Understanding the Value of Real-World Evidence: Focus on Stroke Prevention in Atrial Fibrillation with Rivaroxaban

A. John Camm[1]   Craig I. Coleman[2]   Torben Bjerregaard Larsen[3]   Peter Brønnum Nielsen[3]
CAPT Sally Tamayo[4]

[1] Division of Clinical Sciences, St George's University of London and Imperial College, London, United Kingdom
[2] University of Connecticut School of Pharmacy, Connecticut, United States
[3] Department of Cardiology, Aalborg University Hospital, Aalborg, Denmark
[4] Naval Medical Center, Portsmouth, Virginia, United States

Address for correspondence A. John Camm, MD, Cardiovascular Sciences Research Centre, Division of Clinical Sciences, St George's, University of London, London SW17 0RE, United Kingdom (e-mail: jcamm@sgul.ac.uk).

Abstract

Real-world data are a well-recognized component within the drug lifecycle, and such data are generated from a range of sources and study designs, including claims databases, electronic health records, non-interventional studies (NIS) and registries. While this information can be of vital clinical importance, there may be challenges in understanding the relevance of the differing study designs, endpoints and populations. Here, we summarize the value of real-world evidence and considerations pertinent to their use in clinical research. Owing to the variety of analyses being conducted using real-world data, it is important for researchers and clinicians to have a clear understanding of the nature and origin of those data, and to ensure they are valid, reliable and robust in terms of extrapolating meaningful findings. There are crucial questions to address when evaluating real-world studies, and we introduce a checklist to meet these objectives. In addition to advice for appraising data quality and study designs, several updates will be covered from real-world studies of rivaroxaban for stroke prevention in patients with atrial fibrillation (AF): the nationwide Danish cohort study, U.S. Department of Defense Military Health System database, retrospective claim database study REAFFIRM and a pooled analysis from the global NIS XArelto on preveNtion of sTroke and non-central nervoUS system systemic embolism in patients with non-valvular atrial fibrillation (XANTUS). Real-world studies consistently show that rivaroxaban is an effective treatment option with acceptable safety when used for stroke prevention in a large number of patients with AF across the globe.

Keywords
► non-interventional study
► database
► registries
► rivaroxaban

Introduction

Real-world studies are now regarded as a vital resource for evaluating the long-term safety and effectiveness of drugs in routine clinical practice, including patients who may not be represented in randomized controlled trials (RCTs).[1,2] Clinical evidence stems from a wide variety of sources and it can prove challenging to understand the clinical impact when considering the differing study designs, study endpoints and populations.[1] RCTs are regarded as gold standard in terms of investigating drug efficacy and safety, although they can sometimes lack external validity and the findings may not

be generalizable to the general population. Moreover, RCTs are costly and may have insufficient study duration and population size to thoroughly assess the impact of a treatment, particularly in the long term. In comparison, real-world studies show potential for high generalizability. They are cheaper to perform and may emulate experience in clinical practice, increasing external validity and, in turn, complementing clinical findings from RCTs.[2]

Real-world data are being generated from claims databases, electronic health records (EHRs), non-interventional studies (NIS) and registries.[3] Claims databases and EHRs are often used to provide insights on under-represented groups (e.g. those with low or high CHA_2DS_2-VAS_C scores, the elderly, those with specific combinations of comorbidities) in addition to practical clinical management issues and outcomes, using large patient cohorts. NIS, such as XArelto on preveNtion of sTroke and non-central nervoUS system systemic embolism in patients with non-valvular atrial fibrillation (XANTUS),[4] have extended the safety and effectiveness of rivaroxaban that were previously shown in Phase III RCTs to the general population. This review will provide an update on some examples of real-world analyses including several that focus on the direct, oral factor Xa inhibitor rivaroxaban in atrial fibrillation (AF). We will also discuss how to properly evaluate these data and understand their clinical relevance and limitations.

How Can We Decipher Real-World Data Accurately and Effectively?

It is clear that data can be generated from an assortment of study designs, inclusion/exclusion criteria, patient characteristics, endpoints and statistical analyses, and the results may tend to vary based on the choice of methods used.[1] Generally, it is inappropriate to directly compare Phase III and real-world data, and indeed, rather than compare the two designs we should understand that these study types are complementary to each other. RCTs are more restricted in their design, with inclusion and exclusion criteria that may limit the ability to answer clinically relevant questions. Management approaches and patient characteristics in RCTs are likely to substantially differ from those receiving treatment in routine clinical practice. Furthermore, patients enrolled in RCTs may receive better overall care and may possibly be exposed to more intense monitoring than those in real-world practice. Real-world studies are essential to document benefits and harms of therapy in a wider population with a range of comorbidities and conditions.

Whether a study is prospective or retrospective is another important consideration, as event rates can vary considerably.[3] Prospective studies set out to collect data in real time; retrospective studies, however, analyse data that have already been collected. Prospective cohort data can answer many important clinical questions and test many possible determinants, even factors that may not have been originally considered. However, prospective studies can be time consuming and costly.[3] Although retrospective studies are less costly and take less time to perform than prospective studies

(since data are already available for research), their data are not designed to be used in a study[3] and there may be challenges correcting for bias and confounding, if important covariates were not available in the data. Several reporting biases exist, such as language bias, but the most relevant reporting biases for the NOACs in real-world analyses are selective outcome reporting, selective analysis reporting and selective reporting of a subgroup of participants.

We need to ensure that data are sufficiently valid, reliable and robust to extrapolate meaningful clinical findings. There are several crucial questions that must be considered when we evaluate the quality of real-world data, which encompass the hypothesis, data source, study population, treatment administered, follow-up length, outcomes assessed and availability of sensitivity analyses data. These questions have been incorporated into a checklist tool, which may be useful to help researchers and stakeholders to understand the relevance of studies, and consider the pros and cons of their design (►**Table 1**). Additional checklists have been developed and established by the Good Research for Comparative Effectiveness (GRACE) initiative[5] and REporting of studies Conducted using Observational Routinely-collected Data (RECORD) collaborative.[6] Generally, we suggest that clinicians should always refer to the PICO (Population, Intervention, Comparator and Outcome) principles and ascertain whether there is a clear research question.[7] The next section focuses on the growing use of databases in clinical research.

Non-interventional Studies: A Vital Component in Real-World Research

According to the European Union Clinical Trials NIS Definition, NIS are studies in which the medicinal product(s) is (are) prescribed in the usual manner according to the terms of the marketing authorization.[8] Therefore, assigning the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice. Consequently, the prescription of the medicine is clearly separated from the decision to include a patient in a study. Additionally, no further diagnostic or monitoring procedures should be applied to the patients and epidemiological methods should be used for analysing the collected data.[8] To be of the utmost clinical relevance, data generated from NIS must be of high quality and there are various guidelines in place to help with quality assurance measures—which, when used, should lead to generation of robust, scientific findings.[9] Any possible bias can also be avoided by using adequate study design and data analysis, along with ensuring authenticity, completeness and validity of the data and promptly resolving any concerns. High-quality NIS are also defined by streamlined planning and suitable methods for data analysis.[9]

Phase IV NIS such as XArelto in the prophylaxis of post-surgical VTE after elective Major Orthopaedic Surgery of hip or knee (XAMOS),[10] XANTUS[4] and XArelto for Long-term and Initial Anticoagulation in venous thromboembolism (XALIA)[11] have been fundamental in providing real-world

Table 1 Real-world assessment checklist

		Additional information
1. Research question		
1.1. Is the research question clinically relevant and clearly defined in the paper?	[Yes/No]	The research question should be clinically relevant and stated explicitly at the start of the paper. Assessment of the paper is dependent on how adequately the posed research question has been answered.
2. Data source		
2.1. What type of data source was used?	[Administrative claims/EHRs/ patient registries/ others (please note)]	The data source of the study does not necessarily affect the quality of the study; however, the strengths and limitations of the data source depend on how the underlying data were generated and analysed. Both assessment of study design and interpretation of results depend on the data source used.
3. Study population		
3.1. In relation to the research question, was the initially defined study population selected appropriately?	[Yes/No]	Based on inclusion and exclusion criteria, is the population comparable to: I. The patient population studied in the RCT? II. Patients treated in the real world? Both can be appropriate if the research question requires the patient characteristics to mimic either populations reflective of routine clinical practice or specific RCT patient populations. The population size should also be appropriate to permit the planned analyses and to answer the specific research question posed.
3.2. Was the study population restricted to new initiators of any treatment for this condition (treatment-naive patients)?	[Yes/No, but there was a valid rationale/No]	Standard pharmacoepidemiological approach restricts the patient population to new users when comparing outcomes between treatment groups. However, the research question may necessitate the inclusion of patients who switch from one treatment to another.
4. Treatment exposure		
4.1. Are the treatment regimens compared in the study adequately described?	[Yes/No]	Treatment arms need to be clearly defined to ensure the analyses and outcomes are appropriate to answer the posed research question.
4.2 Does the study account for drug dosage?	[Yes/No]	The impact of dose adjustment on outcomes must be considered and reported when drawing conclusions.
4.3 Is the classification of exposed and unexposed person-time free of 'immortal time bias'?	[Yes/No]	Immortal time in epidemiology refers to a period of cohort follow-up time during which death or an outcome that determines end of follow-up cannot occur. As an example: Readmission is a key quality metric, as it is a frequent and costly adverse event for patients. The American College of Surgeons National Surgical Quality Improvement Program (NSQIP) defines immortal time for readmissions as 30 d following surgery. The standard period for readmission, however, is 30 d after hospital discharge. This means patients cannot be readmitted before hospital discharge and are therefore 'immortal' for this outcome until they leave the hospital, creating immortal time bias.

(Continued)

Table 1 (*Continued*)

Question	Response	Description
4.4. Is adherence/persistence discussed?	[Yes/No]	The degree of persistence/adherence across a patient population can affect treatment outcomes. This should be considered, and stated, when data are analysed and interpreted.
4.5. Is the comparator concurrent?	[Yes/No, but the use of a historical comparison group was justified/No]	Data on comparators are concurrent when they are collected during the same time period as the treatment group. Concurrent comparators should be used whenever possible. Reasonable justifications should be clearly stated if a historical comparator was used.
5. Follow-up		
5.1. Is the follow-up period long enough to capture the outcomes of interest?	[Yes/No]	The length of the follow-up period can vary, depending on the research question, data source, structure of the database and the outcomes being assessed. The length of the follow-up period should be clearly stated and it should be appropriate to assess the outcomes of interest.
5.2. Is there informative censoring?	[Yes/No]	Informative censoring occurs when participants are not included in the final data analysis due to discontinuation of treatment, as a result of an established risk of the treatment. If informative censoring occurred, then it should be stated in the paper, as it can lead to bias outcomes.
6. Outcomes		
6.1. What outcomes were assessed?	[Safety and effectiveness/ Effectiveness only/Safety only]	To provide a more complete set of information on a treatment, studies should assess both safety and effectiveness, and not either one alone.
6.2. Does the paper provide clear and transparent definitions of the outcome measures studied, so that the study can be replicated?	[Yes/No]	All published studies should be replicable by other teams.
6.3. Were the primary effectiveness and safety outcomes studied equivalent in their severity?	[Yes/No]	The primary effectiveness and safety outcomes should be equivalent in their severity, in the context of the disease being treated, for a balanced risk–benefit assessment.
6.4. Were the primary outcomes of interest present before the index date?	[Yes/No]	Usually patients should not have had the event of interest before the index date. To ensure this, patients with an event during the baseline period are excluded.
6.5. Have the outcome definitions been validated or adjudicated?	[Validated/ Adjudicated/No]	Validation is the process by which outcome definitions are assessed, tested and proven to be sound and accurate. Adjudication is the process by which an independent panel of experts assess the events identified as potential endpoints.
6.6. Was the primary clinical outcome measured objectively?	[Yes/No/NA, if outcomes are not clinical]	Objective clinical outcomes are measures, or values, that are not influenced by HCPs' personal opinions. Examples include measures such as occurrence of ischaemic stroke, ICH, hospitalization, mortality (all-cause vs. CV-related mortality). Clinical opinion or judgement would not be considered an objective outcome measure. For example, clinical opinion about whether the patient's condition has improved.

Table 1 (*Continued*)

6.7 Were outcomes also expressed in absolute rates/numbers, as opposed to relative risk/hazard ratio/odds ratio only?	[Yes/No]	To be considered clinically meaningful, and interpreted as such, reported changes in outcome measures should be considered in a clinically relevant context.
7. Sensitivity analysis		
7.1. Were any sensitivity analyses conducted to test uncertainty in the results?	[Yes, and primary results did not substantially change/Yes, but primary results changed dramatically/No]	Sensitivity analysis is performed to recalculate outcomes under alternative assumptions and computations. However, not all study aspects can be evaluated through sensitivity analysis, e.g. the magnitude of confounding due to variables that have not been measured or available for analysis is difficult to assess.
8. Bias		
8.1. Were there potential sources of bias in the study? 8.1.1. Confounding	[Yes/No]	Confounding occurs when one or more prognostic variables (variables that predict the outcome of interest) also predict whether an individual receives one or the other of the interventions of interest. One example is confounding by indication.
8.1.2. Selection bias	[Yes/No]	Selection bias arises when there is exclusion of some eligible participants, or exclusion of the initial follow-up time of some participants, or exclusion of some outcome events that are related to both intervention and outcome.
8.1.3. Misclassification bias	[Yes/No]	Differential misclassification happens when the information errors differ between groups. In other words, the bias is different for exposed and non-exposed, or between those who receive a treatment and those who don't.
8.1.4. Bias due to missing data	[Yes/No]	Data may be missing for exposure, outcomes or covariates. Differences between intervention groups in the extent of and reasons for missing data are key. If the proportion of missing data is low and the reasons for missing data are similar across intervention groups, then the risk of bias is likely to be low. As the proportion of missing data rises, differences in treatment response between available and missing participants may increase the potential for bias.
8.1.5. Reporting bias	[Yes/No]	Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results.
8.2. Were important confounding and effect-modifying variables taken into account in the design and/or analyses?	[Yes/No]	The level of impact of covariates is dependent on the treatment and/or outcome of interest. For example, when the study outcome is all-cause mortality, the impact of conditions such as active, or historical cancer on mortality risk should be considered and adjusted for. Other comorbidities or characteristics that affect patient well-being or robustness should also be considered.
8.3. Is there a discussion in the paper about study limitations and how various types of biases could have affected the results, including the direction?	[Yes/No]	The 'Discussion' section should include comment from the authors on potential biases that may have impacted the observed results, especially if these are deviated from what was reported in the RCTs.

evidence confirming the safety of rivaroxaban in routine practice settings.

Current Findings and Analyses from NIS of Rivaroxaban in AF

A Comparison of Real-World versus RCT Outcomes in Similar Populations from XANTUS and ROCKET AF

Differences in patient characteristics can prove to be a common barrier to comparing clinical outcomes efficiently across studies that differ in design. Patients in ROCKET AF and XANTUS exhibited different baseline stroke risk (the mean [standard deviation] CHADS$_2$ score in XANTUS was 2.0 [1.3] vs. 3.5 [0.9] in the rivaroxaban arm of ROCKET AF).[12,13] Adjusted incidence rates for outcomes in XANTUS were compared with those from ROCKET AF in a statistical analysis, using the Matched Adjusted Indirect Comparison (MAIC) method. Patients with CHADS$_2$ scores of 0 and 1 were excluded from the XANTUS population to align with the ROCKET AF population, and the balancing weights were calculated. The major bleeding rates were 3.10%/year and 3.60%/year for XANTUS and ROCKET AF, respectively (MAIC rate ratio: 0.86 [95% CI: 0.67–1.12]). Stroke and non-central nervous system (CNS) SE rates were 1.54%/year and 1.70%/year, respectively (MAIC rate ratio: 0.91 [95% CI: 0.62–1.32]). The incidence rates of myocardial infarction and vascular death (restricted to deaths due to cardiovascular causes or intracranial/extracranial bleeding) were also similar between groups, while all-cause death was higher in XANTUS than in ROCKET AF (3.22%/year vs. 1.87%/year, respectively; MAIC rate ratio: 1.72 [95% CI: 1.31–2.27]). The higher rate of all-cause mortality in XANTUS is probably related to the inclusion of patients with serious comorbidities such as cancer who are routinely excluded from RCTs, and the routine nature of the medical care available to XANTUS patients compared with the more intense care patterns within a RCT. A sensitivity analysis (adjusted for different baseline characteristics) also supported the results from the primary analysis. Comparing the outcomes of the appropriately adjusted populations from each study supports the validity of the ROCKET AF trial in a high-risk population, and emphasizes robust data collection from the XANTUS real-world study. Although XANTUS and ROCKET AF were analysed using a thorough statistical methodology, it is important to underline that the results are based on a model and associated limitations apply.[12]

Pooled Analysis of the XANTUS Studies: Exploring the Global Safety Profile of Rivaroxaban in AF

The XANTUS pooled analysis included unselected patients with AF newly starting rivaroxaban (prescribed in accordance with country label) for stroke prevention and included three prospective, observational studies in the program: XANTUS;[4] Xarelto for Prevention of Stroke in Patients With Atrial Fibrillation in Latin-America and EMEA Region (XANTUS EL)[14] and Xarelto for Prevention of Stroke in Patients With Atrial Fibrillation in Asia (XANAP).[15] Patients were followed up for 1 year, at approximately 3-month

intervals, or for ≥30 days after permanent discontinuation (if <1 year). Primary outcomes were treatment-emergent adverse events (AEs) or serious AEs (SAEs), including major bleeding, all-cause mortality and any other AEs/SAEs. Secondary outcomes included symptomatic thromboembolic events, non-major bleeding and treatment persistence. Treatment-emergent major outcomes were adjudicated by one Central Adjudication Committee to minimize reporting bias.[16]

In total, 16,187 patients were screened between June 2012 and December 2014, of whom 11,121 patients were included in the analysis. Patients from 47 countries were grouped as follows: Western Europe/Canada/Israel (47.5%), Eastern Europe (23.2%), East Asia (20.1%), the Middle East/Africa (6.2%) and Latin America (3.0%).[16]

In total, 8,540 (76.8%) patients received rivaroxaban and were followed up for at least the full 12-month study period; treatment was prematurely ended in 2,566 (23.1%) patients, of whom 724 (6.5%) discontinued rivaroxaban due to AEs. More than 96% of the pooled XANTUS population did not experience any treatment-emergent major bleeding, stroke/non-CNS SE or all-cause death. There were 190 treatment-emergent major bleeding events in 172 patients (1.7 events/100 patient-years; 95% CI: 1.5–2.0). The incidence rates of fatal bleeding, critical organ bleeding and intracranial haemorrhage (ICH) were 0.2 events/100 patient-years (95% CI: 0.1–0.3), 0.6 events/100 patient-years (95% CI: 0.5–0.8) and 0.4 events/100 patient-years (95% CI: 0.3–0.6), respectively. Major gastrointestinal (GI) bleeding (the most common bleeding site) occurred in 71 patients (0.7 events/100 patient-years; 95% CI: 0.6–0.9). Symptomatic thromboembolic events occurred at a rate of 1.8 events/100 patient-years (95% CI: 1.6–2.1), and the incidence of ischaemic and haemorrhagic stroke was similarly low. All-cause mortality occurred in 187 (1.7%) patients (1.9 events/100 patient-years; 95% CI: 1.6–2.2). The rate of treatment persistence at 1 year (defined as day 360) was 77.4%.[16]

XANTUS pooled results showed that outcomes with rivaroxaban were consistent across different regions and patient populations worldwide, based on each of the individual XANTUS studies. Results were also consistent with those of the Phase III ROCKET AF study.[13,16]

The Growing Use of Claims Databases to Inform Clinical Practice

Claims data are collected by payers to track and assure reimbursement for healthcare services provided.[17] The data may be longitudinal or cross-sectional and are always retrospective in nature. National databases that contain a wide-ranging assortment of patient information are available across many countries and regions, such as the MarketScan Commercial Claims and Encounter and Medicare in the United States; IMS LifeLink Disease Analyzer in France and Germany; the Clinical Practice Research Datalink in the United Kingdom and Danish and Swedish national databases. Details of these and other databases are summarized in ►Table 2.

Table 2 Overview of a selection of global real-world evidence databases[56]

Database Information	Dept. of Defense MHS EHR Database	MarketScan Commercial Claims and Encounter (CCE)	Medicare	Optum Insight Clinformatics Data Mart	The Health Improvement Network (THIN)	Clinical Practice Research Datalink (CPRD)	Système National d'Information Inter-Régimes de l'Assurance Maladie (SNIIRAM): France	IMS RWD EMR [IMS LifeLink Disease Analyzer]: France	IMS RWD EMR [IMS LifeLink Disease Analyzer]: Germany	Danish NPR	Danish Medical Product Register (MPR)	Danish COD Register	Swedish NPR	Swedish PDR	Swedish COD Register
Country	USA	USA	USA	USA	UK	UK	France	France	Germany	Denmark	Denmark	Denmark	Sweden	Sweden	Sweden
Further information	www.7health.mil	http://www.truvenhealth.com/your_healthcare_focus/pharmaceutical_and_medical_device/	www.resdac.umn.edu	http://www.optum.ca/content/dam/optum/resources/product-Sheets/Clinformatics-Data-Mart.pdf	https://www.ucl.ac.uk/pcph/research-groups/themes/thin-pub/database	http://www.cprd.com/	https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/sniiram/finalites-du-sniiram.php	http://www.imshealth.com	www.imshealth.de	http://esundhed.dk/sundhedsregistre/LPR/Sider/LPR.aspx	www.dkma.dk	http://www.esundhed.dk/sundhedsregistre/DAR01/Sider/DAR01.aspx	http://www.socialstyrelsen.se/english	http://www.socialstyrelsen.se/register/dodsorsaksregistret	http://www.socialstyrelsen.se/register/dodsorsaksregistret
Region	North America, USA; military bases with medical facilities	National (all regions of USA)	National (all regions of USA)	National (all regions of USA)	National (all regions of UK)	National (for primary care data)	National (all regions of France)	National (all regions of France)	National (all regions of Germany)	National (excl. Greenland & the Faroe Islands)	National (all regions of Denmark)	National (Denmark, the Faroe Islands & Greenland)	National (all regions of Sweden)	National (all regions of Sweden)	National (all regions of Sweden)
Database type	LPD, EMR, inpatient, outpatient	LPD, DD data, MPIC, inpatient, outpatient	LPD, DD data, MPIC, inpatient, outpatient	LPD, DD data, MPIC, inpatient, outpatient	LPD, DD data, EMR, MPIC, outpatient	LPD, DD data, EMR, inpatient, outpatient	LPD, DD data, inpatient, outpatient	LPD, DD data, EMR, outpatient	LPD, DD data, EMR, outpatient	LPD, DD data, EMR, inpatient, outpatient	LPD, DD data, outpatient, prescription registry	Death registry	LPD, DD data, EMR (inpatient and outpatient (outpatients non-primary care)	LPD, drug data only, MPIC outpatient	Death registry
Database source	EMRs	MIC	MIC	MIC	EMRs	EMRs (primary care), death registry, medical records, cancer records	MPIC (individual patient data limited to 3 years prior to request. Specific request possible for larger access)	EMRs from GP visits	EMRs, incl. disease & therapy profiles	EMR or medical records	Prescriptions for individuals in primary care	Death certificates, incl. details of events leading to death	EMR or medical records	EMR or medical records	Death certificates
Active population size	10 m active lives	~47 m patients	55.5 m beneficiaries (2015)	~18 m with current eligibility in any given year	~4 m patients (Nov 2016)	>24 m patients active in primary care database (March 2017)	57 m	~1.8 m patients	13.6 m anonymized EMRs (≥1 visit in past 3 y)	5.4 m patients/year	>5m patients	~54,000 deaths per year (varies per year)	1.6 m discharges (inpatients) 11.6 m visits (outpatients)	6.5 m (2016)	~90,000 deaths per year

(Continued)

Table 2 (Continued)

Database Information	Dept. of Defense MHS EHR Database	MarketScan Commercial Claims and Encounter (CCE)	Medicare	Optum Insight Clinformatics Data Mart	The Health Improvement Network (THIN)	Clinical Practice Research Datalink (CPRD)	Système National d'Information Inter-Régimes de l'Assurance Maladie (SNIIRAM); France	IMS RWD EMR [IMS LifeLink Disease Analyzer]; France	IMS RWD EMR [IMS LifeLink Disease Analyzer]; Germany	Danish NPR	Danish Medicinal Product Register (MPR)	Danish COD Register	Swedish NPR	Swedish PDR	Swedish COD Register
Approx. % participants aged >65 y	20%	0% (97% in MarketScan Supplemental and COB database aged >65 y)	84%	11%	24.7% (aged 60+ y)	17.7% (based on March 2011 UK census data)	Unknown	18%	25.5%	~20% (1.3 m patients aged >65 y annually)	16%	83.7% (in 2015)	41% inpatients 29% outpatients	29%	88%
Patient type	Inpatient, outpatient, emergency	Inpatient, outpatient, emergency	Inpatient, outpatient, emergency, prescription drug claims	Inpatient, outpatient	Outpatient/non-institutionalized (primary care records)	Inpatient, outpatient, with full hospital records	Inpatient, outpatient	Outpatient/non-institutionalized	Outpatient/non-institutionalized	Inpatient, outpatient, emergency, psychiatric, hospital exam & treatment	Outpatient	Inpatient, outpatient, emergency, nursing homes	Inpatient, outpatient (outpatient do not include primary care)	Outpatient	Decedents
Death recorded	Yes[a]	Yes	Yes	No	Yes	Yes	Yes (date only—cause not available)	No	No	No; death data available via COD Register with PRN	No; death data available via COD Register with PRN	Yes	Yes	Yes	Yes
Diagnosis coded	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM, DRG	ICD-9-CM, DRG	ICD-9-CM, DRG	READ, (complemented by internally created AHD codes, records variety of clinical measures)	READ, ICD-10 in linked datasets (e.g. death & hospitalization)	ICD-10	ICD-10	ICD-10	ICD-8, ICD-10, +K37	N/A	ICD-8 (through 1993), ICD-10 (1994 +)	ICD-10	N/A	ICD-6 (1952–1957), ICD-7 (1958–1968), ICD-8 (1969–1986), ICD-9 (1987–1996), ICD-10 (1997–2015)
Procedure coded	ICD-9-CM, CPT, HCPCS	CPT, HCPCS, ICD-10, ICD-9, LIONC	CPT, HCPCS	CPT-4, HCPCS, ICD-9, LIONC	READ, AHD codes	ICD-10, OPCS, classification surgical ops & procedures, (v4), READ	Yes (code unknown)	N/A, procedure data not routinely collected	N/A, procedure data not routinely collected	Danish NCSP (1977–1995), NOMESCO (1996 +)	N/A	ICD-8 (through 1993), ICD-10 (1994 +)	ICD-10	N/A	N/A
Laboratory information	Yes	Yes	No	Yes	Yes	Yes	Yes (no results)	Yes	Yes	No	N/A	Yes	No	N/A	Yes
Info on drugs administered in hospital?	No	Yes	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	N/A

Table 2 (*Continued*)

Database Information	Dept. of Defense MHS EHR Database	MarketScan Commercial Claims and Encounter (CCE)	Medicare	Optum Insight Clinformatics Data Mart	The Health Improvement Network (THIN)	Clinical Practice Research Datalink (CPRD)	Système National d'Information Inter-Régimes de l'Assurance Maladie (SNIIRAM): France	IMS RWD EMR [IMS LifeLink Disease Analyzer]: France	IMS RWD EMR [IMS LifeLink Disease Analyzer]: Germany	Danish NPR	Danish Medicinal Product Register (MPR)	Danish COD Register	Swedish NPR	Swedish PDR	Swedish COD Register
Drug data: pharmacy or hospital?	Pharmacy only	Pharmacy & hospital	Pharmacy & hospital	Pharmacy (mostly)	Pharmacy only	No	Pharmacy only (hospital: limited)	N/A, only covers GP prescriptions	N/A, info only covers GP prescriptions	N/A	Pharmacy only	No	Hospital (pharmacy data available from Swedish PDR)	Pharmacy only	N/A
Drug regimen	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A	No	No	N/A (see Swedish PDR)	Yes	N/A
Drug route	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	No	N/A (see Swedish PDR)	Yes	N/A
Drug dosage	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	No	N/A (see Swedish PDR)	Yes	N/A
Drug days' supply	Yes	Yes	Yes	Yes	No; can be calculated	Yes (limited)	Yes	Yes	Yes	N/A	Yes	No	N/A (see Swedish PDR)	Yes	N/A
Direct access to medical records	Yes (limited)	No	Yes	Yes, in collaboration with vendor	Yes	No	No	No	No	No	No	No	N/A	Yes	Yes
Linkage to other databases	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Linkage capabilities	Yes	Yes	Yes	Yes (additional fee)	Yes	Yes	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes
Database access	Third-party access	Third-party access	Direct or third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Third-party access	Direct or third-party access	Direct or third-party access	Direct or third-party access

[a]Death recorded is not always precise. There may be a delay in the records indicating the date of death. Eventually, there is a connection in the U.S. Department of Defense system of the members, death once they lose their medical benefits and retirement pay.

Abbreviations: AHD, additional health data; COB, coordination of benefits; COBRA, The Consolidated Omnibus Budget Reconciliation Act; COD, cause of death; CPT(-4), Current Procedural Terminology (fourth edition); DD, drug and diagnosis; DRG, diagnosis-related group; HER, electronic health records; EMR, electronic medical records; GP, general practitioner; HCPCS, Healthcare Common Procedure Coding System; ICD(-9[-CMJ]/10[-CMJ]), International Statistical Classification of Diseases and Related Health Problems (9th Revision [Clinical Modification]/10th Revision [Clinical Modification]); LOINC, Logical Observation Identifiers Names and Codes; LPD, longitudinal population database; MHS, Military Health Service; MIC, medical insurance claim; MPIC, medical and pharmacy insurance claims; MS-DRG, Medicare Severity-Diagnosis Related Group; N/A, not applicable; NCSP, Nordic Classification of Surgical Procedures; NOMESCO, Nordic Medico-Statistical Committee Classification of Surgical Procedures; NPR, National Patient Register; OPCS, Office of Population Censuses and Surveys; PDR, Prescribed Drug Register; PRN, (Danish) personal registry number; READ, Read Coding System for Clinical Terminology (UK); RWD, real-world data.

Recorded data come from a variety of sources, including EHRs, death certificates, prescriptions and insurance claims. Some claims databases are single payer, such as Medicare,[18] the U.S. Department of Defense (DoD) Military Health System (MHS)[19] and national healthcare databases, such as the Danish Nationwide Database.[20–22] Other claims data are drawn from multiple different payers and are assembled by specific health analytic companies, such as Truven MarketScan[23] and Optum Insight.[24] The Danish Nationwide Database encompasses several different databases (e.g. medicine claims from one database and medical records from another).[20–22] A distinguishing factor of the Danish databases is that every person in the country is given a number at birth that follows them throughout life to include prescriptions, medical procedures and other vital health-related information. The U.S. DoD MHS database is similar to this in that it permits longitudinal follow-up data for many years. These databases may be suitable sources of data for diseases and conditions that require several years' observation time, such as recurrent venous thromboembolism (VTE). However, most commercial U.S. databases do not have such a long follow-up period. Moreover, for those that are linked to insurance claims, the data may not be as cohesive because as people move jobs, datasets may contain only 2 to 3 years' worth of information per insurance company. It could be assumed that there may be lower risk of loss to follow-up with lifetime databases compared with those that link to insurance claims.

Databases predominantly include data needed to facilitate reimbursement of healthcare services, such as basic demographics, types and number of healthcare encounters, diagnostic codes (International Classification of Diseases [ICD]-9 or 10) and other billing codes associated with each healthcare encounter and prescription drug fill records.[25,26] Each healthcare encounter is usually associated with multiple codes.[27] The first or primary diagnosis code listed is generally assumed to depict the main reason for the encounter. However, there are cases where this may not be true. A good example relating to anticoagulation is pulmonary embolism. The primary diagnosis code may be listed as mechanical ventilation or thrombolysis/thrombectomy, as these may be associated with better reimbursement. Secondary or non-primary codes are used to demonstrate patient complexity/acuity (and possible justification for higher reimbursement) by designating comorbidities. Based on review of medical records for an encounter, diagnostic codes are often chosen by a clerical coder/biller or they may be initially reviewed by a healthcare professional, and subsequently reviewed by a clerical secretary upon discharge (in the case of the Nordic registries).

Diagnostic coding (the presence of a code, its position and the type of encounter it is associated with) and prescription fill records are used to determine if patients have the disease state of interest (e.g. ICD-9 of 427.31/ICD-10 of 148 suggests AF), whether comorbidities of interest are present (e.g. CHA_2DS_2-VASc or HAS-BLED scores to clinically predict the risk of stroke or bleeding, respectively, in patients with AF) within some historical time frame (often

6–12 months prior to oral anticoagulant [OAC] initiation or an index event), the occurrence of an outcome of interest (ischaemic stroke or major bleeding, e.g. intracranial or GI bleeding occurred); the drug therapies and doses that have been used are known (anticoagulants, antiplatelet agents [excluding over-the-counter medications such as acetylsalicylic acid (ASA)]); persistence to index therapies (non-vitamin K antagonist [VKA] oral anticoagulant [NOAC] and VKAs) and 'on-treatment' status are established.[26,28] The next section will explore several recent country-specific databases that have investigated the safety and effectiveness of NOACs including rivaroxaban in real-world management of AF.

The Value and Limitations of Database Analyses

Database analyses provide insights on practical clinical management issues and patient outcomes from large cohorts.[3] Results from a 2015 online survey of 1,003 UK-based general practitioners showed that 30% of respondents felt that real-world data helped them identify how best to fit a treatment into their current practice, while 27% said it helped validate Phase III trial data and 25% said it indicated how medicines are being used by their peers in clinical practice.[29]

Despite the utility of database analyses, there are several limitations. Differential misclassification bias is caused by the inaccurate or insufficient classification of patients, and can be caused by errors in diagnostic coding by the medical coder (or, in fact, the coder may not be medically trained); 'tactical' coding (upcoding) or poor coding detail (e.g. there is a diagnosis code for heart failure (HF), but ejection fraction and New York Heart Association classification cannot be determined).[26,30] For this reason, use of classification methods could be considered to increase validity, such as the U.S. Food and Drugs Administration (FDA) Mini-Sentinel coding system[31] and the Cunningham algorithm for major bleeding.[32] Confounding bias is the inherent imbalance between characteristics associated with the outcome and exposure.[33] Importantly, confounding can only be taken into account if the data source provides information on relevant covariates, while other sources of confounding from unobserved or unsuspected covariates may remain. Also, the methods used to adjust or balance patients on appropriate characteristics may not be ideal, such as regression and/or propensity scoring (owing to adjustment, inverse weighting, 'greedy matching' or high-dimensional matching).[33] Therefore, characteristics adjusted/matched upon should take into account not only the disease state but also the endpoints of interest (e.g. matching on CHA_2DS_2-VASc and HAS-BLED criteria may be sufficient for stroke or major bleeding in NVAF, but this may not be the case for all-cause mortality). Furthermore, residual confounding will always exist in the absence of randomization.[33]

In addition, it is worth remembering that these data were not originally intended for research purposes and may be missing vital information such as laboratory or diagnostic test results (e.g. creatinine clearance, serum creatinine

or computed tomography results); vital signs or other clinical characteristics (heart rate, body weight); out-of-hospital mortality data (while included in government databases, access is often restricted and is not reliably reported); qualitative data (quality of life, patient satisfaction, explanations for treatment decisions); listing of medical history or comorbidities.[25,26] Additionally, the information included may not fully represent prescribing habits, that is, the reasons for choosing one type of therapy over another. This may be especially pertinent for NOACs, as the agents were licensed for stroke prevention in AF at different times. This in turn may present an important channeling bias (patients may have been more likely to switch from VKA to dabigatran, then to rivaroxaban and apixaban) in patients who have a history of anticoagulation therapy. Despite these potential limitations, healthcare claims and EHR databases are regularly used to uncover key information on NOAC use in daily practice. Recent results from several noteworthy analyses focused on rivaroxaban across the United States and Europe will be covered in the next section.

Rivaroxaban in AF: Real-World Findings from National Datasets

Effectiveness and Safety of Reduced Dose NOACs and Warfarin in Patients with NVAF: A Propensity-Weighted Nationwide Cohort Study from Denmark

A large propensity-weighted study of more than 60,000 patients included in a Danish observational nationwide cohort showed that apixaban, dabigatran and rivaroxaban could be used effectively as alternatives to warfarin in a clinical care setting.[34] NOACs were comparable to warfarin with regards to the overall incidence of ischaemic stroke, although certain differences were noted. Rivaroxaban was associated with a lower risk of ischaemic stroke or systemic embolism (combined endpoint) than warfarin (no significant differences were noted for dabigatran and apixaban). A significantly lower risk of death, any bleeding and major bleeding were noted with apixaban and dabigatran, compared with rivaroxaban or warfarin, of which both had similar bleeding risk profiles.[34]

The use of lower NOAC doses is common in routine clinical practice, and may occur for a variety of reasons such as renal insufficiency or use of concomitant medications. In a study of almost 15,000 patients with AF using a large U.S. administrative database, Yao et al provided evidence of inappropriate dose reductions in the absence of renal indications, and conversely, failure to reduce the dose when there was a clear need to do so.[35] Moreover, apixaban dose reduction in patients with no severe renal impairment was related to reduced effectiveness for stroke prevention without any safety benefit.[35]

NOAC overdosing and underdosing may compromise drugs' safety without improving their effectiveness and further evidence around the impact of alternative dosing on outcomes is lacking. A subsequent Danish nationwide observational study published findings on the safety and effectiveness of reduced-dose NOACs compared with warfarin in patients with NVAF.[36] The clinical effectiveness and safety of apixaban 2.5 mg twice daily (bid), dabigatran 110 mg bid and rivaroxaban 15 mg once daily (od) were each compared with warfarin in patients with NVAF who had not previously taken OACs (initial OAC prescription was between August 2011 and February 2016). A propensity score was calculated for the four treatment alternatives. Based on the propensity of receiving either treatment, an inverse probability treatment weighting approach was applied to account for any baseline differences in the study population. Patients were followed up from treatment onset. The primary effectiveness outcomes were ischaemic stroke and systemic embolism (SE); the primary safety outcomes were any bleeding events.

There were 55,644 patients with NVAF (mean age: 73.9 years) who met the inclusion criteria—of these, 69.9% received warfarin ($n = 38,893$), 7.9% received apixaban ($n = 4,400$), 15.9% received dabigatran ($n = 8,875$) and 6.3% received rivaroxaban ($n = 3,476$).[36] The highest prevalence of renal disease was shown in the apixaban and rivaroxaban groups (9.5 and 9.1%, respectively), compared with 3.9 and 8.3% in the dabigatran and warfarin groups, respectively. Patients receiving apixaban had more comorbidities and a higher mean CHA$_2$DS$_2$-VASc score than patients in other treatment groups (4.3 vs. 3.8, 3.6 and 3.0 for dabigatran, rivaroxaban and warfarin, respectively). A total of 1,779 patients experienced ischaemic stroke or SE events during the first year of follow-up. The weighted event rates of ischaemic stroke/SE for apixaban, dabigatran, rivaroxaban and warfarin were 4.8, 3.3, 3.5 and 3.7%, respectively. For the NOAC versus warfarin comparison, the effectiveness outcomes hazard ratios (HRs) from the inverse probability of treatment-weighted analysis were 1.19 for apixaban (95% confidence interval [CI]: 0.95–1.49), 0.89 for dabigatran (95% CI: 0.77–1.03) and 0.89 for rivaroxaban (95% CI: 0.69–1.16). The incidence of any bleeding events for NOACs versus warfarin was 0.96 for apixaban (95% CI: 0.73–1.27), 0.80 for dabigatran (95% CI: 0.70–0.92) and 1.06 for rivaroxaban (95% CI: 0.87–1.29).[36]

The cohort stratified by indication for potential dose reduction (age $\geq$80 and/or renal disease: $n = 21,949$) exhibited worse outcomes than those in the main analysis. Apixaban 2.5 mg bid was associated with higher rates of ischaemic stroke/SE compared with warfarin, while rivaroxaban 15 mg od and dabigatran 110 mg bid showed a trend toward lower thromboembolic rates—but neither comparison was statistically significant. The incidence of any bleeding events was significantly lower for dabigatran versus warfarin, but not for apixaban or rivaroxaban versus warfarin. While the data provide valuable evidence on the use of reduced-dose NOACs in clinical practice, unmeasured residual confounding and selective prescribing behaviour may have been limiting factors, and misclassification and miscoding of diagnoses and outcomes may have been possible. The authors recommended future studies on effectiveness and safety with respect to label

adherence, particularly for those patients with renal impairment.[36]

Effectiveness and Safety of NOACs versus Warfarin in Patients with NVAF and Previous Stroke or Transient Ischaemic Attack: The REAFFIRM Study

NVAF patients experiencing a prior ischaemic stroke or transient ischaemic attack (TIA) are at increased risk of recurrent stroke and bleeding. REAFFIRM evaluated the real-world effectiveness and safety of apixaban, dabigatran and rivaroxaban versus warfarin in NVAF patients with a prior history of ischaemic stroke or TIA, using U.S. Truven MarketScan claims from January 2012 to June 2015.[37] Inclusion criteria included ≥ 2 ICD-9 diagnosis codes for AF (without codes suggesting valvular disease); age ≥ 18 years; a history of prior ischaemic stroke or TIA; recent initiation of apixaban, dabigatran, rivaroxaban or warfarin and ≥ 180 continuous days of prior medical/prescription coverage (baseline period). Excluded patients were those with a transient cause of NVAF, VTE, hip or knee arthroplasty, malignant cancer, pregnancy or those prescribed greater than 1 OAC. Each eligible apixaban, dabigatran and rivaroxaban user was 1:1 propensity score-matched (caliper of 1%) to a warfarin user. Residual differences between matched cohorts were assessed via standardized differences (<10% was considered well balanced).[37]

The numbers of patients in each of the three independent matched cohorts were as follows: apixaban versus warfarin: $n = 2,514$; dabigatran versus warfarin: $n = 1,962$; rivaroxaban versus warfarin: $n = 5,208$ (mean durations [$\pm$ standard deviation] of follow-up: 0.5 [$\pm$ 0.5], 0.6 [$\pm$ 0.6] and 0.6 [$\pm$ 0.6] years, respectively). A total of 20.8, 17.7 and 26.2% of patients received reduced doses of apixaban (<5 mg bid), dabigatran (<150 mg bid) and rivaroxaban (<20 mg od), respectively. Results from these independent statistical analyses of each NOAC versus warfarin were relatively consistent with their respective Phase III trials and prior stroke/TIA subgroup analyses. Using Cox regression, neither apixaban nor dabigatran reduced the combined primary endpoint of ischaemic stroke or intracranial haemorrhage (ICH; HR: 0.70; 95% CI: 0.33–1.48 and HR: 0.53; 95% CI: 0.26–1.07, respectively) and had a nonsignificant effect on hazards of major bleeding (HR: 0.79; 95% CI: 0.38–1.64 and HR: 0.58; 95% CI: 0.26–1.27, respectively) versus warfarin. Rivaroxaban reduced the combined endpoint of ischaemic stroke or ICH (HR: 0.45; 95% CI: 0.29–0.72) without an effect on major bleeding (HR: 1.07; 95% CI: 0.71–1.61). ICH occurred at rates of 0.16 to 0.61 events per 100 person-years in the three NOAC analyses, with no significant difference for any NOAC versus warfarin.[37]

Limitations of the analysis included misclassification bias, underpowered analyses and residual confounding, along with short follow-up, insufficient reporting of clinical/laboratory data and confirmation of whether patients were still taking the OAC. These results should be reassuring to clinicians prescribing a NOAC to patients with a prior ischaemic stroke or TIA. However, the authors discourage any type of cross-comparison between oral anticoagulation analyses due to the distinct populations in each of the three cohorts.[37]

U.S. DoD MHS Post-Marketing Safety Surveillance Study

The protocol for the U.S. DoD MHS Post-Marketing Safety Surveillance study (PMSS) was designed by Janssen in collaboration with the U.S. FDA and is evaluating the safety profile of rivaroxaban in a large number of patients with NVAF.[19] Data are from the U.S. DoD MHS, which, compared with other U.S. claims databases, can follow up military personnel and their relatives for many years. Initial data published from 27,467 patients with NVAF who were treated with rivaroxaban showed that the real-world rate of major bleeding with rivaroxaban—assessed using the validated Cunningham algorithm—was 2.86 per 100 person-years (496 major bleeding events in 478 patients: 95% CI: 2.61–3.13).[19] Of the patients with major bleeding, 63.2% had been treated with rivaroxaban 20 mg od (vs. 32.2 and 4.6% who received rivaroxaban 15 and 10 mg od, respectively). Compared with the non-major bleeding group, patients with major bleeding were older (mean age: 78.4 vs. 75.7 years), and had a higher prevalence of comorbidities, including coronary artery disease (64.2 vs. 36.7%), HF (48.5 vs. 23.7%) and hypertension (95.6 vs. 75.8%). Most bleeding events were GI (88.5%), followed by intracranial (7.5%). Fatal bleeding outcomes were low: 0.08 per 100 person-years (95% CI: 0.05–0.14).[19]

Subsequent data from 44,793 patients in the U.S. DoD MHS PMSS showed a strong relationship between CHA_2DS_2-VASc scores and the risk of major bleeding in rivaroxaban-treated patients with NVAF.[38] Among known high-risk populations, the highest rate of major bleeding by CHA_2DS_2-VASc component was 5.69 (95% CI: 5.18–6.24) per 100 person-years for patients with vascular disease. The greatest incidence of ICH was noted in those with previous stroke, TIA or VTE, who were at greater risk of mortality. Limitations included capture of drug-dispensing information only; a temporal association being established between rivaroxaban and a bleeding event, rather than a causal association; possible inaccurate reporting of major bleeding events (potentially due to false positives or under-reporting owing to database limitations) and lack of International Normalized Ratio (INR) data, which prevented precise calculation of HAS-BLED scores.[38]

Diabetes mellitus (DM) is a major risk factor for stroke and increases the incidence of AF. An additional analysis of 44,793 patients revealed that patients with DM had more comorbidities and a higher incidence of major bleeding than those without DM.[39] A total of 12,039 (26.9%) had DM, and were more likely to be male, younger, and with more comorbidities and higher CHA_2DS_2-VASc scores. Major bleeding incidence in the DM versus the non-DM group was 3.68 (95% CI: 3.37–4.03) versus 2.51 (95% CI: 2.34–2.69), respectively per 100 person-years (ROCKET AF incidence of major bleeding: 3.79 in the DM group vs. 3.47 per 100 person-years in the non-DM group). Intracranial bleeding incidence in the DM versus the non-DM group was 0.19

(95% CI: 0.13–0.28) and 0.25 (95% CI: 0.20–0.31) per 100 person-years, respectively. Fatal outcomes were rare for both cohorts (0.09 per 100 person-years). The limitations of this study are the same as the others described earlier. It is also worth noting that the definition of DM was based on diagnosis codes, not on laboratory values. This means that important clinical indicators, such as glycated haemoglobin levels and their relation to major bleeding, were not included in the analysis. Similarly, the role of glycaemic management and its potential effect on major bleeding remains unknown within this population.[39]

Real-World Comparative Effectiveness of Oral Anticoagulants: Focus on Bleeding Risk

Bleeding is a key consideration when prescribing any anticoagulant. In the past, higher GI bleeding risks have been reported with NOACs over warfarin.[40,41] One notable large retrospective analysis of medical and pharmacy administrative claims from a diverse U.S. database (Optum Labs Data Warehouse) of more than 100 million enrolees published data on the real-world risk of GI bleeding with dabigatran, rivaroxaban and warfarin in AF.[42] A total of 92,816 patients were taking anticoagulants during the study period (1 November 2010 and 30 September 2013; dabigatran: 9.2% [$n = 8,578$], rivaroxaban: 17.5% [$n = 16,253$] and warfarin: 73.2% [$n = 67,985$]).[42] In AF patients, the incidence of GI bleeding per 100 patient-years with dabigatran versus warfarin was 2.29 (95% CI: 1.88–2.79) and 2.87 (95% CI: 2.41–3.41), respectively. For rivaroxaban versus warfarin in AF patients, the incidence of GI bleeding per 100 patient-years was 2.84 (95% CI: 2.30–3.52) and 3.06 (95% CI: 2.49–3.77), respectively. In non-AF patients, the incidence of GI bleeding per 100 patient-years for dabigatran versus warfarin was 4.10 (95% CI: 2.47–6.80) and 3.71 (95% CI: 2.16–6.40), respectively; and for rivaroxaban versus warfarin it was 1.66 (95% CI: 1.23–2.24) and 1.57 (95% CI: 1.25–1.99), respectively. In propensity score-matched models, the risk of GI bleeding with NOACs versus warfarin was similar in both AF patients (dabigatran vs. warfarin, HR: 0.79 [95% CI: 0.61–1.03]; rivaroxaban vs. warfarin, HR: 0.93 [95% CI: 0.69–1.25]) and non-AF patients (dabigatran vs. warfarin, HR: 1.14 [95% CI: 0.54–2.39]; rivaroxaban vs. warfarin, HR: 0.89 [95% CI: 0.60–1.32]). However, by the age of 76 years, the risk of GI bleeding was higher than with warfarin in AF patients on dabigatran (HR: 2.49 [95% CI: 1.61–3.83]) and in patients with and without AF on rivaroxaban (2.91 [95% CI: 1.65–4.81] and 4.58 [95% CI: 2.40–8.72], respectively). The authors therefore exercised caution when prescribing NOACs to elderly patients, particularly those older than 75 years.[42] The data generally support the safety of NOACs in the context of GI bleeding risk and emphasize the importance of matching and adjusting for the correct variables. This is especially pertinent as GI bleeding can be influenced by non-CV factors such as nonsteroidal anti-inflammatory drug and ASA use, and pre-existing GI disease.[43]

A recent analysis using the VA pharmacy database (803 patients on NOACs and 6,263 on warfarin) showed that the incidence of clinically significant GI bleeding in hospitalized patients was approximately four times higher in those receiving warfarin than those receiving NOACs.[44] Further complexity when interpreting data stems from the fact that differential outcomes have been noted across several database studies investigating NOACs compared with warfarin, favouring one NOAC over another in relation to effectiveness and/or safety.[45–48] This may be a reflection of the heterogeneity of patient characteristics noted in clinical practice and in each respective database, the intrinsic properties of the NOACs and/or the doses used, as well as the methodology used.

Meta-analyses are a valuable form of comparative effectiveness research because they emphasize the magnitude of intervention effects rather than relying on tests of statistical significance among primary studies.[49] The contribution made to the totality of the evidence in systematic reviews by studies with non-significant results is as important as those with statistically significant results. Many meta-analyses have investigated NOACs in AF patients and while studies of this kind are valuable, the approach must be robust to minimize bias and provide a strong basis on which to draw conclusions.[50] Meta-analyses have also produced conflicting findings in relation to bleeding risk when evaluating NOACs compared with VKAs. One meta-analysis of 32 RCTs and observational studies in AF patients showed more favorable outcomes for apixaban, dabigatran and edoxaban in relation to major bleeding and ICH versus VKAs, which was not seen with rivaroxaban.[51] Another meta-analysis of 28 observational nationwide or health insurance databases reporting matched or adjusted results comparing apixaban, dabigatran or rivaroxaban with VKAs in AF patients concluded that while all NOACs reduced ICH, apixaban was associated with the fewest GI and major bleeding events versus VKAs.[52] A large systematic review and meta-analysis of 43 RCTs ($N = 166,289$ patients) concluded that the risk of major GI bleeding between NOACs and conventional anticoagulation was generally similar, although findings suggested that dabigatran and rivaroxaban may be associated with increased odds of major GI bleeding. The authors concluded that further high-quality studies are needed to characterize GI bleeding risk among NOACs.[53]

While many studies have been undertaken to compare NOACs with warfarin in the setting of AF, direct comparisons between NOACs are less common. One analysis compared the risks of thromboembolic stroke, ICH, major extracranial bleeding (including major GI bleeding) and mortality in patients with NVAF who initiated dabigatran (150 mg bid) or rivaroxaban (20 mg od) for stroke prevention.[54] Patients were enrolled in fee-for-service Medicare and hence were restricted to 65 years or older. Differences in baseline characteristics were adjusted using stabilized inverse probability of treatment weights based on propensity scores. Of 52,240 dabigatran-treated and 66,651 rivaroxaban-treated patients, rivaroxaban was found to be associated with significantly increased ICH (HR: 1.65; 95% CI: 1.20–2.26; $p = 0.002$; adjusted incidence rate differences [AIRD]: 2.3 excess

cases/1,000 person-years) and major extracranial bleeding (HR: 1.48; 95% CI: 1.32–1.67; $p < 0.001$; AIRD: 13.0 excess cases/1,000 person-years), including major GI bleeding (HR: 1.40; 95% CI: 1.23–1.59; $p < 0.001$; AIRD: 9.4 excess cases/ 1,000 person-years), compared with dabigatran. There was a reduced risk of thromboembolic stroke with rivaroxaban, but this finding was not significant when data were adjusted for underlying cardiovascular disease. There was no significant difference in the risk of mortality between rivaroxaban and dabigatran.[54] The authors acknowledge that the observational nature of the study may be subject to residual confounding, as well as the restrictions of the claims database used (patients ≥65 years). Earlier findings from Graham *et al* using Medicare source data showed that dabigatran was associated with an increased risk of major GI bleeding compared with warfarin—an analysis that has been postulated to be in line with the RE-LY trial data.[55] The magnitude of the effects with dabigatran versus rivaroxaban or warfarin in observational studies may not easily compare with registrational RCTs due to different patient populations and patient monitoring during follow-up. Thus, small differences in observed treatment effects must be cautiously interpreted, along with the likelihood of whether such small differences could have any clinical relevance in routine practice.

We must also not forget the possibility that the bleeding events observed in real-world analyses such as these could simply be attributed to the drugs themselves. In addition, outcomes may be impacted by the conditions in which NOACs are administered, for example the dosing and frequency. There are myriad factors that may influence clinical outcomes observed with NOACs in routine practice, such as patient characteristics, methodology and bleeding definitions. Owing to the discrepancies in findings, therapy may need to be individualized based on patient need.

Summary

All studies contribute vital pieces of information to the totality of evidence toward understanding the true clinical value of a given therapy. One study should not be replaced with another and there is not a 'one-size-fits-all' approach to answering research questions. Nevertheless, data published from a large number of real-world studies performed in various countries and settings have and continue to produce findings that support the effectiveness and safety of NOACs in patients with AF. It is worth noting that inconsistencies across methods used in these studies can substantially impact the findings. Not all questions can be adequately addressed within all types of real-world studies. It is important to consider the therapies being compared, the conclusions being extrapolated and how analyses differ alongside registrational RCTs. It is clear that NOACs do proffer advantages over traditional standard of care—it is the magnitude of these effects that needs to be considered in more detail with appropriately designed studies.

What is known about this topic?

Real-world data have become a cornerstone in the evaluation of therapies, and they complement randomized controlled trials in terms of the knowledge gaps they can address. Real-world evidence can be generated from a variety of sources, such as claims databases, electronic health records, non-interventional studies (NIS) and registries. There are myriad considerations for researchers and clinicians to ensure data are robust and relevant. Real-world data have played a prominent role in supporting the effectiveness and safety the non-vitamin K antagonist oral anticoagulant rivaroxaban in daily practice.

What does this paper add?

This review underlines the value of undertaking real-world studies, and the key questions one must consider to ensure the data are clinically meaningful. We also present crucial questions to address when evaluating real-world studies, which are presented in the form of a checklist. Recent updates from a variety of real-world studies investigating rivaroxaban for stroke prevention in patients with atrial fibrillation are covered, including the nationwide Danish cohort study; US Department of Defense Military Health System database; retrospective claim database study REAFFIRM; and a pooled analysis from the global NIS XArelto on preveNtion of sTroke and non-central nervoUS system systemic embolism in patients with non-valvular atrial fibrillation (XANTUS).

Conflicts of Interest

All authors confirm that they have had full access to data and contributed to drafting of the paper. A.J.C. has acted in the capacity of advisor and investigator for Bayer AG, Boehringer Ingelheim, Pfizer/BMS and Daiichi Sankyo. C.I.C. has received research funding from Bayer AG, Janssen Pharmaceuticals; consulting/honorarium from Bayer AG and Janssen Pharmaceuticals and Boehringer Ingelheim Pharmaceuticals. P.B.N. has received speaking fees from Boehringer Ingelheim, consulting fees from Bayer and BMS/Pfizer and grant support from BMS/Pfizer. T.B.L. has served as an investigator for Janssen Scientific Affairs, LLC, and Boehringer Ingelheim and has served as a speaker

for Bayer, Bristol-Myers Squibb/Pfizer and Boehringer Ingelheim. S.T. has no financial interests to declare.

References

1 Cohen AT, Goto S, Schreiber K, et al. Why do we need observational studies of everyday patients in the real-life setting? Eur Heart J Suppl 2015;17(Suppl D):D2–D8

2 Frieden TR. Evidence for health decision making - beyond randomized, controlled trials. N Engl J Med 2017;377(05):465–475

3 Annemans L, Aristides M, Kubin M. Real-life data: a growing need. ISPOR. Available at: https://www.ispor.org/news/articles/oct07/rld.asp. Accessed February 8, 2018

4 Camm AJ, Amarenco P, Haas S, et al; XANTUS Investigators. XANTUS: a real-world, prospective, observational study of patients treated with rivaroxaban for stroke prevention in atrial fibrillation. Eur Heart J 2016;37(14):1145–1153

5 Dreyer NA, Bryant A, Velentgas P. The GRACE Checklist: a validated assessment tool for high quality observational studies of comparative effectiveness. J Manag Care Spec Pharm 2016;22(10):1107–1113

6 Benchimol EI, Smeeth L, Guttmann A, et al; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015;12(10):e1001885

7 Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. AMIA Annu Symp Proc 2006;2006:359–363

8 European Union. DIRECTIVE 2001/20/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Available at: http://ec.europa.eu/health//sites/health/files/files/eudralex/vol-1/dir_2001_20/dir_2001_20_en.pdf. Accessed February 8, 2018

9 Theobald K, Capan M, Herbold M, Schinzel S, Hundt F. Quality assurance in non-interventional studies. Ger Med Sci 2009;7:Doc29

10 Turpie AG, Haas S, Kreutz R, et al. A non-interventional comparison of rivaroxaban with standard of care for thromboprophylaxis after major orthopaedic surgery in 17,701 patients with propensity score adjustment. Thromb Haemost 2014;111(01):94–102

11 Ageno W, Mantovani LG, Haas S, et al. Safety and effectiveness of oral rivaroxaban versus standard anticoagulation for the treatment of symptomatic deep-vein thrombosis (XALIA): an international, prospective, non-interventional study. Lancet Haematol 2016;3(01):e12–e21

12 Camm AJ, Amarenco P, Haas S, et al. Real-world versus randomized trial outcomes in similar populations of rivaroxaban-treated patients with non-valvular atrial fibrillation in ROCKET AF and XANTUS. Presented at the American College of Cardiology (ACC) 66th Annual Scientific Session and Expo, Washington, DC, March 17–19, 2017, Poster 084

13 Patel MR, Mahaffey KW, Garg J, et al; ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 2011;365(10):883–891

14 ClinicalTrials.gov. Xarelto for prevention of stroke in patients with atrial fibrillation in Latin America and EMEA Region (XANTUS-EL). NCT01800006. Available at: https://clinicaltrials.gov/ct2/show/NCT01800006. Accessed February 8, 2018

15 ClinicalTrials.gov. Xarelto for prevention of stroke in patients with atrial fibrillation in Asia (XANAP). NCT01750788. Available at https://clinicaltrials.gov/ct2/show/NCT01750788. Accessed February 8, 2018

16 Kirchhof P, Radaideh G, Kim YH, et al. Safety analysis of rivaroxaban: a pooled analysis of the global XANTUS programme. (real world, prospective, observational studies for stroke prevention in patients with atrial fibrillation). Eur Heart J 2017;38(Suppl):768–769

17 Riley GF. Administrative and claims records as sources of health care cost data. Med Care 2009;47(07, Suppl 1):S51–S55

18 Centers for Medicare & Medicaid Services. Medicare. Available at: https://www.cms.gov/Medicare/Medicare.html. Accessed February 8, 2018

19 Tamayo S, Frank Peacock W, Patel M, et al. Characterizing major bleeding in patients with nonvalvular atrial fibrillation: a pharmacovigilance study of 27 467 patients taking rivaroxaban. Clin Cardiol 2015;38(02):63–68

20 Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health 2011;39(7, Suppl):30–33

21 Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health 2011;39(7, Suppl):38–41

22 Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011;39(7, Suppl):22–25

23 Hansen L. The MarketScan® Databases for Life Sciences Researchers. White Paper, May 2016. Available at: http://content.truvenhealth.com/rs/699-YLV-293/images/%7B87d8921a-c27c-4382-bd88-d8ec9011de70%7D_2016_Truven_Health_MarketScan_white_paper_for_Life_Sciences.pdf?aliId=1274509. Accessed February 8, 2018

24 Wallace P. Optum Labs - Partners. Data and Design. Available at: http://www.himssasiapac.org/sites/default/files/HIMSSAP_AP15Slides_UnitingDataAndPartnershipToConfrontAndSolve-Healthcare-sBigChallenges.pdf. Accessed February 8, 2018

25 Humedica. Real Health Data. Healthcare Database Information. Available at: https://hinora.uncc.edu/sites/hinora.uncc.edu/files/media/Database%20Narrative_19mar2013.pdf. Accessed February 8, 2018

26 Gandhi S, Salmon JW, Kong SX, et al. Administrative databases and outcomes assessment: an overview of issues and potential utility. J Manag Care Spec Pharm 1999;5(03):215–222

27 ProfNet Medical. Questions and answers on ICD10 coding. Available at: http://www.profnetmedical.co.za/media/1080/qa-on-icd-10-v2.pdf. Accessed February 8, 2018

28 Coleman CI, Antz M, Bowrin K, et al. Real-world evidence of stroke prevention in patients with nonvalvular atrial fibrillation in the United States: the REVISIT-US study. Curr Med Res Opin 2016;32(12):2047–2053

29 MHP Communications. The tricky second album: From RCTs to RWE, does industry have the right mix for clinicians' tastes? Available at https://issuu.com/mhpcommunications/docs/mhp_health_tricky_second_album. Accessed February 8, 2018

30 Research Data Assistance Center. Strengths and limitations of CMS administrative data in research. Available at: https://www.resdac.org/resconnect/articles/156. Accessed February 8, 2018

31 Mini-Sentinel. Mini-sentinel medical product assessment a protocol for assessment of dabigatran. Version 3, March 27, 2015. Available at: https://www.sentinelsystem.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Protocol-for-Assessment-of-Dabigatran_0.pdf. Accessed February 8, 2018

32 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(06):560–566

33 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46(03):399–424

34 Larsen TB, Skjøth F, Nielsen PB, Kjældgaard JN, Lip GY. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. BMJ 2016;353:i3189

35 Yao X, Shah ND, Sangaralingham LR, Gersh BJ, Noseworthy PA. Non-vitamin K antagonist oral anticoagulant dosing in patients with atrial fibrillation and renal dysfunction. J Am Coll Cardiol 2017;69(23):2779–2790

36 Nielsen PB, Skjøth F, Søgaard M, Kjældgaard JN, Lip GY, Larsen TB. Effectiveness and safety of reduced dose non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. BMJ 2017;356:j510

37 Coleman CI, Peacock WF, Bunz TJ, Alberts MJ. Effectiveness and safety of apixaban, dabigatran, and rivaroxaban versus warfarin in patients with nonvalvular atrial fibrillation and previous stroke or transient ischemic attack. Stroke 2017;48(08):2142–2149

38 Peacock WF, Tamayo S, Patel M, Sicignano N, Hopf KP, Yuan Z. CHA2DS2-VASc scores and major bleeding in patients with non-valvular atrial fibrillation who are receiving rivaroxaban. Ann Emerg Med 2017;69(05):541–550.e1

39 Peacock WF, Tamayo S, Sicignano N, Hopf KP, Yuan Z, Patel M. Comparison of the incidence of major bleeding with rivaroxaban use among nonvalvular atrial fibrillation patients with versus without diabetes mellitus. Am J Cardiol 2017;119(05):753–759

40 Holster IL, Valkhoff VE, Kuipers EJ, Tjwa ET. New oral anticoagulants increase risk for gastrointestinal bleeding: a systematic review and meta-analysis. Gastroenterology 2013;145(01):105–112.e15

41 Miller CS, Grandi SM, Shimony A, Filion KB, Eisenberg MJ. Meta-analysis of efficacy and safety of new oral anticoagulants (dabigatran, rivaroxaban, apixaban) versus warfarin in patients with atrial fibrillation. Am J Cardiol 2012;110(03):453–460

42 Abraham NS, Singh S, Alexander GC, et al. Comparative risk of gastrointestinal bleeding with dabigatran, rivaroxaban, and warfarin: population based cohort study. BMJ 2015;350:h1857

43 McCarthy DM. Prevention and treatment of gastrointestinal symptoms and complications due to NSAIDs. Best Pract Res Clin Gastroenterol 2001;15(05):755–773

44 Cangemi DJ, Krill T, Weideman R, Cipher DJ, Spechler SJ, Feagins LA. A comparison of the rate of gastrointestinal bleeding in patients taking non-vitamin K antagonist oral anticoagulants or warfarin. Am J Gastroenterol 2017;112(05):734–739

45 Deitelzweig S, Luo X, Gupta K, et al. Comparison of effectiveness and safety of treatment with apixaban vs. other oral anticoagulants among elderly nonvalvular atrial fibrillation patients. Curr Med Res Opin 2017;33(10):1745–1754

46 Deitelzweig S, Bruno A, Trocio J, et al. An early evaluation of bleeding-related hospital readmissions among hospitalized patients with nonvalvular atrial fibrillation treated with direct oral anticoagulants. Curr Med Res Opin 2016;32(03):573–582

47 Lip GY, Keshishian A, Kamble S, et al. Real-world comparison of major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban, or warfarin. A propensity score matched analysis. Thromb Haemost 2016;116(05):975–986

48 Yao X, Abraham NS, Sangaralingham LR, et al. Effectiveness and safety of dabigatran, rivaroxaban, and apixaban versus warfarin in nonvalvular atrial fibrillation. J Am Heart Assoc 2016;5(06): e003725

49 Conn VS, Ruppar TM, Phillips LJ, Chase JA. Using meta-analyses for comparative effectiveness research. Nurs Outlook 2012;60(04): 182–190

50 Camm AJ, Fox KAA, Peterson E. Challenges in comparing the non-vitamin K antagonist oral anticoagulants for atrial fibrillation-related stroke prevention. Europace 2018;20(01):1–11

51 Almutairi AR, Zhou L, Gellad WF, et al. Effectiveness and safety of non-vitamin K antagonist oral anticoagulants for atrial fibrillation and venous thromboembolism: a systematic review and meta-analyses. Clin Ther 2017;39(07):1456–1478.e36

52 Ntaios G, Papavasileiou V, Makaritsis K, Vemmos K, Michel P, Lip GYH. Real-world setting comparison of nonvitamin-K antagonist oral anticoagulants versus vitamin-K antagonists for stroke prevention in atrial fibrillation: a systematic review and meta-analysis. Stroke 2017;48(09):2494–2503

53 Miller CS, Dorreen A, Martel M, Huynh T, Barkun AN. Risk of gastrointestinal bleeding in patients taking non-vitamin K antagonist oral anticogulants: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2017;15(11): 1674–1683.e3

54 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

55 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(02):157–164

56 DGI, LLC (2017). B.R.I.D.G.E. TO DATA® (Online database). Available at: www.bridgetodata.org. Accessed February 8, 2018