PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2016; **25**(Suppl. 1): 88–102
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3862

ORIGINAL REPORT

Understanding inconsistency in the results from observational pharmacoepidemiological studies: the case of antidepressant use and risk of hip/femur fractures

Patrick C. Souverein[1*,†], Victoria Abbing-Karahagopian[1†], Elisa Martin[2], Consuelo Huerta[2], Francisco de Abajo[3,4], Hubert G. M. Leufkens[1,5], Gianmario Candore[6], Yolanda Alvarez[6], Jim Slattery[6], Montserrat Miret[7], Gema Requena[4], Miguel J. Gil[2], Rolf H. H. Groenwold[1,8], Robert Reynolds[9], Raymond G. Schlienger[10], John W. Logie[11], Mark C. H. de Groot[1], Olaf H. Klungel[1,8], Tjeerd P. van Staa[1,12], Toine C. G. Egberts[1,13], Marie L. De Bruin[1,4] and Helga Gardarsdottir[1,13]

[1] *Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, The Netherlands*
[2] *BIFAP Research Unit, Division of Pharmacoepidemiology and Pharmacovigilance, Agencia Española de Medicamentos y Productos Sanitarios (AEMPS), Madrid, Spain*
[3] *Clinical Pharmacology Unit, University Hospital Príncipe de Asturias, Madrid, Spain*
[4] *Department of Biomedical Sciences, School of Medicine and Health Sciences, University of Alcalá, Spain*
[5] *MEB, Medicines Evaluation Board, Utrecht, The Netherlands*
[6] *EMA, European Medicines Agency, London, United Kingdom*
[7] *Merck KGaA, Geneva, Switzerland*
[8] *Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands*
[9] *Department of Epidemiology, Pfizer, New York, United States*
[10] *Global Clinical Epidemiology, Novartis Pharma AG, Basel, Switzerland*
[11] *Institute of Population Health, University of Manchester, Manchester, United Kingdom*
[12] *Worldwide Epidemiology, GlaxoSmithKline, Stockley Park, United Kingdom*
[13] *Department of Clinical Pharmacy, Division of Laboratory and Pharmacy, University Medical Center Utrecht, Utrecht, The Netherlands*

ABSTRACT

Purpose Results from observational studies on the same exposure–outcome association may be inconsistent because of variations in methodological factors, clinical factors or health care systems. We evaluated the consistency of results assessing the association between antidepressant use and the risk of hip/femur fractures in three European primary care databases using two different study designs.

Methods Cohort and nested case control studies were conducted in three European primary care databases (Spanish BIFAP, Dutch Mondriaan and UK THIN) to assess the association between use of antidepressants and hip/femur fracture. A common protocol and statistical analysis plan was applied to harmonize study design and conduct between data sources.

Results Current use of antidepressants was consistently associated with a 1.5 to 2.5-fold increased risk of hip/femur fractures in all data sources with both designs, with estimates for SSRIs generally higher than those for TCAs. In general, risk estimates in Mondriaan, the smallest data source, were higher compared to the other data sources. This difference may be partially explained by an interaction between SSRI and age in Mondriaan. Adjustment for GP-recorded lifestyle factors and matching on general practice had negligible impact on adjusted relative risk estimates.

Conclusion We found a consistent increased risk of hip/femur fracture with current use of antidepressants across different databases and different designs. Applying similar pharmacoepidemiological study methods resulted in similar risks for TCA use and some variation for SSRI use. Some of these differences may express real (or natural) variance in the exposure-outcome co-occurrences. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS—antidepressants; electronic healthcare record databases; hip/femur fracture; methodology; observational studies; confounding; adjustment; pharmacoepidemiology

Received 7 October 2014; Revised 27 July 2015; Accepted 29 July 2015

*Correspondence to: Patrick Souverein, Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, PO BOX 80082, 3508 TB Utrecht, The Netherlands. E-mail: P.C.Souverein@uu.nl
†contributed equally

INTRODUCTION

Antidepressant drugs are widely used for a range of indications.[1] Data from five European countries

indicate that the prevalence of antidepressant drug use ranged between 40 and 90/1000 persons in 2008.[2] Antidepressants have been associated with falls[3] and lower bone mineral density.[4] In observational epidemiological studies, antidepressant use has been associated with a 3–7% contribution to the population rate of hip fractures.[5] The magnitude of reported relative risk estimates varies considerably between studies, reflecting, at least in part, differences in study design, inclusion/exclusion criteria, exposure and outcome definitions and availability and classification of data on potential confounders. Data from several meta-analyses from observational studies[6–10] have shown that relative risk estimates for fractures range from 1.01[11] to 2.40[12] for selective serotonin re-uptake inhibitors (SSRIs) and from 1.21[11] to 2.40[13] for tricyclic antidepressants (TCAs). The potential inconsistency of study findings on the same exposure–outcome association has fuelled the debate on the validity and value of observational evidence.[14]

In this light, the aim of this study was to evaluate the impact of study design and confounder adjustment strategies on the association between use of antidepressants and the risk of risk of hip/femur fractures and compare results across databases.

METHODS

Setting

Data were obtained from three different primary care databases in European countries. Details on these data sources, as well as the background of the PROTECT project have been described elsewhere.[15] In short, we used data from the Dutch Mondriaan project (http://www.projectmondriaan.nl) with two databases: Netherlands Primary Care Research Database (NPCRD: maintained by NIVEL www.nivel.nl), and Almere Health Care group ((AHC) http://www.zorggroepalmere.nl), the Spanish "Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria" (BIFAP) (http://www.bifap.org) and The Health Improvement Network (THIN) (http://www.thin-uk.com). Three of these databases are nationwide primary care databases covering a part of their country population: 2% (Mondriaan NPCRD), 5.7% (THIN) and 6.8% (BIFAP). Mondrian-AHC is a primary care regional database representing 170 thousand patients (90.3% of citizens in 2008) from the city of Almere in the Netherlands containing GP diagnoses and prescriptions, as well as drugs prescribed by specialists that were dispensed from AHC community pharmacies. All participating databases fulfill quality standards for pharmacoepidemiological research. A common protocol and data specification document approved by all study participants and by an external committee was used for this study. The study protocol has been registered in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) study registry (http://www.encepp.eu/). A full list of drug, outcome and confounders codes that were used can be found online at the ENCePP website (http://www.encepp.eu/encepp/viewResource.htm?id=3667).

Study base, follow-up and case ascertainment

In each data source, patients with at least one prescription for a SSRI or TCA during the study period 2001–2009 were identified. Patients were eligible for inclusion if they had at least one year of prior enrolment with the general practitioner (GP) and were ≥18 years old at the date of the first antidepressant prescription (new user design). Patients were followed from first antidepressant prescription (start date) until the end of the study period or censoring (a recording code of hip/femur fracture (case) as identified by International Classification Primary Care (ICPC)-2 codes and string text diagnosis in BIFAP, by ICPC-2 codes in Mondriaan and by Read codes in THIN, death or transferring out of practice), whichever came first. Patients with a history of a hip/femur fracture in the year prior to the start date were excluded from the analysis. Using this study base, both cohort studies with time-varying exposure measurements and nested case–control studies were conducted.

Exposure definition

Exposure assessment for both the cohort and nested case–control study was based on the calculation of treatment episodes, defined by subsequent prescriptions for antidepressants, independent of switching of type and dose change and constructed according to the method of Gardarsdottir *et al.*[16] The duration of a prescription was based on the amount of tablets dispensed and the prescribed dosage regimen. In Mondriaan, a fixed duration of 90 days for each prescription was used because of absence of dosage instructions. A new treatment episode was assumed when an interval of 30 days or more occurred between the theoretical end date of a prescription and the dispensing date of the subsequent prescription for the same patient. Subsequently, a 30-day washout period was added to each treatment episode. These treatment episodes were the basis for

DOI: 10.1002/pds

exposure classification within both study designs, as described below.

Cohort study

Based on the treatment episodes, exposure status for each patient during follow-up was categorized in time-varying periods of current, recent and past use of antidepressants. Recent use included the period between 1 and 60 days after current use, while past use included the period following recent use until a new prescription was issued, the outcome occurred or the end of follow-up was reached. Current and past use exposure periods longer than 182 days were subdivided in 182-day time windows to enable updates of exposure categories and confounder status. For each current use time window, we assessed the type of antidepressant (SSRI, TCA or both) and duration of current use (1–30 days, 31–182 days, 183–365 days and >365 days). Cases were all patients with a first record/diagnosis of hip/femur fracture identified during follow-up.

Nested case–control study

Cases were all patients with a first record/diagnosis of hip/femur fracture identified during follow-up. The date of the first diagnosis of a hip/femur fracture was considered the index date. Up to four controls were matched on index date, sex, age (±2 years) and duration of follow-up time from the start date (±6 months) using risk-set sampling. The preference was for an age difference of zero years, with controls progressively being selected by relaxing time day by day up to a maximum of 6 months. Because of the matching process, the number of cases included in the nested case–control analysis does not necessarily match exactly the number of events in the cohort study. We categorized cases and controls according to exposure status on the day before index date as current, recent or past users. Current use was defined when the index date was between the start and end of a treatment episode. For patients not being current users, we made a distinction between recent use (treatment episode ended 1–60 days before the index date) and past use (treatment episode ended more than 60 days before the index date). We stratified current use by type of antidepressant (SSRI, TCA or both) and duration of use (1–30 days, 31–182 days, 183–365 days and >365 days).

Confounder definition

Besides age and sex, we defined three groups of potential confounders: (i) well-established risk factors for fracture (e.g. history of fractures, use of glucocorticoids), (ii) risk factors immediately related to the outcome (e.g. history of osteoporosis, use of bisphosphonates) and (iii) other risk factors that have been associated with fractures in previous studies (e.g. use of benzodiazepines, history of epilepsy). Comorbidities were identified as ever before the start date of each exposure time window (cohort study) or index date (nested case–control study), while confounders based on medication use were assessed in a 182-day time window prior to such dates.

Data analyses

In the cohort design, time-dependent Cox proportional hazard analysis was used to study the strength of the association between current use of antidepressants and the risk of hip/femur fracture, expressed as hazard ratios (HR) with 95% confidence intervals (95% CI). In the nested case–control analysis, conditional logistic regression analysis was used to calculate odds ratios (OR) with 95% CI. In both analyses past use was the reference category.

Confounder adjustment in both designs was conducted using the same statistical analysis plan for all three data sources. After calculating the crude relative risk estimates, multivariable models were fitted with incremental addition of confounders, as indicated in Box 1. Because of the small number of events in Mondriaan, adjustment according to model D was not possible. Therefore, results only indicate model C for Mondriaan analyses.

Box 1: Definition of models for incremental adjustment for confounders in cohort and case control study on antidepressant use and risk of hip/femur fractures

Model A: age and sex

Model B: variables in model A and previous fracture, systemic glucocorticoid use and rheumatoid arthritis and lifestyle factors (smoking, alcohol use and body mass index). Lifestyle factors were not available for Mondriaan and information on alcohol use was not available in BIFAP.

Model C: variables in model B and history of osteoporosis, history of other bone diseases (Paget's disease and osteogenesis imperfect), previous use of bisphosphonate or any of the other bone protecting drugs: raloxifene, strontium ranelate, parathyroid hormone, calcium and vitamin D, calcitonin, calcitriol

DOI: 10.1002/pds

Model D: variables in model C and co-medications: benzodiazepines, antidepressants other that TCAs or SSRIs, antipsychotics/lithium, anti-Parkinson drugs, anticonvulsants, inhaled glucocorticoids, bronchodilators (including Beta-2-adrenoceptors agonist and anticholinergics), anti-arrhythmics, sedating antihistamines, antihypertensive drugs (including ACE inhibitors, angiotensin II antagonists, Beta blocking agents, calcium channel blockers, other antihypertensives), diuretics, estrogen-containing hormone replacement therapy (HRT), thyroid hormones, antithyroid drugs, disease-modifying anti-rheumatic drug (DMARD), 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 and co-morbidities: anaemia, seizures, syncope, ischaemic heart disease, cerebrovascular disease, malignant neoplasms, inflammatory bowel disease, obstructive airway disease, liver disease, impaired renal function, mental disorders and dementia and/or Alzheimer.

Model D-LS: variables in model D without lifestyle factors (smoking, alcohol use and body mass index). Model D and Model D-LS are the same in Mondriaan.

All statistical analyses were conducted using SAS® v9.2 for THIN and Mondriaan, and Stata®-11 in BIFAP. Each database owner performed all analyses locally. A blinding procedure was maintained until all results were available at the coordinator center (at Utrecht University, the Netherlands).

Sensitivity analyses

Several sensitivity analyses were conducted to study their impact on risk estimation.

(1) The 90-day duration assumption for a single prescription was varied to 30 and 60 day in the Mondriaan cohort study
(2) The effect of additionally adjusting for life-style factors was studied in both the cohort study (THIN/BIFAP) and the nested case–control study (THIN)
(3) In the nested case–control study, the impact of including the general practice in the matching algorithm was studied in THIN.

RESULTS

Cohort study

The study population comprised 252 203, 22 954 and 587 637 new users of antidepressants in BIFAP, Mondriaan and THIN, respectively. The characteristics of the three cohorts are displayed in Table 1. The mean age was similar in all data sources, ranging from 49 to 51 years, while the proportion of women was 10% higher in BIFAP compared to Mondriaan and THIN. The distribution of baseline covariates varied between the cohorts (Table 1). The median duration of an antidepressant treatment episode was 90 days in BIFAP, 176 days in Mondriaan and 86 days in THIN.

Current use of antidepressants was associated with an increased risk of hip hip/femur fractures in all three data sources, with similar crude HRs between 2.16 and 2.37 (Table 2). Subsequent adjustment for confounding yielded final adjusted HRs of 1.59 (95% CI: 1.42–1.77) in BIFAP, 1.55 (95% CI: 1.45–1.67) in THIN and 2.61 (95% CI: 1.61–4.29, Model C) in Mondriaan. These analyses showed that age and sex were the most important confounders, as the subsequent blocks of confounders only had marginal impact on risk estimates (Table 2). In all cohorts, relative risk estimates were higher for current use of SSRIs than for current use of TCAs, although confidence intervals overlap. In general, observed effect sizes were higher in Mondriaan than in the other two cohorts (Table 2). When exploring this disprepancy to other data sources in more detail, we found some evidence of an interaction between age and SSRI use ($p = 0.07$). Without taking interaction into account, the overall effect of SSRI use on hip/femur fracture was 3.05 (95% CI: 1.83–5.09). Allowing for interaction between age and SSRI use in the adjusted model yielded a HR of 1.49 (95% CI: 0.57–3.93) for subjects aged 50.9 years (the mean age in BIFAP—Table 1). No interaction between age and TCA use was present ($p = 0.85$). There was also no evidence for interaction between age and antidepressant use in BIFAP and THIN. Additional analyses (results not shown) on non-linearity of age and hip/femur fracture in Mondriaan showed that the relation between age and the outcome was linear and adding e.g. age squared to the model did not change the effect estimates of both SSRI and TCA use.

Nested case–control study

We identified 1525, 79 and 3756 cases of hip/femur fracture in BIFAP, Mondriaan and THIN, respectively. Current use of antidepressants was associated with an increased risk of hip fracture in all three data sources

DOI: 10.1002/pds

Table 1. Patient characteristics of cohort and nested case–control studies, stratified by data source

	Cohort study			Nested case–control study					
	BIFAP	Mondriaan	THIN	BIFAP		Mondriaan		THIN	
				Cases	Controls	Cases	Controls	Cases	Controls
Patient characteristics	N = 252 203	N = 22 954	N = 587 637	N = 1525	N = 6137	N = 79	N = 305	N = 3756	N = 15 017
Age (years) mean ± SD	50.9 ± 16.9	48.8 ± 17.2	49.7 ± 18.5	78.2 ± 11.4	78.2 ± 11.3	75.3 ± 16.0	74.6 ± 15.9	78.1 ± 13.1	78.1 ± 13.0
Range	18–106	18–104	18–106	22–100	22–100	29–96	28–97	19–103	19–103
Females (%)	72.7	63.6	63.7	81.4	81.5	83.5	83.3	76.8	76.8
Median AD treatment episode (days)	90	176	86						
Range	1–3278	1–2920	1–3104						
Time in database before index date (days) mean ± SD				818 ± 667	818 ± 666	1713 ± 742	1769 ± 727	3880 ±1800	3899 ± 1792
Previous fractures (%)	5.0	4.8	18.6	20.7	12.7	20.3	14.1	38.0	24.1
Body mass index (kg/m^2) (%)									
<25	14.2	NA	39.8		NA		NA	49.0	37.9
≥25	32.1		49.0					45.7	47.7
Missing	53.7		11.2					15.3	14.4
Smokers (%)									
Never	33.4	NA	48.5		NA		NA	52.9	57.0
Ex	1.6		20.1					26.7	27.7
Current	15.5		30.0					16.6	11.1
Missing	49.5		1.4					3.8	4.2
Alcohol (%)									
Users	NA	NA	71.5		NA		NA	55.7	58.2
Non users			16.5					30.0	27.1
Missing			12.0					14.2	14.7
Co-morbidities (%)									
Rheumatoid arthritis	0.6	0.0	1.6	2.1	1.1	0.0	0.0	4.9	3.4
Osteoporosis	5.4	1.8	2.5	20.0	15.5	7.6	11.5	21.9	11.7
Paget's disease	0.1	NA	0.1	0.8	0.5	NA	NA	0.4	0.5
Anaemia	7.4	4.6	7.0	17.7	13.5	19.0	12.8	20.7	14.9
Epilepsy/seizure	0.9	0.7	2.3	2.3	1.3	5.1	1.6	4.1	2.2
Syncope	6.5	3.2	3.6	15.4	14.6	10.1	5.2	10.3	7.7
Ischaemic heart disease	4.8	4.4	6.9	12.6	12.8	12.7	17.0	21.9	21.1
Cerebrovascular disease	3.8	2.3	3.9	13.2	9.8	15.2	7.5	19.5	14.6
Malignant neoplasms	6.5	5.9	8.1	12.6	11.5	13.9	15.1	25.1	20.8
Inflammatory bowel disease	0.4	0.6	1.1	0.6	0.3	0.0	0.7	1.8	1.4
Obstructive airway disease	2.7	5.1	7.3	8.5	6.6	11.4	10.8	18.2	15.0
Liver disease	1.7	0.4	0.8	2.1	1.8	0.0	0.0	1.6	1.0
Chronic renal failure	0.9	0.0	0.5	5.5	4.4	0.0	0.0	2.4	1.9
Mental disorders (except depression)	2.1	4.1	2.4	2.8	1.9	8.9	3.6	4.0	2.8
Dementia and/or Alzheimer	1.2	0.6	0.6	12.9	7.8	11.4	4.3	10.0	4.7
Co-medication (%)									
Glucocorticoids (oral)	0.2	4.7	4.7	7.4	4.9	15.2	8.2	11.3	8.5
Bisphosphonate	2.5	1.6	1.6	9.2	7.9	3.8	7.5	10.5	7.2
Raloxifene	0.9	0.0	0.1	0.8	0.7	0.0	0.0	0.3	0.2
Strontium ranelate	0.1	0.0	0.0	0.5	0.4	0.0	0.0	0.5	0.3

(*Continues*)

DOI: 10.1002/pds

Table 1. (Continued)

	Cohort study			Nested case–control study					
	BIFAP	Mondriaan	THIN	BIFAP		Mondriaan		THIN	
				Cases	Controls	Cases	Controls	Cases	Controls
Patient characteristics	N = 252 203	N = 22 954	N = 587 637	N = 1525	N = 6137	N = 79	N = 305	N = 3756	N = 15 017
Parathyroid hormone	0.0	0.0	0.0	0.4	0.2	0.0	0.0	0.0	0.0
Calcium and vitamin D	4.2	1.6	0.2	13.4	12.3	5.1	5.2	0.7	0.5
Calcitonin	0.4	0.0	0.0	1.8	1.4	0.0	0.0	0.1	0.0
Benzodiazepines	39.6	34.5	16.4	60.1	56.4	43.0	43.0	24.5	19.1
Antidepressants *	3.5	3.3	2.6	10.8	7.8	5.1	3.6	5.9	4.2
Antipsychotics/lithium	5.5	2.6	5.0	14.3	10.1	7.6	5.9	12.8	8.2
Anti-Parkinson drugs	0.8	0.7	0.6	7.0	3.1	1.3	1.6	4.6	1.7
Anticonvulsants	4.2	3.2	2.8	13.8	9.8	7.6	5.6	7.9	5.1
Inhaled glucocorticoids	1.9	4.0	5.6	2.5	3.3	2.5	4.6	6.8	6.4
Bronchodilators	6.8	9.5	11.7	12.6	12.8	16.5	14.4	17.7	15.9
Anti-arrhythmic	0.6	0.4	0.5	2.9	2.1	0.0	0.7	1.9	1.3
Sedating antihistamines	0.9	0.0	0.4	2.0	2.1	0.0	0.0	0.9	0.7
Antihypertensive drugs**	18.9	19.2	22.9	50.7	51.3	36.7	40.7	50.1	52.1
Diuretics	9.2	9.6	13.2	33.6	32.9	40.5	33.1	40.7	39.7
Estrogen-containing HRT	2.2	3.2	6.2	0.1	0.4	2.5	3.0	2.2	3.2
Thyroid hormones	3.5	3.0	5.1	5.3	5.4	12.7	5.6	11.0	11.5
Anti-thyroid drugs	0.2	0.0	0.2	0.3	0.5	0.0	0.0	0.5	0.5
DMARDs	0.6	1.0	1.0	1.4	0.5	2.5	1.6	2.4	1.6
Thiazolidinediones	0.2	0.2	0.5	0.7	0.6	0.0	0.3	1.1	0.8
Other anti-diabetics	5.4	5.8	4.6	17.9	15.2	29.1	13.1	10.1	8.7
Metoclopramide	2.6	4.1	2.6	3.4	2.7	10.1	5.9	5.6	3.6
Anticoagulants	2.8	10.8	1.9	14.1	9.5	41.8	31.8	5.8	5.4
Morphine/opiate	6.8	8.6	11.3	19.1	12.9	21.5	10.8	22.9	15.4
NSAIDs (≥2 prescriptions)	16.6	11.4	19.7	20.4	22.4	11.4	13.4	16.6	16.7
Statins	9.2	9.0	11.5	18.0	22.4	17.7	18.4	26.8	28.7
Proton pump inhibitors	23.1	15.8	14.5	54.3	49.1	39.2	33.1	35.0	29.5
Aromatase Inhibitors	0.2	0.2	0.3	0.5	0.6	0.0	1.0	1.0	0.8

COPD = Chronic Obstructive Pulmonary Disease, HRT = Hormone Replacement Therapy DMARD = Disease-modifying Anti-rheumatic Drug, NSAID = Nonsteroidal Anti-inflammatory Drug.
*Antidepressant other than selective serotonin reuptake inhibitors or tricyclic antidepressants
**Includes ACE inhibitors, beta blockers, calcium channel blockers and angiotensin-2 antagonists

DOI: 10.1002/pds

Table 2a. Crude and adjusted hazard ratios of antidepressant use and risk of hip/femur fracture with incremental adjustment for confounders in the PROTECT cohort study, stratified by data source

		BIFAP			Mondriaan			THIN		
Model	Exposure	E	PY	HR (95%CI)	E	PY	HR (95%CI)	E	PY	HR (95%CI)
Crude	Past use	526	409 303	1.00 (reference)	27	36 421	1.00 (reference)	1751	1 401 610	1.00 (reference)
	Recent use *	137	57 617	1.91 (1.58-2.30)	7	3388	2.57 (1.11–5.93)	239	134 840	1.42 (1.23–1.64)
	Current use *	872	293 268	2.37 (2.12-2.65)	48	25 465	2.39 (1.47–3.86)	1766	651 516	2.16 (2.01–2.32)
	Type									
	TCA	93	37 135	2.01 (1.61 –2.50)	13	18 360	2.52 (1.30–4.92)	597	199 712	2.39 (2.17–1.63)
	SSRI	740	236 974	2.48 (2.22–2.78)	35	6620	2.40 (1.44–4.00)	1118	425 123	2.10 (1.94–2.27)
	SSRI +TCA	39	19 159	1.64 (1.18–2.27)	0	485	NA	51	26 681	1.52 (1.15–2.02)
	Duration (days)									
	0–30	83	58 203	1.44 (1.14–1.82)	7	2408	3.75 (1.61–8.71)	246	96 757	1.61 (1.32–1.97)
	31–182	302	105 401	2.04 (1.76–2.35)	24	10 077	3.05 (1.73–5.38)	600	254 163	1.74 (1.55–1.94)
	183–365	153	48 300	2.46 (2.06–2.95)	8	5030	1.99 (0.90–4.42)	290	108 087	2.05 (1.79–2.34)
	>365	334	81 364	3.21 (2.80–3.68)	9	7950	1.50 (0.70–3.19)	639	192 509	2.69 (2.46–2.95)
Model A	Recent use *	137	57 617	1.71 (1.42–2.07)	7	3388	2.60 (1.12–6.01)	239	134 840	1.38 (1.19–1.60)
	Current use *	872	293 268	1.63 (1.46–1.82)	48	25 465	2.53 (1.56–4.09)	1766	651 516	1.70 (1.59–1.60)
	Type									
	TCA	93	37 135	1.33 (1.06–1.66)	13	18 360	1.83 (0.94–3.56)	597	199 712	1.37 (1.25–1.52)
	SSRI	740	236 974	1.70 (1.52–1.90)	35	6620	3.05 (1.83–5.09)	1118	425 123	1.96 (1.81–2.12)
	SSRI +TCA	39	19 159	1.37 (0.99–1.89)	0	485	NA	51	26 681	1.42 (1.08–1.88)
	Duration (days)									
	0–30	83	58 203	1.20 (0.95–1.52)	7	2408	3.63 (1.56–8.44)	246	96 757	1.61 (1.31–1.97)
	31–182	302	105 401	1.64 (1.42–1.90)	24	10 077	3.14 (1.78–5.52)	600	254 163	1.76 (1.57–1.96)
	183–365	153	48 300	1.73 (1.44–2.07)	8	5030	2.29 (1.03–5.06)	290	108 087	1.73 (1.51–1.98)
	>365	334	81 364	1.71 (1.49–1.97)	9	7950	1.59 (0.75–3.39)	639	192 509	1.67 (1.53–1.84)
Model B	Recent use *	137	57 617	1.72 (1.43–2.08)	7	3388	2.63 (1.14–6.09)	239	134 840	1.38 (1.19–1.60)
	Current use *	872	293 268	1.61 (1.44–1.80)	48	25 465	2.61 (1.61–4.29)	1766	651 516	1.69 (1.57–1.81)
	Type									
	TCA	93	37 135	1.33 (1.06–1.66)	13	18 360	1.91 (0.98–3.73)	597	199 712	1.38 (1.26–1.53)
	SSRI	740	236 974	1.67 (1.49–1.87)	35	6620	3.13 (1.87–5.22)	1118	425 123	1.92 (1.77–2.07)
	SSRI +TCA	39	19 159	1.37 (0.99–1.90)	0	485	NA	51	26 681	1.39 (1.05–1.84)
	Duration (days)									
	0–30	83	58 203	1.13 (0.89–1.43)	7	2408	3.76 (1.62–8.75)	246	96 757	1.60 (1.31–1.96)
	31–182	302	105 401	1.54 (1.33–1.79)	24	10 077	3.23 (1.83–5.68)	600	254 163	1.76 (1.57–1.97)
	183–365	153	48 300	1.72 (1.44–2.07)	8	5030	2.36 (1.06–5.22)	290	108 087	1.70 (1.49–1.95)
	>365	334	81 364	1.77 (1.54–2.03)	9	7950	1.65 (0.78–3.52)	639	192 509	1.66 (1.52–1.83)
Model C	Recent use *	137	57 617	1.72 (1.43–2.08)	7	3388	2.63 (1.14–6.09)	239	134 840	1.38 (1.19–1.60)
	Current use *	872	293 268	1.61 (1.44–1.80)	48	25 465	2.61 (1.61–4.29)	1766	651 516	1.68 (1.57–1.81)
	Type									
	TCA	93	37 135	1.33 (1.06–1.66)	13	18 360	1.91 (0.98–3.73)	597	199 712	1.38 (1.25–1.52)
	SSRI	740	236 974	1.67 (1.49–1.87)	35	6620	3.13 (1.87–5.22)	1118	425 123	1.92 (1.78–2.08)
	SSRI +TCA	39	19 159	1.36 (0.98–1.89)	0	485	NA	51	26 681	1.39 (1.05–1.84)
	Duration (days)									
	0–30	83	58 203	1.13 (0.89–1.43)	7	2408	3.76 (1.62–8.75)	246	96 757	1.60 (1.31–1.96)
	31–182	302	105 401	1.55 (1.33–1.80)	24	10 077	3.23 (1.83–5.68)	600	254 163	1.76 (1.58–1.97)
	183–365	153	48 300	1.72 (1.43–2.06)	8	5030	2.36 (1.06–5.22)	290	108 087	1.70 (1.48–1.95)
	>365	334	81 364	1.77 (1.54–2.03)	9	7950	1.65 (0.78–3.52)	639	192 509	1.67 (1.52–1.83)

(Continues)

Pharmacoepidemiology and Drug Safety, 2016; **25**(Suppl. 1): 88–102
DOI: 10.1002/pds

Table 2a. (Continued)

Model	Exposure	BIFAP E	BIFAP PY	BIFAP HR (95%CI)	Mondriaan E	Mondriaan PY	Mondriaan HR (95%CI)	THIN E	THIN PY	THIN HR (95%CI)
Model D–LS	Recent use *	137	57 617	1.68 (1.39–2.03)				239	134 840	1.34 (1.16–1.56)
	Current use *	872	293 268	1.59 (1.42–1.77)				1766	651 516	1.55 (1.45–1.67)
	Type									
	TCA	93	37 135	1.30 (1.04–1.62)				597	199 712	1.31 (1.19–1.45)
	SSRI	740	236 974	1.66 (1.48–1.86)				1118	425 123	1.75 (1.61–1.90)
	SSRI +TCA	39	19 159	1.23 (0.89–1.71)				51	26 681	1.23 (0.93–1.63)
	Duration (days)									
	0–30	83	58 203	1.19 (0.94–1.50)				246	96 757	1.58 (1.29–1.93)
	31–182	302	105 401	1.62 (1.40–1.87)				600	254 163	1.71 (1.53–1.92)
	183–365	153	48 300	1.65 (1.37–1.98)				290	108 087	1.55 (1.36–1.78)
	>365	334	81 364	1.66 (1.45–1.91)				639	192 509	1.54 (1.40–1.69)

E = events; PY = person years; HR = hazard ratio; CI = confidence interval; NA = not applicable.

Table 2b. Crude and adjusted odds ratios of antidepressant use and risk of hip/femur fracture with incremental adjustment for confounders in the PROTECT nested case control study, stratified by data source

Model		BIFAP (1525 cases/6137 controls) Cases (%)	BIFAP Controls (%)	BIFAP Odds ratio (95% CI)	Mondriaan (79 cases/305 controls) Cases (%)	Mondriaan Controls (%)	Mondriaan Odds ratio (95% CI)	THIN (3756 cases/15 017 controls) Cases (%)	THIN Controls (%)	THIN Odds ratio (95% CI)
Crude (age and sex adjusted)	Past use	34.3	43.5	1.00 (reference)	32.9	49.2	1.00 (reference)	46.6	58.2	1.00 (reference)
	Recent use	8.9	7.4	1.65 (1.32–2.06)	8.9	7.9	2.31 (0.83–6.45)	6.4	5.7	1.59 (1.35–1.87)
	Current use	56.8	49.1	1.55 (1.36–1.76)	58.2	43.0	2.44 (1.36–4.40)	47.0	36.2	1.79 (1.65–1.94)
	Type									
	TCA	6.1	6.0	1.37 (1.06–1.76)	15.2	15.4	1.68 (0.77–3.68)	15.9	15.1	1.45 (1.30–1.62)
	SSRI	48.2	40.8	1.59 (1.39–1.81)	43.0	26.6	3.03 (1.59–5.80)	29.8	19.8	2.06 (1.88–2.26)
	SSRI +TCA	2.5	2.4	1.40 (0.97–2.01)	0	1.0	NA	1.4	1.3	1.37 (1.00–1.87)
	Duration (days)									
	0-30	5.4	5.7	1.23 (0.89–1.70)	8.9	5.6	4.04 (1.36–11.96)	6.2	5.5	1.87 (1.49–2.34)
	31-182	19.7	18.0	1.50 (1.24–1.82)	29.1	17.4	5.92 (2.43–14.41)	16.0	13.2	1.78 (1.57–2.03)
	183-365	10.0	8.5	1.49 (1.21–1.85)	10.1	10.5	2.37 (0.80–7.05)	7.7	5.4	1.88 (1.61–2.20)
	>365	21.8	16.9	1.65 (1.41–1.94)	10.1	9.5	1.25 (0.52–2.99)	17.0	12.2	1.74 (1.56–1.94)
Model B	Past use	34.3	43.5	1.00 (reference)	32.9	49.2		46.6	58.2	1.00 (reference)
	Recent use	8.9	7.4	1.64 (1.31-2.05)	8.9	7.9	2.36 (0.76–7.34)	6.4	5.7	1.52 (1.29–1.799)

(Continues)

DOI: 10.1002/pds

Table 2b. (Continued)

Model		BIFAP (1525 cases/6137 controls)			Mondriaan (79 cases/305 controls)			THIN (3756 cases/15 017 controls)		
		Cases (%)	Controls (%)	Odds ratio (95% CI)	Cases (%)	Controls (%)	Odds ratio (95% CI)	Cases (%)	Controls (%)	Odds ratio (95% CI)
	Current use	56.8	49.1	1.55 (1.36-1.76)	58.2	43.0	2.80 (1.46–5.37)	47.0	36.2	1.76 (1.623–1.911)
	Type									
	TCA	6.1	6.0	1.34 (1.04-1.73)	15.2	15.4	1.40 (0.54–3.59)	15.9	15.1	1.43 (1.28–1.59)
	SSRI	48.2	40.8	1.59 (1.39-1.82)	43.0	26.6	3.91 (1.90–8.04)	29.8	19.8	2.04 (1.86–2.24)
	SSRI +TCA	2.5	2.4	1.37 (0.95-1.98)	0	1.0	NA	1.4	1.3	1.33 (0.97–1.84)
	Duration (days)									
	0–30	5.4	5.7	1.23 (0.89–1.70)	8.9	5.6	6.05 (1.79–20.53)	6.2	5.5	1.89 (1.51–2.38)
	31–182	19.7	18.0	1.51 (1.25–1.82)	29.1	17.4	7.62 (2.73–21.25)	16.0	13.2	1.73 (1.52–1.97)
	183–365	10.0	8.5	1.48 (1.19–1.85)	10.1	10.5	2.67 (0.81–8.82)	7.7	5.4	1.86 (1.59–2.18)
	>365	21.8	16.9	1.66 (1.41-1.95)	10.1	9.5	1.39 (0.54-3.58)	17.0	12.2	1.72 (1.55–1.92)
Model C	Past use	34.3	43.5	1.00 (reference)	32.9	49.2		46.6	58.2	1.00 (reference)
	Recent use	8.9	7.4	1.64 (1.31-2.05)	8.9	7.9	2.45 (0.74–8.08)	6.4	5.7	1.49 (1.26–1.77)
	Current use	56.8	49.1	1.56 (1.37-1.78)	58.2	43.0	2.66 (1.37–5.17)	47.0	36.2	1.76 (1.62–1.91)
	Type									
	TCA	6.1	6.0	1.33 (1.03-1.72)	15.2	15.4	1.48 (0.58–3.77)	15.9	15.1	1.41 (1.26–1.58)
	SSRI	48.2	40.8	1.61 (1.41-1.84)	43.0	26.6	3.67 (1.77–7.64)	29.8	19.8	2.05 (1.87–2.25)
	SSRI +TCA	2.5	2.4	1.38 (0.95-1.99)	0	1.0	NA	1.4	1.3	1.35 (0.98–1.87)
	Duration (days)									
	0–30	5.4	5.7	1.24 (0.90–1.72)	8.9	5.6	5.45 (1.58–18.84)	6.2	5.5	1.88 (1.49–2.37)
	31–182	19.7	18.0	1.51 (1.25–1.83)	29.1	17.4	6.83 (2.39–19.55)	16.0	13.2	1.73 (1.51–1.97)
	183–365	10.0	8.5	1.49 (1.20–1.86)	10.1	10.5	2.35 (0.68–8.09)	7.7	5.4	1.88 (1.61–2.21)
	>365	21.8	16.9	1.67 (1.42-1.97)	10.1	9.5	1.47 (0.58-3.77)	17.0	12.2	1.72 (1.54–1.92)
Model D-LS	Past use	34.3	43.5	1.00 (reference)				46.6	58.2	1.00 (reference)
	Recent use	8.9	7.4	1.56 (1.24–1.96)				6.4	5.7	1.47 (1.24–1.75)
	Current use	56.8	49.1	1.52 (1.33–1.74)				47.0	36.2	1.60 (1.47–1.74)
	Type									
	TCA	6.1	6.0	1.27 (0.98–1.66)				15.9	15.1	1.35 (1.21–1.52)
	SSRI	48.2	40.8	1.58 (1.37–1.81)				29.8	19.8	1.81 (1.64–1.99)
	SSRI +TCA	2.5	2.4	1.23 (0.83–1.80)				1.4	1.3	1.19 (0.86–1.66)
	Duration (days)									
	0–30	5.4	5.7	1.26 (0.90–1.76)				6.2	5.5	1.70 (1.34–2.16)
	31–182	19.7	18.0	1.49 (1.22–1.81)				16.0	13.2	1.59 (1.39–1.82)
	183–365	10.0	8.5	1.43 (1.14–1.79)				7.7	5.4	1.74 (1.48–2.05)
	>365	21.8	16.9	1.62 (1.37–1.91)				17.0	12.2	1.54 (1.37–1.72)

CI = confidence interval; TCA = tricyclic antidepressants; SSRI = selective serotonin re-uptake inhibitors; NA = not applicable (no events)

DOI: 10.1002/pds

(Table 2b). Although confidence intervals were overlapping, the point estimates of current antidepressant use were the highest in Mondriaan (adjusted OR 2.66, 95% CI: 1.37–5.17), whereas the magnitude of the adjusted ORs in BIFAP and THIN were comparable to each other (adjusted OR 1.52, 95% CI: 1.33–1.74 and adjusted OR 1.60, 95% CI: 1.47–1.74, respectively). The difference between crude and adjusted models in BIFAP and THIN was relatively small. In Mondriaan, adjustment using model C resulted in a 9% increase of the OR compared to the crude estimate. In Mondriaan, increased risk of hip/femur fracture was found in the first 6 months of current antidepressant use, whereas no effect of duration was found in BIFAP and THIN. Noting that confidence intervals overlap, effect sizes seemed higher for SSRIs than for TCAs in all three sources, with the strongest effect observed for SSRI use in Mondriaan as well. The effect for TCAs was similar in all three data sources.

Sensitivity analyses

Varying the pre-defined duration of an antidepressants prescription in Mondriaan resulted in a shorter antidepressant treatment episode length, (median durations changing from 176 to 134 and 120 days for the 60 and 30-day prescription durations, respectively). However, this did not result in notable changes in risk estimates associated with current use of antidepressants overall, nor for SSRI and TCA separately (Supplementary Table C1).

Figure 1 (numbers in supplementary Table C2) shows the results of inclusion of lifestyle factors in the multivariable analysis in BIFAP (cohort study) and THIN (cohort study and nested case–control study). Despite observed differences at baseline/index dates in both study designs, the additional adjustment for lifestyle factors had negligible effect on the observed association between antidepressant use and hip/femur fractures, as effect sizes for the fully adjusted model with and without lifestyle factors were virtually identical for all exposure contrasts studied.

The effect of additional case–control matching on GP practice was studied in THIN. Inclusion of this matching criterion resulted in a smaller number of controls that could be matched to cases (n = 9031, vs. n = 15 017 in the non-GP matched THIN-dataset). The prevalence of comorbidities and co-medication was not different between controls included in the GP matched and the non-GP matched sets (data not shown). Crude and adjusted ORs for current use (overall, as well as for SSRI and TCA separately) were not different between the GP and non-GP matched datasets. For recent use, point estimates were consistently lower in the GP-matched cohort, though differences were not significant (Figure 2, numbers in

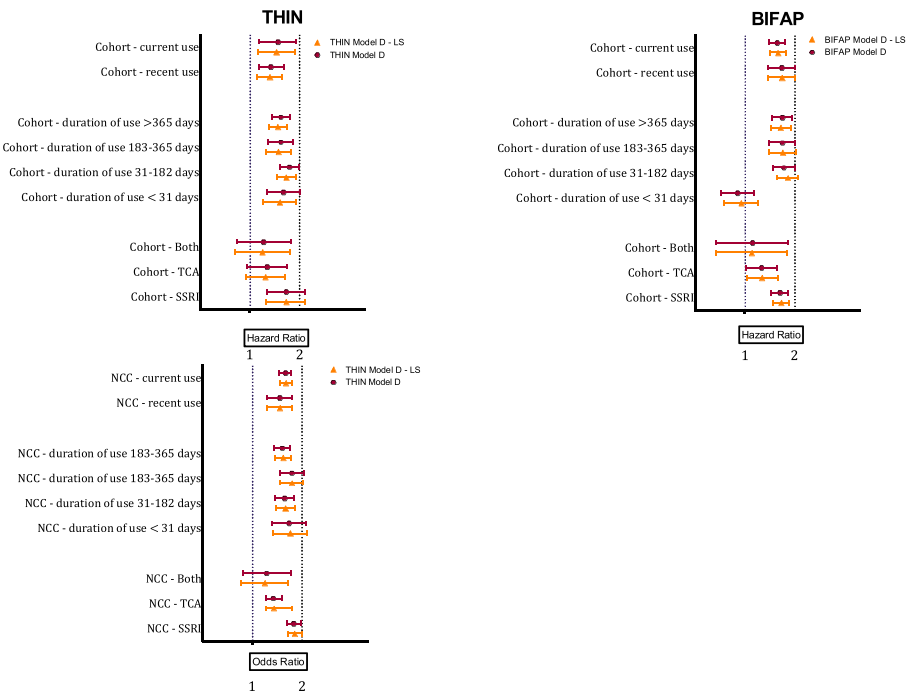

Figure 1. Hazard ratios and odds ratios with 95% confidence intervals for the association between use of antidepressants and risk of hip/femur fracture in THIN and BIFAP (cohort studies in top panels, nested case control study in bottom panel) for assessing the impact of additional adjustment for GP-recorded lifestyle factors

DOI: 10.1002/pds

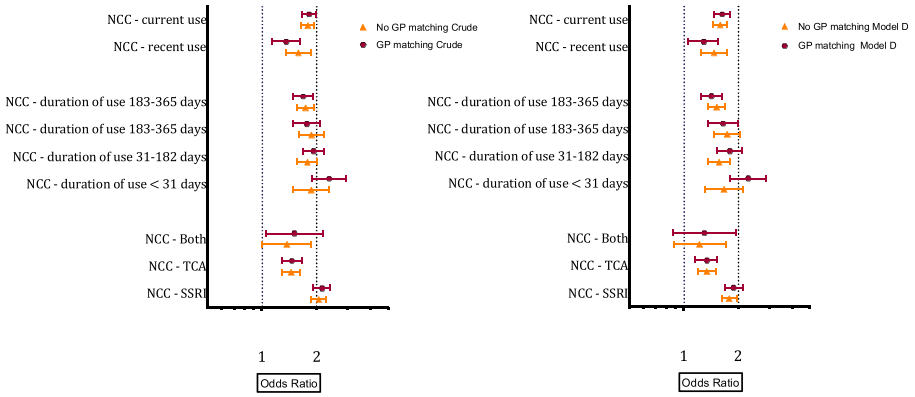

Figure 2. Crude (left panel) and adjusted odds ratios (right panel) with 95% confidence intervals for the association between use of antidepressants and risk of hip/femur fracture in THIN to assess impact of including general practice (GP) in the matching algorithm

supplementary Table C3). In the duration of use analysis, the risk estimates in the GP-matched analysis were more supportive for a higher effect among shorter duration of current use, but confidence intervals overlapped.

DISCUSSION

Using three data sources and two study designs, we found that antidepressants were consistently associated with an increased risk of hip/femur fractures. Effect sizes were similar between the cohort and nested case–control designs, which is reassuring as the same study base was used for both designs. In general, risk estimates tended to be higher in the Dutch Mondriaan database, despite harmonizing the study protocol and data specification. It should be noted that crude cohort results were actually very comparable between data sources, further suggesting that this difference only becomes apparent after age-adjustment, which is in line with the age-interaction finding in Mondriaan. While making note of overlapping confidence intervals, point estimates for SSRIs were higher compared to TCAs both in the cohort and nested case–control studies. The magnitude of the risk estimates obtained was in accordance with existing literature,[6–10] with only the effect sizes of SSRIs in Mondriaan being relatively high, but similar to estimates found by Van den Brand *et al.* in a Dutch study using data from the PHARMO RLS.[17]

Strengths of this study are its uniform definition of exposure, outcome and confounders using a common study protocol and data specification that facilitates both reproducibility and consistency. All studies were population-based and made use of routinely collected electronic health care data. However, some limitations remain. First, the impact of confounder adjustment, as well as adjustment for lifestyle factors as recorded by GPs depends on the completeness and quality of this information. Therefore, any influence of unmeasured or inadequately measured lifestyle factors on the effect estimates cannot be completely excluded. Second, there was variation in the coding systems used in the data sources used (Read codes in THIN vs. ICPC-2 codes in Mondriaan and BIFAP for diagnoses), although it seems unlikely that this would explain any substantial differences in results between data sources. If there would have been differences in coding practices between countries this might have affected results, although this effect is difficult to measure and quantify. Third, there was no exact dosage instruction available in Mondriaan and a fixed duration of use of each prescription was used. This assumption could have resulted in misclassification of the timing of antidepressants use, but a sensitivity analysis conducted in the cohort study showed no difference in the risk estimate for the association between antidepressant drugs and hip/femur fracture when using shorter or longer fixed durations. Fourth, assessment of medical diagnoses ever before the index date depends on the duration of history that was available prior to the index date for each patient. In THIN and BIFAP the period of registered data was longer than in Mondriaan, but this seems unlikely to affect risk estimation within each data source. Overall, we found comparable incidences of fracture and a minimal effect of adjustment of confounders on risk estimates besides the effect of adjusting for age and sex in all data sources. The list of the adjusted confounders was extensive, limiting the possibility of a substantial influence of residual confounding on our results. In the cohort analyses, age was the main confounding factor attenuating the crude risk estimates. In the nested case–control design, age was taken into account by matching. Finally, we were unable to distinguish between hip and femur fractures, the proportion of which may differ by age. No linkages to hospital data were performed in any of the data sources to prevent

DOI: 10.1002/pds

discrepancies in the level of outcome ascertainment. An extra case ascertainment step using free text information was performed in BIFAP as the only divergent step in the study protocol. This step excluded 31.5% of initially identified cases with hip/femur fractures based on ICPC-codes, as the event could not be validated in the clinical patient profile. We did not estimate the HRs including these fracture cases or explore the distribution of these over the different exposure states, but we believe exclusion should not be differential among exposures. Despite this additional ascertainment in BIFAP, risk estimates were comparable between BIFAP and THIN. Also, exclusion of traumatic hip/femur cases from the analysis, being a 'known' cause of the event,[18] did not yield in different results.

The higher risk estimate for hip/femur fracture found among current users of antidepressants in Mondriaan may be explained by factors other than study methods alone. The interaction between age and SSRI use observed in Mondriaan warrants further study to investigate variation of risk estimates along exposure and age continuum within larger study populations to disentangle the effect of small sample size on the presence of the effect modifier and stability of the risk estimates. In an earlier study,[2] it was found that the Netherlands has the lowest and most stable prevalence of antidepressant prescribing during the study period compared with Spain and the UK. Moreover, we compared incidence rates of hip/femur fracture during the study period and found in 2008 an incidence of hip/femur fracture per 10 000 person-years of 10.6 in Spain, 7.3 (NPCRD) and 8.9 (AHC) in the Netherlands and 8.7 in the UK.[19] These rates were comparable to rates in other European countries. The duration of antidepressant use was longer in Mondriaan compared with BIFAP and THIN. Sensitivity analyses showed that the higher point estimates for current use obtained in Mondriaan were not because of the difference in defining prescription duration. There are two previously published studies on antidepressant use and fractures conducted in different Dutch databases. These studies (cohort design[20] and case–control design[21]), reported higher risk estimates for SSRI use compared with results from the published studies selected in our study (but lower than found in Mondriaan). Furthermore, a study in the UK CPRD reported odds ratios for fracture outcomes ≥43 days after the first prescription of TCA (OR 1.15, 95% CI: 1.08–1.23) and SSRI (OR 1.32, 95% CI: 1.19–1.48).[22] These risk estimates, despite study design differences, are comparable to our results in THIN. Whether this country-level difference, specifically the higher risk of fracture among antidepressant users in the Netherlands, is because of specific clinical factors deserves further exploration.

Our approach of harmonizing study methods allowed us to minimize methodological differences and explain possible non-methodological factors. In pharmacoepidemiology, there are several consorted efforts focusing on the use of different methods and data sources for improving drug safety systems. A recently published study[23] by the Observational Medical Outcomes Partnership (OMOP)[24] is worth contrasting against our study. The OMOP study examined 53 drug and adverse-event associations in nine different databases applying cohort and self-controlled case series designs. In the OMOP study, authors developed a common data model, which was applied to different data sources and subsequently performed a uniform analysis to estimate risks. Although keeping study designs constant, the heterogeneity (I^2 index) remained substantial. Our approach of extensive harmonization, from definitions of variables to analyses step, minimized heterogeneity because of study method differences and allowed to investigate other possible factors introducing heterogeneity. The common-protocol approach in which data sources are analysed separately instead of a priori pooling of data sources allows the investigation of additional sources of variability that would otherwise have been lost in the pooling process as in the OMOP study.[23]

Lack of information on potential confounders is an important consideration in epidemiological research, as residual confounding can lead to spurious risk estimates. In analyses in BIFAP/THIN, we explored the difference in risk estimates that were obtained by adjusting for lifestyle factors and matching on general practice (THIN). The latter accounts for within-practice differences in registration and prescribing preferences and is an indirect marker for region, accessibility of health care and socioeconomic status, factors that are usually unmeasured. We found no effect of adjusting for lifestyle factors in THIN compared to a full model without these factors. Schneeweiss *et al.* found in a study on SSRI use and hip fractures in Medicare that excluding adjustment for body mass index, smoking, activities of daily living score, cognitive impairment and physical impairment resulted in considerable overestimation of the risk estimate for individual confounders, resulting in a net confounding of +9.6%. However, the association after adjusting for these additional confounders still yielded a statistically significant association.[25] In a similar study assessing the association between NSAIDs and myocardial infarction, the net bias was only 1%, indicating that absence of information on non-measured confounders was unlikely to cause important bias.[26] Also, Groenwold *et al.* concluded that

DOI: 10.1002/pds

the potential for unmeasured confounding when studying the association between influenza vaccination and health outcomes was small for non-randomized intervention studies conducted within extensive and reliable databases.[27] Therefore, lack of information on lifestyle factors does not necessarily lead to biased risk estimates, which is potentially reassuring when using data sources that does not hold such information. Adjustment for commonly available variables in electronic health care databases might act as a proxy for such unmeasured factors, possibly limiting the importance of lifestyle data itself. However, replication studies in other settings and for other exposure-outcome pairs are important to determine generalisability of this finding, as there are several alternative interpretations for the observed lack of an effect of lifestyle variables in this study. First, lifestyle factors might not be confounders for this specific exposure-outcome pair under investigation in our data sources, whereas this could be different in other settings and other exposure-outcome pairs. Second, it could be that lifestyle variables are confounders in our data sources, but are 'controlled for' by design (restriction to users of antidepressants with comparison to past use). Third, it could be that these factors are indeed confounders, but were so poorly measured that adjustment did not make a difference and the result remains biased.

Risk estimates for the GP-matched analyses were not obviously different compared to the analyses without GP-matching. Matching on general practice is common in many pharmacoepidemiological studies, but with its aggregated nature it might not always be adequate to use at the individual patient level. Within the catchment area of a general practice there might be heterogeneity of relevant confounders measured at the individual level, as shown by Movahedi *et al.* in the UK for socioeconomic status by Townsend score.[28] This is in line with findings from Thorogood *et al.*, who also found that in a study on myocardial infarction in young women, general practice controls were more likely to represent the general distribution of the population and not necessarily the level of deprivation of the cases.[29] On the other hand, GP-matching will deal with practice variation in prescribing and recording clinical information that might mitigate such drawbacks.

The conduct of multi-database studies constitutes a trade-off between optimal model selection in each data source and comparability of results between data sources. Within the PROTECT project, we choose to fit the same multivariable models to increase comparability between data sources. The selection strategy of the confounders was based on the association of confounders with hip/femur fracture in the literature, in other words being known risk factors. Inherently this means that variables were included in the model that may not have been a true confounder in the sense that they changed the association between antidepressant use and hip/femur fracture. Such unnecessary adjustment for confounding is theoretically bias-neutral, but affects precision.[30] In Mondriaan, the number of cases and controls was substantially lower compared to BIFAP and THIN, which limited the number of confounder that could be included in the model.

In conclusion, we observed an increased risk of hip/femur fracture in users of antidepressants in a multi-country database study using two study designs. Applying similar pharmacoepidemiological study methods to different populations and data sources resulted in similar risks for TCA use and some variation for SSRI use. Some of these differences may express real (or natural) variance in the exposure-outcome co-occurrences. However, consistently similar methods also enable the identification of relevant effect modifiers.

CONFLICT OF INTEREST

OK, MdG and PS have received unrestricted funding for pharmacoepidemiological research from the Dutch private-public funded Top Institute Pharma.

MM is an employee of Merck.

JL is an employee and stockholder of GlaxoSmithKline.

RS is a full-term Novartis employee and owns Novartis shares

RR is an employee and stockholder of Pfizer, Inc.

KEY POINTS

- Current use of antidepressants was consistently associated with increased hip/femur fracture risk in three European electronic health care data sources using both cohort and nested case–control designs
- Results show that risk of hip/femur fracture was higher in users of SSRIs than those using TCAs.
- Despite harmonization of study design, protocol and data specification, risk estimates for SSRI use were higher in the Dutch Mondriaan database compared to Spanish BIFAP and British THIN databases. Analysis shows that this could partly be explained by an interaction between exposure to SSRI and age in Mondriaan.
- Additional adjustment for GP recorded lifestyle factors, as well as matching on general practice had negligible impact on risk estimates in this study.

DOI: 10.1002/pds

ACKNOWLEDGEMENTS

The research leading to these results was conducted as part of the PROTECT consortium (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium) which is a public-private partnership coordinated by the European Medicines Agency.

FUNDING

The PROTECT project is supported by the Innovative Medicine Initiative Joint Undertaking (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 companies' in kind contribution. In the context of the IMI Joint Undertaking (IMI JU), the Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, also received a direct financial contribution from Pfizer. The University of Alcalá received a direct financial contribution from AstraZeneca. The views expressed are those of the authors only and not of their respective institution or company.

MEMBERS OF WP2 OF PROTECT (FRAMEWORK FOR PHARMACOEPIDEMIOLOGY STUDIES)

Y. Alvarez, G. Candore, J. Durand, J. Slattery (European Medicines Agency); J. Hasford, M. Rottenkolber (Ludwig-Maximilians-Universität-München); S. Schmiedl (Witten University); F. de Abajo Iglesias, M. Gil, C. Huerta Alvarez, B. Oliva, G. Requena, E. Martín, R. González, A. Alvarez, D Montero (Agencia Espanola de Medicamentos y Productos Sanitarios—Universidad de Alcalá); R. Brauer, G. Downey, M. Feudjo-Tepie, M. Schoonen (Amgen NV); S. Johansson (AstraZeneca); J. Robinson, M. Schuerch, I. Tatt (Roche); L.A. Garcia, A. Ruigomez (Fundación Centro Español de Investigación Farmacoepidemiológica); J. Campbell, A. Gallagher, E. Ng, T. Van Staa (Clinical Practice Research Datalink); O. Demol (Genzyme); N. Boudiaf, K. Davis, J. Logie, J. Pimenta, D. Webb (GlaxoSmithKline Research and Development LTD); L. Bensouda-Grimaldi (L.A. Sante Epidemiologie Evaluation Recherche); U. Hesse, P. F. Rønn (Lægemiddelstyrelsen (Danish Medicines Agency); M. Miret (Merck KGaA); J. Fortuny, P. Primatesta, E. Rivero, R. Schlienger (Novartis); A. Bate, N. Gatto, R. Reynolds (Pfizer); E. Ballarin, P. Ferrer, L. Ibañez, J.R. Laporte, M. Sabaté (Fundació Institut Català de Farmacologia); V. Abbing-Karahagopian, S. Ali, D. de Bakker, S. Belitser, A. De Boer, M.L. de Bruin, A.C.G. Egberts, L. van Dijk, H. Gardarsdottir, R.H. Groenwold, M. De Groot, A.W. Hoes, O.H. Klungel, H.G.M.. Leufkens, W. Pestman, K.C.B. Roes, P. Souverein, F. Rutten, J. Uddin, H.A.. Van den Ham, E. Voogd, F. de Vries (Universiteit Utrecht).

REFERENCES

1. Gardarsdottir H, Heerdink ER, van Dijk L, Egberts AC. Indications for antidepressant drug prescribing in general practice in the Netherlands. *J Affect Disord* 2007; **98**(1-2): 109–115.
2. Abbing-Karahagopian V, Huerta C, Souverein PC, *et al.* Antidepressant prescribing in five European countries: application of common definitions to assess the prevalence, clinical observations, and methodological implications. *Eur J Clin Pharmacol* 2014; **70**(7): 849–857.
3. Darowski A, Chambers SA, Chambers DJ. Antidepressants and falls in the elderly. *Drugs Aging* 2009; **26**(5): 381–394.
4. Rizzoli R, Cooper C, Reginster JY, *et al.* Antidepressant medications and osteoporosis. *Bone* 2012; **51**(3): 606–613.
5. Prieto-Alhambra D, Petri H, Goldenberg JS, *et al.* Excess risk of hip fractures attributable to the use of antidepressants in five European countries and the USA. *Osteoporos Int* 2014; **25**(3): 847–855.
6. Eom CS, Lee HK, Ye S, Park SM, Cho KH. Use of selective serotonin reuptake inhibitors and risk of fracture: a systematic review and meta-analysis. *J Bone Miner Res* 2012; **27**(5): 1186–1195.
7. Oderda LH, Young JR, Asche CV, Pepper GA. Psychotropic-related hip fractures: meta-analysis of first-generation and second-generation antidepressant and antipsychotic drugs. *Ann Pharmacother* 2012; **46**(7-8): 917–928.
8. Rabenda V, Nicolet D, Beaudart C, Bruyere O, Reginster JY. Relationship between use of antidepressants and risk of fractures: a meta-analysis. *Osteoporos Int* 2013; **24**(1): 121–137.
9. Wu Q, Bencaz AF, Hentz JG, Crowell MD. Selective serotonin reuptake inhibitor treatment and risk of fractures: a meta-analysis of cohort and case–control studies. *Osteoporos Int* 2012; **23**(1): 365–375.
10. Wu Q, Qu W, Crowell MD, Hentz JG, Frey KA. Tricyclic antidepressant use and risk of fractures: a meta-analysis of cohort and case–control studies. *J Bone Miner Res* 2013; **28**(4): 753–763.
11. Diem SJ, Blackwell TL, Stone KL, *et al.* Use of antidepressant medications and risk of fracture in older women. *Calcif Tissue Int* 2011; **88**(6): 476–484.
12. Liu B, Anderson G, Mittmann N, To T, Axcell T, Shear N. Use of selective serotonin-reuptake inhibitors or tricyclic antidepressants and risk of hip fractures in elderly people. *Lancet* 1998; **351**(9112): 1303–1307.
13. Lewis CE, Ewing SK, Taylor BC, *et al.* Predictors of non-spine fracture in elderly men: the MrOS study. *J Bone Miner Res* 2007; **22**(2): 211–219.
14. Vandenbroucke JP. Balancing benefits and harms in health care: observational data on harm should complement systematic reviews of benefit. *BMJ* 2003; **327**(7417): 750.
15. Abbing-Karahagopian V, Kurz X, de Vries F, *et al.* Bridging differences in outcomes of pharmacoepidemiological studies: design and first results of the PROTECT project. *Curr Clin Pharmacol* 2014; **9**(2): 130–138.
16. 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; **63**(4): 422–427.
17. van den Brand MW, Pouwels S, Samson MM, *et al.* Use of anti-depressants and the risk of fracture of the hip or femur. *Osteoporos Int* 2009; **20**(10): 1705–1713.
18. Rothman KJ, Ray W. Should cases with a 'known' cause of their disease be excluded from study? (commentary). *Pharmacoepidemiol Drug Saf* 2002; **11**(1): 11–14.
19. 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; **94**(6): 580–589.
20. Ziere G, Dieleman JP, van der Cammen TJ, Hofman A, Pols HA, Stricker BH. Selective serotonin reuptake inhibiting antidepressants are associated with an increased risk of nonvertebral fractures. *J Clin Psychopharmacol* 2008; **28**(4): 411–417.
21. Verdel BM, Souverein PC, Egberts TC, van Staa TP, Leufkens HG, de Vries F. Use of antidepressant drugs and risk of osteoporotic and non-osteoporotic fractures. *Bone* 2010; **47**(3): 604–609.
22. Hubbard R, Farrington P, Smith C, Smeeth L, Tattersfield A. Exposure to tricyclic and selective serotonin reuptake inhibitor antidepressants and the risk of hip fracture. *Am J Epidemiol* 2003; **158**(1): 77–84.
23. Madigan D, Ryan PB, Schuemie M, *et al.* Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol* 2013; **178**(4): 645–651.

24. Health FftNIo. Observational medical outcomes partnership [6 April, 2015]. Available from: http://archive-omop.fnih.org/.
25. Schneeweiss S, Wang PS. Association between SSRI use and hip fractures and the effect of residual confounding bias in claims database studies. *J Clin Psychopharmacol* 2004; **24**(6): 632–638.
26. Schneeweiss S, Glynn RJ, Tsai EH, Avorn J, Solomon DH. Adjusting for unmeasured confounders in pharmacoepidemiologic claims data using external information: the example of COX2 inhibitors and myocardial infarction. *Epidemiology* 2005; **16**(1): 17–24.
27. Groenwold RH, Hoes AW, Nichol KL, Hak E. Quantifying the potential role of unmeasured confounders: the example of influenza vaccination. *Int J Epidemiol* 2008; **37**(6): 1422–1429.
28. Movahedi M, Bishop T, Barrett J, Law G. The evaluation of matching in a case–control study of colorectal cancer using general practice lists. *Iranian J Public Health* 2008; **37**(4): 26–31.
29. Thorogood M, Arscott A, Walls P, Dunn NR, Mann RD. Matched controls in a case–control study. Does matching by doctor's list mean matching by relative deprivation? *Int J Soc Res Methodol* 2002; **5**(2): 165–172.
30. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009; **20**(4): 488–495.

SUPPORTING INFORMATION

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

DOI: 10.1002/pds