matched	1 [Reference]	0.35 (0.17-0.71)

Abbreviation: PS, propensity score.

ARTICLE INFORMATION

Accepted for Publication: August 15, 2017.

Published Online: November 20, 2017. doi:10.1001/jamainternmed.2017.3919

Author Affiliations: Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Fralick, Kesselheim, Avorn, Schneeweiss); Program on Regulation, Therapeutics, and Law (PORTAL), Division of

Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Fralick, Kesselheim, Avorn); Eliot Phillipson Clinician-Scientist Training Program, University of Toronto, Toronto, Ontario, Canada (Fralick).

Author Contributions: Dr Schneeweiss had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: All authors.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Fralick.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Fralick, Schneeweiss.
Obtained funding: Kesselheim, Avorn, Schneeweiss.
Administrative, technical, or material support: Fralick, Avorn.
Study supervision: Kesselheim, Avorn, Schneeweiss.

Conflict of Interest Disclosure: Dr Kesselheim reported receiving grants from the US Food and Drug Administration Office of Generic Drugs and Division of Health Communication (2013-2016) unrelated to the topic of this article. Dr Schneeweiss reported receiving support from grants from the US Food and Drug Administration and the Patient Centered Outcomes Research Institute; serving as a consultant to WHISCON, LLC, and Aetion Inc, a software manufacturer of which he also owns equity; and being a principal investigator of investigator-initiated grants to the Brigham and Women's Hospital from Genentech, Bayer, and Boehringer Ingelheim (not directly related to the topic of this article). No other disclosures were reported.

Funding/Support: This study was funded by the Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, and the Laura and John Arnold Foundation. Dr Fralick was funded by the University of Toronto Clinician Scientist Training Program, Toronto, Ontario, Canada. Dr Kesselheim's work is additionally supported by the Engelberg Foundation and the Harvard Program in Therapeutic Science.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. Kesselheim AS, Avorn J. New "21st century cures" legislation. *JAMA*. 2017;317(6):581-582.

2. 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(23):2293-2297.

3. Schneeweiss S. Developments in post-marketing comparative effectiveness research. *Clin Pharmacol Ther*. 2007;82(2):143-156.

4. Fralick M, Kesselheim AS. Periodic benefit-risk evaluation reports have substantial promise to guide patient care and should be made publicly available. *Pharmacoepidemiol Drug Saf*. 2017;26(5):597-599.

5. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915-920.

6. Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. *Pharmacoepidemiol Drug Saf*. 2010;19(8):858-868.

7. Goodman SN, Schneeweiss S, Baiocchi M. Using design thinking to differentiate useful from misleading evidence in observational research. *JAMA*. 2017;317(7):705-707.

8. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492-499.

9. Wang B, Kesselheim AS. Characteristics of efficacy evidence supporting approval of supplemental indications for prescription drugs in United States, 2005-14: systematic review. *BMJ*. 2015;351:h4679.

10. Berndt ER, Cockburn IM, Grépin KA. The impact of incremental innovation in biopharmaceuticals: drug utilisation in original and supplemental indications. *Pharmacoeconomics*. 2006;24(suppl 2):69-86.

11. Yusuf S, Teo KK, Pogue J, et al; ONTARGET Investigators. Telmisartan, ramipril, or both in patients at high risk for vascular events. *N Engl J Med*. 2008;358(15):1547-1559.

12. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS One*. 2014;9(3):e92286.

13. McCormick N, Bhole V, Lacaille D, Avina-Zubieta JA. Validity of diagnostic codes for acute stroke in administrative databases: a systematic review. *PLoS One*. 2015;10(8):e0135834.

14. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS One*. 2014;9(8):e104519.

15. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-337.

16. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083-3107.

17. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and reproducibility of observational cohort studies using large healthcare databases. *Clin Pharmacol Ther*. 2016;99(3):325-332.

18. Fralick M, Schneeweiss S, Patorno E. Risk of diabetic ketoacidosis after initiation of an sglt2 inhibitor. *N Engl J Med*. 2017;376(23):2300-2302.

19. Kim SC, Solomon DH, Rogers JR, et al. Cardiovascular safety of tocilizumab versus tumor necrosis factor inhibitors in patients with rheumatoid arthritis: a multi-database cohort study. *Arthritis Rheumatol*. 2017;69(6):1154-1164.

20. Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments, part 1: ethical and scientific issues. *Ann Intern Med*. 2000;133(6):455-463.

21. Institute of Medicine of the National Academies. Ethical and scientific issues in studying the safety of approved drugs. The national academy of press website. May 2012. https://www.nap.edu/catalog/13219/ethical-and-scientific-issues-in-studying-the-safety-of-approved-drugs. Accessed June 7, 2017.

22. Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Ther*. 2017;102(6):924-933. doi:10.1002/cpt.857

23. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. *N Engl J Med*. 2008;358(8):771-783.

24. Choi N-K, Solomon DH, Tsacogianis TN, Landon JE, Song HJ, Kim SC. Comparative safety and effectiveness of denosumab versus zoledronic acid in patients with osteoporosis: a cohort study. *J Bone Miner Res*. 2017;32(3):611-617.

25. Patorno E, Goldfine A, Schneeweiss S, Glynn R, Liu J, Kim S. Safety of canagliflozin vs. other antidiabetic agents in routine care. Paper presented at: 77th Scientific Sessions of the American Diabetes Association; June 9, 2017; San Diego, California.

26. Graham DJ, Reichman ME, Wernecke M, et al. Stroke, bleeding, and mortality risks in elderly medicare beneficiaries treated with dabigatran or rivaroxaban for nonvalvular atrial fibrillation. *JAMA Intern Med*. 2016;176(11):1662-1671.

27. Gagne JJ, Choudhry NK, Kesselheim AS, et al. Comparative effectiveness of generic and brand-name statins on patient outcomes: a cohort study. *Ann Intern Med*. 2014;161(6):400-407.

28. Wang PS, Schneeweiss S, Avorn J, et al. Risk of death in elderly users of conventional vs. atypical antipsychotic medications. *N Engl J Med*. 2005;353(22):2335-2341.

29. Ray WA, Chung CP, Murray KT, Hall K, Stein CM. Atypical antipsychotic drugs and the risk of sudden cardiac death. *N Engl J Med*. 2009;360(3):225-235.

30. Finucane FF, Madans JH, Bush TL, Wolf PH, Kleinman JC. Decreased risk of stroke among postmenopausal hormone users: results from a national cohort. *Arch Intern Med*. 1993;153(1):73-79.

31. Blais L, Desgagné A, LeLorier J. 3-Hydroxy-3-methylglutaryl coenzyme A reductase inhibitors and the risk of cancer: a nested case-control study. *Arch Intern Med*. 2000;160(15):2363-2368.

32. Wassertheil-Smoller S, Hendrix SL, Limacher M, et al; WHI Investigators. Effect of estrogen plus progestin on stroke in postmenopausal women: the Women's Health Initiative: a randomized trial. *JAMA*. 2003;289(20):2673-2684.

33. Dale KM, Coleman CI, Henyan NN, Kluger J, White CM. Statins and cancer risk: a meta-analysis. *JAMA*. 2006;295(1):74-80.

34. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*. 2001;286(7):821-830.

35. Gruber S, Chakravarty A, Heckbert SR, et al. Design and analysis choices for safety surveillance evaluations need to be tuned to the specifics of the hypothesized drug-outcome association. *Pharmacoepidemiol Drug Saf*. 2016;25(9):973-981.

36. Moride Y, Abenhaim L. Evidence of the depletion of susceptibles effect in

non-experimental pharmacoepidemiologic research. *J Clin Epidemiol*. 1994;47(7):731-737.

37. Schneeweiss S, Patrick AR, Stürmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Med Care*. 2007;45(10)(suppl 2):S131-S142.

38. Setoguchi S, Warner Stevenson L, Stewart GC, et al. Influence of healthy candidate bias in assessing clinical effectiveness for implantable cardioverter-defibrillators: cohort study of older patients with heart failure. *BMJ*. 2014;348:g2866.

39. D'Agostino RB Sr, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics. *Stat Med*. 2003;22(2):169-186.

40. Snapinn SM. Noninferiority trials. *Curr Control Trials Cardiovasc Med*. 2000;1(1):19-21.

41. Agoritsas T, Merglen A, Shah ND, O'Donnell M, Guyatt GH, et al. Adjusted analyses in studies addressing therapy and harm users' guides to the medical literature clinical scenario. *JAMA*. 2017;317(7):748-759.

42. Fonarow GC. Randomization—there is no substitute. *JAMA Cardiol*. 2016;1(6):633-635.

43. Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ*. 2013;347:f6409.

44. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.

45. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. In: Bero L, ed. *Cochrane Database of Systematic Reviews*. Chichester, England: John Wiley & Sons; 2014:MR000034.

46. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887-1892.

47. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ*. 2016;352:i493.

48. Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? a systematic comparison of studies on acute coronary syndromes. *Eur Heart J*. 2012;33(15):1893-1901.

Invited Commentary

Comparison of Observational Data and the ONTARGET Results for Telmisartan Treatment of Hypertension

Bull's-eye or Painting the Target Around the Arrow?

Robert M. Califf, MD

In this issue of *JAMA Internal Medicine,* Fralick and colleagues[1] create a straw man to demonstrate that observational treatment comparisons could be useful for expanding indications for medical products. The authors modeled the Ongoing Telmisartan Alone and in Combination with Ramipril Global End-point Trial (ONTARGET),[2] which compared the angiotensin receptor antagonist telmisartan and the angiotensin-converting enzyme inhibitor ramipril for the treatment of hypertension. That trial,[2] published in 2008, found that telmisartan was equally effective to ramipril, with fewer incidences of angioedema. Participants who received both drugs experienced more adverse events but no increase in benefits.

Related article page 55

ONTARGET was a good choice for this demonstration: telmisartan, with a toxicity and adverse effect profile similar to ramipril, was already approved for treating hypertension, and many angiotensin receptor antagonists are noninferior to angiotensin-converting enzyme inhibitors for preventing cardiovascular events. The study by Fralick et al[1] is valuable and technically excellent; however, it examines only 1 drug-indication pair of many. Thus, it is open to the criticism that generalizing from 1 positive finding to a vast field of potential treatment comparisons with observational data is analogous to painting the target around the arrow, especially considering the high probability that the telmisartan-ramipril comparison would work.

Theory and experience have shown randomization to be the key element of high-quality evidence when drawing causal inferences about therapeutic effects and when making the case for regulatory approval. The classic construct invokes a hierarchy of evidence in which randomized clinical trials (RCTs) occupy the apex of the evidence pyramid, with observational analyses relegated to lower levels. An accompanying body of folklore known as good clinical practice has accumulated around organizational and operational aspects of RCTs. Such trials, however, cannot answer every clinical question, and bureaucracy engendered by common interpretations of good clinical practice has driven the costs of traditional regulatory RCTs to such levels that many important questions are effectively unanswerable within the existing clinical research ecosystem. For example, many regulated RCTs expend substantial resources auditing data that may not be essential to the result of the trial at a cost that far exceeds the value in obtaining a reliable answer to the primary questions posed by the trial.[3]

The shortcomings of traditional regulatory RCTs have long been debated.[4] During the past few decades, however, alternative approaches for understanding the effects of specific therapies have evolved. Recently, Frieden[5] pointed out that as analytical methods continue to improve, confidence in the value of observational analyses should correspondingly increase.

The evidentiary standard for initial marketing approval for drugs, biologics, or medical devices is a high bar generally construed as 2 traditional RCTs demonstrating benefit in terms of

non-experimental pharmacoepidemiologic research. *J Clin Epidemiol*. 1994;47(7):731-737.

37. Schneeweiss S, Patrick AR, Stürmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. *Med Care*. 2007;45(10)(suppl 2):S131-S142.

38. Setoguchi S, Warner Stevenson L, Stewart GC, et al. Influence of healthy candidate bias in assessing clinical effectiveness for implantable cardioverter-defibrillators: cohort study of older patients with heart failure. *BMJ*. 2014;348:g2866.

39. D'Agostino RB Sr, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics. *Stat Med*. 2003;22(2):169-186.

40. Snapinn SM. Noninferiority trials. *Curr Control Trials Cardiovasc Med*. 2000;1(1):19-21.

41. Agoritsas T, Merglen A, Shah ND, O'Donnell M, Guyatt GH, et al. Adjusted analyses in studies addressing therapy and harm users' guides to the medical literature clinical scenario. *JAMA*. 2017;317(7):748-759.

42. Fonarow GC. Randomization—there is no substitute. *JAMA Cardiol*. 2016;1(6):633-635.

43. Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ*. 2013;347:f6409.

44. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.

45. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. In: Bero L, ed. *Cochrane Database of Systematic Reviews*. Chichester, England: John Wiley & Sons; 2014:MR000034.

46. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887-1892.

47. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ*. 2016;352:i493.

48. Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? a systematic comparison of studies on acute coronary syndromes. *Eur Heart J*. 2012;33(15):1893-1901.

Invited Commentary

Comparison of Observational Data and the ONTARGET Results for Telmisartan Treatment of Hypertension

Bull's-eye or Painting the Target Around the Arrow?

Robert M. Califf, MD

In this issue of *JAMA Internal Medicine,* Fralick and colleagues[1] create a straw man to demonstrate that observational treatment comparisons could be useful for expanding indications for medical products. The authors modeled the Ongoing Telmisartan Alone and in Combination with Ramipril Global End-point Trial (ONTARGET),[2] which compared the angiotensin receptor antagonist telmisartan and the angiotensin-converting enzyme inhibitor ramipril for the treatment of hypertension. That trial,[2] published in 2008, found that telmisartan was equally effective to ramipril, with fewer incidences of angioedema. Participants who received both drugs experienced more adverse events but no increase in benefits.

Related article page 55

ONTARGET was a good choice for this demonstration: telmisartan, with a toxicity and adverse effect profile similar to ramipril, was already approved for treating hypertension, and many angiotensin receptor antagonists are noninferior to angiotensin-converting enzyme inhibitors for preventing cardiovascular events. The study by Fralick et al[1] is valuable and technically excellent; however, it examines only 1 drug-indication pair of many. Thus, it is open to the criticism that generalizing from 1 positive finding to a vast field of potential treatment comparisons with observational data is analogous to painting the target around the arrow, especially considering the high probability that the telmisartan-ramipril comparison would work.

Theory and experience have shown randomization to be the key element of high-quality evidence when drawing causal inferences about therapeutic effects and when making the case for regulatory approval. The classic construct invokes a hierarchy of evidence in which randomized clinical trials (RCTs) occupy the apex of the evidence pyramid, with observational analyses relegated to lower levels. An accompanying body of folklore known as good clinical practice has accumulated around organizational and operational aspects of RCTs. Such trials, however, cannot answer every clinical question, and bureaucracy engendered by common interpretations of good clinical practice has driven the costs of traditional regulatory RCTs to such levels that many important questions are effectively unanswerable within the existing clinical research ecosystem. For example, many regulated RCTs expend substantial resources auditing data that may not be essential to the result of the trial at a cost that far exceeds the value in obtaining a reliable answer to the primary questions posed by the trial.[3]

The shortcomings of traditional regulatory RCTs have long been debated.[4] During the past few decades, however, alternative approaches for understanding the effects of specific therapies have evolved. Recently, Frieden[5] pointed out that as analytical methods continue to improve, confidence in the value of observational analyses should correspondingly increase.

The evidentiary standard for initial marketing approval for drugs, biologics, or medical devices is a high bar generally construed as 2 traditional RCTs demonstrating benefit in terms of

clinical outcomes important to patients. Recent guidances, regulations, and statutes, however, make it clear that the law supports the position of the US Food and Drug Administration (FDA) to apply a standard that encompasses evidence judged convincing by qualified experts[6]; this evidence can be produced by various methods, although it is typically anchored by at least 1 traditional RCT.[6] After a drug is marketed, it may be prescribed for medical conditions beyond those approved, and manufacturers typically fund studies to obtain further information within the labeled indication(s) or to evaluate the intervention for other possible indications.

Some have expressed concerns that when the FDA sustains the high standard for additional indications or labeling changes, the agency creates unfortunate incentives. For clinical researchers, the excessive bureaucracy pushes them to gravitate away from performing the labor-intensive, expensive, and rigorous studies designed to meet approval criteria or to avoid participation in clinical trials altogether. For manufacturers, the incentive is to leverage marketing and thought leader influence to encourage off-label use without bringing study data before regulators. The result is that the US clinical research system is failing to answer many questions that undergird clinical practice.[7]

Given such circumstances, it is understandable that many physicians and patients would endorse off-label use of drugs and devices based on the best recommendations possible with available evidence. The 21st Century Cures Act, enacted in 2016, and user-fee agreements included in the FDA Reauthorization Act of 2017 encourage use of real-world evidence. They offer industry a lower cost of applications for additional indications to provide more medical evidence for the drug label and encourage development of methods for these purposes.

From a technical perspective, the article by Fralick et al[1] offers a cogent summary of the care, insight, and expertise that can be applied to the daunting scientific problem of observational treatment comparisons. The authors chose their cohort and database carefully to match ONTARGET's conditions, studied patients who newly initiated treatment to avoid biases involved in starting with current users, used powerful propensity score matching after adjusting for 73 patient characteristics, and used validated outcomes. They also performed sensitivity analyses, used a positive outcome control (angioedema) to demonstrate that their methods could detect a known difference in adverse effects between treatments, and reproduced their results in a larger population using broader entry criteria.

One issue not addressed by Fralick et al[1] is a common concern in research based on electronic health records: accounting for death in the analysis. The composite end point used in their study has been validated independently and is widely used in cardiovascular outcome trials. Although 93% concordance for myocardial infarction and 80% to 85% concordance for heart failure and stroke are good, the myocardial infarction component is misleading. More than half of myocardial infarctions that occur outside the hospital are fatal. Such out-of-hospital deaths may or may not be captured by electronic health records, and health systems do not routinely assure the quality of such data in the electronic record. The result of this wrinkle in data collection is that Fralick et al[1] used a different end point—the composite of myocardial infarction, stroke, or heart failure admission—than ONTARGET, which used all cardiovascular deaths, heart failure admission, nonfatal myocardial infarction, or nonfatal stroke.

Of interest, the authors note that from 2005 to 2014, manufacturers and other sponsors filed only 290 supplemental applications with the FDA.[1] This is a small proportion of all indications explored by industry and academic researchers and a fraction of the indications adopted by professional society guidelines and clinicians in practice. It thus seems reasonable to conclude that concerns about shunting of effort and resources from trials capable of supporting regulatory approval to off-label development may be valid. The hope is that recent elimination of user fees for supplemental indications will encourage industry to include more information in product labels. Of the 138 applications for new indications filed in this period[1] most were based on biomarkers or composite end points, meaning that the outcome could not be reliably identified in electronic health record or claims data. For most new indications, other approaches may be needed to leverage real-world evidence.

Critics often note that few clinicians actually read product labels. This observation, however, overlooks the label's core value, namely, generating derivative information for other important applications, including informing internet resources, clinical decision support, and reimbursement decisions. Solving these methodologic issues in defining the use (and limitations) of real-world evidence and eliminating perverse incentives that fuel off-label development of drugs should be key priorities.

Nevertheless, the study by Fralick et al[1] points toward a more fluid future. Given the provisions of the 21st Century Cures Act and FDA Reauthorization Act, the efficacy standard for initial marketing approval for new drugs is unlikely to change soon. The combination of traditional development pathways and an array of accelerated pathways provide the FDA with considerable flexibility to encourage manufacturers to match the level of evidence with the clinical indication (while still maintaining the RCT as the cornerstone). Such flexible use of real-world evidence could lead to the incorporation of many more indications into labeling and boost efforts to optimize the evidence base for health and health care.

The increasing use of observational treatment comparisons reflects another element of the health care ecosystem's broader evolution into a learning health system. Regardless of data sources or other factors, randomization should be used whenever feasible; there is no substitute for randomization when we need to be confident that a difference in outcome is caused by a difference in therapy. In many circumstances, however, observational analyses will supplement RCTs for new indications and provide deeper knowledge about real-world use within labeled indications. Despite the need for more examples and robust efforts to guide the use of different methods for different circumstances, observational analyses have an important place in the continuum of clinical evidence.

ARTICLE INFORMATION

Author Affiliations: Department of Medicine, Duke University School of Medicine, Durham, North Carolina; Verily Life Sciences (Alphabet), South San Francisco, California.

Corresponding Author: Robert M. Califf, MD, Duke University School of Medicine, 100 Trent Dr, Ste 401, Green Zone, Davison Bldg, Durham, NC 27710 (robert.califf@duke.edu).

Published Online: November 20, 2017. doi:10.1001/jamainternmed.2017.5091

Conflict of Interest Disclosures: Dr Califf reported serving as the Commissioner of Food and Drugs, US Food and Drug Administration (FDA), from February 2016 to January 2017. Before his appointment to the FDA as Deputy Commissioner for Medical Products and Tobacco in February 2015, Dr Califf reported receiving research grant funding from the Patient-Centered Outcomes Research Institute, the National Institutes of Health, the FDA, Amylin, and Eli Lilly and Company; receiving research grants and consulting payments from Bristol-Myers Squibb, Janssen Research and Development, Merck, and Novartis; receiving consulting payments from Amgen, Bayer Healthcare, BMEB Services, Genentech, GlaxoSmithKline, Heart.org–Daiichi Sankyo, Kowa, Les Laboratoires Servier, Medscape/Heart.org, Regado, and Roche; and holding equity in N30 Pharma and Portola. He currently receives consulting payments from Merck and is employed as a scientific adviser by Verily Life Sciences (Alphabet).

REFERENCES

1. Fralick M, Kesselheim AS, Avorn J, Schneeweiss S. Use of health care databases to support supplemental indications of approved medications [published online November 20, 2017]. *JAMA Intern Med*. doi:10.1001/jamainternmed.2017.3919

2. Yusuf S, Teo KK, Pogue J, et al; ONTARGET Investigators. Telmisartan, ramipril, or both in patients at high risk for vascular events. *N Engl J Med*. 2008;358(15):1547-1559.

3. Eisenstein EL, Collins R, Cracknell BS, et al. Sensible approaches for reducing clinical trial costs. *Clin Trials*. 2008;5(1):75-84.

4. Califf RM, Pryor DB, Greenfield JC Jr. Beyond randomized clinical trials: applying clinical experience in the treatment of patients with coronary artery disease. *Circulation*. 1986;74(6):1191-1194.

5. Frieden TR. Evidence for health decision making: beyond randomized, controlled trials. *N Engl J Med*. 2017;377(5):465-475.

6. Sherman RE, Davies KM, Robb MA, Hunter NL, Califf RM. Accelerating development of scientific evidence for medical products within the existing US regulatory framework. *Nat Rev Drug Discov*. 2017;16(5):297-298.

7. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA*. 2009;301(8):831-841.