PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2016; **25**(Suppl. 1): 66–78
Published online 23 June 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3816

ORIGINAL REPORT

Hip/femur fractures associated with the use of benzodiazepines (anxiolytics, hypnotics and related drugs): a methodological approach to assess consistencies across databases from the PROTECT-EU project

Gema Requena[1†], Consuelo Huerta[2*,†], Helga Gardarsdottir[3,4], John Logie[5], Rocío González-González[2], Victoria Abbing-Karahagopian[3], Montserrat Miret[6], Cornelia Schneider[7], Patrick C. Souverein[3], Dave Webb[5], Ana Afonso[3], Nada Boudiaf[5], Elisa Martin[2], Belén Oliva[2], Arturo Alvarez[2], Mark C. H. De Groot[3], Andrew Bate[8], Saga Johansson[9], Raymond Schlienger[10], Robert Reynolds[11], Olaf H. Klungel[4,12] and Francisco J. de Abajo[1,13]

[1] *Pharmacology Section, Department of Biomedical Sciences, School of Medicine and Health Sciences, University of Alcalá, Madrid, Spain*
[2] *BIFAP Research Unit, Division of Pharmacoepidemiology and Pharmacovigilance, Medicines for Human Use Department, Spanish Agency for Medicines and Medical Devices (AEMPS), Madrid, Spain*
[3] *Utrecht Institute for Pharmaceutical Sciences (UIPS), Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, Utrecht, The Netherlands*
[4] *Department of Clinical Pharmacy, Division Laboratory and Pharmacy, University Medical Center Utrecht (UMCU), Utrecht, The Netherlands*
[5] *Worldwide Epidemiology, GlaxoSmithKline Research and Development, Uxbridge, Middlesex, UK*
[6] *Merck KGaA, Geneva, Switzerland*
[7] *Division of Clinical Pharmacy & Epidemiology, University of Basel, Basel, Switzerland*
[8] *Epidemiology, Pfizer Ltd, Tadworth, UK*
[9] *AstraZeneca AB, Mölndal, Sweden*
[10] *Global Clinical Epidemiology, Novartis Pharma AG, Basel, Switzerland*
[11] *Epidemiology, Medical Division, Pfizer, New York, NY, USA*
[12] *Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht (UMCU), Utrecht, The Netherlands*
[13] *Clinical Pharmacology Unit, University Hospital Príncipe de Asturias, Madrid, Spain*

ABSTRACT

Background Results from observational studies may be inconsistent because of variations in methodological and clinical factors that may be intrinsically related to the database (DB) where the study is performed.

Objectives The objectives of this paper were to evaluate the impact of applying a common study protocol to study benzodiazepines (BZDs) (anxiolytics, hypnotics, and related drugs) and the risk of hip/femur fracture (HFF) across three European primary care DBs and to investigate any resulting discrepancies.

Methods To measure the risk of HFF among adult users of BZDs during 2001–2009, three cohort and nested case control (NCC) studies were performed in Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria (BIFAP) (Spain), Clinical Practice Research Datalink (CPRD) (UK), and Mondriaan (The Netherlands). Four different models (A–D) with increasing levels of adjustment were analyzed. The risk according to duration and type of BZD was also explored. Adjusted hazard ratios (cohort), odds ratios (NCC), and their 95% confidence intervals were estimated.

Results Adjusted hazard ratios (Model C) were 1.34 (1.23–1.47) in BIFAP, 1.66 (1.54–1.78) in CPRD, and 2.22 (1.55–3.29) in Mondriaan in cohort studies. Adjusted odds ratios (Model C) were 1.28 (1.16–1.42) in BIFAP, 1.60 (1.49–1.72) in CPRD, and 1.48 (0.89–2.48) in Mondriaan in NCC studies. A short-term effect was suggested in Mondriaan, but not in CPRD or BIFAP. All DBs showed an increased risk with the concomitant use of anxiolytic and hypnotic drugs.

Conclusions Applying similar study methods to different populations and DBs showed an increased risk of HFF in BZDs users but differed in the magnitude of the risk, which may be because of inherent differences between DBs. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS—benzodiazepines; hip fractures; cohort; nested case control; databases; pharmacoepidemiology

Received 25 September 2014; Revised 13 May 2015; Accepted 20 May 2015

INTRODUCTION

The increasing availability of large electronic healthcare databases (DBs) has led to a rapid growth in the number of epidemiological studies addressing

*Correspondence to: C. Huerta, División de Farmacoepidemiología y Farmacovigilancia, Subdirección General de Medicamentos de Uso Humano, Agencia Española de Medicamentos y Productos Santiarios-AEMPS, Calle Campezo 1 Edificio 8 28022, Madrid, Spain. E-mail: chuerta_fcsai@bifap.aemps.es
[†]Authors contributed equally.

adverse effects of medicinal products.[1] The potential for inconsistencies in findings for the same exposure-outcome association has challenged the validity and value of observational evidence.[2–5]

The association of use of benzodiazepines (BZDs) with hip fractures has been widely studied, and inconsistencies in results are common in the literature. Risk estimates from cohort designs ranged from 0.77 to 1.60,[6–9] and for case control designs from 0.59 to 2.50.[10] In two recent meta-analysis, the pooled relative risk found with different designs showed 35–40% increase of risk fractures among BZDs users.[10,11]

Although meta-analysis is a well-established method of reviewing evidence, it should incorporate a careful investigation of potential sources of heterogeneity.[12,13] The reasons for discrepancies can be due to methodological features, such as the study design, bias control, outcome and exposure definitions, or population selection criteria. But also they may be related to clinical aspects, such as population characteristics or health systems factors that may be intrinsically related to the DBs where the study is performed. There is a need, therefore, to gain an understanding of how the characteristics of a DB might be a source of discrepancy in the results when using several DBs to analyze the same drug-event association. This has recently been explored in the Observational Medical Outcomes Partnership Project,[14] and it is the main purpose of the present research performed within the framework of the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium (PROTECT) project (http://www.imi-protect.eu/). We placed particular emphasis on methodological questions rather than on the clinical aspects of the association under investigation. The primary objective for the present paper was to compare the results across DBs and secondly across designs.

PATIENTS AND METHODS

A common protocol was developed and registered in The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, ENCePP,[15] for both a cohort and a nested case control (NCC), to be performed in three primary care DBs: "Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria" (BIFAP) from Spain; the Clinical Practice Research Datalink (CPRD, formerly GPRD) from the UK; and the Dutch Mondriaan project, which contains health care data from various sources; in this study, two DBs were used: the Netherlands Primary Care Research Database and the Almere Health Care Group Database (AHC). Information contained in all DBs originates from general practitioners (GPs) and is described in detail elsewhere.[16]

Cohort definition and selection of cases and controls

For each DB, the source population included patients aged 18 years or older, with at least 1 year of registration with a GP, and who received at least one BZD prescription within the study period from 1 January 2001 to 31 December 2009. The start date was the date of the first BZD prescribed to patients without a hip/femur fracture (HFF) in the previous 12 months (new cases) and no BZD prescription in the 6 months prior to start date (new users). This washout period was selected to ensure that patients returned to a naïve state following that period. Patients were then followed from start date until the earliest of one of the following dates: a first HFF, death, the patient left the practice, the practice left the DB, or 31 December 2009. The HFF was identified using the International Classification of Primary Care (ICPC-2) codes and free text in BIFAP, READ codes in CPRD, and ICPC-2 codes in Mondriaan (Supporting Information Table S1).

An NCC study was nested in the cohort using all cases match up to four controls randomly selected from the pool of eligible person-time (risk-set sampling). Cases were matched to controls by sex, age (±2 years), and follow-up time within the cohort (±6 months), defined as the time from study entry to the index date. The index date of each control was the fracture date of the matched case. A sensitivity analysis with additional matching by GP practice was also performed in CPRD.

Exposure definition

Among BZDs, we included those classified as anxiolytics, hypnotics, and related drugs in the Anatomical Therapeutic Chemical classification[17] (Supporting Information Table S2). Related drugs (Z-drugs) were introduced to the market as an alternative therapy to BZDs in the treatment of insomnia; however, they are not exempted of HFF risk[18] and were therefore included in this study. For all patients, treatment episodes of BZDs were constructed by summing consecutive prescriptions.[19] The expected duration of each prescription was derived based on the number of tablets and the prescribed dosage regimen; if the time span between the theoretical end date of a prescription supply and the prescribing date of the subsequent prescription exceeded 30 days, a new treatment episode was considered. Exposure was divided into the following periods: *current* use, from a first prescription until

DOI: 10.1002/pds

30 days after the estimated end date of the treatment episode; *recent* use, up to 60 days after current use; and *past* use, after recent use until the patient became a current user again, or the end of follow-up. For both designs, periods of current use were further stratified according to the following: (i) *duration* (of each treatment episode, not cumulatively over follow-up) in 1–30 (short term use), 31–60, 61–182,183–365, >365 days, and (ii) *type of BZD*, as single use of anxiolytics (N05BA), single use of hypnotics (N05CD, N05CF, and N05CM02), concomitant use or use of both within the same treatment episode but not concomitantly.

In the BIFAP, CPRD, and Mondriaan-Netherlands Primary Care Research DBs, the prescription of the drug of interest was the indicator of exposure, while in the Mondriaan-AHC, DB exposure included all BZDs prescribed by the GP and BZDs prescribed by specialists, which were dispensed from the AHC pharmacy (<5% of all total exposure).

Potential confounders

Potential confounders included age, gender, comorbidities, and medication[20–25] (Table 1). In the cohort analysis, baseline comorbidities were measured any

Table 1. Baseline demographics and characteristics of patients initiating prescription of anxiolytic or hypnotic medication

	BIFAP	CPRD	MONDRIAAN
	N = 558 599	*N* = 663 894	*N* = 50 464
Duration of follow-up (after start date) years	1 695 045	2 430 138	146 455
	n (%)	*n* (%)	*n* (%)
Age, year Mean (SD)	55.14 (18.7)	51.1 (18.4)	48.7 (16.6)
18–29	75 264 (13.5)	85 154 (12.8)	6138 (12.2)
30–39	105 270 (18.8)	117 464 (17.7)	9983 (19.8)
40–49	107 332 (19.2)	130 498 (19.7)	12 329 (24.4)
50–59	98 620 (17.7)	114 670 (17.3)	9672 (19.2)
60–69	71 976 (12.9)	90 272 (13.6)	5649 (11.2)
70–79	62 765 (11.2)	71 939 (10.8)	4172 (8.3)
80+	37 372 (6.7)	53 897 (8.1)	2521 (5.0)
Sex			
Male	192 519 (34.5)	261 661 (39.4)	21 358 (42.3)
Female	366 080 (65.5)	402 233 (60.6)	29 106 (57.7)
Smoking			
Yes	84 224 (15.1)	187 547 (28.3)	—
No	173 543 (31.1)	276 239 (41.6)	—
Past	8240 (1.5)	150 232 (22.6)	—
Unknown	292 592 (52.4)	49 876 (7.5)	—
Alcohol use			
Yes		431 614 (65.0)	—
No		133 628 (20.1)	—
Unknown		98 652 (14.9)	—
BMI (kg/m^2)			
<18.5	3838 (0.7)	14 816 (2.2)	—
18.5–24.9	69 796 (12.5)	228 910 (34.5)	—
25–30	92 072 (16.5)	180 089 (27.1)	—
>30	74 693 (13.4)	115 113 (17.3)	—
Unknown	318 200 (57.0)	124 966 (19.8)	—
Co-morbidity			
Previous fractures	24 635 (4.4)	129 519 (19.5)	2212 (4.4)
Rheumatoid arthritis	2886 (0.5)	10 363 (1.6)	0 (0)
Osteoporosis	22 494 (4.0)	16 710 (2.5)	596 (1.2)
Paget's disease	432 (0.1)	690 (0.1)	
Anemia	34 196 (6.1)	48 698 (7.3)	1676 (3.3)
Epilepsies/ seizures	5013 (0.9)	20 553 (3.1)	493 (1.0)
Syncope	28 322 (5.1)	26 521 (4.0)	1188 (2.4)
Ischemic heart disease	26 919 (4.8)	51 550 (7.8)	1714 (3.4)
Cerebrovascular disease	17 294 (3.1)	27 637 (4.1)	800 (1.6)
Malignant neoplasms	34 367 (6.2)	66 896 (10.1)	3304 (6.6)
Inflammatory bowel disease	2396 (0.4)	7976 (1.2)	294 (0.6)
COPD	15 383 (2.8)	54 813 (8.3)	1863 (3.7)
Liver disease	8699 (1.6)	6246 (0.9)	155 (0.3)
Chronic renal failure	4688 (0.8)	3709 (0.6)	0 (0.0)
Mental disorders*	8005 (1.4)	21 985 (3.3)	1057 (2.1)
Dementia and/or Alzheimer	5732 (1.0)	8640 (1.30)	222 (0.4)
Co-medication			
Oral glucocorticoids†	2671 (0.5)	9645 (1.5)	1963 (3.9)
Bisphosphonate	9855 (1.8)	13 981 (2.1)	557 (1.1)
Raloxifene	3907 (0.7)	601 (0.1)	8 (0.0)
Strontium ranelate	478 (0.1)	231 (0.0)	4 (0.0)
Parathyroid hormone	65 (0.0)	0 (0.0)	0 (0.0)
Vitamin D (calcium) and analogs	17 153 (3.1)	6982 (1.1)	461 (0.9)
Calcitonin	1450 (0.3)	34 (<0.01)	0 (0.0)
Antidepressants	39 868 (7.1)	127 818 (19.3)	2211 (4.4)
Antipsychotics/ lithium	19 050 (3.4)	39 880 (6.0)	746 (1.5)
Anti-Parkinson drugs	2991 (0.5)	6252 (0.9)	207 (0.4)
Anticonvulsants	12 498 (2.2)	21 494 (3.2)	881 (1.7)
Inhaled glucocorticoids	9175 (1.6)	50 974 (7.7)	1700 (3.4)
Bronchodilators†	34 043 (6.1)	71 241 (10.7)	3760 (7.5)
Antihypertensive drug‡	93 745 (16.8)	147 809 (22.3)	8596 (17.0)
Diuretics	43 976 (7.9)	93 868 (14.1)	4145 (8.2)
Anti-arrhythmic	3209 (0.6)	9645 (1.5)	195 (0.4)
Sedating antihistamines	3891 (0.7)	2318 (0.4)	0 (0.0)
HRT	6789 (1.2)	30 974 (4.7)	1205 (2.4)
Thyroid hormones	14 966 (2.7)	33 264 (5.0)	1185 (2.4)
Antithyroid drugs	946 (0.2)	1316 (0.2)	0 (0.0)
DMARDs	2730 (0.5)	6411 (1.0)	487 (1.0)
Thiazolidinediones	845 (0.2)	2729 (0.4)	83 (0.2)

(*Continues*)

DOI: 10.1002/pds

Table 1. (Continued)

	BIFAP	CPRD	MONDRIAAN
	N = 558 599	*N* = 663 894	*N* = 50 464
Other antidiabetics	26 009 (4.7)	27 249 (4.1)	2301 (4.6)
Antiemetic	11 307 (2.0)	11 231 (1.7)	1397 (2.8)
Anticoagulants	14 119 (2.5)	14 107 (2.1)	4857 (9.6)
Morphine/opiate	23 164 (4.2)	156 874 (23.6)	2552 (5.1)
NSAIDs[§]	62 915 (11.3)	57 190 (8.6)	3845 (7.6)
Statins	45 161 (8.1)	75 392 (11.4)	3902 (7.7)
Proton pump inhibitors	97 741 (17.5)	92 309 (13.9)	5661 (11.2)
Aromatase inhibitors	806 (0.1)	1585 (0.2)	56 (0.1)

SD, standard deviation; BMI, body mass index; COPD, chronic obstructive pulmonary diseases; NSAIDs, non-steroidal anti-inflammatory drugs; HRT, hormonal replacement therapy; DMARDs, disease-modifying anti-rheumatic drugs; BIFAP, Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria; CPRD, Clinical Practice Research Datalink.

*Depression was not included.

[†]Including beta-2-adrenoceptors agonist and anticholinergics.

[‡]Including angiotensin-converting enzyme inhibitors, angiotensin II antagonists, diuretics, calcium antagonists, beta blocking agents, calcium channel blockers, and other antihypertensives.

[§]Two or more prescriptions of NSAIDs.

time before start date and medications within a 6-month period before start date. Then, comorbidities and medications were updated whenever exposure status changed or every 182 days when the exposure was stable. In the NCC, comorbidities were measured any time prior to the index date, and medications were assessed within the 182 days before index date.

Statistical analysis

Confounders were added sequentially to the previous model as follows: age (referred to as age-adjusted model); sex (model A) (in the NCC, age and sex were matching variables); well-established risk factors (Model B); risks factors related to HFF (Model C); and all remaining comorbidities and co-medications (Model D). Information on alcohol use, smoking, and body mass index was not available in Mondriaan, and not systematically available in BIFAP, therefore, all analyses were performed without these variables. To evaluate the impact of these three factors, a sensitivity analysis was performed in CPRD. The low number of cases in Mondriaan precluded the use of fully adjusted model D. Therefore, we used the model C for main comparative analyses across three DBs.

The associations were expressed as follows: (i) hazard ratios (HRs) with 95% confidence intervals (CI) using time-dependent Cox proportional hazard models for the cohort design and (ii) odds ratios (ORs) with 95% CI using conditional logistic regression for the NCC. Past use was considered as the reference group in both designs. Statistical analyses were conducted locally by each DB owner using SAS® in CPRD and Mondriaan and Stata®-11 in BIFAP. A blinding procedure was maintained until all results were available at the coordinating center (at the Utrecht University, the Netherlands).

A pooled HR and OR and 95% CIs from PROTECT studies were calculated assuming a random-effects model.[26] To assess consistency among the results, a comparison was performed between our results and published studies from two recent meta-analysis.[10,11]

RESULTS

Cohort studies

Cohort analyses included 558 599 from BIFAP, 663 894 from CPRD, and 50 464 patients from Mondriaan. Patients in BIFAP were older (mean age: 55) than in CPRD (mean: 51) and Mondriaan (mean: 49). Baseline co-morbidity and co-medication use varied across DBs, with the major differences found in the prevalence of previous fractures, and the use of antidepressants, vitamin D, inhaled corticoids, hormone replacement therapy (HRT), oral anticoagulants, and opioids (Table 1).

In BIFAP, cases were identified through free text (in addition to codes). For that reason, a review of all cases was carried out for validation. Out of 3992 potential cases detected, 34% were excluded (of them, 13% due to high-energy trauma, 60% due to other fractures (i.e., pelvis), and the remaining patients did not have a clear date of the event). Such a revision was not feasible in the other DBs.

We identified 2459 cases of HFF in BIFAP, 4469 in CPRD, and 151 in Mondriaan. Crude HRs (95% CI) of HFF for current use of BZDs as compared with past use were similar for all DBs: 2.83 (2.60–3.09), 3.32 (3.10–3.56), and 3.32 (2.31–4.75) for BIFAP, CPRD, and Mondriaan, respectively (Table 2). Age adjustment resulted in around 50% decrease in risk estimates in BIFAP (1.39; 1.28–1.52), and CPRD (1.69; 1.57–1.81), and 34% decrease in Mondriaan (2.18; 1.52–3.13) (Table 2). Models A, B, and C did not yield substantial additional changes in risk estimates (Table 2).

In the fully adjusted model (Model D), we observed an additional decrease in HRs in both BIFAP (−14%) (1.19; 1.08–1.30) and CPRD (−10%) (1.51; 1.41–1.63) (Table 2). Adding lifestyle and body mass index variables to the full model did not materially change the risk estimates in CPRD (HR = 1.46; 1.36–1.57). Median duration of current BZD use periods was 60 days in BIFAP and Mondriaan and 57 days in

DOI: 10.1002/pds

Table 2. Cohort study

	BIFAP							
	Cases (2459)	Person-years	Crude HR (95%CI)	Age-adjusted HR (95%CI)	Model A* HR (95%CI)	Model B† HR (95%CI)	Model C‡ HR (95%CI)	Model D§ HR (95%CI)
Past (reference category)	851	980 993	1	1	1	1	1	1
Recent	268	164 513	1.89 (1.64–2.17)	1.38 (1.20–1.58)	1.36 (1.19–1.56)	1.35 (1.18–1.55)	1.35 (1.18–1.55)	1.27 (1.11–1.46)
Current	1340	549 539	2.83 (2.60–3.09)	1.39 (1.28–1.52)	1.37 (1.25–1.49)	1.34 (1.23–1.47)	1.34 (1.23–1.47)	1.19 (1.08–1.30)
Among current users only								
Duration (days)								
1–30	184	117 683	1.80 (1.53–2.11)	1.14 (0.97–1.34)	1.12 (0.96–1.32)	1.08 (0.92–1.27)	1.08 (0.92–1.27)	0.98 (0.84–1.16)
31–60	179	100 433	2.06 (1.75–2.43)	1.28 (1.08–1.50)	1.26 (1.07–1.48)	1.21 (1.03–1.42)	1.21 (1.03–1.42)	1.10 (0.94–1.30)
61–180	267	117 810	2.59 (2.26–2.98)	1.25 (1.09–1.44)	1.24 (1.07–1.42)	1.19 (1.03–1.36)	1.19 (1.03–1.36)	1.05 (0.91–1.21)
181–365	212	67 097	3.64 (3.13–4.23)	1.54 (1.32–1.79)	1.52 (1.30–1.76)	1.49 (1.28–1.74)	1.49 (1.28–1.74)	1.29 (1.10–1.50)
>365	498	146 516	3.93 (3.52–4.39)	1.59 (1.42–1.78)	1.55 (1.39–1.73)	1.55 (1.38–1.73)	1.55 (1.38–1.73)	1.34 (1.19–1.50)
ATC subgroup								
Anxiolytics and hypnotics¶	9	3825	2.73 (1.42–5.27)	1.32 (0.68–2.54)	1.32 (0.68–2.55)	1.28 (0.67–2.48)	1.28 (0.67–2.48)	1.06 (0.55–2.05)
Concomitant anxiolytics and hypnotics	150	52 225	3.33 (2.80–3.96)	1.65 (1.38–1.96)	1.61 (1.35–1.92)	1.58 (1.33–1.89)	1.58 (1.33–1.89)	1.24 (1.04–1.48)
Single use of anxiolytics	894	418 896	2.48 (2.25–2.73)	1.37 (1.24–1.50)	1.34 (1.22–1.47)	1.31 (1.19–1.44)	1.31 (1.19–1.44)	1.18 (1.07–1.30)
Single use of hypnotics	287	74 593	4.46 (3.90–5.11)	1.36 (1.19–1.56)	1.36 (1.18–1.55)	1.34 (1.17–1.53)	1.34 (1.17–1.53)	1.17 (1.02–1.35)
	CPRD							
	Cases, (4469)	Person-years	Crude HR (95% CI)	Age-adjusted HR (95% CI)	Model A* HR (95% CI)	Model B† HR (95% CI)	Model C‡ HR (95% CI)	Model D§ HR (95% CI)
Past (reference category)	2627	1,03 431	1	1	1	1	1	1
Recent	434	191 277	1.75 (1.56–1.97)	1.35 (1.20–1.50)	1.34 (1.20–1.50)	1.33 (1.18–1.49)	1.31 (1.17–1.47)	1.13 (1.01–1.27)
Current	1408	335 430	3.32 (3.10–3.56)	1.69 (1.57–1.81)	1.68 (1.57–1.81)	1.66 (1.54–1.78)	1.66 (1.55–1.78)	1.51 (1.41–1.63)
Among current users only								
Duration (days)								
1–30	420	130 463	2.31 (2.03–2.64)	1.40 (1.23–1.60)	1.39 (1.22–1.59)	1.38 (1.21–1.57)	1.37 (1.20–1.56)	1.27 (1.11–1.45)
31–60	334	92 842	2.90 (2.55–3.30)	1.61 (1.41–1.84)	1.60 (1.41–1.83)	1.59 (1.39–1.81)	1.58 (1.39–1.81)	1.47 (1.29–1.68)
61–180	403	71 311	4.15 (3.73–4.63)	1.79 (1.61–2.00)	1.79 (1.60–1.99)	1.76 (1.58–1.96)	1.77 (1.59–1.97)	1.62 (1.45–1.81)
181–365	156	23 518	4.8 (4.09–5.65)	1.96 (1.67–2.31)	1.95 (1.66–2.29)	1.90 (1.61–2.24)	1.91 (1.62–2.25)	1.68 (1.42–1.97)
>365	95	17 293	3.98 (3.24–4.89)	2.14 (1.74–2.62)	2.11 (1.72–2.60)	2.09 (1.70–2.56)	2.08 (1.70–2.56)	1.81 (1.47–2.22)
ATC subgroup								
Anxiolytics and hypnotics¶	44	9890	3.40 (2.58–4.48)	2.04 (1.54–2.68)	2.19 (1.62–2.95)	2.13 (1.58–2.87)	2.14 (1.59–2.89)	1.68 (1.27–2.21)
Concomitant anxiolytics and hypnotics	52	11 752	3.29 (2.44–4.43)	2.22 (1.64–2.99)	2.02 (1.53–2.67)	1.99 (1.51–2.62)	1.98 (1.50–2.61)	1.71 (1.27–2.31)
Single use of anxiolytics	436	129 857	2.70 (2.42–3.00)	1.97 (1.77–2.20)	1.93 (1.73–2.15)	1.92 (1.73–2.14)	1.91 (1.72–2.13)	1.71 (1.53–1.91)
Single use of hypnotics	876	183 929	3.69 (3.41–4.00)	1.56 (1.44–1.69)	1.56 (1.43–1.69)	1.53 (1.41–1.66)	1.54 (1.42–1.67)	1.43 (1.32–1.55)

(*Continues*)

DOI: 10.1002/pds

Table 2. (Continued)

	MONDRIAAN[‖]						
	Cases (151)	Person-years	Crude HR (95%CI)	Age-adjusted HR (95%CI)	Model A* HR (95%CI)	Model B[†] HR (95%CI)	Model C[‡] HR (95%CI)
Past (reference category)	80	112 608	1	1	1	1	1
Recent	21	12 360	2.24 (1.49–3.95)	2.10 (1.29–3.40)	2.09 (1.29–3.40)	2.15 (1.32–3.49)	2.14 (1.32–3.48)
Current	50	21 487	3.32 (2.31–4.75)	2.18 (1.52–3.13)	2.18 (1.52–3.13)	2.22 (1.55–3.19)	2.22 (1.55–3.19)
Among current users only							
Duration (days)							
1–30	22	7610	4.14 (2.56–6.71)	3.16 (1.95–5.11)	3.15 (1.94–5.10)	3.27 (2.02–5.30)	3.27 (2.02–5.29)
31–60	12	7398	2.32 (1.26–4.27)	1.78 (0.97–3.29)	1.78 (0.97–3.28)	1.83 (0.99–3.38)	1.83 (0.99–3.37)
61–180	9	3567	3.57 (1.79–7.12)	2.06 (1.03–4.13)	2.06 (1.03–4.13)	2.12 (1.06–4.25)	2.13 (1.06–4.26)
181–365	3	1256	3.42 (1.08–10.84)	1.49 (0.47–4.73)	1.48 (0.47–4.71)	1.51 (0.47–4.79)	1.50 (0.47–4.78)
>365	4	1656	3.42 (1.25–9.33)	1.39 (0.51–3.80)	1.39 (0.51–3.80)	1.31 (0.48–3.59)	1.32 (0.48–3.63)
ATC subgroup							
Anxiolytics and hypnotics[¶]	6	1429	5.94 (2.59–13.63)	3.85 (1.67–8.85)	3.85 (1.68–8.86)	3.97 (1.73–9.14)	4.00 (1.74–9.19)
Concomitant anxiolytics and hypnotics	2	534	5.33 (1.31–21.72)	3.85 (0.95–15.70)	3.86 (0.95–15.73)	3.93 (0.96–16.00)	4.00 (0.98–16.30)
Single use of anxiolytics	20	11 502	2.48 (1.51–4.07)	2.12 (1.29–3.49)	2.12 (1.29–3.48)	2.20 (1.34–3.61)	2.20 (1.33–3.61)
Single use of hypnotics	22	8022	3.89 (2.42–6.26)	1.92 (1.19–3.10)	1.91 (1.18–3.10)	1.92 (1.18–3.10)	1.92 (1.18–3.10)

Risk of HFF in association with benzodiazepines according to categories of exposure, duration, and class of drug for the different models of adjustment of confounders.
HR, hazard ratio; CI, confidence interval; BIFAP, Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria; CPRD, Clinical Practice Research Datalink.
*Model A: age and sex.
[†]Model B: Model A plus previous fracture, systemic glucocorticoids and rheumatoid arthritis.
[‡]Model C: Model A plus Model B plus history of osteoporosis, or Paget's disease, and use of bisphosphonate, raloxifene, strontium ranelate, calcitonin, parathyroid hormone, calcium, and vitamin D.
[§]Model D: Model A plus Model B plus Model C plus the following co-medications: antidepressants, antipsychotics/lithium, anti-Parkinson drugs, anticonvulsants, inhaled glucocorticoids, bronchodilators (including beta-2-adrenoceptors agonist and anticholinergics), anti-arrhythmic, sedating antihistamines, antihypertensive drugs (including ACE inhibitors, angiotensin II antagonists, beta blocking agents, calcium channel blockers, and other antihypertensives), diuretics, estrogen-containing hormone replacement therapy (HRT), thyroid hormones, antithyroid drugs, disease-modifying anti-rheumatic drugs, thiazolidinediones, other antidiabetics, antiemetic (metoclopramide), anticoagulants, morphine/opiate, two or more prescriptions for a non-steroidal anti-inflammatory drug (NSAIDs), statins, proton pump inhibitors, and aromatase inhibitors; plus the following co-morbidities: anemia, seizures, syncope, ischemic heart disease, cerebrovascular disease, malignant neoplasms, inflammatory bowel disease, obstructive airway disease, liver disease, impaired renal function, mental disorders, and any form of dementia (including Alzheimer's disease).
[¶]Use of anxiolytics and hypnotics within the current period, but not in concomitant use.
[‖]Because of the number of outcomes in Mondriaan, Model D is not presented in this table.

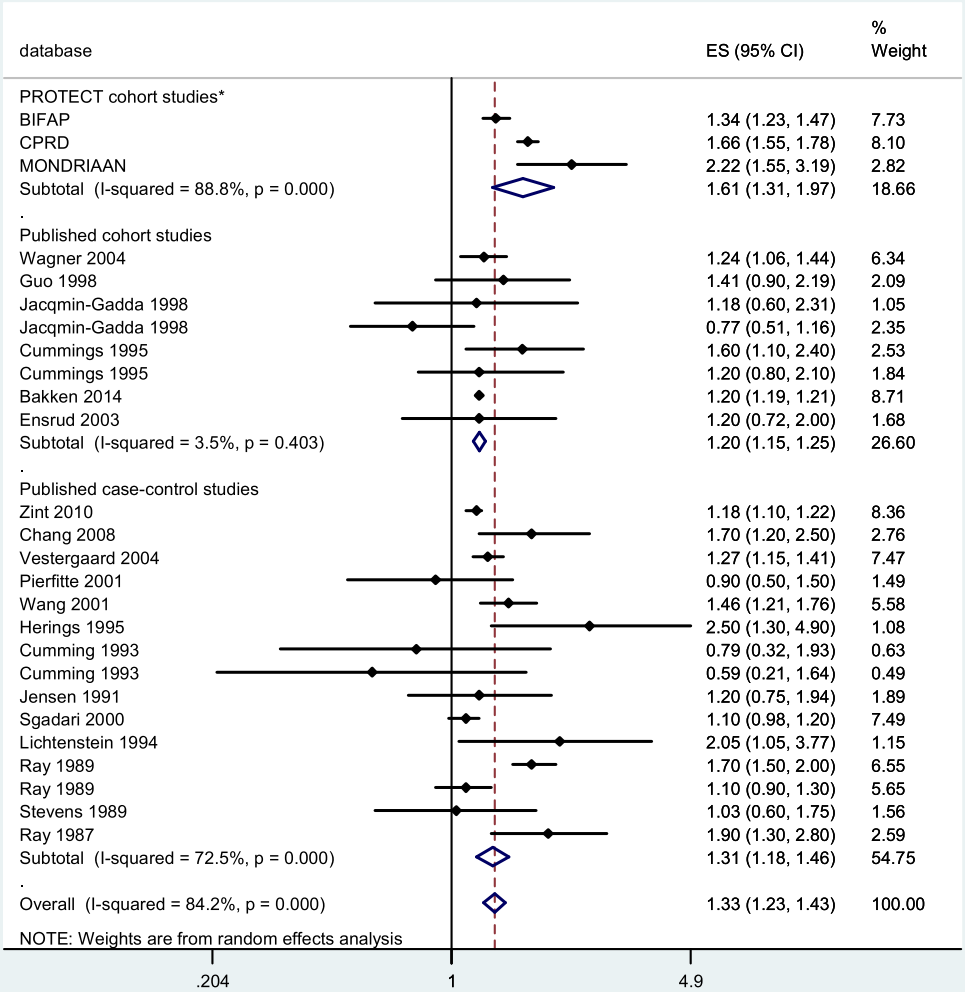
CPRD. Median duration of individual BZD prescriptions was 30 in BIFAP and Mondriaan and 28 days in CPRD.

In both CPRD and BIFAP, the HR appeared to increase with duration of use, whereas in Mondriaan, the highest risk was found at the beginning of the current period (Table 2). In CPRD and Mondriaan, point estimates of HR with single use of anxiolytics versus hypnotics were slightly greater (Model C) (CPRD: 1.91; 1.72–2.13 vs. 1.54; 1.42–1.67 and Mondriaan: 2.20; 1.33–3.61 vs. 1.92; 1.18–3.10) respectively, though confidence intervals overlapped. No differences were found in BIFAP (Table 2). In all DBs, the highest risk was found when both anxiolytics and hypnotics were used concomitantly (Table 2).

The pooled HR of HFF for the three cohorts using Model C was 1.61 (1.31–1.97) (I^2 = 88.8%). Pooled risk of eight published cohort studies[10,11] was 1.20 (1.15–1.25) (I^2 = 3.5%) (Figure 1).

Nested case-control studies

Overall, 2459 cases from BIFAP, 5966 from CPRD, and 148 cases from Mondriaan were matched to at least one control and included in the analyses. The number of cases in CPRD for the NCC differed from



* In BIFAP, CPRD and Mondriaan risk estimates were calculated from the Model C.

Figure 1. Meta-analysis of the association between several studies of benzodiazepines and hip/femur fracture including the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium cohort studies

DOI: 10.1002/pds

those detected in the cohort design as two different versions of the DB were used, with the most updated version used for the NCC study that included a larger number of patients.

Models B and C did not yield substantial additional changes in risk estimates as occurred in the cohort design. Adjusted OR (Model C) (95% CIs) for current use were 1.28 (1.16–1.42) in BIFAP, 1.60 (1.49–1.72) in CPRD, and 1.48 (0.89–2.48) in Mondriaan (Table 3).

In the cohort studies, no short-term effects (1–30 days) were observed in BIFAP and CPRD (Table 3) while the risk in Mondriaan was 3.12 (1.42–6.86) (Model C). In all data sources, patients using both anxiolytics and hypnotics appeared at greater risk than those using either drug alone (Table 3). In CPRD, risk estimates using Model D were slightly higher when an additional matching by GP practice was performed (OR with no matching = 1.32; (1.23–1.43)z; and OR with matching = 1.47 (1.35–1.59).

In our NCC studies, the pooled OR using Model C was 1.44 (1.19–1.75) with $I^2 = 83.9\%$. The pooled risk for 13 published case control studies[10,11] was 1.31 (1.18–1.46) with $I^2 = 72.5\%$ (data not shown).

The pooled risk of 22 studies, including published cohort and case-control studies[10,11] and the results from Model C of our three cohort studies, was 1.33 (1.23–1.43) with $I^2 = 84.2\%$ (Figure 1). This pooled risk was 1.28 (1.21–1.37) with $I^2 = 75.4\%$ when only BIFAP and CPRD studies were included using full model estimates (Model D) (Supporting Information Figure S1).

DISCUSSION

Using harmonized methods and procedures for the cohort and the NCC studies, *a priori*, we would not expect different results across DBs because of inconsistencies in the methodological approach. Similarities in findings across designs included the direction of the effect and the greater risk associated with the concomitant use of anxiolytics and hypnotics. However, differences were seen in the magnitude of overall risk and the effect of current duration, which suggests that other sources of heterogeneity, distinct from the methodological ones, have an important role, at least for this particular drug-event pair.

In cohort analyses, age was the main confounder in all DBs (Table 2). The rest of covariates included in the analysis had little impact, most probably because many of these variables are associated with age and because all subjects were BZD users and many comorbidities and chronic treatments associated with the use of BZDs were automatically controlled by design, minimizing the confounding by indication. Nevertheless, for BIFAP and CPRD, the model D yielded consistent lower estimates than the observed with the other models (A–C), probably indicating the existence of a residual confounding in model C.

We found an increased risk in all DBs when both anxiolytic and hypnotic drugs were used concomitantly. Other authors also found a dose effect,[27] suggesting that high doses of BZD were associated with a higher risk of falling and thus a higher risk of HFF. The increased risk associated with concomitant use is a reflection of a dose effect as BZDs (either hypnotics or anxiolytics) have the same mechanism of action, and taking two BZDs would be equivalent to an increase in dose. In addition, associations are relatively small and residual confounding (i.e., patients taking both medications are likely sicker than patients on monotherapy) may also play a role in explaining these associations. In a twin study on the same drug-event pair in which case-only designs were performed,[28] stronger associations were found for the concomitant use, and it was suggested that this might be because of a better control for intrinsic confounding factors that are difficult or impossible to control for in other designs.

Risk for HFF from BZDs has been described as "immediate and transient";[27,29] thus, a positive association with a recent use would not be expected, or at least, we would presume a gradient between current and recent periods of use. Although all DBs detected an increased risk in the recent use period, a gradient in current versus recent users was only seen in CPRD for both designs. Differences could be partly explained by a residual effect because of spare pills; however, we did not find differences in median duration of current periods nor median prescription duration among the DBs. Drug withdrawal might play a role[30] as well, so further research is worthwhile.

The timing to reach the peak effect remains an important issue, although this was not found in BIFAP and CPRD; in Mondriaan, the results suggested a short-term effect. Similar inconsistencies have also been found in published studies.[8,27] From a pharmacological point of view, we would expect a greater sedating effect at the beginning of treatment triggering falls and fractures, but such an effect is dose dependent and, for that reason, guidelines advise to prescribe, particularly in the elderly, lower doses at the start and to escalate doses according to the response.[31] Such a dose scaling might explain the lack of a short-term effect in some studies. This issue was not specifically explored in the present research,

DOI: 10.1002/pds

Table 3. Nested case control study

	BIFAP					
	Cases (2459) *N* (%)	Controls (9840) *N* (%)	Model A* OR (95%CI)	Model B† OR (95% CI)	Model C‡ OR (95% CI)	Model D§ OR (95%CI)
Past users (reference category)	851 (34.61)	3949 (40.17)	1	1	1	1
Recent users	268 (10.90)	934 (9.50)	1.39 (1.18–1.63)	1.36 (1.15–1.60)	1.35 (1.15–1.60)	1.30 (1.09–1.54)
Current users	1340 (54.49)	4947 (50.33)	1.30 (1.17–1.43)	1.28 (1.16–1.42)	1.28 (1.16–1.42)	1.13 (1.02–1.26)
Among current users only						
Duration (days)						
1–30	184 (13.73)	786 (15.89)	1.06 (0.86–1.32)	1.10 (0.88–1.36)	1.09 (0.88–1.36)	1.03 (0.82–1.30)
31–60	179 (13.36)	686 (13.87)	1.26 (1.02–1.56)	1.27 (1.03–1.57)	1.27 (1.02–1.57)	1.18 (0.95–1.48)
61–182	267 (19.93)	1058 (21.39)	1.20 (1.02–1.41)	1.18 (1.00–1.39)	1.18 (1.00–1.39)	1.04 (0.88–1.24)
183–365	212 (15.82)	660 (13.34)	1.51 (1.27–1.80)	1.48 (1.24–1.76)	1.48 (1.24–1.77)	1.30 (1.08–1.56)
>365	498 (37.16)	1757 (35.52)	1.34 (1.17–1.53)	1.32 (1.16–1.51)	1.31 (1.15–1.50)	1.14 (0.99–1.31)
ATC subgroup						
Use of both anxiolytics and hypnotics¶	161 (12.01)	407 (8.23)	1.87 (1.53–2.28)	1.82 (1.49–2.22)	1.82 (1.49–2.23)	1.39 (1.12–1.72)
Single use of anxiolytics	894 (66.72)	3476 (70.26)	1.23 (1.10–1.37)	1.22 (1.09–1.36)	1.22 (1.09–1.36)	1.11 (0.99–1.25)
Single use of hypnotics	285 (21.27)	1064 (21.51)	1.28 (1.09–1.49)	1.26 (1.08–1.48)	1.27 (1.08–1.48)	1.10 (0.94–1.30)
	CPRD					
	Cases (5966) *N* (%)	Controls (21 806) *N* (%)	Model A* OR (95%CI)	Model B† OR (95%CI)	Model C‡ OR (95%CI)	Model D§ OR (95%CI)
Past users (reference category)	3111 (52.15)	13 218 (60.62)	1	1	1	1
Recent users	466 (7.81)	1781 (8.17)	1.20 (1.06–1.35)	1.20 (1.06–1.36)	1.19 (1.05–1.35)	1.08 (0.95–1.22)
Current users	2389 (40.04)	6807 (31.22)	1.61 (1.50–1.72)	1.60 (1.49–1.71)	1.60 (1.49–1.72)	1.32 (1.23–1.43)
Among current users only						
Duration (days)						
1–30	379 (15.86)	1339 (19.67)	1.41 (1.16–1.70)	1.39 (1.15–1.69)	1.37 (1.12–1.66)	1.29 (1.06–1.59)
31–60	276 (11.55)	940 (13.81)	1.37 (1.13–1.65)	1.39 (1.15–1.68)	1.39 (1.15–1.68)	1.28 (1.05–1.57)
61–182	480 (20.09)	1280 (18.80)	1.78 (1.56–2.03)	1.76 (1.54–2.01)	1.77 (1.54–2.03)	1.45 (1.26–1.67)
183–365	382 (15.99)	903 (13.27)	1.96 (1.70–2.26)	1.97 (1.71–2.28)	2.00 (1.73–2.30)	1.64 (1.41–1.91)
>365	872 (36.50)	2345 (34.45)	1.52 (1.39–1.67)	1.50 (1.37–1.65)	1.51 (1.37–1.65)	1.20 (1.08–1.32)
ATC subgroup						
Use of both anxiolytics and hypnotics¶	173 (7.24)	286 (4.20)	2.70 (2.21–3.28)	2.70 (2.21–3.29)	2.69 (2.20–3.29)	1.87 (1.51–2.31)
Single use of anxiolytics	664 (27.79)	1757 (25.81)	1.81 (1.63–2.01)	1.80 (1.62–2.01)	1.80 (1.61–2.00)	1.45 (1.30–1.63)
Single use of hypnotics	1552 (64.96)	4764 (69.99)	1.47 (1.36–1.59)	1.46 (1.35–1.58)	1.47 (1.36–1.58)	1.24 (1.15–1.35)
	MONDRIAAN‖					
	Cases (148) *N* (%)	Controls (580) *N* (%)	Model A* OR (95%CI)	Model B† OR (95%CI)	Model C‡ OR (95%CI)	
Past users (reference category)	79 (53.4)	337 (58.1)	1	1	1	
Recent users	21 (14.2)	84 (14.5)	1.21 (0.64–2.30)	1.19 (0.61–2.30)	1.22 (0.62–2.39)	
Current users	48 (32.4)	159 (27.4)	1.46 (0.89–2.38)	1.53 (0.92–2.54)	1.48 (0.89–2.48)	
Among current users only						
Duration (days)						
1–30	20 (41.7)	41 (25.8)	2.96 (1.40–6.26)	2.97 (1.37–6.44)	3.12 (1.42–6.86)	
31–60	12 (25.0)	60 (37.7)	1.07 (0.50–2.27)	1.12 (0.53–2.39)	0.99 (0.45–2.16)	
61–182	9 (18.8)	31 (19.5)	1.45 (0.60–3.50)	1.50 (0.61–3.69)	1.55 (0.62–3.86)	
183–365	3 (6.2)	11 (6.9)	1.18 (0.33–4.27)	1.28 (0.34–4.86)	1.35 (0.35–5.16)	
>365	4 (8.3)	16 (10.1)	1.11 (0.37–3.38)	1.26 (0.40–3.96)	1.13 (0.35–3.66)	
ATC subgroup						
Use of both anxiolytics and hypnotics¶	7 (14.6)	10 (6.3)	3.22 (1.17–8.86)	3.40 (1.17–9.92)	3.48 (1.18–10.26)	
Single use of anxiolytics	19 (39.6)	84 (52.8)	1.08 (0.57–2.03)	1.11 (0.58–2.13)	1.06 (0.54–2.06)	
Single use of hypnotics	22 (45.8)	65 (40.9)	1.60 (0.86–2.95)	1.73 (0.92–3.26)	1.67 (0.87–3.19)	

Risk of HFF in association with BZDs according to categories of exposure, duration, and class of drug for the different models of adjustment of confounders OR: odds ratio; CI, confidence interval; BZDs, benzodiazepines; BIFAP, Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria; CPRD, Clinical Practice Research Datalink.

*Model A: age and sex.

†Model B: Model A plus previous fracture, systemic glucocorticoid, and rheumatoid arthritis.

‡Model C: Model A plus Model B plus history of osteoporosis, or Paget's disease and use of bisphosphonate, raloxifene, strontium ranelate, calcitonin, parathyroid hormone, calcium and vitamin D.

§Model D: Model A plus Model B plus Model C plus the following co-medications: antidepressants, antipsychotics/lithium, anti-Parkinson drugs, anticonvulsants, inhaled glucocorticoids, bronchodilators (including beta-2-adrenoceptors agonist and anticholinergics), anti-arrhythmic, sedating antihistamines, antihypertensive drugs (including ACE inhibitors, angiotensin II antagonists, beta blocking agents, calcium channel blockers, and other antihypertensives), diuretics, estrogen-containing hormone replacement therapy, thyroid hormones, antithyroid drugs, disease-modifying anti-rheumatic drugs, thiazolidinediones, other antidiabetics, antiemetic (metoclopramide), anticoagulants, morphine/opiate, two or more prescriptions for a non-steroidal anti-inflammatory drug (NSAIDs), statins, proton pump inhibitors, and aromatase inhibitors; plus the following co-morbidities: anemia, seizures, syncope, ischemic heart disease, cerebrovascular disease, malignant neoplasms, inflammatory bowel disease, obstructive airway disease, liver disease, impaired renal function, mental disorders, and any form of dementia (including Alzheimer's disease).

¶Use of anxiolytics and hypnotics within the current period, regardless of whether or not they were in concomitant use.

‖Because of the number of outcomes in Mondriaan, Model D is not presented in this table.

DOI: 10.1002/pds

but it will be addressed in future investigations. Van der Hooft *et al.,*[32] have also hypothesized that clinicians might be more vigilant or alternatively patients might be more alert when initiating BZD therapy, and such vigilance may decrease with time. Differences observed from one country to another could stem from different practices in vigilance, but also from other factors including the different pattern of drugs used, as noted in a companion descriptive study.[33] In any case, results obtained in Mondriaan have to be interpreted cautiously because of the small number of cases found in this DB.

We found differences in the magnitude of risk between DBs. Many factors can contribute to explain these results such as the background risk of HFF of the population,[34] the multiple comorbidities and concomitant use of other drugs that may interact in complex ways with BZDs, the living conditions of the populations, and the healthcare system. The differences found in the pattern of individual drugs used in the respective countries may contribute as well.[33] Although BZDs share the mechanism of action, they have different pharmacokinetics (e.g., generation of active metabolites), and we cannot assume that all active ingredients have similar risks. Moreover, it is possible that some handling factors unique to each DB may still play a role. Although we used a common protocol with as much as possible harmonization across DBs, we accommodated specific features of individual DBs. For instance, in BIFAP, the search for potential cases included free text, in addition to the diagnostic codes. Other databases did not perform this strategy. In BIFAP, the search for free text raises the sensitivity to detect cases but has the potential to increase the rate of false positives, so it is part of the quality control of this DB to perform a case validation of all potential cases (detected either through codes or free text) in order to eliminate false positives. Assuming that, after this validation, there is no false positives in BIFAP, and that the rate of false negatives is probably non-differential with respect to the exposure, we should not expect a relevant impact on the measures of association. Therefore, it is unlikely that a misclassification of cases can explain the lower relative risk estimate observed in BIFAP as compared with the ones of the other DBs. In a parallel study of antidepressants and risk of HFF,[35] we found similar results between DBs, suggesting that these data management differences across DBs were less important than population characteristics and their drug use.

The two analytical designs evaluated in this study were expected to give comparable results. As expected, small differences were observed in CPRD and BIFAP. However, marked differences in results between the two designs were observed in Mondriaan. We cannot provide a clear explanation for these results. In principle, we may postulate that matching on time in a case-control study nested in a cohort of new BZDs users will lower the risk estimates, as cases that experienced an HFF shortly after inclusion were matched with controls with equal amount of time from initiating BZD use. As a consequence, such controls were more likely to be exposed at that moment, which would result in an overall lower risk estimates in the NCC as compared with the cohort analysis. However, if this explanation were correct, we should have observed the most important difference between designs in the early period after exposure (first 30 days), particularly in Mondriaan in which most cases occurred in such a period, but this was not the case. However, we should not overemphasize the differences between designs in Mondriaan, as this DB contributed with small number of cases, and the confidence intervals of the risk estimate associated with current use in NCC and cohort study overlap considerably.

The design of a single cohort of new users, in which the comparison is made between current use periods with past use periods of the same group of patients, was selected in order to minimize the confounding by indication, which will more likely arise if an external cohort of either non-users or new users of non-psychotropic drugs had been chosen. The option to compare a cohort of BZDs new users with a cohort of Z-drugs new users was discarded as Z-drugs have also been associated with a greater risk of hip fractures.[18] However, the design selected has the limitation to increase a time dependent within person confounding similar to case-only designs.

The pooled estimates from our studies were consistent with those reported in the meta-analyses by Xing *et al.*,[10] and Khong *et al.*,[11] although the heterogeneity (I^2 index) in our studies remained substantial, being slightly higher in the cohort studies than in the NCC studies. However, the interpretation of the I^2 value has to be taken with caution because, in addition to many factors, the stability of I^2 index is a factor of the number of studies included in the meta-analysis, which was only three for our studies.[36,37]

This drug-event pair was also explored in the Observational Medical Outcomes Partnership Project project applying cohort and self-controlled case series designs in ten different DBs.[14] Despite using a common data model and common methods, important discrepancies were found. However, this study

was not specific to BZD-HFF and explored 53 drug-event pairs under a surveillance perspective rather than specifically addressing this pair in a formal hypothesis testing study.

An important objective of PROTECT, in addition to comparing the results across DBs, was to assess the feasibility of a multi-site collaboration using a decentralized approach. Multisite studies are ever more often being used in pharmacoepidemiology as they allow for assessment of consistency of drug-event associations in different settings (DBs and countries). In addition, using multisite studies allows for increasing the sample size and helps to identify sources of heterogeneity that may point to potential effect modifiers. Although a decentralized approach demands greater coordination efforts than a centralized one, we have shown that it is feasible and has the advantage of taking explicitly into account the intrinsic differences among DBs (including the differences in population and drug utilization patterns).

This study has potential limitations that need to be considered. The purpose of applying a common protocol was to ensure a consistent analytical approach to all DBs; however, specific requirements for DBs were present, and minor differences in data processing and analyses could explain some of the differences noted across sites. Also, the use of different versions of the CPRD for both designs provided in the NCC study a population one-third larger than used for the cohort study, so the number of HFF occurred in the cohort study would be lower. Finally, model C instead of model D was chosen as primary analysis because of the relatively small size of Mondriaan, but the results from Model D in CPRD and BIFAP show that in Model C there is some residual confounding.

It is also possible that other factors of potential importance remained unmeasured, such as information on socioeconomic status, physical frailty, or cognitive impairment. Over the counter delivery is not expected to affect results because BZDs are prescription drugs under strict dispensing control in all participating countries. However, we cannot exclude the possibility that patients did not collect the drug from the pharmacy and that the extent of this might differ between DBs because of differences in country-specific reimbursement rules for the BZDs. We found differences between DBs with regard to prior risk factors (for example, previous hip fractures), which are to a large extent due to the differences in the time to look back for each patient in the corresponding DB.

Another aspect of using a common protocol is the selection of a fixed set of confounders in the multivariable models regardless of their impact, which would need further investigation. Finally, comparability of these findings with others should be restricted to studies performing new-users design.

In conclusion, in this study, the risk of HFF in BDZs users was tested in a multi-DB setting, and a consistently increased risk was confirmed. However, some relevant discrepancies across DBs were also obtained. Before embarking into a muti-DB study, the potential differences between DBs should be carefully considered. These include, among others, the type of drug usage, background incidence of the outcome, characteristics of the underlying populations, availability of information on potential confounders, and an estimation of the DB sample size in order to allow for a proper and homogenous confounding control. The outcome in such studies should be defined, potential misclassification of outcome should be discussed, and whenever possible clinical validation of the outcome should be performed. Methodological studies applying different designs and using multiple DBs like the present one are needed to learn more about the potential caveats and strengths of using electronic administrative databases in pharmacoepidemiologic studies.

CONFLICT OF INTEREST

G.R., C.H., H.G., V.A., B.O., R.G.G., A.Afonso, A.A., E.M., and F.d.A. declare that they have no conflict of interest.

O.K. and M.d.G. have received unrestricted funding for pharmacoepidemiological research from the Dutch private–public funded Top Institute Pharma.

P.C.Souverein has received unrestricted funding for pharmacoepidemiological research from the private-public funded Top-Institute PHARMA (www.tipharma.nl, includes co-funding from universities, government and industry) and the EU Innovative Medicines Initiative (IMI).

J.L., D.W., and N.B. are employees and stockholders of GlaxoSmithKline.

M.M. was an employee of Merk Serono SA.

C.S. has received unconditional grant from Merck Serono SA – Geneva.

S.J. is an employee of Astra-Zeneca.

R.S. is a full-term Novartis employee and owns Novartis shares.

A.B. and R.R. are employees and stockholders of Pfizer, Inc.

DOI: 10.1002/pds

KEY POINTS

- Results from three uniformly conducted cohort studies and three NCC designs consistently showed that the use of BZDs (anxiolytics, hypnotics, and related drugs) is associated with a moderate increased risk of hip/femur fracture.
- Consistent findings included the greater risk associated with the concomitant use of anxiolytics and hypnotic drugs.
- Differences were found in the magnitude of overall risk and the timing to reach the peak effect, which suggests that other sources of heterogeneity, distinct from the methodological ones, have an important role.
- With this study, we have shown that performing multi-site studies, using a common protocol in a decentralized cooperative effort, is feasible and may contribute to a better assessment of drug safety.

ETHICS STATEMENT

Study protocols were approved by institutional review boards responsible for each individual database.

ACKNOWLEDGEMENTS

The paper is on behalf of WP2 and WP6. The research leading to these results was conducted as part of the PROTECT consortium (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium, www.imi-protect.eu), which is a public–private partnership coordinated by the European Medicines Agency.

Authors would like to thank the excellent collaboration of physicians in the participating countries, whose contribution in recording their professional practice with high-quality standards makes possible the availability of databases used in this research.

B. O. was employed at the BIFAP Research Unit, and M. M. was employed by Merck Serono SA when this study was conducted, but they no longer work there.

The PROTECT project has received support from the Innovative Medicines Initiative Joint Undertaking (IMI JU) (www.imi.europa.eu) under grant agreement no. 115004, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA (European Federation of Pharmaceutical Industries and Association) companies' in-kind contribution. As a special form of the IMI JU grant, Utrecht University and University of Alcalá received a direct financial contribution from Pfizer and Astra-Zeneca, respectively. M. M., A. B., R. S., S. J., and R. R. belong to EFPIA (European Federation of Pharmaceutical Industries and Association) member companies in the IMI JU, and costs related to their part in the research were carried by the respective company as in-kind contribution under the IMI JU scheme. The views expressed are those of the authors only and not of their respective institution or company.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design. The corresponding authors responsible for each database performed data extraction and raw data analysis. C. H., G. R., J. L., and F. d. A. wrote the first draft, and all authors contributed with critical comments to the final version.

REFERENCES

1. Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. *Eur J Clin Pharmacol*. 2012 Feb; **68**(2): 123–129. PubMed PMID: 21808989.
2. Cardwell CR, Abnet CC, Veal P, Hughes CM, Cantwell MM, Murray LJ. Exposure to oral bisphosphonates and risk of cancer. *Int J cancer*. 2012 Sep 1; **131**(5): E717–E725. PubMed PMID: 22161552. Pubmed Central PMCID: 3517893 2012.
3. Green J, Czanner G, Reeves G, Watson J, Wise L, Beral V. Oral bisphosphonates and risk of cancer of oesophagus, stomach, and colorectum: case-control analysis within a UK primary care cohort. *BMJ* 2010; **341**: c4444. PubMed PMID: 20813820. Pubmed Central PMCID: 2933354.
4. Meier CR, Schlienger RG, Kraenzlin ME, Schlegel B, Jick H. HMG-CoA reductase inhibitors and the risk of fractures. *JAMA*. 2000 Jun 28; **283**(24): 3205–3210. PubMed PMID: 10866867. Epub 2000/06/24. eng.
5. van Staa TP, Wegman S, de Vries F, Leufkens B, Cooper C. Use of statins and risk of fractures. *JAMA*. 2001 Apr 11; **285**(14): 1850–1855. PubMed PMID: 11308398.
6. Wagner AK, Zhang F, Soumerai SB, *et al.* Benzodiazepine use and hip fractures in the elderly: who is at greatest risk? *Arch Intern Med*. 2004 Jul 26; **164**(14): 1567–1572. PubMed PMID: 15277291. Epub 2004/07/28. eng.
7. Cummings SR, Nevitt MC, Browner WS, *et al.* Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. *N Engl J Med* 1995 Mar 23; **332**(12): 767–773. PubMed PMID: 7862179. Epub 1995/03/23. eng.
8. Guo Z, Wills P, Viitanen M, Fastbom J, Winblad B. Cognitive impairment, drug use, and the risk of hip fracture in persons over 75 years old: a community-based prospective study. *Am J Epidemiol*. 1998 Nov 1; **148**(9): 887–892. PubMed PMID: 9801019.
9. Jacqmin-Gadda H, Fourrier A, Commenges D, Dartigues JF. Risk factors for fractures in the elderly. *Epidemiology*. 1998 Jul; **9**(4): 417–423. PubMed PMID: 9647906.
10. Xing D, Ma XL, Ma JX, Wang J, Yang Y, Chen Y. Association between use of benzodiazepines and risk of fractures: a meta-analysis. *Osteoporos Int*. 2014 Jan; **25**(1): 105–120. PubMed PMID: 24013517.
11. Khong TP, de Vries F, Goldenberg JS, *et al.* Potential impact of benzodiazepine use on the rate of hip fractures in five large European countries and the United States. *Calcif Tissue Int*. 2012 Jul; **91**(1): 24–31. PubMed PMID: 22566242. Pubmed Central PMCID: 3382650. Epub 2012/05/09. eng.
12. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549. PubMed PMID: 21310794.
13. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003 Sep 6; **327**(7414): 557–560. PubMed PMID: 12958120. Pubmed Central PMCID: 192859.
14. Madigan D, Ryan PB, Schuemie M, *et al.* Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol*. 2013 Aug 15;

178(4): 645–651. PubMed PMID: 23648805. Pubmed Central PMCID: 3736754.
15. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance ENCePP. [Oct 2014]. Available from: http://www.encepp.eu/encepp/studySearch.htm
16. Abbing-Karahagopian V, Kurz X, de Vries F, *et al*. Abstracts: bridging differences in outcomes of pharmacoepidemiological studies: design and first results of the PROTECT project. *Curr Clin Pharmacol*. 2014 May; **9**(2): 130–138.
17. WHO Collaborating Centre for Drug Statistics Methodology. Anatomical Therapeutic Chemical (ATC) Classification System. [Cited 2014 September]; [Last updated: 2013-12-19];Available from: http://www.whocc.no/atc_ddd_index/.
18. Finkle WD, Der JS, Greenland S, *et al*. Risk of fractures requiring hospitalization after an initial prescription of zolpidem, alprazolam, lorazepam, or diazepam in older adults. *J Am Geriatr Soc* 2011; **59**: 1883–1890.
19. Gardarsdottir H, Souverein PC, Egberts TC, Heerdink ER. Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *J Clin Epidemiol*. 2010 Apr; **63**(4): 422–427. PubMed PMID: 19880282. Epub 2009/11/03. eng.
20. Ray WA, Griffin MR, Downey W. Benzodiazepines of long and short elimination half-life and the risk of hip fracture. *JAMA*. 1989 Dec 15; **262**(23): 3303–3307. PubMed PMID: 2573741. Epub 1989/12/15. eng.
21. Ray WA, Griffin MR, Schaffner W, Baugh DK, Melton LJ, 3rd. Psychotropic drug use and the risk of hip fracture. *N Engl J Med*. 1987 Feb 12; **316**(7): 363–369. PubMed PMID: 2880292. Epub 1987/02/12. eng.
22. Huang AR, Mallet L, Rochefort CM, Eguale T, Buckeridge DL, Tamblyn R. Medication-related falls in the elderly: causative factors and preventive strategies. *Drugs Aging*. 2012 May 1; **29**(5): 359–376. PubMed PMID: 22550966.
23. Ensrud KE, Blackwell T, Mangione CM, *et al*. Central nervous system active medications and risk for fractures in older women. *Arch Intern Med*. 2003 Apr 28; **163**(8): 949–957. PubMed PMID: 12719205. Epub 2003/04/30. eng.
24. Kanis J. Assessment of osteoporosis at the primary health care level: on behalf of the World Health Organization Scientific Group (WHO); 2007 [Nov 2012]. Available from: http://www.shef.ac.uk/FRAX/pdfs/WHO_Technical_Report.pdf.
25. Tebe C, del Rio LM, Casas L, *et al.* Risk factors for fragility fractures in a cohort of Spanish women. *Gac Sanit*. 2011 Nov-Dec; **25**(6): 507–512. PubMed PMID: 21955641. Epub 2011/10/01. Factores de riesgo de fracturas por fragilidad en una cohorte de mujeres espanolas. spa.
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986 Sep; **7**(3): 177–188. PubMed PMID: 3802833.
27. Zint K, Haefeli WE, Glynn RJ, Mogun H, Avorn J, Sturmer T. Impact of drug interactions, dosage, and duration of therapy on the risk of hip fracture associated with benzodiazepine use in older adults. *Pharmacoepidemiol Drug Saf*. 2010 Dec; **19**(12): 1248–1255. PubMed PMID: 20931664.
28. Requena G, Logie J, Martín E, *et al*. Do case-only designs yield consistent results across design and different databases? A case study of hip fractures and benzodiazepines. *Pharmacoepidemiol Drug Saf.* 2016; **25**(Suppl. 1): 79–87.
29. Wang PS, Bohn RL, Glynn RJ, Mogun H, Avorn J. Hazardous benzodiazepine regimens in the elderly: effects of half-life, dosage, and duration on risk of hip fracture. *Am J Psychiatry*. Jun 2001; **158**(6): 892–898.
30. Chang CM, Wu EC, Chang IS, Lin KM. Benzodiazepine and risk of hip fractures in older people: a nested case-control study in Taiwan. *Am J Psychiatry*. 2008 Aug; **16**(8): 686–692. PubMed PMID: 18669947. Epub 2008/08/02. eng.
31. Ashton PCH. Guidelines for the rational use of benzodiazepines 2000 [26/07/2012]. Available from: http://www.benzo.org.uk/asgr.htm.
32. van der Hooft CS, Schoofs MW, Ziere G, *et al.* Inappropriate benzodiazepine use in older adults and the risk of fracture. *Br J Clin Pharmacol*. 2008 Aug; **66**(2): 276–282. PubMed PMID: 18460035. Pubmed Central PMCID: 2492915.
33. Huerta C, Abbing-Karahagopian V, Requena G, *et al.* Exposure to benzodiazepines (anxiolytics, hypnotics and related drugs) in seven European electronic healthcare databases: a cross-national descriptive study from the PROTECT-EU Project. Submitted to Pharmacoepidemiology and Drug Safety. Accepted for publication, PDS-14-0405.R2 2015.
34. Requena G, Abbing-Karahagopian V, Huerta C, *et al.* Incidence rates and trends of hip/femur fractures in five European countries: comparison using e-healthcare records databases. *Calcif Tissue Int*. 2014 Jun; **94**(6):580–589. doi: 10.1007/s00223-014-9850-y. Epub 2014 Apr 1. PMID: 24687523.
35. Abbing-Karahagopian V, Souverein PC, Martin E, *et al.* Understanding inconsistency in the results from observational pharmacoepidemiological studies: the case of antidepressant use and risk of hip fracture. Submitted to Pharmacoepidemiology and Drug Safety. Under review. PDS-14-0413 2014.
36. Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008 Nov 27; **8**: 79. doi:10.1186/1471-2288-8-79. PubMed PMID: 19036172; PubMed Central PMCID: PMC2648991.
37. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy*. 2002 Jan; **7**(1): 51–61. Review. PubMed PMID:11822262.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site.