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 suite of products resulting of the integration of the different health systems of information in the Navarre region, 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 approximately 1.5 per thousand of the European population. The BARDENA has been developed by the health department of the Navarre region government with public funding. As part of the Spanish National Health System (HS), the public health system of Navarre (SNS-O) provides an universal need based free health care, where pharmaceuticals dispensed out of hospitals are copaid according to the income fo the individual. (1) and , has a coverage over 99% of all citizens,[ Asistencia sanitaria. Cifras relativas. Instituto Nacional Estadística (INE), 2018. https://www.ine.es/jaxi/Tabla.htm?tpx=47935&L=0.] Care delivery is mainly undertaken through a network of publicly owned, staffed and operated inpatient and outpatient centres. Each regional NHS is geographically organised into Primary Healthcare Districts (among 5000–25000 people served by one Primary Care Centre), which in turn are embedded into Healthcare Departments where the coverage is divided into areas served by public Hospitals. In Navarre, there are three public hospitals: Pamplona (~477,000 inhabitants), Tudela (~109,000 inhabitants) and Estella (~64,000 inhabitants). Since each regional HS develops and operates its own information systems, the resulting products presented are unequal, being BARDENA among the best in terms of richness and quality of data, exploitation capabilities, and the integration of different products. TheBARDENA Suite, besides the databases, which conform to the BARDENA Core, comprises elements which allow linkage, interaction, analysis, knowledge diffusion and accounting. .

The BARDENA Core is the backbone of the BARDENA, where the different data sources are stored in the backend and then linked in a star schema design (See Fig. 1). The data comes from multiple sources, covering all the population of the Navarre region that have any encounter with the HS. They provide exhaustive longitudinal information 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 costs are also incorporated following a sequence of top-down and bottom-up approaches, and this information is analysed in the analytic accounting component (SECA). Moreover, it includes a set of associated population level databases and registries of significant care areas such as acute stroke or diabetes. All the information in the BARDENA Core databases is linked at the individual level through a single personal and pseudonymized 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 byall 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 standardisation using as common data model the Observational Medical Outcomes Partnership (OMOP).

Data are sourced from a variety of datasets owned by the HS of the Navarre region. 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 and supplementary table.

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- and other administrative data. LAKORA information module is paramount to BARDENA as it is the source of the - 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. 2).

The Primary care module (ATENEA) was implemented in 2003 as the electronic medical record (EMR) for primary care, reaching 95% in 2008. ATENEA includes paediatric 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 two 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. 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 2012 and provides comprehensive information covering all areas of specialised care from admission, outpatient consultations, hospitalisation, 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 hospitalisation, 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 (approximately 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 (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 weekly as 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. In the integration of the different databases into the BARDENA Core all the information is linked with the CIPNA number. In addition, from the BARDENA Core, when some variable contains any value that are clearly incorrect, such as wrong dates, phone numbers, CIPNA number, etc., this value is flagged and sent to the origin source for their revision, but the data never is transformed in the process of generation of the BARDENA Core.

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 pseudonymization 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 pseudonymization codes to allow the successive incorporation of information into the cohort.

Data resource use

In recent years the BARDENA Suite 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 .

Discussion

The 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 paediatric 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, the data quality in most of the databases is high, such as LAKORA data, the pharmaceutical module, the MBDS (admissions data) or the vaccines registry, providing insight into a population of 0.65 million inhabitants. Fourth, the availability of cost data per patient and activity allows economic evaluation studies through the BARDENA SECA. Five, as a source of population records, it can contribute to pragmatic trials by providing the arm of patients on whom routine practice is performed, and the cost of developing research and the timing of access to the data is considerably lower than in experimental designs, such clinical trials. Six, in pharmacoepidemiology, the possibility of linking prescription and dispensation data at the individual level allows for an accurate analysis of drug utilisation, such as medication adherence studies. Seven, the BARDENA Core belongs to a suite that can interact with other integrated tools or with other information systems besides health, such as education, open government and transparency. Therefore, the information stored in the Bardena Core could be accessed by researchers via the Bardena diffusion product, which allows researchers to access the Bardena Core through a Tableau server, obtaining analysis tables for their research and producing the visuals of the key information of the BARDENA Core. Finally, the Bardena Suite could perform data mining of their data through theBARDENA Milenia tool, which permits performing data mining at the same time that the data is being populated. Its main functions are: discover processes that allow for the application of improved protocols, check the current processes, detect and remove bottlenecks, and simulate possible changes in the process estimating how it affects the system.

The Bardena Suite also has some weaknesses. Some of the databases that comprise the BARDENA Core 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 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 (pseudoonymized, 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 published articles.

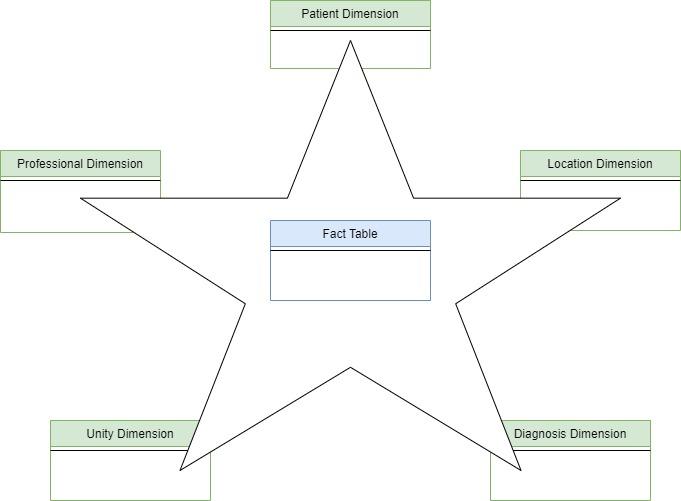


Figure 1. The star schema.

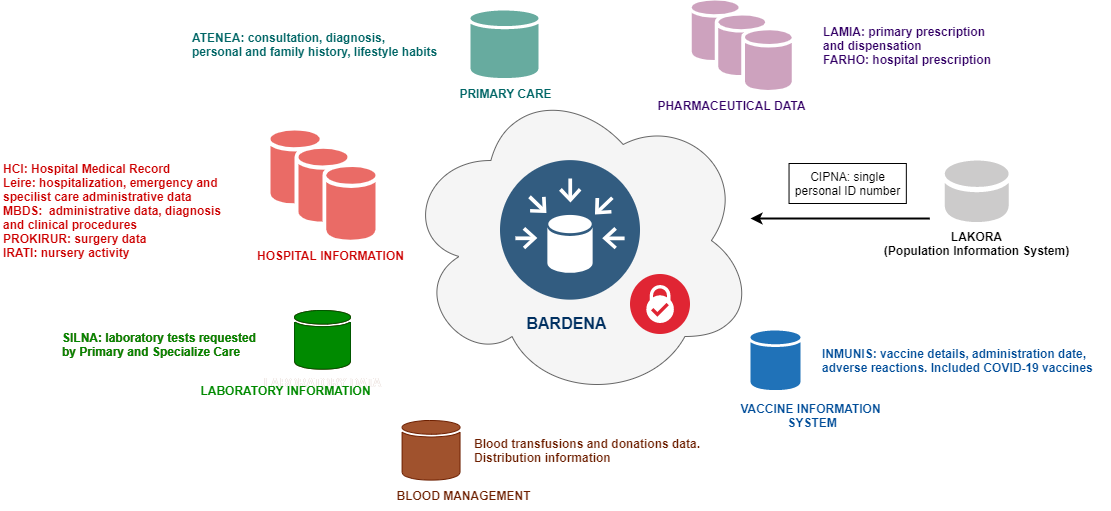


Figure 2. The BARDENA Core data sources.