Received: 18 December 2019 | Revised: 5 March 2020 | Accepted: 29 March 2020

DOI: 10.1002/pds.5006

ORIGINAL REPORT

WILEY

Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria (BIFAP): A data resource for pharmacoepidemiology in Spain

Miguel-Angel Maciá-Martínez | Miguel Gil | Consuelo Huerta | Elisa Martín-Merino | Arturo Álvarez | Verónica Bryant | Dolores Montero | on behalf of the BIFAP Team

Pharmacoepidemiology and Pharmacovigilance Division, Medicines for Human Use Department, Spanish Agency for Medicines and Medical Devices (AEMPS), Madrid, Spain

Correspondence
Miguel-Angel Maciá-Martínez, Spanish Agency for Medicines and Medical Devices (AEMPS). Calle Campezo 1, Edificio 8, E-28022 Madrid, Spain.
Email: mmacia@aemps.es

Abstract

Purpose: *Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria* (BIFAP) is a population based database administered by the AEMPS (Spanish Agency for Medicines) of longitudinal electronic medical records (EMR) of patients attended in primary care. Its main purpose is to serve as source of information for independent studies on drug safety and support of medicines regulation activities. This article aim is to describe the characteristics of BIFAP, how to access the database and a summary of its potential for research.

Methods: Health problems are registered by primary care physicians as episodes of care and include socio-demographic data, results of diagnostic procedures, lifestyle data, general data, and interventions. A proportion of data on hospitalizations and specialist care are currently available through linkage with other data sources. EMRs of the Spanish healthcare system are provided by the regional administrations. Specific data extraction and standardization processes are performed.

Results: BIFAP includes data from 12 million patients starting in 2001 and updated annually. Validation of drug and diagnosis definitions has been ascertained. Participation in international collaborative projects and a number of articles in peer reviewed journals reflect its contribution to the knowledge of the risks associated with medicines and drug utilization patterns.

Conclusions: BIFAP is a useful tool for generating scientific evidence on medicines related issues, helping regulatory decision making in Europe. The main strengths of BIFAP are related to large sample size, population-based, longitudinal nature and annual update of data. BIFAP shares common challenges with similar data sources including accurate and efficient identification of health outcomes and of treatment exposure.

1 | INTRODUCTION

BIFAP (*Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria*) is a longitudinal population-based database of electronic medical records (EMRs) from patients attended in primary care facilities of the SNS (*Sistema Nacional de Salud*), the Spanish National Health System, and located in one of the participating regions throughout Spain.

BIFAP is a nonprofit program administered and fully financed by the Spanish Agency for Medicines and Medical Devices (AEMPS) that started in 2001 when primary care EMRs where fully implemented throughout Spain.[1] Ten Autonomous Communities are committed to provide data through bilateral agreements (http://bifap.aemps.es). The main use of BIFAP is to serve as real world data source for research projects with the aim of the studying adverse and beneficial effects and utilization patterns of medicines in the general population.

This article describes BIFAP characteristics, organization and governance, as well as a summary of its potential as a relevant data source for pharmacoepidemiology research with public health impact.

2 | DATABASE CONTENT

2.1 | Database in the context of the Spanish National Health System

The SNS is configured as a coordinated set of health services from the Central Government Administration and the Autonomous Communities (regions) that integrates all healthcare functions and benefits for which public authorities are legally responsible. Each of the 17 regions has its own Health Service, which is the administrative and management body responsible for health centers, services and facilities in its region.[2]

Regional Health Services integrated in the SNS provide universal access to healthcare to the entire Spanish population regardless their income. While regional variation on covered services and practice guidance are allowed, a portfolio of covered services and health benefits for the Spanish population are adopted for the whole SNS. This includes decisions on medicines financing, price and out-of-pocket payments. Regional Health Services are health service providers for all levels of care in their territories. There is a 16% of the Spanish population that contracts with private-for-profit health insurance companies, with an important regional variation.[3] In general private health schemes do not cover outpatient prescription medicines and these patients usually are also registered and attend SNS for healthcare and reimbursement.

According to the organization of the primary care level in the SNS, each primary care team or practice is basically formed by a primary care physician (PCP, general practitioner or pediatrician) and a nurse with 1000 to 3000 patients assigned. These teams are organized in primary care centers and grouped in areas covering a defined population. Pediatric patients are transferred to a general practitioner at the age of 14.

PCPs at SNS primary care centers have a central role. The primary care system is the gate to the SNS. PCPs act as gatekeepers to the system coordinating and registering all the clinical encounters with the patients. They drive the consultations to other levels of the health system and ensure the continuity of care. Also, patients are usually appointed to visit their primary care team following specialist consultation, hospitalization or emergency service visits, and they usually register the outcome of these other contacts with the healthcare system. They prescribe all medicines for the health problems they directly attend as well as subsequent prescriptions for treatments started in other levels of care. The result is that practically the entire population (98.9%) is registered with a PCP.[4]

KEY POINTS

- BIFAP stands for "Base de datos para la investigacion Farmacoepidemiologica en Atencion Primaria", a population based database of longitudinal electronic medical records of patients attended in primary care in Spain.
- BIFAP is a non-profit program of the AEMPS (Spanish Agency for Medicines and Medical Devices). Regional health administrations provide data via bilateral agreements with the AEMPS.
- BIFAP includes data from 12 million patients from 7 regions of Spain up to the end of 2018. Available data start in 2001 and are updated annually. Quality is verified through established validation processes.
- BIFAP governance framework restricts access to research projects led by investigators without conflict of interest in medicinal products.
- A number of published articles reflect BIFAP contribution to the knowledge of the risks associated with medicines, drug utilization patterns or the incidence/prevalence of health problems, supporting medicines regulation.

What's already known about this topic?

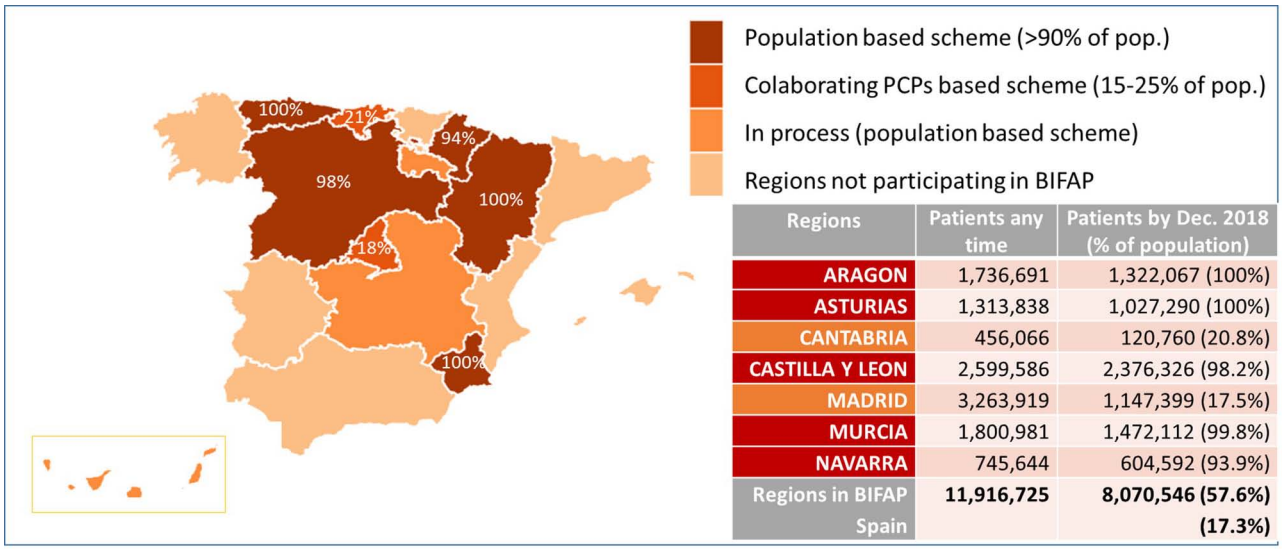
- Pharmacoepidemiological studies are increasingly performed to provide real-world evidence on drugs and multiple data sources may be used in the regulatory decision making.
- Electronic medical records generated in primary care are a powerful piece of longitudinal and comprehensive information.
- In order to be useful for medicines regulation quality of data and its correct analysis and interpretation are essential elements.

What does this study add?

- A review of BIFAP, a data resource managed by the Spanish medicines regulatory agency (AEMPS), is presented: how the data are generated, their coverage, governance, and actual examples of use.
- With longitudinal electronic medical records of a total of 12 million patients attended in primary care in Spain, BIFAP is a major resource for independent studies on drug safety in Europe.
- BIFAP has shown its potential as a relevant data source for supporting medicines regulatory decision making.

2.2 | Population covered

Following the scheme and experience of the British CPRD (Clinical Practice Research Datalink),[5] EMRs from patients attended in primary care facilities of the Spanish SNS are the fundamental source of data for BIFAP.



Regions	Patients any time	Patients by Dec. 2018 (% of population)
ARAGON	1,736,691	1,322,067 (100%)
ASTURIAS	1,313,838	1,027,290 (100%)
CANTABRIA	456,066	120,760 (20.8%)
CASTILLA Y LEON	2,599,586	2,376,326 (98.2%)
MADRID	3,263,919	1,147,399 (17.5%)
MURCIA	1,800,981	1,472,112 (99.8%)
NAVARRA	745,644	604,592 (93.9%)
Regions in BIFAP	**11,916,725**	**8,070,546 (57.6%)**
Spain		**(17.3%)**

FIGURE 1 Population coverage of BIFAP (up to 2018) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Data available in BIFAP

Type of patient data	Data available
Administrative data	Dates patients were registered at or have left a PCP's practice, including date of death.
Socio-demographic data	Age, Gender
Lifestyle data / Additional health data	Smoking, Alcohol intake, Weight, Height, Blood pressure. Recorded in structured format
Health problems in primary care	Diseases/symptoms leading to patient consultation (episodes of care) Diagnosis.
Results of diagnostic procedures	Laboratory, Imaging Recorded either in structured format or as free text.
Diagnosis / Referrals to other levels of care	Referrals to specialists and to emergency services. Essential data derived from referral (new diagnoses, interventions, results of specialized tests, etc.) are recorded either in structured format or as free text.
Hospital discharge diagnosis	Present coverage: 50% (see Figure S6, Supplementary Material)
Interventions: Medicines	Medicines prescriptions Dispensations: coverage in 2018:75%(see Figure S5, Supplementary Material) Vaccination records.
Interventions: Other	Recorded either in structured format or as free text.

Up to the end of 2018, BIFAP includes data from 7566 PCPs (6419 general practitioners and 1147 pediatricians). This yields a total of 12 million (2.3 million pediatric) patients for studies. Out of them, 8 million contains up-to-date information the year 2018 ("active" patients), representing 57.6% of the participating regions (14 million) and 17% of the total Spanish population (46.6 million).

Geographical distribution of patient data contributing to BIFAP within Spain is shown in Figure 1. In all but two participating regions, the EMRs from the whole population attended in primary care are included in BIFAP. In two regions the original scheme when BIFAP was launched in the 2000s is still followed and an agreement to extract the EMRs is needed from volunteer "collaborating" PCPs; coverage in these two regions is 15%-20% of the population.

Valid study period in BIFAP starts in year 2001. BIFAP population reflects the population receiving healthcare in the Spanish SNS. The age of patients in BIFAP is somewhat older than the general population, explained by a more frequent use of the primary care system in people over 50 years old[4] (Figure S1, Supplementary material).

BIFAP yearly uptake and withdrawal data from 2001 to 2018 depend on dates of introduction of EMRs in primary care practices from participating regions fulfilling current quality criteria. In terms of length of follow-up 102 million person-years are included. The mean patient-time of follow-up is 8.6 years. Additional details on characteristics of BIFAP population are presented in the appendix as supplementary material.

2.3 | Data available

Most, but not all participating regions share the same EMRs software. Data structure and diagnostic coding systems also differ. Consequently, there is heterogeneity in how the information is provided to BIFAP and information is harmonized into a unified data structure to be used for research purposes.

Information collected by PCPs includes administrative data, socio-demographic data, lifestyle, and other general data, health problems, results of diagnostic procedures, and interventions. Table 1 summarizes availability of patient data in BIFAP.

All information on prescriptions of medicines by the PCP is incorporated. This Information is linked by the PCP to a health problem (episode of care) that generally informs on the clinical indication of the medicinal product. Data include the number of packages, intended duration, dosage and strength and is organized following Anatomical Therapeutic Chemical (ATC) codes. Product names can be traced to national codes for medicinal products and linked to Centro de Información online de Medicamentos de la AEMPS (CIMA), the national database of medicinal products of the AEMPS (https://cima.aemps.es).

Additionally, information on dispensation of medicines at pharmacies is being extracted from the e-prescription system which is widely implemented in Spain,[6] representing 75% of all prescriptions in BIFAP issued in 2018. The e-prescription allows electronic transmission of prescription data to retail pharmacies for dispensing the medication to the patient. In general, dispensing information is chosen for research if available. However, if e-prescription is not available paper based prescription information is used.

According to the structure of BIFAP the PCP couples the relevant registered information to a health problem as an episode of care following International Classification of Primary Care (ICPC)[7] or International Classification of Diseases (ICD-9)[8] terminologies. A variable proportion of information is registered in "clinical notes" in free text fields in the EMR.

Events or disorders giving rise to admission are also usually recorded by PCPs in the primary care EMRs. Additionally, information on hospital discharge diagnoses coded in ICD terminology is being linked to patients included in BIFAP. To date, it is only available for a subset of periods and regions participating in BIFAP, representing around half of the BIFAP population in the five last years.

Additional details on data available in BIFAP are presented in the appendix as supplementary material.

3 | STANDARDIZATION AND QUALITY REVIEW

Annually, the database is updated following a data gathering process that starts with extraction of data and pseudonymization of the EMRs in participating regions. Once transferred to BIFAP, this is followed by a data cleaning excluding any patient EMR not fulfilling minimum quality requirements is excluded. Consistency and quality control checks are then performed. Additional details are presented in the appendix as supplementary material.

Standardization of diagnostic data transferred to BIFAP takes place after each annual update of the database. Two coding systems with different levels of granularity coexist in BIFAP: ICPC and ICD-9. The ICPC is the coding system for most of participant regions and is less granular than ICD-9 (1300 vs 13 000 codes, respectively).

To help PCPs to identify and register the episode of care of interest, the EMR software in regional healthcare systems contains an internal thesaurus, where a list of descriptors of diseases, signs, or symptoms is linked to the different dictionary codes. Often, these descriptors provide more detailed information than the corresponding code. Likewise, only for ICPC-based EMR software, new descriptors can be included at the local level, and the PCPs can also modify or add information to the selected descriptor. This results in a huge number of different descriptors in the BIFAP database (8.5 million).

To standardize this BIFAP has developed its own research dictionary (ICPC-BIFAP) by adding, for the most frequently used descriptors, a fourth digit to the original three-digit ICPC code, increasing its granularity. In the last version (2017) the ICPC-BIFAP dictionary included 5110 indexed terms. This allows rising up to 90% the coverage of coded diagnosis. Concerning the regions coding with ICD-9 and given the high granularity of the medical terms dictionary information received in BIFAP is already normalized and no further actions are performed.

To ease the management of event definitions in pharmacoepidemiological studies specific case-finding algorithms (CFA) have been developed for an increasing number of clinical/diagnostic events frequently used in studies as outcome variables or covariates. CFAs include related codes in both ICD-9 and ICPC-BIFAP dictionaries and might also include laboratory test results, additional health data, or text mining strategies. Text mining strategies are usually included to build CFAs in order to capture all cases of a specific event.

The quality review of the CFAs is addressed both by reviewing a sample of cases (internal validity/precision evaluation) and by comparing epidemiological measures (ie, incidence, prevalence) with the external available evidence. Investigators can use or adapt CFAs to their study specific diagnostic definitions and can additionally access free text in the EMRs for confirmation and validation of cases according to BIFAP data access governance.[9]

Additionally, a network of more than 1500 collaborating PCPs from the participating regions are involved in crowdsourcing tasks aimed to standardize the clinical events in BIFAP and are committed to follow quality procedures of registry in the EMRs from their own practice. Also, all investigators have to commit to share any definitions and algorithms created for their own projects to other future investigators' projects.

For the standardization of medicines data, a drug dictionary is maintained on the basis of ATC[10] codes up to active substance level. Duration of prescription and posology are either directly recorded by the physician and/or calculated from the information associated to the prescriptions in the database (dosage and strength). Medicines

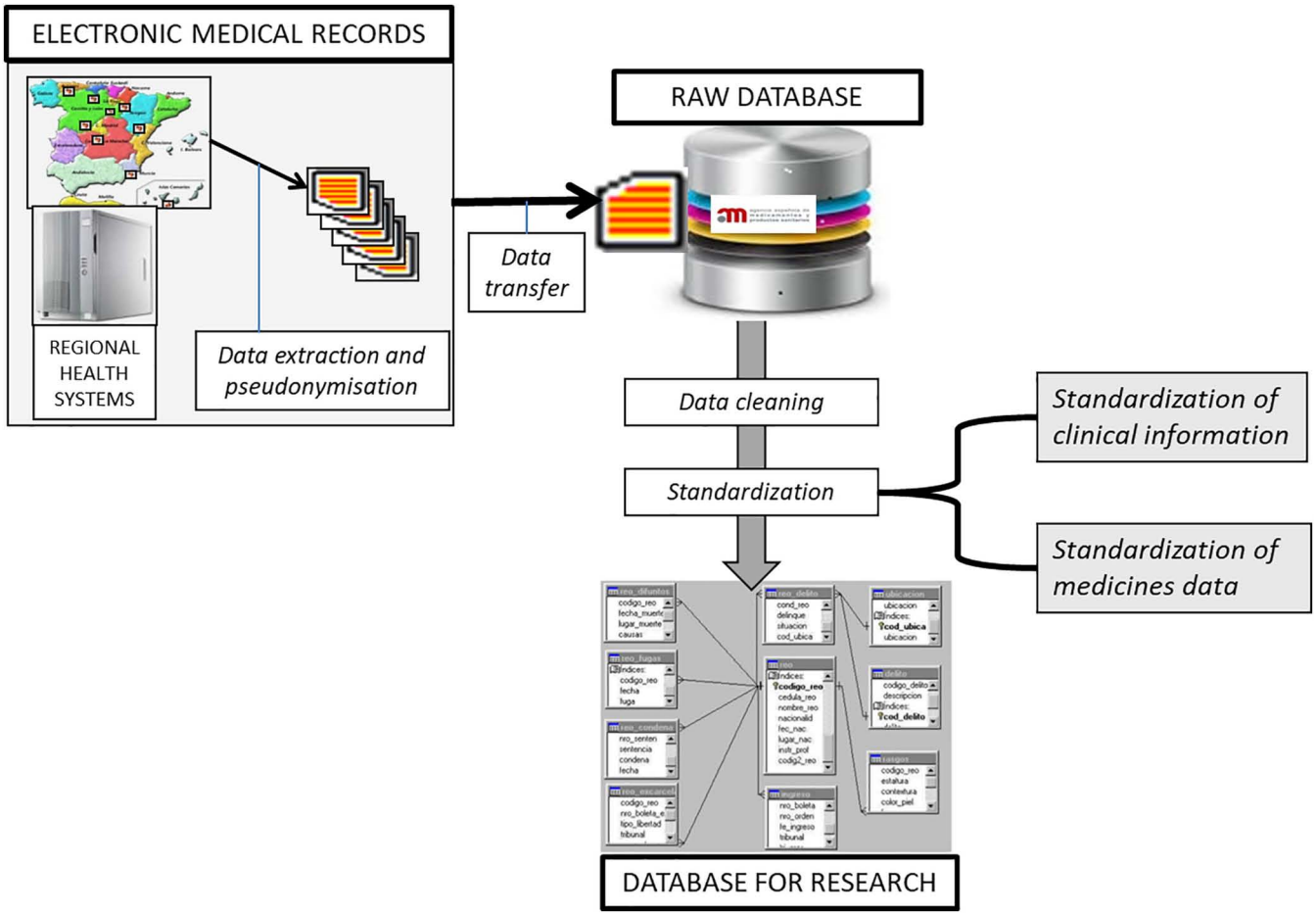


FIGURE 2 Annual process of creation and update of the database for research [Colour figure can be viewed at wileyonlinelibrary.com]

data are estimated using an algorithm with different duration metrics for calculations and quality control checks.

A summary of the data extraction and standardization process is showed in Figure 2.

Validation of drug and diagnostic case definitions in BIFAP has been ascertained through several studies and presented in the next section.

4 | STUDIES PERFORMED IN BIFAP

A number of articles published in peer reviewed journals reflect the BIFAP contribution to the knowledge of the risks associated with medicines, either in the evaluation of potential safety signals or in risk quantification. In addition, drug utilization studies performed to describe the patterns and determinants of the use of medicines including evaluation of the impact of risk minimization activities have been performed. Table 2 summarizes the different areas of research to date.

In addition, incidence or prevalence of health problems such as community acquired pneumonia,[11–14] dementia,[15] hip/femur fractures,[16] inflammatory bowel disease,[17] Guillain-Barre syndrome, colorectal cancer,[18] and acute liver injury[19] have been obtained from BIFAP including internal validation of clinical events by manual review of anonymized free text. Manual review of free text has also been performed as part of the case definition procedure within drug-event association studies (see Table 2). Furthermore, comparisons of incidence/prevalence of clinical events and of exposure to medicines in BIFAP with external figures has been evaluated in benchmarking processes through international multi-database projects[16,20,21] and with population-based registries.[18,22]

Regarding medicines exposure data non-standardized information registered in the EMRs on vaccinations has been validated by manual review of patient profiles.[23] In addition, an algorithm to identify pregnancies in BIFAP has been developed to allow studying the use and effects of drugs during pregnancy.[24]

BIFAP is fully integrated in the AEMPS roles and activities as a useful tool for generating scientific evidence on medicines related issues, helping regulatory decision making. This includes participation in international collaborative projects with other data sources worldwide in European Union financed programs.[20,25,26] Studies in BIFAP are performed by the in-house team and also by independent investigators from public organizations following specific data access procedures. An up-to-date list of ongoing and finalized studies and published articles with BIFAP is available in http://bifap.aemps.es.

TABLE 2 Areas (drugs and diagnosis) of research in BIFAP

Drug class (ATC)	Drug-event[a] association studies	Drug utilization studies
Drugs used in diabetes (A10B)	Cancer (bladder, pancreas, etc)[20] Cardiovascular risk[20] Pancreatitis[20] Pneumonia[37]	Use[20 b]
Calcium supplements (A12A)	Ischemic stroke[38]	
Antithrombotic agents (B01A)	Bleeding, stroke[21b]	Trends[39 b]
Low-ceiling diuretics, thiazides (C03A)	Skin cancer[c]	
Lipid modifying agents (C10)		Indication[40,41]
Topical calcineurin inhibitors (D11A)		Impact of RMM[42]
Estrogens (G03C, G03F)		Trends[43]
Antiandrogens (G03H)	Meningioma[44]	
Antibacterials (J01)	Acute liver injury[45,46 b]	Use[47 b] Indication (urinary tract infection)[c]
Vaccines (J07)	Pertussis vaccine adverse events[48]	Coverage[22,49,50]
Anti-inflammatory drugs (M01)	Myocardial infarction[51] Colorectal cancer[52] Gastrointestinal bleeding[53] Ischemic stroke[54]	Indication[55]
Antigout drugs (M04)	Myocardial infarction[56,57]	
Anti-osteoporotic drugs (M05B)	Fractures[58,59] Venous thromboembolism[60 b]	Trends[61] Trends[62 b] Use[63]
Opioids (N02A)		Impact of RMM (codeine)[64 b] Impact of RMM (fentanyl)[c]
Anti-epileptics (N03A)	Herpes virus infection[65]	Use[66 b]
Antipsychotics (N05A)		Indication (dementia)[c]
Anxiolytics, hypnotics and sedatives (N05B, N05C)	Fractures[67,68 b] SCAR[69]	Trends[70 b] Adherence[71]
Antidepressants (N06A)	Fractures[72 b]	Use[71 b] Use[73]
Psychostimulants, agents used for ADHD and nootropics (N06B)	Valvular heart disease[74]	Trends[75 b]
Anti-dementia drugs (N06D)		Trends[76]
Adrenergics, inhalants (R03A)	COPD exacerbations[77]	Trends[78 b]
Overall	SCAR[79]	Use (polypharmacy)[80]

Abbreviations: COPD, chronic obstructive pulmonary disease; RMM, risk minimization measures; SCAR, Severe cutaneous adverse reactions.
[a]Include manual validation of events.
[b]Multi-database studies.
[c]On-going projects.

5 | BIFAP ORGANIZATION AND DATA ACCESS GOVERNANCE

A BIFAP Advisory Committee with members of the AEMPS and the participating regions oversees the correct functioning of the program and is consulted on governance policies and improvements of the program.

BIFAP relies on a Scientific Review Board to provide scientific and feasibility advice on protocols to conduct research with BIFAP data and to ensure that all the studies with BIFAP have a focus on public health.

The use of BIFAP follows a data governance framework[9] approved by its Advisory Committee which is consistent with the applicable principles of the General Data Protection Regulation.[27] Before EMR data are transferred to the AEMPS, there is a process of pseudonymization of personal identifiers followed by additional procedures[28] which ensure the anonymization of the data available for investigators. In addition, appropriate technical and organizational measures are in place to guarantee personal data protection.

BIFAP governance framework agreed with data providers (regional authorities) restricts access to research projects led by investigators without any vested interest in the medicinal products being

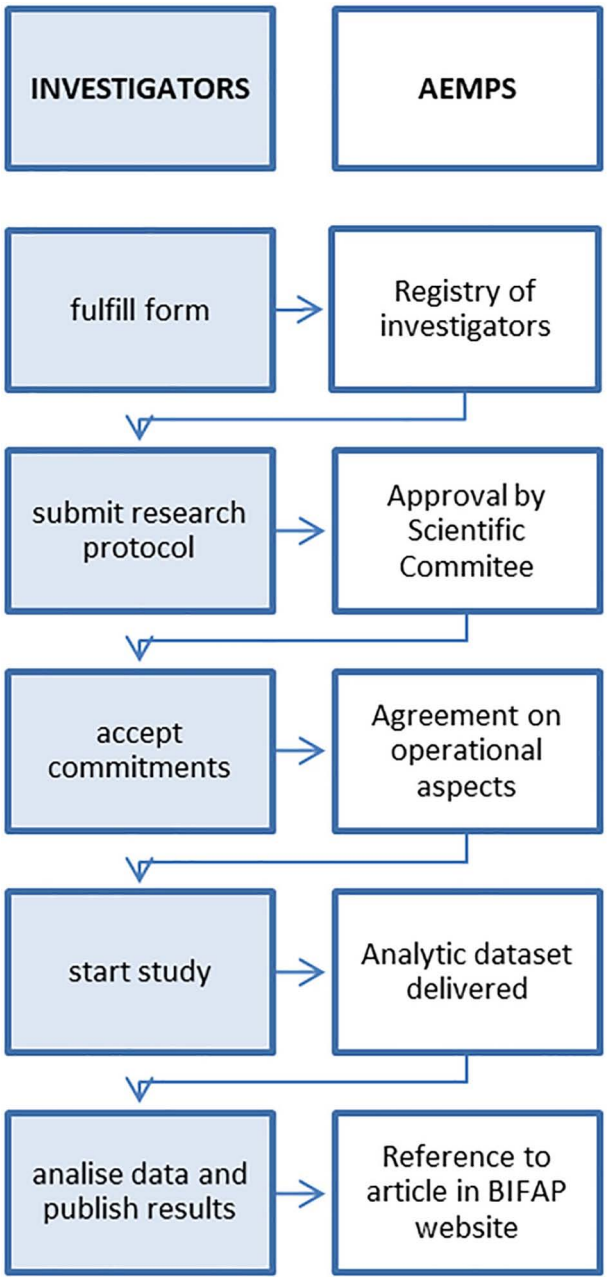


FIGURE 3 Steps to perform a study in BIFAP: role of investigators and AEMPS [Colour figure can be viewed at wileyonlinelibrary.com]

studied. Access to data is free of charge and industry funded studies are not accepted. Specifically, access is restricted to in-house investigators, healthcare professionals from the SNS and academics in public universities and research centers. All of them have to declare any conflict of interest along with the submission of the study protocol to the Scientific Review Board. International projects are accepted provided that a principal investigator from Spain in the consortium is appointed. The steps needed to perform a study in BIFAP are summarized in Figure 3. In the process, registered investigators receive training and additional support on a case by case basis before the delivery of the analytical dataset.

6 | STRENGTHS AND WEAKNESSES

The main strengths of BIFAP are related to its large sample size (12 million patients) and population-based nature. The whole population is available for most of the participating regions with plans to expand this model to other regions. The longitudinal nature of the database and the availability of data for a number of years are also relevant.

BIFAP includes life-style data (see Table 1) recorded by PCPs as well as their clinical notes registered during routine clinical care which can be analyzed and used for validation of study variables. This is an advantage over claims databases where concerns about validity of administrative data should generally be addressed using external data.[29]

Validity and quality of the diagnostic information has been assessed in published studies and the availability of complete and detailed information of characteristics of prescriptions of medicines allows for drug utilization studies and for analytical studies where drug use is the exposure of interest. Moreover, the database is annually updated.

Concerning the limitations of BIFAP data, it is important to note that data are collected primarily for clinical and routine use rather than for research purposes.

No structured information is available on socioeconomic status and ethnicity of patients that could be needed as co-factor variables. In relation to outcomes, only a proportion of data on hospitalizations and specialist care are currently available through linkage with other specific data sources. Results of referrals to hospitals or to specialists are available as entered by the PCPs. Cause of death is not yet available systematically although plans for linking information to the Spanish national statistics system are under way.

Laboratory or imaging test results are not always registered systematically in a structured way, being more likely to be entered if abnormal.

With regard to information on medicines, despite prescribed medicines are linked to a medical event record, this linkage may not reflect the prescription's indication accurately. Drugs prescribed by hospital doctors, other specialists or in the private health care setting are not systematically registered in BIFAP, neither are medications dispensed without a prescription. Dispensation data are available since 2011 enabling evaluation of primary adherence.

BIFAP shares common development areas and challenges with other databases used for epidemiological research including the accurate identification of health outcomes of interest,[30] the precise measurement of the treatment exposure and methodological issues related to study designs, missing imputation, outliers treatment etc. BIFAP is also dealing with the evolving role of healthcare databases in pharmacovigilance.[31] Among them, the implementation of multi-database studies using Common Data Models (CDM)[32] is one of the challenges of the participation in international collaborative projects. Among others, this has been explored in the context of monitoring of vaccination benefits and risks with BIFAP participation in ADVANCE,[25] a scenario where the need to expedite multi-database studies is particularly relevant.[33] Also, BIFAP is one of the EMR databases contributing to define strategies to implement a CDM at European level.[34]

BIFAP is exploring the possibilities of linkage of the database with other health care data sources, particularly with hospital databases and mortality records.

Another challenge is the harmonization processes of diagnostic information from different data structures in the regions participating in BIFAP. Meta-ontology UMLS "Unified Medical Language System"[35] is being explored to unify terminologies. Improvement of the processing of information stored in natural language from free text in electronic patient records is also being addressed.[36]

In conclusion, BIFAP has become a relevant resource for independent pharmacoepidemiology research. The availability of longitudinal and comprehensive clinical records from a relevant proportion of the Spanish population has enabled a significant number of studies including multi-database projects. Continued efforts are necessary to deal with current limitations and upcoming challenges.

ETHICS STATEMENT

The authors state that no ethical approval was needed.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the excellent collaboration of the primary care practitioners and pediatricians, and also the support of the regional authorities participating in the database (Aragón, Asturias, Canarias, Cantabria, Castilla y León, Castilla-La Mancha, La Rioja, Comunidad de Madrid, Murcia y Navarra).

Funding information: The "Base de datos para la investigación Farmacoepidemiológica en Atención Primaria" (BIFAP) is fully financed by the Spanish Agency on Medicines and Medical Devices (AEMPS).

Disclaimer

This document expresses the opinion of the authors of the article, and may not be understood or quoted as being made on behalf of or reflecting the position of the AEMPS.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Miguel-Angel Maciá-Martínez https://orcid.org/0000-0002-1471-4376

Miguel Gil https://orcid.org/0000-0003-3827-9564

Elisa Martín-Merino https://orcid.org/0000-0002-3576-8605

REFERENCES

1. Salvador Rosa A, Moreno Pérez JC, Sonego D, García Rodríguez LA, de Abajo Iglesias FJ. The BIFAP project: database for pharmacoepidemiological research in primary care. *Aten Primaria*. 2002;30(10):655-661. https://doi.org/10.1016/s0212-6567(02)79129-4.
2. Bernal-Delgado E, Garcia-Armesto S, Oliva J, et al. Spain: health system review. *Health Syst Transit*. 2018;20(2):1-179.
3. Encuesta Nacional de Salud. MSSSI/INE. Indicadores clave [Internet]. Sistema Nacional de Salud; 2018 Jun. http://inclasns.msssi.es/main.html. Accessed December 12, 2019.
4. Key Indicators. Spanish National Health System. http://inclasns.msssi.es/?lang=EN. Accessed December 12, 2019.
5. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836. https://doi.org/10.1093/ije/dyv098.
6. OECD/European Union. Health at a Glance: Europe 2018: State of Health in the EU Cycle, OECD Publishing, Paris/European Union, Brussels; 2018. https://doi.org/10.1787/health_glance_eur-2018-en
7. ICPC-2. *International Classification of Primary Care (Oxford Medical Publications) World Organization of Family Doctors (WONCA) International Classification Committee*. Oxford, UK: Oxford University Press; 1998. ISBN 13: 9780192628022.
8. ICD-ICD-9-CM-International Classification of Diseases, Ninth Revision, Clinical Modification. https://www.cdc.gov/nchs/icd/icd9cm.htm. Published March 1, 2019. Accessed December 12, 2019.
9. Gobernanza del Acceso a los Datos de BIFAP. Agencia Española de Medicamentos y Productos Sanitarios. http://bifap.aemps.es/gobernanza.php Accessed February 27, 2020.
10. WHO Collaborating Centre for Drug Statistics Methodology. *Guidelines for ATC Classification and DDD Assignment 2020*. Oslo, Norway: Norwegian Institute of Public Health; 2019.
11. Saiz LC, Garjón J, Gorricho J, Erviti J, Gil-García MJ, Martín-Merino E. Validation and incidence of community-acquired pneumonia in patients with type 2 diabetes in the BIFAP database. *Epidemiol Infect*. 2017;145(14):3056-3064. https://doi.org/10.1017/S0950268817001868.
12. Chacón García A, Ruigómez A, García Rodríguez LA. Incidence rate of community acquired pneumonia in a population cohort registered in BIFAP. *Aten Primaria*. 2010;42(11):543-549. https://doi.org/10.1016/j.aprim.2010.05.004.
13. Rivero-Calle I, Pardo-Seco J, Aldaz P, et al. Incidence and risk factor prevalence of community-acquired pneumonia in adults in primary care in Spain (NEUMO-ES-RISK project). *BMC Infect Dis*. 2016;16(1):645. https://doi.org/10.1186/s12879-016-1974-4.
14. Rivero-Calle I, Cebey-López M, Pardo-Seco J, et al. Lifestyle and comorbid conditions as risk factors for community-acquired pneumonia in outpatient adults (NEUMO-ES-RISK project). *BMJ Open Respir Res*. 2019;6(1):e000359. https://doi.org/10.1136/bmjresp-2018-000359.
15. de Hoyos-Alonso MC, Bonis J, Tapias-Merino E, Castell MV, Otero A. Estimated prevalence of dementia based on analysis of drug databases in the Region of Madrid (Spain). *Neurologia*. 2016;31(1):1-8. https://doi.org/10.1016/j.nrl.2014.08.008.
16. 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. https://doi.org/10.1007/s00223-014-9850-y.
17. Martin-Merino E, Molina Infante J, Martin-Perez M, Montero-Corominas D. The recording and incidence of inflammatory bowel disease in Girls' primary care medical Records in Spain. *Pharmacoepidemiology and Drug Safety*. *Vol 28*. Hoboken, NJ: Willey; 2019: 162.
18. Gil M, Rodríguez-Miguel A, Montoya-Catalá H, et al. Validation study of colorectal cancer diagnosis in the Spanish primary care database, BIFAP. *Pharmacoepidemiol Drug Saf*. 2019;28(2):209-216. https://doi.org/10.1002/pds.4686.
19. Ruigómez A, Brauer R, Rodríguez LAG, et al. Ascertainment of acute liver injury in two European primary care databases. *Eur J Clin Pharmacol*. 2014;70(10):1227-1235. https://doi.org/10.1007/s00228-014-1721-y.
20. Community Research and Development Information Service (CORDIS), European Commission. Final Report Summary—SAFEGUARD (Safety Evaluation of Adverse Reactions in Diabetes). Final Publishable Summary Report. Erasmus Universitair Medisch Centrum Rotterdam, Netherlands. 2015. https://cordis.europa.eu/docs/results/282/282521/final1-safeguard_final-publishable-summary-report.pdf Accessed December 12, 2019.
21. van den Ham H, Souverein P, Klungel O, et al. Risk of major bleeding associated with the use of individual direct oral anticoagulants compared to vitamin K antagonists in patients with non-valvular atrial fibrillation: a meta-analysis of results from multiple population-based cohort studies using a common protocol in Europe and Canada. *Pharmacoepidemiology and Drug Safety*. Vol 28. Hoboken, NJ: Willey; 2019:410-411.

22. Braeye T, Bauchau V, Sturkenboom M, et al. Estimation of vaccination coverage from electronic healthcare records; methods performance evaluation—a contribution of the ADVANCE-project. *PLoS One*. 2019; 14(9):e0222296. https://doi.org/10.1371/journal.pone.0222296.
23. Martín-Merino E, Llorente-García A, Montero-Corominas D, Huerta C. The recording of human papillomavirus (HPV) vaccination in BIFAP primary care database: a validation study. *Pharmacoepidemiol Drug Saf*. 2019;28(2):201-208. https://doi.org/10.1002/pds.4674.
24. Sanchez Ortiz S, Llorente García A, Astasio P, Huerta C, Cea SL. An algorithm to identify pregnancies in BIFAP primary care database in Spain: results from a cohort of 155 419 pregnancies. *Pharmacoepidemiol Drug Saf*. 2019;29:57-68. https://doi.org/10.1002/pds.4910.
25. ADVANCE – Accelerated development of vaccine benefit-risk collaboration in Europe. http://www.advance-vaccines.eu.
26. PROTECT (The Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium). http://www.imi-protect.eu.
27. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
28. Article 29 Data Protection Working Party 0829/14/EN WP216 Opinion 05/2014 on Anonymisation Techniques Adopted on April 10, 2014. http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.
29. Rawson NSB, D'Arcy C. Healthcare databases for drug safety research: data validity assessment remains crucial. *Drug Saf*. 2018;41 (9):829-833. https://doi.org/10.1007/s40264-018-0673-z.
30. Lanes S, Brown JS, Haynes K, Pollack MF, Walker AM. Identifying health outcomes in healthcare databases. *Pharmacoepidemiol Drug Saf*. 2015;24(10):1009-1016. https://doi.org/10.1002/pds.3856.
31. Trifirò G, Sultana J, Bate A. From big data to smart data for pharmacovigilance: the role of healthcare databases and other emerging sources. *Drug Saf*. 2018;41(2):143-149. https://doi.org/10.1007/s40264-017-0592-4.
32. Trifirò G, Coloma PM, Rijnbeek PR, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? *J Intern Med*. 2014;275(6):551-561. https://doi.org/10.1111/joim.12159.
33. Sturkenboom M, Braeye T, van der Aa L, et al. ADVANCE database characterisation and fit for purpose assessment for multi-country studies on the coverage, benefits and risks of pertussis vaccinations. *Vaccine*. Available online 12 February 2020:S0264410X20301535. https://doi.org/10.1016/j.vaccine.2020.01.100.
34. Pacurariu A, Plueschke K, McGettigan P, et al. Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. *BMJ Open*. 2018;8(9):e023090. https://doi.org/10.1136/bmjopen-2018-023090.
35. Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Res*. 2004;32(Database issue):D267-D270. https://doi.org/10.1093/nar/gkh061.
36. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. *J Am Med Inform Assoc*. 2011;18(5):544-551. https://doi.org/10.1136/amiajnl-2011-000464.
37. Gorricho J, Garjón J, Alonso A, et al. Use of oral antidiabetic agents and risk of community-acquired pneumonia: a nested case-control study. *Br J Clin Pharmacol*. 2017;83(9):2034-2044. https://doi.org/10.1111/bcp.13288.
38. de Abajo FJ, Rodríguez-Martín S, Rodríguez-Miguel A, Gil MJ. Risk of ischemic stroke associated with calcium supplements with or without vitamin D: a nested case-control study. *J Am Heart Assoc*. 2017;6(5): pii:e005795. https://doi.org/10.1161/JAHA.117.005795.
39. Ibáñez L, Sabaté M, Vidal X, et al. Incidence of direct oral anticoagulant use in patients with nonvalvular atrial fibrillation and characteristics of users in 6 European countries (2008-2015): a cross-national drug utilization study. *Br J Clin Pharmacol*. 2019;85(11):2524-2539. https://doi.org/10.1111/bcp.14071.
40. Macías Saint-Gerons D, de la Fuente HC, Montero Corominas D, Gil MJ, de Andrés-Trelles F, Catalá-López F. Standard and intensive lipid-lowering therapy with statins for the primary prevention of vascular diseases: a population-based study. *Eur J Clin Pharmacol*. 2014; 70(1):99-108. https://doi.org/10.1007/s00228-013-1586-5.
41. Macías Saint-Gerons D, de la Fuente HC, González Bermejo D, et al. Clinical-epidemiological profile of patients initiating intensive statin therapy for the secondary prevention of vascular disease in Spain. *Rev Esp Salud Publica*. 2015;89(2):159-171. https://doi.org/10.4321/S1135-57272015000200005.
42. Oliva B, Gil M, Montero D, et al. Patterns of topical Calcineurin inhibitor drug use. Impact of regulatory actions in off-label use: 332. *Pharmacoepidemiol Drug Saf*. 2012;21(SUPPL.3):156.
43. Baladé Martínez L, Montero Corominas D, Macías S-GD. Utilization of hormone replacement therapy in Spain: trends in the period 2000–2014. *Med Clin (Barc)*. 2016;147(7):287-292. https://doi.org/10.1016/j.medcli.2016.05.023.
44. Gil M, Oliva B, Timoner J, Maciá MA, Bryant V, de Abajo FJ. Risk of meningioma among users of high doses of cyproterone acetate as compared with the general population: evidence from a population-based cohort study. *Br J Clin Pharmacol*. 2011;72(6):965-968. https://doi.org/10.1111/j.1365-2125.2011.04031.x.
45. Brauer R, Douglas I, Garcia Rodríguez LA, et al. Risk of acute liver injury associated with use of antibiotics. Comparative cohort and nested case-control studies using two primary care databases in Europe. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):29-38. https://doi.org/10.1002/pds.3861.
46. Udo R, Tcherny-Lessenot S, Brauer R, et al. The risk of acute liver injury associated with the use of antibiotics—evaluating robustness of results in the pharmacoepidemiological research on outcomes of therapeutics by a European consortium (PROTECT) project. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):47-55. https://doi.org/10.1002/pds.3841.
47. Brauer R, Ruigómez A, Downey G, et al. Prevalence of antibiotic use: a comparison across various European health care data sources. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):11-20. https://doi.org/10.1002/pds.3831.
48. Weibel D, Dodd C, Mahaux O, et al. ADVANCE system testing: can safety studies be conducted using electronic healthcare data? An example using pertussis vaccination. *Vaccine*. 2019 Oct 31. pii: S0264-410X(19)30802-3. https://doi.org/10.1016/j.vaccine.2019.06.040.
49. Emborg H-D, Kahlert J, Braeye T, et al. ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: an example. *Vaccine*. 2019 Oct 31. pii: S0264-410X(19)30930-2. https://doi.org/10.1016/j.vaccine.2019.07.039.
50. Martín-Merino E, Llorente-García A, Corominas DM, Huerta-Álvarez C. The longitudinal incidence of human papillomavirus vaccination in Spanish primary care in the first 6 years after approval. *Pharm Med*. 2019;33(2):135-144.
51. de Abajo FJ, Gil MJ, García Poza P, et al. Risk of nonfatal acute myocardial infarction associated with non-steroidal antiinflammatory drugs, non-narcotic analgesics and other drugs used in osteoarthritis: a nested case-control study. *Pharmacoepidemiol Drug Saf*. 2014;23 (11):1128-1138. https://doi.org/10.1002/pds.3617.
52. Rodríguez-Miguel A, García-Rodríguez LA, Gil M, Barreira-Hernández D, Rodríguez-Martín S, de Abajo FJ. Population-based case-control study: chemoprotection of colorectal cancer with non-aspirin nonsteroidal anti-inflammatory drugs and other drugs for pain control. *Aliment Pharmacol Ther*. 2019;50(3):295-305. https://doi.org/10.1111/apt.15333.
53. de Abajo FJ, Gil MJ, Bryant V, Timoner J, Oliva B, García-Rodríguez LA. Upper gastrointestinal bleeding associated with NSAIDs, other drugs and interactions: a nested case-control study in a new general practice database. *Eur J Clin Pharmacol*. 2013;69(3): 691-701. https://doi.org/10.1007/s00228-012-1386-3.

54. García-Poza P, de Abajo FJ, Gil MJ, Chacón A, Bryant V, García-Rodríguez LA. Risk of ischemic stroke associated with non-steroidal anti-inflammatory drugs and paracetamol: a population-based case-control study. *J Thromb Haemost*. 2015;13(5):708-718. https://doi.org/10.1111/jth.12855.
55. Rafael S-J, Julio B. Estudio descriptivo sobre el uso de antiinflamatorios no esteroideos por vía intramuscular para el tratamiento de la lumbalgia aguda en las consultas de Atención Primaria en España durante 2002-2011. *Rev Clin Med Fam [Internet]*. 2015;8(2):103-109.
56. de Abajo FJ, Gil MJ, Rodríguez A, et al. Allopurinol use and risk of non-fatal acute myocardial infarction. *Heart*. 2015;101(9):679-685. https://doi.org/10.1136/heartjnl-2014-306670.
57. Rodríguez-Martín S, de Abajo FJ, Gil M, et al. Risk of acute myocardial infarction among new users of allopurinol according to serum urate level: a nested case-control study. *J Clin Med*. 2019;8(12):pii:E2150. https://doi.org/10.3390/jcm8122150.
58. Erviti J, Alonso A, Oliva B, et al. Oral bisphosphonates are associated with increased risk of subtrochanteric and diaphyseal fractures in elderly women: a nested case-control study. *BMJ Open*. 2013;3(1):pii:e002091. https://doi.org/10.1136/bmjopen-2012-002091.
59. Erviti J, Alonso A, Gorricho J, López A. Oral bisphosphonates may not decrease hip fracture risk in elderly Spanish women: a nested case-control study. *BMJ Open*. 2013;3(2):pii:e002084. https://doi.org/10.1136/bmjopen-2012-002084.
60. Martín-Merino E, Petersen I, Hawley S, et al. Risk of venous thromboembolism among users of different anti-osteoporosis drugs: a population-based cohort analysis including over 200,000 participants from Spain and the UK. *Osteoporos Int*. 2018;29(2):467-478. https://doi.org/10.1007/s00198-017-4308-5.
61. Leon Vazquez F, Bonis J, Bryant Cerezo V, Herrero Hernandez S, Jamart Sanchez L, Diaz HA. Prevencion de fractura osteoporotica en España: uso de farmacos antes y despues de una fractura de cadera. *Revista de Osteoporosis y Metabolismo Mineral*. 2015;7:54-62. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1889-836X2015000200003&nrm=iso.
62. Martín-Merino E, Huerta-Álvarez C, Prieto-Alhambra D, Álvarez-Gutiérrez A, Montero-Corominas D. Secular trends of use of anti-osteoporotic treatments in Spain: a population-based cohort study including over 1.5million people and more than 12years of follow-up. *Bone*. 2017;105:292-298. https://doi.org/10.1016/j.bone.2017.08.031.
63. Martín-Merino E, Huerta-Álvarez C, Prieto-Alhambra D, Montero-Corominas D. Cessation rate of anti-osteoporosis treatments and risk factors in Spanish primary care settings: a population-based cohort analysis. *Arch Osteoporos*. 2017;12(1):39. https://doi.org/10.1007/s11657-017-0331-6.
64. Hedenmalm K, Blake K, Donegan K, et al. A European multicentre drug utilisation study of the impact of regulatory measures on prescribing of codeine for pain in children. *Pharmacoepidemiol Drug Saf*. 2019;28(8):1086-1096. https://doi.org/10.1002/pds.4836.
65. Gil M, González-González R, Vázquez-Calvo A, et al. Clinical infections by herpesviruses in patients treated with valproic acid: a nested case-control study in the Spanish Primary Care Database, BIFAP. *J Clin Med*. 2019;8(9):pii:E1442. https://doi.org/10.3390/jcm8091442.
66. de Groot MCH, Schuerch M, de Vries F, et al. Antiepileptic drug use in seven electronic health record databases in Europe: a methodologic comparison. *Epilepsia*. 2014;55(5):666-673. https://doi.org/10.1111/epi.12557.
67. Requena G, Logie J, Martin 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. https://doi.org/10.1002/pds.3822.
68. Requena G, Huerta C, Gardarsdottir H, et al. 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. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):66-78. https://doi.org/10.1002/pds.3816.
69. Martín-Merino E, de Abajo FJ, Gil M. Risk of toxic epidermal necrolysis and Stevens-Johnson syndrome associated with benzodiazepines: a population-based cohort study. *Eur J Clin Pharmacol*. 2015;71(6):759-766. https://doi.org/10.1007/s00228-015-1850-y.
70. 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. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):56-65. https://doi.org/10.1002/pds.3825.
71. Martin-Perez M, Bryant V, Martin-Merino E, Montero-Corominas D, Huerta C. Are patients adherent to benzodiazepine treatment? Results of a retrospective cohort study in primary care in Spain. *Pharmacoepidemiol Drug Saf*. 2019:Abstract 993;28(Suppl 2):453. https://doi.org/10.1002/pds.4864.
72. Souverein PC, Abbing-Karahagopian V, Martin E, et al. Understanding inconsistency in the results from observational pharmacoepidemiological studies: the case of antidepressant use and risk of hip/femur fractures. *Pharmacoepidemiol Drug Saf*. 2016;25(Suppl 1):88-102. https://doi.org/10.1002/pds.3862.
73. Macías Saint-Gerons D, Huerta Álvarez C, García Poza P, Montero Corominas D, de la Fuente Honrubia C. Trazodone utilization among the elderly in Spain. A population based study. *Rev Psiquiatr Salud Ment*. 2018;11(4):208-215. https://doi.org/10.1016/j.rpsm.2016.11.003.
74. Saiz LC. Use of methylphenidate and risk for valvular heart disease: a case-control study nested in the BIFAP cohort. *Pharmacoepidemiol Drug Saf*. 2020;29(3):288-295. https://doi.org/10.1002/pds.4954.
75. Raman SR, Man KKC, Bahmanyar S, et al. Trends in attention-deficit hyperactivity disorder medication use: a retrospective observational study using population-based databases. *Lancet Psychiatry*. 2018;5(10):824-835. https://doi.org/10.1016/S2215-0366(18)30293-1.
76. Bonis J, de Hoyos MC, Llorente A, Gil MJ, Montero D, de Abajo FJ. Prevalencia de uso de fármacos para el tratamiento de la enfermedad de Alzheimer y su evolución temporal: un estudio descriptivo con la base de datos de atención primaria BIFAP. *Alzheimer Real Invest Demenc*. 2013;54:20-27.
77. Aguilar-Shea AL, Bonis J. COPD from an everyday primary care point of view. *J Family Med Prim Care*. 2019;8(8):2644-2650. https://doi.org/10.4103/jfmpc.jfmpc_477_19.
78. Rottenkolber M, Voogd E, van Dijk L, et al. Seasonal changes in prescribing of long-acting beta-2-agonists-containing drugs. *Respir Med*. 2015;109(7):828-837. https://doi.org/10.1016/j.rmed.2015.01.010.
79. Rodríguez-Martín S, Martín-Merino E, Lerma V, et al. Active surveillance of severe cutaneous adverse reactions: a case-population approach using a registry and a health care database. *Pharmacoepidemiol Drug Saf*. 2018;27(9):1042-1050. https://doi.org/10.1002/pds.4622.
80. Hernández-Rodríguez MÁ, Sempere-Verdú E, Vicens-Caldentey C, et al. Evolution of polypharmacy in a Spanish population (2005-2015): a database study. *Pharmacoepidemiol Drug Saf*. 2020 Jan 6. https://doi.org/10.1002/pds.4956. [Epub ahead of print]

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maciá-Martínez M-A, Gil M, Huerta C, et al. *Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria* (BIFAP): A data resource for pharmacoepidemiology in Spain. *Pharmacoepidemiol Drug Saf*. 2020;29:1236–1245. https://doi.org/10.1002/pds.5006