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 integrated, multiple, public, population-wide electronic databases 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 coverage of the public health system in navarra is over 99%, x% being exclusively public and y% mixed. (Ref). The BARDENA database provides an exhaustive longitudinal information of the users of the public health system (SNS-O), including sociodemographic and administrative data (sex, age, nationality, etc.), clinical (diagnoses, procedures, diagnostic tests, 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. Activity-based cost information is also incorporated following a sequence of top-down and bottom-up approaches. Besides, it includes a set of associated population databases and registries of significant care areas such as acute stroke or diabetes. All the information in the BARDENA databases can be linked at the individual level through a single personal identification code. The different databases that conform BARDENA, where progresively incorporated (see details in the Data collected section), providing a comprehensive individual-level data fed by all the databases from 2012 to date. Currently, it is in the process of being integrated into the European federated network of data sources (European Health Data & Evidence Network -EHDEN-) through its standardization using as common data model the Observational Medical Outcomes Partnership (OMOP).

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. This provided an universal need based free health care, where pharmaceuticals dispensed out of hospitals are copaid according to the income fo the individual. (1) Each one of the 17 regional Health Systems (HS)is mainly undertaken through a regional network of publicly owned, staffed and operated inpatient and outpatient centers, geographically subdivided into Primary Healthcare Districts (around 5000–25000 people served by one Primary Care Centre), which in turn are embedded into Healthcare Areas (about 150000–250000 people served by one public hospital). The information needed to provide this public health care is regionally stored and operated, based on its own information systems. Consequently, the secondary use of health data and the research based in Real World Data is heterogoneus, BARDENA being highly positioned in terms of data availability and the linkage capacity of databases at a population level.

Data collected

Data are sourced from a wide variety of datasets owned by the Health Department of the Foral Community of Navarre, all differing inthe strcutures, values andiesAll data included in the databases is obtained at individual level. 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/ disentitlement, insurance modality, pharmaceutical copayment status, assigned Healthcare Department, Primary Healthcare District and primary care physician, etc.) and sociodemographic data such as sex, date of birth, nationality, country of birth, 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 linking patient level information across the multiple databases in the network (see Fig. 1).

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

The pharmaceutical data record consists of 2 modules, LAMIA with primary prescriptions and dispensations, and FARHO with hospital prescriptions. Both tools use 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 provide detailed information on prescriptions issued by physicians, such as the duration of treatment and dosage. Additionally, LAMIA includes a comprehensive e-prescription paper-free system connected to all community pharmacies in the region, allowing the linkage of individual prescriptions and dispensations through a specific prescription identification.

The Hospital Medical Record (HCI)begun in 2001 ~~??~~ Alonso, A., Iraburu, M., Saldaña, M.L., & Pedro, M.T. de. (2004). Implantación de una historia clínica informatizada: encuesta sobre el grado de utilización y satisfacción en un hospital terciario. Anales del Sistema Sanitario de Navarra, 27(2), 233-239. and provides comprehensive information covering all areas of specialized care from admission, outpatient consultations, hospitalization, emergencies, diagnostic services (labs, imaging, microbiology, pathology, etc.), pharmacy and surgical block, including day surgery, critical care, prevention and safety, social work, hospital-at-home, day hospitals, mental health 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) was adopted thereafter.

The Vaccine Information System (Sistema de Informacion Vacunal, INMUNIS) stores all the information on vaccination in Navarre 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.

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, data are 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.

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 pseudo-nymization of data by providing only de-identified 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 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 determinants of health and the consequences of illness and treatments at an individual level in the population. This allows for the inclusion in observational studies of groups usually excluded from experimental designs, such as the elderly, people with multiple chronic diseases or the 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. 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 availability of cost data per patient and activity allow economic evaluation studies. Six, the cost of developing research and the timing of access to the data is considerably lower than in experimental designs, such clinical trials. Seven, as a source of population records, it can contribute to pragmatic trials by providing the arm of patients on whom routine practice is performed. Finally, in pharmacoepidemiology, 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 database are subject to the limitations inherent to routine clinical practice electronic databases. There may be information biases due to absent registration (data incompleteness) 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, although according to the coverage the effect should be minimal. Finally, different datasets cover different time periods and we lack data on specific mortality causes and in-hospital or nursing-home pharmaceutical prescriptions (the latter will be available in forthcoming years as it is currently in the process of being integrated).

Data resource access

Any researcher may request anonymized data from BARDENA. The transfer of this type of data (pseudo-nymized, but with some risk of re-identification, in accordance with European regulations) by BARDENA requires the request to 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 or not pharmaceutical data, (iii) the classification of the study by the Spanish Agency of Medicines (some classifications may warrant additional authorizations). The BARDENA Data Commission reviews these requests, and approves or refuses the access to each specific data transfer for research purposes. An 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, not attempt to re-identify or to cross with other databases, 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 not transfer 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.