

## Introduction

The main purpose of drug epidemiology, or pharmacoepidemiology, is to determine the impact of medicinal products upon the health of the population (1). Clinical trials conducted prior to the authorization of a given drug substance afford quite precise information on its efficacy and most frequent adverse effects. Nevertheless, clinical trials have limitations in defining the risks and benefits of drugs in the context of habitual clinical use, i.e., their safety and efficacy, when prescription is extended to special risk groups (elderly subjects, patients with different diseases, polymedicated individuals, pregnant women and children, etc.), or when drugs are used for more extended periods of time or involving more flexible administration regimens or indications.

### *Pharmacoepidemiology and pharmacovigilance*

The salient contributions of pharmacoepidemiology have been related to the evaluation of drug safety, to the point that pharmacoepidemiological research has been regarded as an essential complement to pharmacovigilance (2). The basic pharmacovigilance methodology comprises the spontaneous reporting of suspected adverse drug reactions (ADRs). In Spain, the usual system for reporting suspected reactions of this kind is the so-called "yellow card", and the professionals who most often report suspected ADRs are primary care doctors, who provide 70% of all reports (3). The Spanish pharmacovigilance system is organized on a decentralized basis: each Autonomous Community has a center in charge of stimulating suspected ADR reporting practices, receiving those submitted by the health care professionals, and introducing them in the FEDRA database provided by the Spanish Medicines Agency (3).

### *Limitations of the reporting system and of field work for pharmacoepidemiological research*

Programs for spontaneous reporting are of great help in generating alarm signals in relation to drug-reaction associations. However, due to their intrinsic structure, such programs pose important limitations for quantifying risk, i.e., quantifying the degree of association between the drug in question and the suspected adverse reaction (a necessary step for epidemiological evaluation of the corresponding causal relation), quantifying absolute risk (to determine its impact upon public health), its relation to treatment dose and duration, and the identification of increased risk groups (with the purpose of adopting measures for reducing the problem, where required). In most instances such information can only be obtained through adequate studies – fundamentally cohort studies and case-control studies. The conduct of such studies according to traditional methods implies not only

considerable consumption of resources but also laborious field work, and is usually not acceptable for providing fundamental answers within a reasonable timeframe (in months, rather than years) in response to an alarm signal generated by the pharmacovigilance system. For this reason, decisions in this terrain are often taken based only on spontaneous reporting data adjusted for drug consumption, which is used as an indirect measure of exposure (4).

### *Computer databases*

The first computer-based health care databases were created in the 1970s in the United States. Their initial purpose was merely administrative, though their efficacy as a source of information for the conduct of epidemiological studies was soon appreciated (5). With these databases it has been possible to study cohorts of hundreds of thousands of individuals – something unthinkable with the "traditional" methodology. This advantage may be particularly applicable to pharmacoepidemiological research, which generally attempts to evaluate infrequently-appearing effects requiring the conduction of very large cohort studies or extensive case-control evaluations. Moreover, studies done with computerized databases afford answers involving timeframes more in line with the decision-making process in pharmacovigilance. However, computerized databases not only simplify the logistics and reduce the costs of pharmacoepidemiological research but they can also contribute to increase validity, thanks to two additional advantages:

1. The information they contain regarding exposure to drugs of interest tends to be more complete and reliable than in traditional interview-based studies, since they prospectively register drug prescription or dispensing instead of relying on the memory of the interviewee.
2. In case-control studies, they afford genuine population-based control selection. In contrast, other procedures (e.g., hospital controls) are unable to guarantee that control selection is independent of exposure to the medication of interest.

### *The GPRD database*

A good number of databases have been developed in the United States and in Europe allowing the conduct of pharmacoepidemiological studies, though special mention should be made of the British GPRD (General Practice Research Database)(6-8), which computer accumulates information from close to 2000 primary care doctors. The salient feature of the GPRD is its "integral" nature, since it contains data relating to prescription, diseases and clinical problems, and results of complementary tests, requiring no link to other databases (i.e., record-linkage) in order to gather the information required for epidemiological studies – the

latter being the field where the GPRD has undoubtedly been most useful. The GPRD contains information corresponding to 35 million person-years, and has allowed the publication of over 200 studies in scientific journals – including particularly studies of ADRs (9-11), as well as studies addressing the possible benefits of drugs in routine clinical practice (drug efficacy)(12,13).

#### *Clinical information and primary care*

The organization of the national health care system in Spain is similar to that found in Great Britain. In our setting, primary care is usually the entry point to the health care system, and is also a common destination for those patients initially seen outside primary care. The primary care setting is also the level at which an enormous number of health problems are resolved. Within the context of the primary care team, the doctors – both general practitioners and pediatricians – are the professionals who are closest to integrally knowing the clinical particulars of their patients. The longitudinal care they provide, the great number of diagnoses and treatments involved, their normalizing function in relation to patient access to other parts of the system, their general view of health problems, and the confidence they frequently establish with their patients all place these doctors in a privileged position for attempting to gather the information generated by this and other health care settings or levels. On the other hand, primary care doctors extend over 80% of all prescriptions made in the Spanish National Health System. The fact that part of these prescriptions correspond to prescriptions originally made by other doctors also gives an idea of the important flow of information between other health care levels and the primary care setting. Based on these considerations, it may be stated that primary care doctors are undoubtedly best positioned to conduct a complete prospective registry of health problems, therapeutic interventions, preventive activities and clinical events. In global terms, the information these professionals deal with habitually can be considered particularly adequate for pharmacoepidemiological research (14), since such information *a priori* contains all the elements required for such research: prescriptions, symptoms and diagnoses, demographic data, anthropometrical information, exploratory data, and the results of complementary tests.

#### *Computerization of primary care*

The computerization of clinical visits and histories is increasingly common in primary care (15-18). The usefulness of this practice need not be limited to healthcare, administrative or management tasks but can also be extended to other activities – including of course pharmacoepidemiological research in large population groups. Based on these premises, the BIFAP Project

(Database for Pharmacoepidemiological Research in Primary Care / *Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria*) has been created.

## The BIFAP Project

### *Definition*

The BIFAP Project has been developed to determine whether it is feasible in Spain to establish a public database containing anonymous clinical information (i.e., without personal identifying data) facilitated by primary care doctors of the Spanish National Health System who habitually make use of computers to document patient information. The database will likewise be named BIFAP, and if viable it will be destined to the conduction of pharmacoepidemiological studies, following due validation as an information source.

### *Objectives*

The aim of the project is to create a quality pharmacoepidemiological research tool essentially intended for the conduction of two types of study:

- Drug safety studies attempting to test causal association hypotheses, including those specifically intended to assess drug-related “alarm signals”, particularly those generated by the Spanish pharmacovigilance system.
- Studies of drug efficacy under habitual conditions of use (the broad population base of the BIFAP would contribute to this).

the information afforded by the BIFAP will only be usable for epidemiological research – never to either individually or collectively evaluate aspects relating to the professional activities or costs generated by the collaborating doctors.

### *Justification*

Apart from the considerations made in the Introduction, other reasons exist for attempting to establish a primary care database in Spain - considering that others already exist in other countries - and use it for pharmacoepidemiological research:

- it is always desirable to have the possibility of contrasting the same hypothesis using different information sources;
- some drugs are marketed only in Spain or are only relevantly used in Spain;
- the effects of drugs should be assessed under their conditions of use, which may not be the same as in other countries; and
- the effects of drugs vary according to genetic and environmental factors, which in turn may differ in terms of the population involved.

## Pilot phase

The pilot phase was started in January 2000 and is expected to extend to the end of 2003. Up until June 2002, the following steps have been taken: *a)* a team has been established, composed of pharmacoepidemiologists, primary care doctors and computer technicians; *b)* the BIFAP Data Processing Center (DPC)(located in the Pharmacoepidemiology and Pharmacovigilance Division of the Spanish Medicines Agency) has been equipped with the required technical means; *c)* authorization has been requested from the different Autonomous Communities for starting the project; *d)* information on the project has begun to be distributed to doctors and health care centers; *e)* an evaluation has been made of the software programs of greatest implantation in the National Health Care System in terms of the possibility of developing a specific data exportation module (the programs must at least allow the coded recording of health problems according to the International Classification of Primary Care, the International Classification of Diseases – Ninth Edition (ICD-9), or posterior version of both; of the prescribed drugs according to the National Pharmaceutical Product Codes; and of the complementary data and observations of each process); *f)* a data exportation module has been developed, compatible with the OMI-AP program in its versions 4 and 5, and tests have been made of the latter - with satisfactory results. The functions of the module are (Fig. 1): to extract the information of interest (Table 1), generate non-identifying patient codes (dissociation), and encrypt the data; *g)* the database structure has been created; and *h)* a registry guide has been edited for OMI-AP using collaborating doctors, the function of which is to provide recommendations concerning different aspects of the registry (e.g., the assignment of dates, the coding of processes and prescriptions, the inactivation of duplicate histories, the existence of imaginary patient histories – often used to perform registry tests and afford training in computer use – and so on), which increase the consistency and validity of the information recorded while preserving maximum program flexibility.

The development of the BIFAP Project is supervised by a scientific committee with representation by the principal scientific societies in the field of primary care, and including general practitioners, primary care pediatricians, specialists in bioethics, specialists in primary care computerization, and pharmacoepidemiologists. Their tasks are to supervise development of the project, provide counseling on specific aspects, contribute suggestions, and ensure adherence to legislation and to current recommendations in matters of data protection.

## Doctor collaboration

In the year 2001, recruitment of the collaboration of general practitioners and primary care pediatricians started in those Autonomous Communities for which authorization had been obtained. The aim is to ensure the cooperation of 300-500 doctors (a figure considered to be convenient for the BIFAP validation studies; see below). Collaboration is of an individual and voluntary nature (with prior approval from the corresponding directing or management body). The essential requirements for cooperating in the pilot phase are: *a)* habitual registration of patient information with the OMI-AP program; and *b)* foreseeable workplace stability. The collaborating doctors will receive formal acknowledgement as such.

None of the activities implied in such collaboration will pose a significant added workload for those doctors with minimal skills in using a computer in the consulting room. These activities are the following (those activities which must be carried out by the collaborating doctor in person are shown in boldface; depending on the contents, other team doctors, residents, primary care computer technicians or administrative personnel may participate in the rest of activities):

### *Activities related to data transmission to the DPC:*

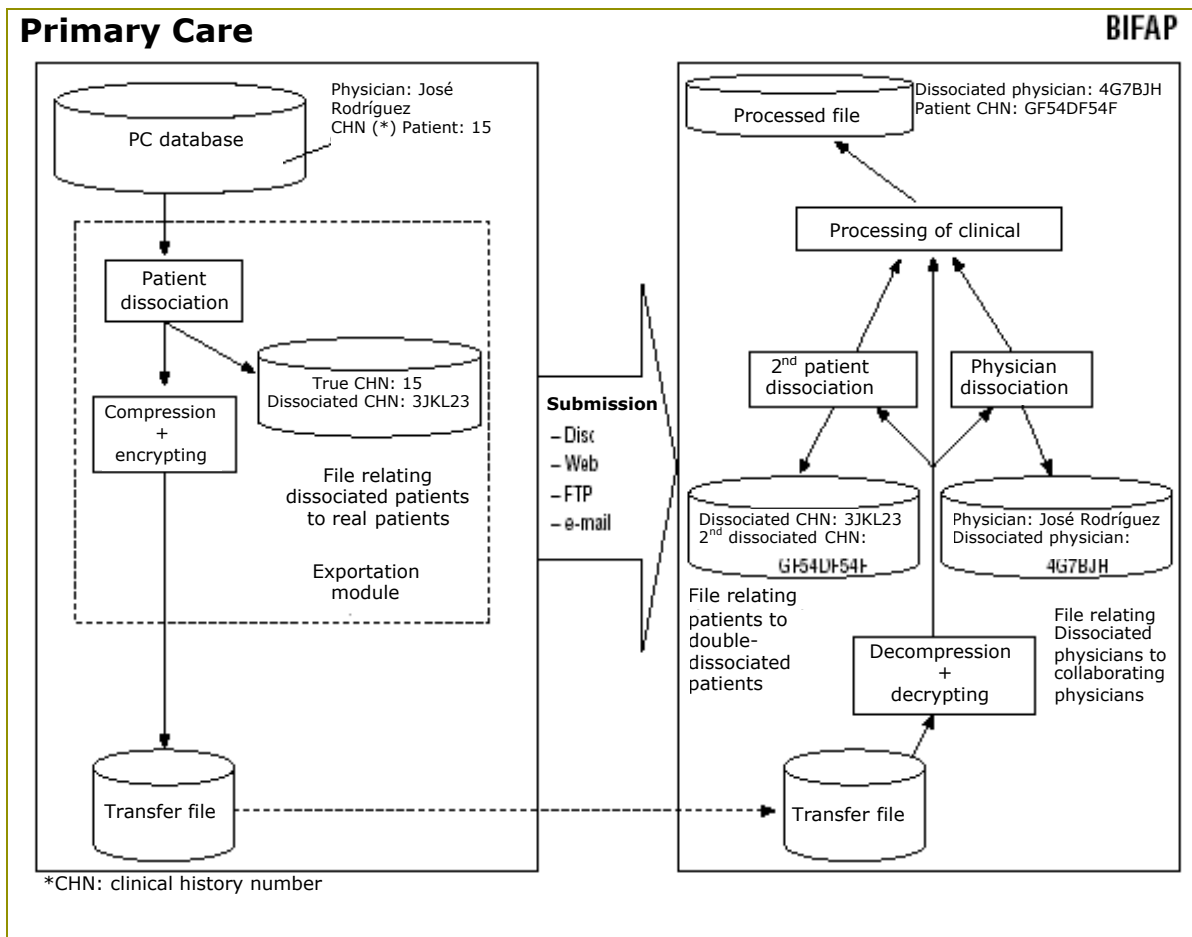
- Installation of the exportation module.
- **Periodic exportations** (every 2-3 months). The initiation of each exportation session will require the use of a secret code established by the collaborating doctor.
- Submission to the DPC of the files generated in the exportation process. The transmissions can be made electronically or by mail (using a computer disc or other appropriate support). The project information flow is represented in Figure 1.

### *Following of the guide by cooperating doctors:*

The aim is to ensure a sufficient registration level in the course of the first 6 months. It is advisable to promote adherence to the guide by the substituting doctors and residents who attend the same quota of patients. The different types of data to be recorded are specified in Table 1.

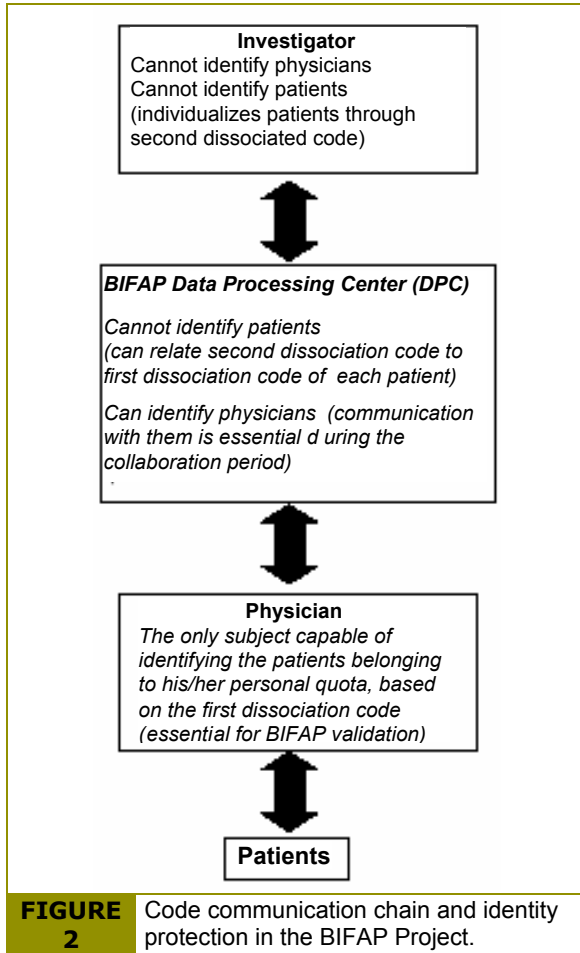
### *Activities related to BIFAP validation to be conducted during the pilot phase (Fig. 2)(See also the section: “BIFAP validation”):*

- Photocopy or print-out of admission reports, specialist reports and complementary tests performed (filed on paper or computer) in small patient samples (to this effect, the DPC will supply the individual dissociation codes of the patients selected). The identity of the



**FIGURE 1.** Information flow in the BIFAP Project and example of the dissociation procedure.

<b>TABLE 1.</b>	<b>BIFAP information</b>
1.	Administrative data (admission/discharge dates) and demographic particulars (sex and date of birth)
2.	Morbidity events
	Diseases/symptoms leading to patient consultation
	Starting date of first diagnosis of chronic and recurrent illnesses
	Significant results of complementary tests
	Indications of the prescribed medication, and reasons for changing the dose or suspending the medication
	Events or disorders giving rise to admission, referral to the Emergency Service or to a specialist, and essential data derived from the latter (new diagnoses, interventions, results of specialized tests, etc.)
3.	Prescriptions
4.	Pregnancy and its outcome
5.	Deaths and their causes
6.	Other data of clinical or epidemiological interest (vaccinations, height, weight, toxic habits)



patient will only be accessible to the collaborating doctor, since the module allows the latter to **reassociate this code** with the corresponding clinical case number (the use of this application also requires a secret code).

- **Inscription of the dissociated patient code in the copies of reports.**
- **Document anonymity** (photocopies or print-outs). This will consist of the elimination of all identifying data – either direct (patient first and last name/s) or indirect (name of the doctors in charge of the reports, hospital, locality, etc.).
- Submission to the DPC of the anonymous reports (See the section: “Validation”).

#### Data protection in the BIFAP Project

The mechanisms guaranteeing data anonymity and confidentiality are the following: *a)* data exportation does not include identifying information (Fig. 1); *b)* double code dissociation is performed, at origin and destination, assigning a random code to each patient, and generating a dissociated code for each doctor at destination (Figs 1 and 2); *c)* high-level encrypting is performed at origin to reinforce data security (Fig. 1); *d)* the DPC follows a series of internal security norms

that affect both physical and logical safety and the procedures of data use; and *e)* the identity of the collaborating doctors is not made known to third persons (with the obligate exception of the corresponding directory/managerial bodies).

#### Controls prior to data loading in BIFAP

The data reaching the DPC will be subjected to automatized controls to verify that the information is minimally coherent from the qualitative and quantitative perspective, taking population indicators and mean values of the data received as reference. In addition, the controls will entail the generation of reports to be distributed individually to each collaborator to specify which areas of the recording process are adequate, which can be improved, and which are defective. These reports of course will be exclusively addressed to the corresponding doctor. The aim of such reports is to ensure a feedback mechanism to secure adequate recording performance and maintain such performance where already achieved.

#### BIFAP validation

The database validation studies will be conducted by the BIFAP Project investigators, once the aggregated information has covered a period of at least 6-12 months. These validation studies will comprise correlation of the BIFAP registries with anonymous copies of clinical reports (Fig. 2)(See the section “Doctor collaboration”), in samples of patients selected by disease or medication code. A concordance of diagnoses and prescribed drugs of over 90% will be considered adequate. For those centers working “without papers”, an alternative validation strategy will be established.

### BIFAP feasibility and perspectives

Only if at the end of the pilot phase a sufficient number of collaborating doctors has been secured and it is concluded that the BIFAP is effectively valid for pharmacoepidemiological research will it make sense to maintain the database as a permanent source of information. In such a case, the collaborating doctors will be requested to accept a stable cooperation commitment, and the initiative will be left open for the incorporation of other software and new collaborating doctors – with the aim of gradually increasing their number. A figure of 2000 collaborating doctors is considered convenient to efficiently investigate the rare effects or effects of infrequently prescribed drugs.

On the other hand, if in future BIFAP becomes established as a research tool, it will be obligate to define a series of standardized operative procedures relating to the use and provision of data to the investigators. These procedures will include scientific

and ethical revision by an independent committee of the proposed studies involving the mentioned data. Among these conditions, it is expected that the actual collaborating doctors will be among the end-users of the database. In sum, these professionals are central to the BIFAP Project, which hopes to make a database available to public health and the scientific community for conducting quality pharmacoepidemiological studies with important time and resource savings.

## References

### (Bibliography)

1. Jick H, García Rodríguez LA, Pérez-Gutthann S. Principles of epidemiological research on adverse and beneficial drug effects. *Lancet* 1998;352:1767-70.
2. De Abajo FJ, Montero D, Cachá A. Pharmacovigilance: goals and strategies. *Methods Find Exp Clin Pharmacol* 2000;22:405-7.
3. Madurga M, De Abajo FJ, Martín-Serrano G, Montero D. El Sistema Español de Farmacovigilancia. En: Grupo IFAS, editores. *Nuevas perspectivas de la farmacovigilancia en España y en la Unión Europea*. Madrid: Jarpyo, 1998; p. 37-61.
4. Arnáiz JA, Carné X, Riba N, Codina C, Ribas J, Trilla A. The use of evidence in pharmacovigilance. Case reports as the reference source for drug withdrawals. *Eur J Clin Pharmacol* 2001;57:89-91.
5. Gardner JS, Park BJ, Stergachis A. Automated Databases in Pharmacoepidemiological Studies. En: Hartzema AG, Porta MS, Tilson HH, editores. *Pharmacoepidemiology*. 3rd ed. Cincinnati: Harvey Whitney Books Company, 1998; p. 368-88.
6. Jick H, Jick SS, Derby L. Validation of information recorded on a general practitioner based computerised data resource in the United Kingdom. *BMJ* 1991;302:766-8.
7. García Rodríguez LA, Pérez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *Br J Clin Pharmacol* 1998;45:419-25.
8. García Rodríguez LA, Pérez-Gutthann S, Jick S. The UK General Practice Research Database. En: Strom BL, editor. *Pharmacoepidemiology*, 3rd ed. Chichester: John Wiley & Sons, Ltd., 2000; p. 375-85.
9. García Rodríguez LA, Stricker BH, Zimmermann HJ. Risk of acute liver injury associated with the combination of amoxicillin and clavulanic acid. *Arch Intern Med* 1996;156:1327-32.
10. De Abajo FJ, García Rodríguez LA. Risk of

ventricular arrhythmias associated with nonsedating antihistamine drugs. *Br J Clin Pharmacol* 1999;47:307-13.

11. De Abajo FJ, García Rodríguez LA, Montero D. Association between selective serotonin reuptake inhibitors and upper gastrointestinal bleeding: population based case-control study. *BMJ* 1999;319:1106-9.
12. García Rodríguez LA, Ruigómez A. Secondary prevention of upper gastrointestinal bleeding associated with maintenance acid-suppressing treatment in patients with peptic ulcer bleed. *Epidemiology* 1999;10:228-32.
13. García Rodríguez LA, Huerta Álvarez C. Reduced risk of colorectal cancer among long-term users of aspirin and nonaspirin nonsteroidal antiinflammatory drugs. *Epidemiology* 2001;12: 88-93.
14. Gómez de la Cámara A. La investigación en atención primaria. El ensayo clínico y los estudios observacionales de productos farmacéuticos. *Aten Primaria* 1999;24:431-5.
15. Grupo de trabajo sobre informatización de la semFYC. Informatización en la atención primaria (I). *Aten Primaria* 2000;26:488-507.
16. Grupo de trabajo sobre informatización de la semFYC. La informatización de atención primaria (y II). *Aten Primaria* 2000;26:559-76.
17. Grupo de trabajo sobre informatización de la semFYC. El desafío de la informatización en atención primaria [editorial]. *Aten Primaria* 2000;26:437-8.
18. Gervas J, Pérez Fernández M. La historia clínica electrónica en atención primaria. Fundamento clínico, teórico y práctico. *Semergen* 2000;26:17-32.

### How to contact:

Further details on the project can be found at the following website: [www.bifap.org](http://www.bifap.org)  
 Doctors interested in participating in the BIFAP Project can request the exportation module and registry guide by contacting Dr. Antonio Salvador via e-mail ([asalvador@bifap.org](mailto:asalvador@bifap.org)), telephone (91 596 78 88), fax (91 596 78 91), or conventional mail:  
 Data Processing Center-BIFAP  
 Spanish Medicines Agency  
 Division of Pharmacoepidemiology and Pharmacovigilance  
 Ctra. Majadahonda-Pozuelo, km 2  
 28220 Majadahonda (Madrid), Spain