Data Resource Profile

Data Resource Profile: Results Analysis Base of Navarre (BARDENA)

†These authors contributed equally to this work.

Data resource basics

The Results Analysis Base of Navarre (BARDENA) is a set of multiple, public, population-wide electronic data-bases for Navarre, a territory in the north of Spain with 0.65 million inhabitants and an annual birth cohort of 5000 newborns, representing 1.4 % of the Spanish population and around 1.5 per thousand of the European population. The BARDENA provides exhaustive longitudinal information including sociodemographic and administrative data (sex, age, nationality, etc.), clinical (diagnoses, procedures, diagnostic tests, ~~imaging,~~ etc.), pharmaceutical (prescription, dispensation) and healthcare utilization data from hospital care, emergency departments, specialized care (including mental and obstetrics care), primary care and other public health services. It also includes a set of associated population databases and registries of significant care areas such as acute stroke, cardiorespiratory arrest, or chronic population. All the information in the BARDENA databases can be linked at the individual level through a single pseudo-anonymized personal identification code. The databases were initiated at different moments in time (see details in the Data collected section), but all in all the BARDENA provides comprehensive individual-level data fed by all the databases from 2012 to date.

The BARDENA in the context of the Spanish National

Health System

The Spanish National Health System (SNHS) is the result of a system consolidation process started in 1978 and leading to the nearly universal coverage of all citizens, providing care based on need and free at the point of delivery, except for a cost-sharing scheme for pharmaceuticals dispensed out of hospitals. (1) Care delivery is mainly undertaken through a network of publicly owned, staffed and operated inpatient and outpatient centers. In 2002 a process of devolution to the 17 regions that comprise the Spanish state was completed. Each regional NHS is geographically organized into Primary Healthcare Districts (around 5000–25000 people served by one Primary Care Centre), which in turn are embedded into Healthcare Departments (about 150000–250000 people served by one public hospital). Each regional NHS develops and operates its own information systems and the development of real-world data (RWD) research capabilities is heterogeneous, the Results Analysis Base of Navarre (BARDENA) in the region of Navarra being among the best in terms of data availability and the linkage capacity of databases at a population level.

Data collected

Data are sourced from a variety of datasets owned by the Health Department of the Foral Community of Navarre. All data included in the databases can be obtained at the individual level. The type of available data, measurements collected and update frequency is different for each dataset. The main characteristics of each dataset are described below and in Fig. 1.

The Population Information System (LAKORA) is a region-wide database that provides basic information on BARDENA coverage (dates and causes of BARDENA entitlement or disentitlement, insurance modality, pharmaceutical copayment status, assigned Healthcare Department, Primary Healthcare District and primary care doctor, etc.) and also some sociodemographic data such as sex, date of birth, nationality, country of origin, previous year income strata]], employment status, risk of social exclusion, geographic location, address and other administrative data. LAKORA information module is paramount to BARDENA as it is the source of the individual, exclusive and permanent identifier number associated to each individual (the CIPNA number) that is then used throughout the rest of the databases, allowing data linkage across the multiple databases in the network (see Fig. 1).

The Primary care module (ATENEA) was implemented in 2003 as the electronic medical record (EMR) for primary, reaching 95% in 2008. ATENEA includes pediatric and adult primary care, as well as providing related social assistance and nursing care. It use for coding diagnoses International Classification of Primary Care (ICPC-2).

The pharmaceutical data record consist in 2 modules, LAMIA with primary prescription and dispensations, and FARHO with hospital prescriptions, using the Anatomical Therapeutic Chemical (ATC) classification system and the National Pharmaceutical Catalogue which allow the identification of the exact content of each dispensation. LAMIA and FARHO provides detailed information on prescriptions issued by physicians, such as the duration of treatment and dosage. LAMIA includes a comprehensive e-prescription paper-free system connected to all community pharmacies in the region that permits the linkage of individual prescriptions and dispensations through a specific prescription identification number.

The Hospital Medical Record (HCI)has been in implementation since ~~2008~~ and provides comprehensive information covering all areas of specialized care from admission, outpatient consultations, hospitalization, emergencies, diagnostic services (labs, imaging, microbiology, pathology, etc.), pharmacy, surgical block including day surgery, critical care, prevention and safety, social work, at-home hospitalization, day hospitalization, mental health care, and prenatal care.

The Minimum Basic Data Set at Hospital Discharge (MBDS) is a synopsis of clinical and administrative information on all hospital admissions and major ambulatory surgery in the Navarre hospitals, including public partnership hospitals (around 450 000 admissions per year in the region). The MBDS includes admission and discharge dates, age, sex, geographical area and zone of residence, main diagnosis at discharge, up to 21 secondary diagnoses (comorbidities or complications), clinical procedures performed during the hospital episode and the Diagnosis Related Groups (DRG) assigned at discharge. The MBDS used the ICD9CM system for coding until December 2015 and the ICD10ES (a Spanish translation of the ICD10CM) thereafter. The MBDS was extended in 2015 to include the ‘present on admission’ (POA) diagnosis.

The Vaccine Information System (Sistema de Informacion Vacunal, INMUNIS) stores all the information on vaccination in the Navarra since 2000, though data are only considered reliable after 2008. Available data include vaccine by type, manufacturer, batch number, number of doses, location and administration date, adverse reactions related to vaccines, rejected vaccinations and, if applicable, risk groups. Besides, COVID-19 vaccine data is included.

The Rare Diseases Information System (Sistema de Información de Enfermedades Raras de la Comunidad Valenciana, SIER-CV) was created in 2012 to provide population-wide epidemiological information on rare disprevalence, patient characteristics, geographical distribu- tion, etc. It includes the Congenital Anomalies Registry, which has provided information from 2007 on the preva- lence of congenital anomalies in the region and the expo- sure to teratogen agents, and allows for research on the aetiology of these diseases, including genetic and environ- mental risk factors and their interaction.

In all databases in BARDENA, individual data are collected weeklyas part of the routine clinical care provided to patients. Accordingly, datasets are updated daily and hence data may be available for research up to the same day data are extracted. Only in some cases, such as the MBDS, are data subject to a consolidation and quality check process before data are available for research, so in these cases data from the last quarter before the data extraction may be missing or non-consolidated.

Database

It is done a process of unification and verification of the data at the individual level of the different data source and a standardization of clinical and medicines data. In this process, we realized different quality of data validation to improve the coherent

Ethical clearance

Ethics approval by an accredited ethical research committee is required to access the data for research purposes (see Data resource access section). The Navarre Government Health Department ensures the anonymization of data by providing only deidentified datasets, unless researchers have the informed consent of patients to access their data. In the case of dynamic cohort studies, it maintains the pseudo-anonymization codes to allow the successive incorporation of information into the cohort.

Data resource use

In recent years BARDENA has allow population research to be conducted in COVID-19 pandemic (1-3), and other infection (4) or mental diseases (5) and to participate in national network as the Atlas of Variations in Medical Practice in the SNHS,33–35,41

Strengths and weaknesses

Strengths

BARDENA has several strengths and some differential features with regard to other information resources. First, it links population-wide healthcare data with sociodemographic and administrative data, which allows the study of the determinants of health and the consequences of illness and treatments at an individual level in the population. This allows for the inclusion in studies of population that is usually excluded from experimental designs, the elderly, people with multiple chronic diseases or pediatric population. Second, it allows for the construction and follow-up of large cohorts of patients over time and the development of longitudinal studies, enabling research on the adoption of technologies and the monitoring of outcomes in the long-term. Third, it is a population-based data network providing insight into a population of 0.65 million inhabitants. ~~This size allows for the analysis of subgroups of population, or the identification of rare events that are not usually captured in clinical trials and other designs based on primary data~~. Fourth, data quality in some of the databases is high, such as LAKORA date, the pharmaceutical module, the MBDS (admissions data) or the vaccines registry. Fifth, the cost of developing research and the timing of access to the data is considerably lower than in experimental designs such as clinical trials. Finally, the possibility of linking prescription and dispensation data at the individual level allows for an accurate analysis of drug utilization, such as medication adherence studies.

Weaknesses

Some of the databases that comprise the BARDENA are subject to the limitations inherent to routine clinical practice electronic databases. There may be information biases due to absent registration (data completeness) or differing data recording practices (data accuracy, misclassification, and heterogeneity) in the electronic databases, although this is an intrinsic problem of any repository using data from routine clinical practice. Data quality may be a strength in some databases, but also a weakness in other repositories or for certain data, such as incompleteness of early data from acute stroke or cardiorespiratory arrest records or coding reliability of diagnostic information in the EMR. In addition, we do not have information about people who are not in contact with the public healthcare service or who are attended to in the private sector. Finally, different datasets cover different periods and we lack data on specific mortality causes and in-hospital pharmaceutical prescription (the latter will be available in forthcoming years as it is currently in the process of being integrated as part of the ~~HCI~~ information system).

Data resource access

Any researcher may request anonymized data from BARDENA. The transfer of this type of data (anonymized, but with some risk of re-identification, in accordance with European regulations) by BARDENA requires that the request be accompanied by: (i) a complete study protocol that explains the planned use of data, (ii) the approval of the project by an ethics committee and if it includes pharmaceutical data, (iii) the classification of the study by the Spanish Agency of Medicines (some classifications may warrant additional authorizations). BARDENA Data Commission reviews these requests, and approves or otherwise each specific data transfer for research purposes. Authorization to access the data under these requirements should be requested electronically from the Management Office of the BARDENA Data Commission.

Following authorization, researchers are required to commit to keeping the data in a secure environment, to not attempting to re-identify or to cross with other databases, to not using the data for purposes or projects other than those specified in the project protocol (although a new authorization may be requested for these purposes) and to not transferring the data to third parties. These latter commitments limit the possibility of storing data in open data repositories or including data as [supplementary material in](https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyz266#supplementary-data) published articles.

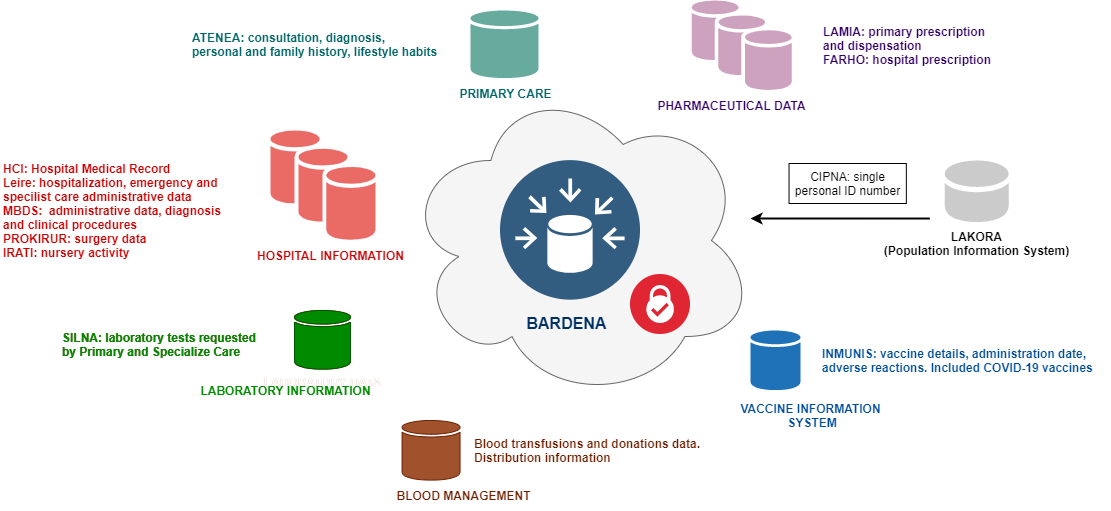


Figure 1. ~~The Valencia Health System Integrated Database (VID); VIS, Vaccine Information System; RedMIVA, Microbiological Surveillance System; CIS, Cancer Information System; SIER-CV, Rare Diseases Information System; CAR, Congenital Abnormalities Registry; BIMCV, Medical Image Bank; CRC, Catalogue of Corporate Resources; MBDS, Minimum Basic hospital Data Set; AED, Accident & Emergency Department record; GAIA, Pharmaceutical Module; SIA, Ambulatory Information System.~~