Received: 18 September 2017 | Revised: 1 February 2018 | Accepted: 2 February 2018

DOI: 10.1002/pds.4416

ORIGINAL REPORT

WILEY

Methodological challenges when evaluating potential off-label prescribing of drugs using electronic health care databases: A case study of dabigatran etexilate in Europe

Miguel Cainzos-Achirica[1] | Cristina Varas-Lorenzo[1] | Anton Pottegård[2] | Joelle Asmar[3] | Estel Plana[1] | Lotte Rasmussen[2] | Geoffray Bizouard[3] | Joan Forns[1] | Maja Hellfritzsch[2] | Kristina Zint[4] | Susana Perez-Gutthann[1] | Manel Pladevall-Vila[1]

[1] Epidemiology, RTI Health Solutions, Barcelona, Spain

[2] Department of Public Health, University of Southern Denmark, Odense, Denmark

[3] IMS Health Information Solutions, France

[4] Global Epidemiology, Boehringer Ingelheim GmbH, Ingelheim, Germany

Correspondence
M. Pladevall-Vila and M. Cainzos-Achirica, RTI Health Solutions, Av. Diagonal 605, 9-1, Barcelona 08028, Spain.
Email: mpladevall@rti.org; mcainzos@rti.org

Funding information
Boehringer Ingelheim; Boehringer Ingelheim Gmb

Abstract

Purpose: To report and discuss estimated prevalence of potential off-label use and associated methodological challenges using a case study of dabigatran.

Methods: Observational, cross-sectional study using 3 databases with different types of clinical information available: Cegedim Strategic Data Longitudinal Patient Database (CSD-LPD), France (cardiologist panel, *n* = 1706; general practitioner panel, *n* = 2813; primary care data); National Health Databases, Denmark (*n* = 28 619; hospital episodes and dispensed ambulatory medications); and Clinical Practice Research Datalink (CPRD), UK (linkable to Hospital Episode Statistics [HES], *n* = 2150; not linkable, *n* = 1285; primary care data plus hospital data for HES-linkable patients). Study period: August 2011 to August 2015. Two definitions were used to estimate potential off-label use: a broad definition of on-label prescribing using codes for disease indication (eg, atrial fibrillation [AF]), and a restrictive definition excluding patients with conditions for which dabigatran is not indicated (eg, valvular AF).

Results: Prevalence estimates under the broad definition ranged from 5.7% (CPRD-HES) to 34.0% (CSD-LPD) and, under the restrictive definition, from 17.4% (CPRD-HES) to 44.1% (CSD-LPD). For the majority of potential off-label users, no diagnosis potentially related to anticoagulant use was identified. Key methodological challenges were the limited availability of detailed clinical information, likely leading to overestimation of off-label use, and differences in the information available, which may explain the disparate prevalence estimates across data sources.

Conclusions: Estimates of potential off-label use should be interpreted cautiously due to limitations in available information. In this context, CPRD HES-linkable estimates are likely to be the most accurate.

KEYWORDS

atrial fibrillation, dabigatran, drug utilization, NOACs, off-label, pharmacoepidemiology

1 | INTRODUCTION

Characterization of off-label use of pharmacotherapies—ie, "use of an approved drug for treatments other than those specified in the product label"[1]—improves our understanding of their use in routine clinical practice and provides valuable context to postauthorization safety findings.

Electronic health care databases represent an opportunity to study potential off-label use of drugs in large, representative populations, reducing bias resulting from interactions between researchers, providers, and patients (including the Hawthorne effect[2]). However, methodological challenges of using electronic health care data sources with different types of clinical information available to estimate potential off-label use have not been explored in detail.

STATEMENT ABOUT PRIOR POSTINGS AND PRESENTATIONS:
An abstract with the key results of this study was accepted for presentation at the International Conference of Pharmacoepidemiology 2017 (August 2017).

Dabigatran etexilate (Pradaxa) is a non-vitamin K antagonist oral anticoagulant.[3,4] The first indication approved (Europe, 2008) was for primary prevention of venous thromboembolic events (VTE) after elective total hip or knee replacement surgery.[5] Subsequently approved indications were prevention of stroke and systemic embolisms in patients with non-valvular atrial fibrillation (AF) with 1 or more risk factors (2011)[5] and treatment and secondary prevention of deep vein thrombosis and pulmonary embolism (2014).

Use of dabigatran and other non-vitamin K antagonist oral anticoagulants has grown rapidly in Europe[6] and the United States.[7] Despite increasing use, population-based studies evaluating potential off-label use of dabigatran in routine clinical practice are scarce. Consequently, comprehensive knowledge of potential off-label use, accounting for changes in the product label over time, is lacking. In addition, the methodological challenges resulting from using data sources with different types of clinical data available have not been explored in detail in the context of this research question.

The study presented in this report was a follow-up measure agreed by the European Medicines Agency and the sponsor in the context of approval of the AF indication for dabigatran,[8] with the primary aim of estimating the potential off-label use of dabigatran in 3 European countries. In this article, we discuss the prevalence estimates of potential off-label use and the methodological challenges of this research.

KEY POINTS

- Availability of detailed clinical information is crucial when assessing potential off-label use of drugs using electronic health care databases.
- Specifically for research on the potential off-label use of oral anticoagulants, availability of both primary care and hospital clinical information is of great importance. Otherwise, overestimation of off-label use is likely to occur.
- In a study including new users of dabigatran from the United Kingdom, Denmark, and France, there were marked differences in the prevalence of recorded approved indications of dabigatran across countries and in the prevalence of potential off-label use of dabigatran. These discrepancies are likely due to differences in the information available and completeness of data across databases.
- The Clinical Practice Research Datalink primary care data linked to Hospital Episode Statistics (UK) held the most complete clinical information and likely provided the most accurate estimates of potential off-label use in this study. Overestimation of potential off-label use is likely to have been present in all data sources, particularly in the Cegedim Strategic Data Longitudinal Patient Database (France).

2 | METHODS

2.1 | Study design and data sources

This was an observational, multinational, European cross-sectional study of new users of dabigatran that characterized the prevalence of approved clinical indications and potential off-label use at the time of the first captured prescription. The study used data collected in 3 electronic health care databases, with different types of clinical information available: the Clinical Practice Research Datalink (CPRD), United Kingdom (UK); the Danish National Health Databases (DNHD), Denmark; and the Cegedim Strategic Data Longitudinal Patient Database (CSD-LPD), France (cardiologist panel and general practitioner [GP] panel). CPRD had primary care data for all patients and data on hospital episodes/procedures (but not hospital prescriptions) for a subset of patients (Hospital Episode Statistics [HES]-linkable data). DNHD had information on hospital episodes (inpatient and outpatient) and medications dispensed at community pharmacies. CSD-LPD had only panel-specific ambulatory care information (Table 1). Details of these databases have been reported elsewhere.[9-12]

2.2 | Study population

The study population comprised patients initiating ambulatory treatment with dabigatran during the study period (including treatments started in the hospital setting and continued after discharge), with no ambulatory use during the previous 12 months ("new users"). Participants had at least 1 year of prior enrollment in the database. No other eligibility criteria were applied.

The study period extended from approval of the indication for prevention of stroke and systemic embolism in patients with non-valvular AF in Europe (01 August 2011) until reaching the target number of new users (approximately 5000 patients per database): DNHD, 30 November 2013; CSD-LPD, 30 June 2014; CPRD, 30 August 2015. In Denmark, a much larger population (up to 28 619 new dabigatran users) was identified at the prespecified time point (November 2013), and all of them were included in the analyses.

2.3 | Assessment of dabigatran use

Dabigatran use was assessed by recorded prescriptions in CPRD and CSD-LPD and by dispensed prescriptions in DNHD (Table 1). The index date was the date of the first recorded prescription (index prescription) for dabigatran for each patient meeting the inclusion criteria.

2.4 | Assessment of clinical indications of dabigatran

The approved indications of dabigatran, changes over time, and applicability in each database are summarized in Table 2. Information on clinical indication was obtained from the electronic databases. In CPRD, the indication for a prescription is usually recorded when a new drug is first prescribed by the GP. Proxies for indication were also created using computer algorithms with appropriate diagnostic, procedural, and medication codes. In DNHD, the indication for prescribing is

TABLE 1 Medication use and clinical information available in each of the 3 electronic health care data sources used for this study

	CPRD, UK			CSD-LPD, France	
	HES-Linkable	Not Linkable	DNHD, Denmark	Cardiologist Panel	GP Panel
Medication use					
Ambulatory prescriptions	Yes	Yes	No	Yes	Yes
Ambulatory dispensing	No	No	Yes	No	No
In-hospital medications	No[a]	No[a]	No[a]	No[a]	No[a]
Clinical information					
Outpatient care (primary or specialized)	Yes	Yes	Only hospital-based outpatient clinics	Yes[b]	Yes[b]
Hospital episodes	Yes	As captured by GP	Yes	As captured by cardiologist	As captured by GP

Abbreviations: CPRD, Clinical Practice Research Datalink; CSD-LPD, Cegedim Strategic Data Longitudinal Patient Database; DNHD, Danish National Health Databases; GP, General Practitioner; HES, Hospital Episode Statistics; UK, United Kingdom.

[a]Only captured if treatments started in the hospital setting were continued after discharge.

[b]Only panel-specific information was available (ie, information generated by physicians from the same panel).

not available from prescription data; indications were identified exclusively by proxies. In CSD-LPD, clinical diagnoses associated with prescriptions are recorded using a French proprietary diagnostic code thesaurus implemented as a prepopulated list in the software. The online appendix contains lists of ICD codes used to identify clinical indications for dabigatran (Supplementary Tables S1 and S2).

A patient was assumed to have only 1 clinical indication for dabigatran. For patients with more than 1 potential indication, algorithms identified the indication most likely to have led to the prescription (online appendix).

2.5 | Assessment of potential off-label use of dabigatran

Two levels of on-label use of dabigatran were defined to estimate potential off-label use. The first level was a broad definition of on-label use that required only the presence of codes for the approved clinical indication (eg, AF). The second level, a subset of the first, was a more restrictive definition of on-label use and excluded patients who had recorded conditions for which the medication is not indicated, eg, valvular heart disease or non-valvular AF but no risk factors recorded as specified in the product label (Figure 1). The definitions accounted for changes in the approved indications during the study period (Table 2). Pediatric use was considered off-label.

2.6 | Assessment of conditions possibly leading to anticoagulant use among potential off-label users of dabigatran

Five prespecified conditions, considered a priori to be those most likely to have led to anticoagulant use among off-label users, were evaluated in all databases. Additional conditions possibly leading to anticoagulant therapy were identified in each database from the conditions recorded most frequently in potential off-label users; these conditions were defined using local dictionaries for each database and identified using algorithms. The final lists varied slightly across countries.

2.7 | Statistical analyses

New users were characterized at the index date. Given the high heterogeneity of the information available, analyses were stratified by database. For CPRD, analyses were performed overall and stratified by HES linkage. For France, because there was no possibility to link patients across panels, analyses were stratified by physician panel.

To estimate the proportion of potential off-label use, the 2 definitions of on-label use were implemented sequentially (Figure 1). At each level, the number of new users with each approved indication was calculated, and the total number of on-label users was calculated as their sum. The prevalence of potential off-label use was calculated as [total number of new users minus on-label users] divided by the total number of new users. Exact 95% confidence intervals were calculated.

The computer algorithms to identify clinical indications and classify use as on- or off-label were validated in a random sample of dabigatran users from CPRD through manual review of computerized patient profiles based on medical charts. Measurements of agreement were calculated under both definitions. All analyses were performed using SAS version 9.2 or later (IMSQuintiles, RTI-HS) or Stata version 14 (DNHD).

3 | RESULTS

3.1 | Study population

Figure 2 displays the attrition of dabigatran users by database. Despite a longer study period, the fewest new users were identified in CPRD (n = 3435). The age and sex distribution of the study participants and the distribution of other characteristics at the index date are described by data source in Supplementary Table S3.

3.2 | Prevalence of approved indications of dabigatran

The prevalence of approved indications among new users of dabigatran in each database is presented in Table 3 and online appendix, Figures S1 to S5. The most frequently recorded indication was AF,

TABLE 2 Approved indications of dabigatran etexilate in Europe by time period, and applicability to the study period in each data source

1 August 2011 to 30 November 2013	1 December 2013 to 30 June 2014	1 July 2014 to Present
Primary prevention of venous thromboembolic events in adult patients who have undergone elective total hip replacement surgery or total knee replacement surgery		
Prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation with 1 or more of the following risk factors:	Prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation with 1 or more risk factors such as:	Prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation with 1 or more risk factors such as:
Previous stroke, transient ischemic attack, or systemic embolism	Prior stroke or transient ischemic attack	Prior stroke or transient ischemic attack
Left ventricular ejection fraction <40%		
Symptomatic heart failure, NYHA class II or higher	Heart failure, NYHA class II or higher	Heart failure, NYHA class II or higher
Age ≥ 75 years	Age ≥ 75 years	Age ≥ 75 years
Age ≥ 65 years associated with diabetes mellitus, coronary artery disease, or hypertension	Diabetes mellitus, hypertension	Diabetes mellitus, hypertension
		Treatment of DVT and PE, and prevention of recurrent DVT and PE in adults
Study period by data source:		
DNHD, Denmark →		
CSD-LPD, France →		
CPRD, UK →		

Abbreviations: AF, atrial fibrillation; CPRD, Clinical Practice Research Datalink; CSD-LPD, Cegedim Strategic Data Longitudinal Patient Database; DNHD, Danish National Health Databases; DVT, deep vein thrombosis; NYHA, New York Heart Association; PE, pulmonary embolism; UK, United Kingdom.

Note: For the indication "prevention of stroke and systemic embolism in patients with non-valvular AF," the list of risk factors to be considered was modified in December 2013 by the European Medicines Agency, and this was taken into account in France and the UK from that time onward. The indication "treatment/secondary prevention of DVT and PE" was effective 03 June 2014. Shaded cells denote no applicability of the period-specific label indications within the study period for each data source.

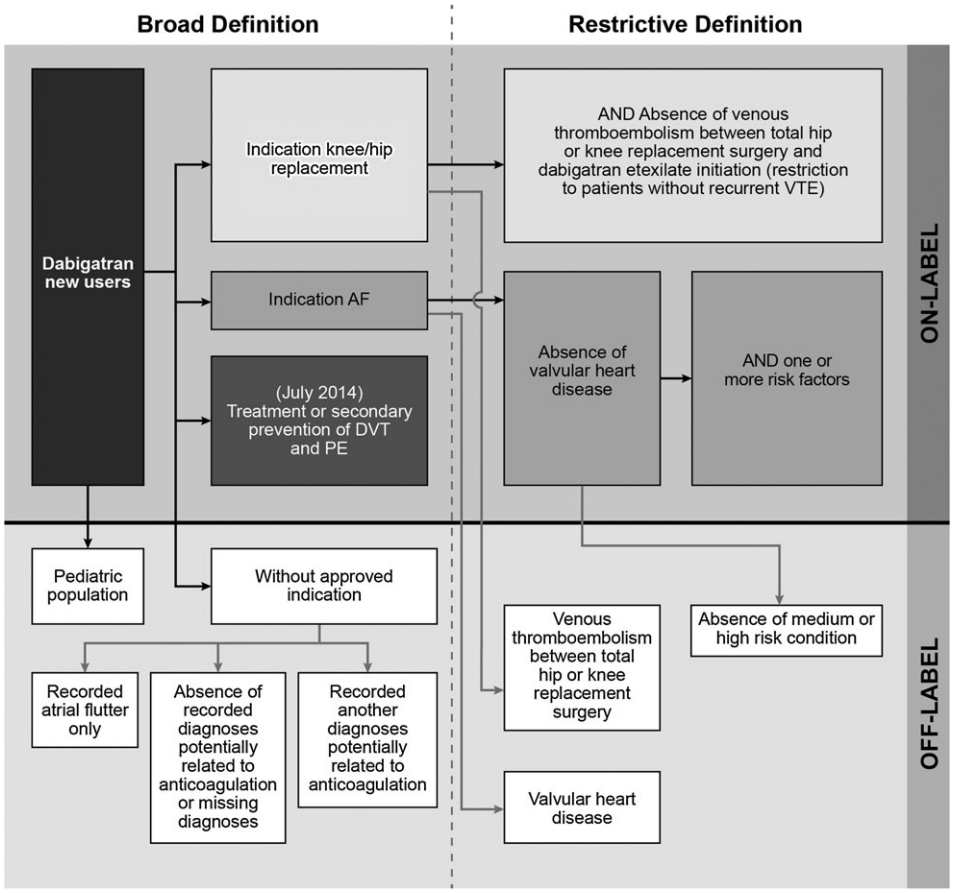

FIGURE 1 Process for identifying recorded diagnoses for labeled indications across the broad and restrictive definitions. AF, atrial fibrillation; DVT, deep vein thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism. Note: The upper half of the figure corresponds to potential on-label uses of dabigatran, and the lower half of the figure to potential off-label uses, under the 2 definitions of on-label use

although marked differences occurred across the databases. Under the broad definition, the highest prevalence of an approved indication was observed in CPRD among HES-linkable patients (88.3%); the lowest in DNHD (59.2%). The other indications were less prevalent. The highest proportion of users with a diagnosis of hip or knee replacement occurred in DNHD (23.7%); the lowest in CSD-LPD (0% cardiologist panel, 1% GP panel). The VTE treatment/secondary prevention indication was evaluated as an on-label indication only in CPRD, starting July 2014 (HES-linkable, 1.0%; nonlinkable, 1.5%).

3.3 | Prevalence of potential off-label use of dabigatran

Table 3 also presents the prevalence of potential off-label use among new users of dabigatran by database. The lowest prevalence occurred in CPRD, lower among HES-linkable, which included the most comprehensive information (broad definition, 5.7%; restrictive definition, 17.4%) than among non-linkable patients (11.5% and 25.6%, respectively). The highest prevalence occurred in CSD-LPD, higher in the GP panel (broad definition, 34.0%; restrictive definition, 44.1%) than in the cardiologist panel (24.1% and 37.5%, respectively). In DNHD, the prevalence of potential off-label use was 17.1% using the broad definition and 29.1% using the restrictive definition.

3.4 | Recorded conditions associated with potential off-label use

Conditions identified as potential clinical reasons for the use of dabigatran among potential off-label users are summarized in Table 4 and online appendix, Tables S4 to S5. The prevalence of each condition varied across countries. In all databases, no conditions potentially related to the use of anticoagulant could be identified in a large proportion of potential off-label users. Specifically, in CPRD, the percentage ranged from 37.7% (CPRD-HES linkable) to 54.7% (CPRD not linkable to HES).

3.5 | Algorithm validation

Patient profiles were reviewed for a random sample of 202 patients from CPRD. Agreement on the on-/off-label classification between automated algorithms and patient profile review by clinicians was very high: broad definition, 98.5%; restrictive definition, 98.1%.

3.6 | Key methodological challenges

The key methodological challenges and limitations identified during this study are summarized in Table 5. Lack of relevant clinical information (primary care data in Denmark, hospital data in France, and CPRD patients not linkable to HES), together with potential under-recording

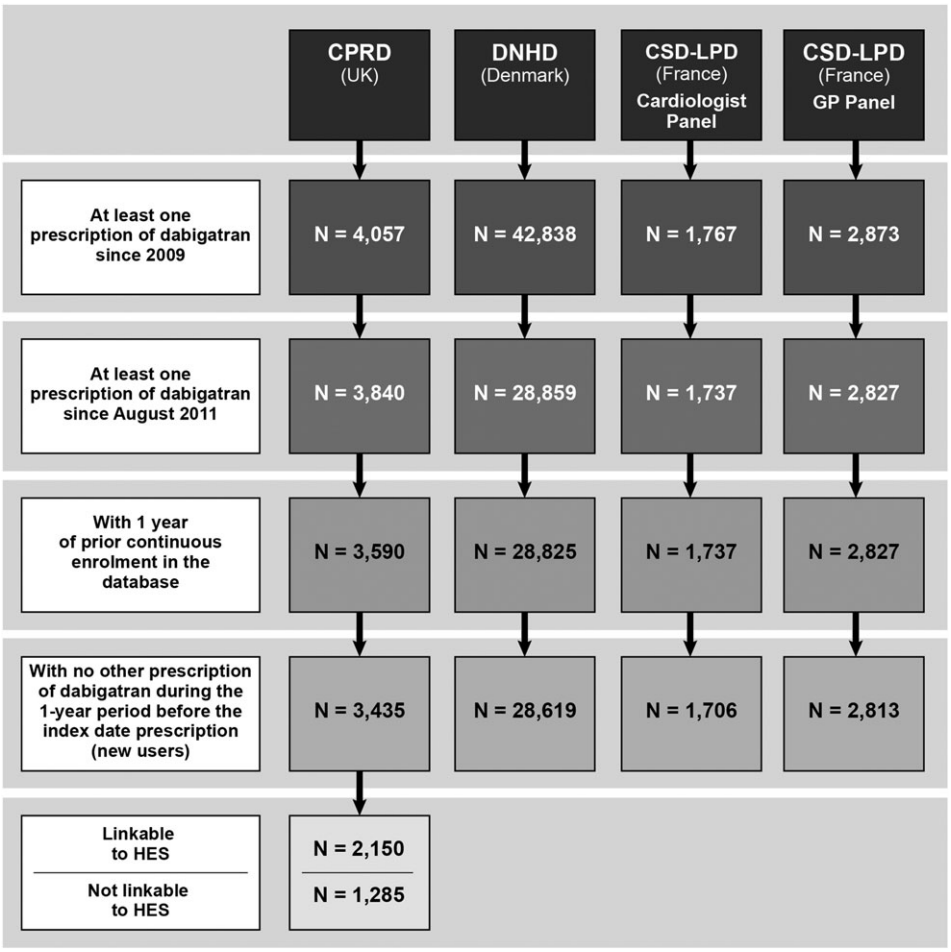

FIGURE 2 Attrition of users of dabigatran etexilate in CPRD (UK), DNHD (Denmark), and CSD-LPD (France, cardiologist and GP panels). CPRD, Clinical Practice Research Datalink; CSD-LPD, Cegedim Strategic Data Longitudinal Patient Database; DNHD, Danish National Health Databases; GP, general practitioner; HES, Hospital Episode Statistics; UK, United Kingdom. Note: Mean (SD) age in CPRD was 73.7 (11.3) years, 43.9% were women. Mean (SD) age in DNHD was 71.8 (10.9) years, 47.4% women. Mean (SD) age in CSD-LPD was 75.5 (10.0) years in the cardiologist panel and 74.0 (10.4) years in the GP panel; proportion of women was 42.1% and 45.2%, respectively. In the cardiologist panel, 282 patients had missing information on sex

of clinical conditions in the electronic databases, were the most important challenges. Both are likely to have led to overestimation of potential off-label use.

4 | DISCUSSION

Among dabigatran new users from UK, Denmark, and France, AF was the most frequently recorded indication. The prevalence of other indications was much lower. Estimates of potential off-label use varied markedly across databases, ranging from 5.7% (CPRD HES-linkable) to 34% (CSD-LPD GP panel) under the broad definition and from 17.4% (CPRD HES-linkable) to 44.1% (CSD-LPD GP panel) under the restrictive definition. The prevalence of conditions associated with potential off-label use also varied across populations. However, the proportion of patients in which no clinical reason for the use of anticoagulant therapy could be established was consistently high in all data sources.

Detailed hospital and primary care clinical information seems crucial for this type of research, and its absence from some data sources may have led to overestimation of potential off-label use. Heterogeneity of the data available across data sources was likely the main driver of the disparate prevalence estimates observed across countries.

4.1 | Strengths

The use of electronic health care databases in general reduces bias related to differential reporting of prescriptions or impacts of contacts with patients and professionals. The use of 3 data sources from different countries allowed evaluation of the research question in data sources with different types of clinical information available and provided insights into the importance, for this type of research, of the availability of detailed information on clinical conditions.

Use of 2 definitions of on-label use also provided complementary information, as they were expected to be impacted to a different extent by some of the study limitations. Finally, the high level of agreement between the computer algorithms and the manual review conducted in CPRD supports the validity of the algorithms used to identify approved clinical indications, at least in the UK.

TABLE 3 Estimated prevalence of approved clinical indications and of potential off-label use of dabigatran etexilate across the 3 study populations, under the broad and the restrictive definitions of on-label use

	UK (CPRD)						Denmark		France (CSD-LPD)			
	Overall (*N* = 3435)		Patients linkable to HES (*N* = 2150)		Patients not linkable to HES (*N* = 1285)		DNHD (*N* = 28 619)		Cardiologist panel (*N* = 1706)		General practitioner panel (*N* = 2813)	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Broad definition												
On-label use												
Atrial fibrillation	2972	86.5	1899	88.3	1073	83.5	16 955	59.2	1295	75.9	1827	64.9
Hip or knee replacement	153	4.5	108	5.0	45	3.5	6783	23.7	0	0	29	1.0
DVT and PE treatment/ secondary prevention[a]	40	1.2	21	1.0	19	1.5	-	-	-	-	-	-
Off-label use (95% CI)	270	7.9 (7.0–8.8)	122	5.7 (4.7–6.7)	148	11.5 (9.8–13.4)	4881	17.1 (16.6–17.5)	411	24.1 (22.1–26.1)	957	34.0 (32.3–35.8)
Restrictive definition												
On-label use												
Atrial fibrillation[b]	2541	74.0	1649	76.7	892	69.4	13 529	47.3	1066	62.5	1544	54.9
Hip or knee replacement	151	4.4	106	4.9	45	3.5	6772	23.7	0	0	29	1.0
DVT and PE treatment/ secondary prevention[a]	40	1.2	21	1.0	19	1.5	-	-	-	-	-	-
Off-label use (95% CI)	703	20.5 (19.1–21.9)	374	17.4 (15.8–19.1)	329	25.6 (23.2–28.1)	8318	29.1 (28.4–29.7)	640	37.5 (35.2–39.8)	1240	44.1 (42.2–45.9)

Abbreviations: CPRD, Clinical Practice Research Datalink; CSD-LPD, Cegedim Strategic Data Longitudinal Patient Database; DNHD, Danish National Health Databases; DVT, deep vein thrombosis; HES, Hospital Episode Statistics; PE, pulmonary embolism; UK, United Kingdom.

[a]Approved indication starting in July 2014.

[b]Non-valvular atrial fibrillation, with at least 1 risk factor listed in Table 1.

TABLE 4 Prevalence of conditions potentially leading to anticoagulant use among potential off-label users of dabigatran at the index date; broad definition of on-label use (CPRD, UK)

	Linkable to HES			Not Linkable to HES		
Condition	***n***	**%**	**95% CI**	***n***	**%**	**95% CI**
General prophylaxis or treatment of a thrombus in any site[a]	25	20.5	13.7–28.7	19	12.8	7.9–19.3
Ischemic stroke, TIA, and occlusion of cerebral arteries[b]	20	16.4	10.3–24.2	24	16.2	10.7–23.2
Treatment/secondary prevention of VTE[ac]	17	13.9	8.3–21.4	6	4.1	1.5–8.6
Heart failure/LVD[b]	11	9.0	4.6 - 15.6	7	4.7	1.9–9.5
Ischemic heart disease[b]	10	8.2	4.0–14.6	< 5	—	—
Anticoagulation for heart valve replacement or stent[a]	8	6.6	2.9–12.5	5	3.4	1.1–7.7
Heart valve disease other than rheumatic, cardiomyopathies, and myocardiopathies[b]	7	5.7	2.3–11.5	6	4.1	1.5–8.6
Peripheral arterial disease[b]	5	4.1	1.3–9.3	0	0.0	0.0–2.5
Only atrial flutter and no other potential off-label diagnoses[b]	<5	—	—	6	4.1	1.5–8.6
Arrhythmias[bd]	<5	—	—	5	3.4	1.1–7.7
General thrombosis prophylaxis in orthopedic surgery[ae]	<5	—	—	0	0.0	0.0–2.5
Hypercoagulability[a]	<5	—	—	< 5	—	—
Conduction disorders[b]	<5	—	—	< 5	—	—
Injuries to the hip and thigh[b]	<5	—	—	0	0.0	0.0–2.5
Pediatric patients (<18 years old)	0	0.0	0.0–3.0	< 5	—	—
Cardiac arrest[b]	0	0.0	0.0–3.0	0	0.0	0.0–2.5
None of the specified conditions (other/unrelated to dabigatran etexilate)	46	37.7	29.1–46.9	81	54.7	46.3–62.9
Total	122	—	—	148	—	—

Abbreviations: CI, confidence interval; CPRD, Clinical Practice Research Datalink; HES, Hospital Episode Statistics; LVD, left ventricular dysfunction; TIA, transient ischemic attack; UK, United Kingdom; VTE, venous thromboembolism.

Note: Due to data protection regulations, counts less than 5, and the corresponding percentages, cannot be provided in cells for CPRD data. The diagnoses are listed by order of decreasing frequency in the group with data linkable to HES.

[a]Prespecified diagnoses considered a priori to be the most likely to lead to off-label use of dabigatran in clinical practice. These conditions were evaluated within a minimum of 1 year before the index date or all available historical data for chronic conditions.

[b]Assessed within 3 months before and after the index prescription.

[c]For CPRD, in July 2014, the indication of treatment or secondary prevention of VTE was added. Therefore, this indication was off-label before July 2014.

[d]Defined as any arrhythmia other than atrial fibrillation or atrial flutter.

[e]Other than hip or knee replacement.

TABLE 5 Key methodological challenges and limitations identified in the study

	CPRD, UK	DNHD, Denmark	CSD-LPD, France
Key challenges/limitations			
Availability of clinical information	For patients not linkable to HES, data on hospital episodes and procedures were available only as captured by GPs	Primary care data were not available when the study was conducted	Data on hospital episodes and procedures were available only as captured by GP or cardiologist Only information generated by a patient's physician was available in the data set
Completeness of clinical information	Underrecording of clinical conditions by GPs is possible, particularly of hospital diagnoses and procedures	Underrecording of clinical conditions is possible, particularly those not relevant to the hospital episode	Physicians record only information relevant to their day-to-day practice Underrecording of clinical conditions is possible, particularly of hospital diagnoses and procedures
Potential consequences			
Estimated prevalence of approved indications of dabigatran	Underestimation, particularly of the hip/knee replacement indication in patients not linkable to HES	Underestimation of the AF indication	Underestimation, particularly of the AF indication in the GP panel, and of the hip/knee replacement indication in both panels
Estimated prevalence of potential off-label use of dabigatran	Overestimation, particularly in patients not linkable to HES	Overestimation. Expected to be large as the prevalence of the most frequent indication of the drug (AF) was likely underestimated	Overestimation in the 2 panels; likely larger than in CPRD and DNHD

Abbreviations: AF, atrial fibrillation; CPRD, Clinical Practice Research Datalink; CSD-LPD, Cegedim Strategic Data Longitudinal Patient Database; DNHD, Danish National Health Databases; GP, general practitioner; HES, Hospital Episode Statistics; UK, United Kingdom.

4.2 | Limitations

A number of limitations must be noted when interpreting the results. When using electronic health care databases, we are dependent on the type of data available and on the completeness of the clinical information recorded in each of the data sources. It is important to note that a major assumption of this type of study is that the absence of a recorded code indicates absence of a condition or risk factor. As a consequence, lack of relevant clinical data and underrecording or misclassification of clinical indications and risk factors in the databases, which may lead to overestimation of potential off-label use, are key limitations of this type of research.

Limitations specific to each database must also be noted. CPRD provided the most complete information, particularly HES-linkable data. Nonetheless, prescriptions for dabigatran issued in the hospital setting or by specialists were not captured in the database unless they were followed by GP prescriptions. Recording by GPs of relevant clinical data, procedures, and tests generated during hospital admissions may have been poor; this limitation is likely to have been more relevant for patients not linkable to HES. Thus, underrecording or misclassification of clinical indications cannot be ruled out in CPRD, particularly in those without HES linkage. Free-text comments from GPs were not available when the study was conducted.

In DNHD, diagnosis and procedure codes were available for hospital and hospital ambulatory care episodes. Although initially primary care data were expected to be available in this data source, this was not the situation when the study was conducted. While DNHD likely captured most hip and knee replacement procedures, the same may have not been true for AF diagnoses. This likely led to underestimation of AF and may explain, at least partially, the lower prevalence of AF in DNHD than in the other databases. This most likely led to overestimation of potential off-label use of dabigatran in Denmark.

In CSD-LPD, by including patients prescribed dabigatran by cardiologists and GPs, we identified first prescriptions issued by 2 key groups of prescribers. Other specialists, such as neurologists, may have issued first prescriptions of dabigatran; these would not have been captured. Information on hospital episodes was not available. Compared to CPRD, primary care information in CSD-LPD was limited; physicians record only conditions that concern their day-to-day medical practice. In CSD-LPD, information generated by each panel is stored separately and cannot be linked to other panels. These features likely explain the high prevalence of potential off-label use observed in France, particularly in the GP panel. Duplicate patients were possible across the 2 physician panels.

Finally, dosage and duration of treatment were not considered in either the broad or restrictive definitions presented in this manuscript due to lack of information in a large number of patients.

4.3 | Interpretation

Estimation of potential off-label use of a drug using electronic health care databases is highly dependent on the databases used, specifically on the availability and completeness of clinical data, which are of utmost importance. Lack of relevant clinical information (eg, approved indications) likely results in the overestimation of the potential off-label use of the drug.

In this study, marked heterogeneity in the information available across databases was probably the main driver of the differences in the prevalence estimates observed across countries. Overestimation of potential off-label use probably occurred in the 3 databases, although the degree was presumably heterogeneous. Because dabigatran may be used to treat conditions typical of both inpatient (eg, VTE prevention) and outpatient (eg, AF) settings, HES-linkable CPRD data, which combine hospital information with detailed primary care data, likely provided the most accurate estimates, while results from the other databases should be interpreted with caution. As most oral anticoagulants are approved for indications in both settings, future studies on the use of these drugs will probably benefit from using databases in which detailed hospital and primary care clinical information is available. For this type of research, it is crucial to take into consideration whether the condition for which the medication is used is managed mainly in the hospital or in the primary care setting, as this may inform database choices.

Other methodological considerations are worth discussing as they may apply to similar studies. Atrial flutter was not considered an on-label indication per information in the product label. However, AF and atrial flutter often coexist, and it is possible that some of those patients might have been misclassified as having atrial flutter *only* (Table 4 and online appendix, Tables S4 and S5). In clinical practice, it is often assumed that antithrombotic therapy recommendations for patients with AF also apply to atrial flutter. This notion was recently supported by the European Society of Cardiology[13] and other consensus documents.[14] Considering atrial flutter as on-label use of dabigatran would have led to slightly lower estimates of potential off-label use, particularly in France.

In CSD-LPD, other arrhythmias and cardioversion were highly prevalent among potential off-label users (online appendix, Table S5). As the main clinical indication for cardioversion is AF,[15] it is possible that an important proportion of these patients (18% of potential off-label users in the cardiologist panel; 27.6%, GP panel) were actually on-label users of dabigatran.

In each database, clinical reasons for using anticoagulant therapy could not be established in a large number of potential off-label users. The lowest proportion was observed in CPRD HES-linkable data, which further supports the notion that this source may have captured clinical indications and risk factors more completely than the others. Evaluation of potential clinical reasons for using the study drug off-label not only improves understanding of the real-world use of the drug in different health care environments, but also informs about the degree of completeness of the databases. In this sense, the proportion of patients without a potential reason for anticoagulation use in CSD-LPD was lower than expected.

Previous studies on the off-label use of dabigatran are scarce. Because these studies used short time periods[16] and/or focused on specific clinical indication groups,[7,17] their findings are not directly comparable. A number of authors have evaluated the methodological limitations inherent to the use of administrative health databases.[18-20] However, studies on the specific challenges of conducting research on off-label use of drugs using electronic databases, or comparing the

strengths and limitations of different databases for this type of research, are scarce. This study can therefore aid the design of future studies on the off-label use of pharmacotherapies, particularly of oral anticoagulants.

5 | CONCLUSIONS

The prevalence of potential off-label prescribing of dabigatran in the study countries ranged from 5.7% in HES-linkable patients in CPRD (UK) to 34% in the French CSD-LPD general practitioner panel (broad definition), and from 17.4% in HES-linkable patients in CPRD (UK) to 44.1% in the French CSD-LPD general practitioner panel (restrictive definition). However, the results regarding potential off-label use need to be interpreted cautiously due to limitations in the available data (no primary care data in Denmark; no hospital data in France). In this context, results from CPRD HES-linkable data are likely to be the most accurate in this study. Availability of detailed clinical information is crucial for studies on off-label use of drugs using electronic health care databases. For research on oral anticoagulants, availability of both detailed primary care and hospital data is of utmost importance. Otherwise, overestimation of off-label use is likely to occur.

These findings may be used to inform the design of future studies on the off-label use of pharmacotherapies, particularly research studies on the use of oral anticoagulant drugs using electronic databases.

ETHICS STATEMENT

The authors state that no ethical approval was needed.

ACKNOWLEDGEMENTS

The authors would like to thank Elena Rivero, MD, MPH, epidemiologist; Nuria Riera, PhD, epidemiologist; Jaume Aguado, MSc, data analyst; Carla Franzoni, BSc, project manager; Adele Monroe, DVM, MSPH, ELS, medical editor; Jason Mathes, BSc, graphic designer (all from RTI Health Solutions); and Morten Olesen, BSc, data scientist (University of Southern Denmark), for their valuable contributions to this study. The authors would also like to thank the general practitioners contributing information to the CPRD in the United Kingdom and Arlene Gallagher and her colleagues from the CPRD research staff. This study was funded by Boehringer Ingelheim GmbH under a contract including independent publication rights for the research team.

CONFLICT OF INTEREST

Miguel Cainzos-Achirica, Estel Plana, Joan Forns, Susana Perez-Gutthann, and Manel Pladevall are fulltime employees of RTI Health Solutions, a unit of RTI International, a non-profit organization that conducts work for government, public, and private organizations, including pharmaceutical companies. Cristina Varas-Lorenzo, now retired, was an employee of RTI Health Solutions when this work was conducted. Anton Pottegård reports participation in research projects funded by Alcon, Almirall, Astellas, Astra-Zeneca, and Servier, all with funds paid to the institution where he was employed (no personal fees) and with no relation to the work reported in this paper. Lotte Rasmussen has nothing to declare. Maja Hellfritzsch has received speaker honoraria from Bristol-Myers Squibb and Pfizer. Joelle Asmar and Geoffray Bizouard declare no conflicts of interest relevant to the present work. Kristina Zint is an employee of Boehringer Ingelheim GmbH.

STATEMENT ABOUT EVALUATION OF PROTOCOLS INVOLVING CPRD DATA:

The study protocol was approved by the CPRD Independent Scientific Advisory Committee (ISAC); protocol number: 15,082. EU PAS Register number: EUPAS7591. clinicaltrials.gov number: NCT02240654.

ORCID

Miguel Cainzos-Achirica http://orcid.org/0000-0002-3805-9089
Anton Pottegård http://orcid.org/0000-0001-9314-5679
Lotte Rasmussen http://orcid.org/0000-0001-5962-6647

REFERENCES

1. Jagger SF. *Prescription drugs: implications of drug labeling and off-label use [testimony before the Subcommittee on Human Resources and Intergovernmental Relations, Committee on Government Reform and Oversight, US House of Representatives]* 1996. http://www.gao.gov/assets/110/106648.pdf (accessed 10 April 2017).
2. Goodwin MA, Stange KC, Zyzanski SJ, Crabtree BF, Borawski EA, Flocke SA. The Hawthorne effect in direct observation research with physicians and patients. *J Eval Clin Pract*. 2017;23(6):1322-1328. https://doi.org/10.1111/jep.12781
3. Eriksson BI, Dahl OE, Buller HR, et al. A new oral direct thrombin inhibitor, dabigatran etexilate, compared with enoxaparin for prevention of thromboembolic events following total hip or knee replacement: the BISTRO II randomized trial. *J Thromb Haemost*. 2005;3(1):103-111. https://doi.org/10.1111/j.1538-7836.2004.01100.x
4. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2009;361(12):1139-1151. https://doi.org/10.1056/NEJMoa0905561
5. European Medicines Agency. Pradaxa: dabigatran etexilate mesilate. Summary of opinion (post authorisation). [EMA/CHMP/304146/2011]. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP), 14 April 2011. http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/000829/WC500105283.pdf (accessed 17 March 2015).
6. Olesen JB, Sørensen R, Hansen ML, et al. Non-vitamin K antagonist oral anticoagulation agents in anticoagulant naive atrial fibrillation patients: Danish nationwide descriptive data 2011-2013. *Europace*. 2015;17(2):187-193. https://doi.org/10.1093/europace/euu225
7. Kirley K, Qato DM, Kornfield R, Stafford RS, Alexander GC. National trends in oral anticoagulant use in the United States, 2007 to 2011. *Circ Cardiovasc Qual Outcomes*. 2012;5(5):615-621. https://doi.org/10.1161/CIRCOUTCOMES.112.967299
8. EMA. Study protocol: evaluation of potential off-label use of dabigatran etexilate in Europe. 2014.
9. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836. https://doi.org/10.1093/ije/dyv098
10. Jouaville SL. Validity and limitations of the longitudinal patient database France for use in pharmacoepidemiological and pharmaco ecocomics studies. ISPOR 20th Annual International Meeting. Philadelphia, PA, USA. May 2015. 2015.
11. Pottegard A, Schmidt SA, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data resource profile: the Danish National Prescription Registry. *Int J Epidemiol*. 2016. https://doi.org/10.1093/ije/dyw213

12. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449-490. https://doi.org/10.2147/CLEP.S91125

13. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J*. 2016;37(38):2893-2962. https://doi.org/10.1093/eurheartj/ehw210

14. Heidenreich PA, Solis P, Estes NA 3rd, et al. 2016 ACC/AHA Clinical performance and quality measures for adults with atrial fibrillation or atrial flutter: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. *J Am Coll Cardiol*. 2016;68(5):525-568. https://doi.org/10.1016/j.jacc.2016.03.521

15. Billionnet C, Alla F, Berigaud E, et al. Identifying atrial fibrillation in outpatients initiating oral anticoagulants based on medico-administrative data: results from the French national healthcare databases. *Pharmacoepidemiol Drug Saf*. 2017;26(5):535-543. https://doi.org/10.1002/pds.4192

16. McCullagh L, Barry M. Prescribing trends for dabigatran etexilate in primary care. *Ir Med J*. 2012;105(5):134-136.

17. Sorensen R, Gislason G, Torp-Pedersen C, et al. Dabigatran use in Danish atrial fibrillation patients in 2011: a nationwide study. *BMJ Open*. 2013;3(5):e002758. https://doi.org/10.1136/bmjopen-2013-002758

18. 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. https://doi.org/10.1016/j.jclinepi.2004.10.012

19. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects—advantages and disadvantages. *Nat Clin Pract Rheumatol*. 2007;3(12):725-732. https://doi.org/10.1038/ncprheum0652

20. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. *J Clin Epidemiol*. 2012;65(2):126-131. https://doi.org/10.1016/j.jclinepi.2011.08.002

SUPPORTING INFORMATION

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

How to cite this article: Cainzos-Achirica M, Varas-Lorenzo C, Pottegård A, et al. Methodological challenges when evaluating potential off-label prescribing of drugs using electronic health care databases: A case study of dabigatran etexilate in Europe. *Pharmacoepidemiol Drug Saf*. 2018;1–11. https://doi.org/10.1002/pds.4416