

INSTITUTO NACIONAL DE ESTADISTICA



Hospital Morbidity Survey Methodology

Year 2015 and previous

Index

1	Background	3
1.1	Legal regulations	4
2	Survey objectives	7
3	Variables: definitions and concepts	8
4	Survey scope	9
5	Sample design	10
5.1	Selection of patients and estimate	11
5.2	Sampling errors	11
6	Information collection	14
7	Use of the information	15
8	Publication of results	16

1 Background

The Hospital Morbidity Survey, which is carried out in hospitals, is a sampling study that began in 1977, having established a working group at the beginning of 1974 at the suggestion of the INE with representatives from the Directorate-General for Health, the National Forecast Institute, the Central Commission for Hospital Coordination and the INE itself.

The meetings held by this work group resulted in the project that has led to the development of this survey since 1977. The objectives proposed in this survey are the following:

- To understand the general morbidity seen in hospitals in Spain according to the definitive diagnosis or the diagnosis on leaving hospital.
- To determine the average stay in hospital by type of diagnosis.
- To act as a basis for determining the scope of influence or degree of attraction of hospitals.
- To serve as a starting point for in-depth studies on specific diseases.

Beginning 1977, the fulfilment of these objectives was guaranteed by the introduction of two basic health documents, the Patient Register and the Patient Record, which systematically recorded all patients who were admitted and discharged from hospital.

In the last few years, and as result of the need to have a good information system that enables health systems to work and, in particular, the need to have a hospital information system adapted to the autonomous and functional reality of the National Health System that is comparable to other EU countries, the concept of the *Basic minimum set of hospital admissions data relating to the patient* (BMSD) has been developed.

This information system represents an ensemble of variables that can be obtained when the patient is admitted to hospital and which meets the needs of multiple users (management staff, planning staff, epidemiological staff, research staff, clinical research staff, etc.). The development of this information nucleus has implied a process of adaptation to the new system by hospitals, which means that today standardised hospital databases are available and provide clinical care, demographic and administrative information, which supports health planning and organisation.

The development of this new hospital information system based on BSMD and obtained from clinical records has been largely consolidated in public hospitals, but to a much lesser extent in private Spanish hospitals, where its introduction has taken longer and where many centres continue to use information based on the Patient Register.

On 8 June 2004, a collaboration agreement was signed between the National Statistics Institute and the Department of Health and Social Security of the Autonomous Government of Cataluña. Via the Catalan Health Service, the latter of these two bodies sends the INE (through its Provincial Delegations) electronic

information corresponding to the BSMD database from public and military hospitals included in its territorial scope and information from private hospitals contained in the aforementioned database.

Lastly, it is important to mention that the additional chapter in this standard is not an exceptional subject, but rather corresponds to a subsection of chapter 4 from the *ESS guidelines on seasonal adjustment*: the dissemination of the metadata from the seasonal adjustment. The INE standard has preferred to separate this matter in a different chapter, in order to grant it greater relevance, given that the INE considers that clarity is of vital importance in this matter.

1.1 Legal regulations

Royal Decree 1360/1976 of 21 May established the obligatory use in all Spanish hospitals, both public and private, of a standardised Patient Register, which required the following minimum data:

Admission order number, correlative for all patients.

Patient clinical record number.

Date and time of admission to the hospital centre.

Surname, name, sex, date of birth, civil status and usual residence of the patient.

Reason for admission to hospital, specifying if it was on doctor's orders, at own request, at family request, a legal or government order or another reason and whether it was urgent or ordinary.

Provisional or admission diagnosis.

Definitive or discharge diagnosis.

Date and reason for discharge, specifying whether it was for cure or healing, transfer to another centre, death or other reason.

Discharge order number.

The character reserved for the data allocated in the Register remained similarly determined when establishing that the responsible authorities would only be able to access the data for statistical or research purposes, maintaining in all cases the complete anonymity of the persons included in it.

This Royal Decree entered into force on 1 January 1977.

Subsequently, the Order of 14 November 1977 from the Ministry of Defence made the use of the Patient Register in military hospitals compulsory. This was used to extend the scope of implementation to all hospitals in the country.

The basic requirement for filling in the Patient Register is the *Patient Form*, which consists of two parts. Part A is filled in by the hospital admissions service and part B, which should be completed by the clinical or managerial department that

discharges the patient with the corresponding data (definitive diagnosis, date, reason), it should be returned to the admissions service when the patient is discharged.

Once the data has been transcribed to the Register, part A is separated from the form, which is filed by order of discharge. Part B is passed on in order to complete the patient's clinical documentation.

The Register continues to be in force in certain hospitals (private), but the need to have available instruments which facilitate managerial control in hospitals has influenced the adaptation of existing information systems to the new situation, incorporating concepts and tools which facilitate continual healing of the quality and similarity of the aforementioned systems.

In 1975, the Information and Documentation Committee for Science and Technology of the European Communities created a working group made up of experts and representatives from different countries in order to understand existing hospital information systems in centres.

Thus, different data (administrative and clinical) are selected from clinical records.

In 1981, the report on the *Basic Minimum Set of Data for Hospital Statistics in the European Communities* (BMSD-E) was presented. It was compiled by different experts on information systems and health statistics from different countries, including Spain.

The BMSD was supported by the EC, WHO-Europe, the EC Hospital Committee and the International Association for Medical IT. Subsequently, the Council of Europe included it as an integral part of the hospital information system.

In Spain on 14 December 1987, the Interterritorial Council from the Ministry for Health and Consumption approved the establishment of a Basic Minimum Set of Data (BMSD) for hospital admissions in agreement with that accepted both by the National Committee on Vital Statistics and Health of the USA and by the European Community Commission on the European Basic Minimum Set of Data.

The first Autonomous Community that established the BMSD was Cataluña by the Order of 23 November 1990, obliging both public and private hospitals to comply with this. The next institution to establish the BMSD for private and public hospitals was the National Institute of Health (INSALUD) by resolution from the General Secretariat of the National Health System on 24 January 1992.

Subsequently the Ministry of Health and Consumption, in collaboration with those Autonomous Communities with transferred health management and INSALUD, agreed to the creation of a Steering Committee for the development of BMSD within the National Health System at a meeting on 22 April 1992. Based on the development of this project, other Autonomous Communities published decrees or resolutions to introduce this information system (C. Valenciana, País Vasco, Galicia, Andalucía, Navarra, Canarias, Madrid, Murcia, Castilla-La Mancha, Castilla y León, Aragón, etc.).

The obligation to undertake a hospital discharge report on patients, which should comply with the BMSD, is established by the Ministry of Health and Consumption Order dated 6 September 1984.

The development of the hospital information system (administrative and clinical) based on the BMSD has made homogeneous information between the different hospitals and administrations available to the same extent that the Patient Register previously did.

Currently there is no general implementation in private hospitals, despite the regulation applying to public and private hospitals. In many of these hospitals the standardised Patient Register is still used, regulated by Royal Decree 1360/1976 of 21 May. Its use was obligatory for both public and private hospitals.

The confidentiality of patient discharge data, both by the BMSD system and through the Register, remains guaranteed by establishing in the respective regulations that the responsible authorities will only have access to the said data for statistical or research purposes. The information allocated will only be used for these purposes and the anonymity of the persons affected will be maintained at all times

At the time of selecting the model, all of the tests provided by the statistical package used must be checked, as well as the graphs from the series and from the residuals.

2 Survey objectives

The objectives of these statistics are the following:

1. To understand general morbidity in hospitals, according to the main diagnosis associated with admission.
2. To determine the average hospital stay according to the main diagnosis when discharged from hospital.
3. To act as a basis for epidemiological studies on certain diagnoses.
4. To understand the geographical distribution of morbidity in hospitals, based on the province of hospitalisation of the patient and their province of residence.
5. To supply a base of standardised information for Spanish hospitals, both public and private, on the basis of selected variables.
6. To act as a basis for national and international comparative studies.

3 Variables: definitions and concepts

The **target variables** are: The number of hospital admissions and average stay per admission.

Discharge is defined as the procedure whereby a patient admitted to a Health Centre or Establishment no longer occupies a hospital bed, as a result of healing, recovery, death, transfer or self-discharge.

For the purposes of the survey, a patient is considered to be any person who has been admitted into a hospital centre in order to be treated, diagnosed and observed as an in-patient and which has resulted in a hospital discharge.

The stay is defined as the discharge date minus the admission date, not computing stays equal to zero.

The **classification variables** used are: main diagnosis, type of hospital admission, reason for the discharge and province of hospitalisation. Regarding patients, age, sex and place of residence are collected.

Place of residence is regarded as where the patient normally lives, identified by the numeric code for the province.

The hospitalisation province is identified with the provincial digits corresponding to the hospital identification.

The main diagnosis is defined as the condition that after the necessary study, establishes that was the cause of the admission to the hospital according to the criterion of the clinical department or doctor who treated the patient, even though significant complications and even independent conditions arose during his/her stay. This definition is established in the Ministry of Health and Consumption Order of 6 September 1984. The classification used is the International Classification of Diseases v9, clinical amendment (ICD CA).

Up until the year 1991, the marital status of the patient has been collected as a variable, and whether the admission is considered a new or a former case, but with changes in the hospital information systems they experienced a high non-response rate, whereby as of said year they are no longer published in the Survey tables. Reason for admission (doctor's orders, own request, etc.) ceased to be published in 1998 for the same reason.

Since 2011 the data broken down are published identifying if the admission has been carried out in a private or in a public hospital according to the classification of functional dependency published in the National Catalogue of Hospitals. This publication is determined by the body or legal entity from whom the hospital depends, that is, the individual or the legal entity who has closest the power or jurisdiction, hierarchical or functional, to the Health Establishment. The classification of functional dependency of the hospitals with management legal status under the Law 15/1997, of 25th April, on Adaptation of new ways of management in the National Health System, and according to Autonomous legal developments, has been assigned to the corresponding Health Services.

4 Survey scope

Territorial scope: The survey scope is national, in other words, it covers all hospitals set up in the whole country. It includes the public and private sectors as well as military hospitals given the significant contingent of civilian personnel who are treated there.

For the purposes of this survey, hospital is considered to be any health centre, regardless of its denomination, whose final objective is to provide medical, surgical or medical-surgical care to patients admitted.

Not included, therefore, are nursing homes, orphanages, crèches, charity homes, psychological re-education institutions, etc., so as not to alter the suggested definition.

Time scope: Hospital discharges are considered when they have taken place during the survey reference year, regardless of the date of admission.

5 Sample design

When carrying out the sample design, the fact that the unit under study is the admitted and discharged patient has been taken into consideration. The latter of these is the point of access to the patient. Therefore, a stratified sample has been carried out in two stages where the hospitals constitute the first stage units and the patients the second stage or elemental units under study.

Hospitals have been stratified with regard to their objective (general, surgical, maternal), given that the patients' main diagnosis, which is the basic characteristic of the survey, is in direct relation to the type of care given, in other words, the objective. For these purposes, a total of 13 objectives have been considered. Hospitals that have more than one objective with an independent admissions department and where an information Register is requested for each one of them are considered as though there were as many hospitals as independent uses.

Conversely, in order to supply information broken down by province, each one of these has been considered as a subpopulation under study. Independent samples were obtained in each province.

The first stage unit sample, in other words, the hospitals, was exhaustive in the provinces where the number of centres is small and where there are different objectives, because it is necessary to rely on a representation of patients in each stratum. Stratum is understood to be a set of hospitals with the same objective within each province.

In those provinces where the number of centres with the same objective is high, a sample of hospitals has been selected according to their importance and the greater or lesser homogeneity of the diagnoses dealt with in each type of centre.

The sample of second stage units has been selected by bearing in mind the volume of patients discharged from hospital during the previous year. Therefore, every year the hospital is required to provide (whether it is included in the sample or not) the total number of discharges during the year and, in hospitals in which there is more than one admissions service, according to specialities, this data needs to be known for each of the independent information Registers.

When the hospital information is collected by means of the BMSD, a sample of patients is not extracted but rather the total number of patients discharged over a given period is collected.

The overall volume of the sample of hospitals collected amounts to 95.5 percent and, at present, the volume of patients considered amounts to 99.5 percent.

Hospital discharges corresponding to certain obligatory declaration diseases which, due to their nature, are either controlled numerically and one by one by the National Epidemiology Centre (botulism, cholera, triquinosis, measles, tetanus, leprosy, congenital rubella, congenital syphilis, neonatal tetanus, recurrent tick fever, anthrax, tularaemia, Q fever) or they are diseases without cases (diphtheria, poliomyelitis, rabies, yellow fever, plague, exanthemas typhus, chicken pox) or they correspond to diseases that are endemic to the general

region (boutonneuse fever, leishmaniasis, equinococosis, anthrax), are not collected.

5.1 Selection of patients and estimate

The selection of patients in each hospital included in the sample is made by applying a systematic sample with random start according to the patient discharge number.

Within each province the total number of patients presenting with a certain characteristic X is estimated by means of the following formula:

$$\widehat{X} = \sum_{j=1}^{13} \frac{N_j}{\sum_{i=1}^{H_j} N_{ij}} \sum_{i=1}^{H_j} \frac{N_{ij}}{n_{ij}} X_{ij}$$

in which sub index i refers to the hospital number and sub index j to the objective of the hospital (general, surgical, maternity, etc.) with:

X_{ij} = Number of patients in the sample that have characteristic X in hospital i with objective j.

N_{ij} = Total number of patients discharged during the year from hospital ij from the sample.

n_{ij} = Annual number of patients studied in the sample in hospital ij.

H_j = Number of hospitals studied in the sample with objective j.

N_j = Total number of patients discharged during the year in the set of population hospitals with objective j.

5.2 Sampling errors

The estimates obtained in any sampling study are affected by sampling errors, which are a result of only using a part of the group under study, as different samples will supply different estimates in general.

Logically, the larger the sample size and the more homogeneous the population under study, the fewer fluctuations there will be due to different possible samples. Or put another way, the fewer fluctuation measurements (sampling errors), the greater the reliability of the estimate.

One of the main advantages of the probable sample, which is used in this survey, in relation to other types of sample, is being able to determine the reliability of the estimates through the calculation of sampling errors.

Given that the sample of hospitals is practically exhaustive in nearly all provinces, the simplification can be made for the calculation of sampling errors that come exclusively from the sample of patients in each hospital. Therefore the absolute sampling error corresponding to a characteristic X has been estimated by means of the formula:

Within each province the total number of patients presenting with a certain characteristic X is estimated by means of the following formula:

$$\widehat{EM}(\widehat{X}) = \sqrt{\sum_{j=1}^{13} \sum_{i=1}^{H_i} \frac{N_{ij}^2}{n_{ij}^2} X_{ij} \left(1 - \frac{X_{ij}}{n_{ij}}\right)}$$

with the notation already established.

When calculating and publishing the sampling error, the statistician provides the user with a means of obtaining a numerical interval that has certain confidence (measured in probability terms) in containing the real value that we wish to estimate. Sample theory determines that, in the interval between the estimate less twice the absolute sampling error, and the estimate plus twice the absolute sampling error, there is 95% confidence in finding the real value or parameter that we want to estimate. From here on this interval is called the 95% confidence interval. It may be interpreted that, on average, of each 100 samples obtained under the same design and general conditions these confidence intervals obtained from each sample will contain the real value 95 times out of 100. From the estimate of a certain characteristic and its sampling error, other confidence intervals may be formed with similar interpretation. For example:

- Estimate \pm one time the sampling error. = 69 percent confidence interval.
- Estimate \pm 2.5 times the sampling error. = 99 percent confidence interval.
- Estimate \pm 3 times the sampling error. = 99.7 percent confidence interval.

As well as the absolute sampling errors, variation coefficients or variation coefficient errors or errors relative to the characteristics studied may be obtained, which express the sampling error as a percentage of the estimated value.

Thus, the variation coefficients provide a measure of the reliability of the estimate, but independently of the numerical value of the estimate (which does not occur with the sampling error), thus facilitating the immediate comparison of the reliability of the different estimates:

$$\widehat{CV}(\widehat{X}) = \frac{\widehat{EM}(\widehat{X})}{\widehat{X}} \cdot 100$$

High variation coefficients combine estimates and confidence intervals that may be so large as to invalidate the use of the data. In any case, it is for the user to determine if data is useful or not for making decisions in accordance with the degree of reliability that is required.

The estimate variation coefficient (or relative sampling error as a percentage of the estimate value) of the number of patients that appears in provincial table 22 and which has been spoken of until now, must not be confused with the population variation coefficient that appears in provincial tables 18 and 19 and national table 21. The population variation coefficient in these tables is a descriptive characteristic that allows us to compare the degree of homogeneity of the age variable of patients discharged with different diagnoses. Thus for example, if the average age variation coefficient of respiratory system patients is less than that for digestive system patients, this means that when the age variable is referred to, the population of patients is more homogeneous in the first group than in the second.

6 Information collection

The majority of hospitals have mechanised information available on discharges, based either on the BMSD system or on the Register. For these hospitals, INE agents electronically collect the necessary information for the HMS based on certain established Register characteristics and designs. Nevertheless, in order to achieve better use of available resources in a high number of hospitals and to not unnecessarily overload the informant hospital, the INE collects HMS information in the same format (BMSD) that the hospital uses and with the same periodicity in order to subsequently transfer it to its own design.

For those hospitals that do not have mechanised information available, the information is collected manually via questionnaire. This questionnaire is filled in directly using the data in the Register and the associated patient file, which is still in operation in some hospitals.

The handmade collection of information is carried out by INE agents through visits to the selected hospitals and the periodicity of the visits depends mostly on the work planning of the agents, as most of the hospital use informatic systems and the burden that this collection means is so little. Also, the periodicity of the visits vary in accordance with the size of the sample to be obtained from the different hospitals.

Therefore, in hospitals with samples of up to 100 patients, one visit a year is carried out; for samples with between 101 and 200 patients, there are two visits a year, one every six months; for samples with between 201 and 400 patients, there is a quarterly visit; in hospitals with a sample between 401 and 1,200 patients, a visit is made once a month and finally, for hospitals with a sample that is greater than 1,200 patients, the visit is fortnightly. This distribution of visits is carried out with the sole purpose of facilitating the information collection work, but any change in the periodicity of the visits does not affect the survey result, as the sample collected is that which is determined.

7 Use of the information

The majority of the information is collected electronically and the data are received from the provincial delegations via a computerised application that carries out a preliminary quality control of the information collected and speeds up the entry of information in central services. Only in those cases when the hospital does not have electronic information data is collected from the questionnaires that are manually completed.

The main diagnosis is published on a 3-digit level from the World Health Organisation International Classification of Diseases (ICD) 9th Clinical Modification Revision. A group of diagnoses has been used in the tables that deal with the social demand for information and the groups used by the OECD, Eurostat and the WHO.

Beginning with the 2002 and 2003 information, the number of diagnoses published increased and a new list of different ICD diagnostic groups was included, broken down into Chapters and large groups of diagnoses.

Likewise, as of 2003, age groups are extended (from 85 to 89 years old, from 90 to 94 years old, 95 years old and over) and new tables are published on a national and regional level for discharged patients, absolute stays and average stays, taking into consideration different *intervals for the stay variable* (from 1 to 30 days, from 31 days to 6 months, over 6 months to 1 year, over 1 year) and/or the sex, age group, main diagnosis and Autonomous Community of hospitalisation variables.

The new tables that consider *different intervals of stay* have been included with a view to separating the effect that patients discharged from long-term hospitals and institutions included in the Survey have on the total number of patients, absolute stays and average stays.

As of Hospital Morbidity Survey 2005, eight new tables were included, making up Chapters III and IV, which correspond to the International Shortlist of diagnostic groups (agreed by Eurostat, OCEC and WHO) for the dissemination of Hospital Morbidity data, dated 24 November 2006. This list allows for the homogenous comparison of figures and diagnoses among countries and includes, for individual diagnoses or groups of them, 149 items with a four-digit level of development according to International Classification of Diseases (ICD-9-CM). For more information, one may view the Eurostat website at <http://www.ec.europa.eu/eurostat>

As of Hospital Morbidity Survey 2007, the tables of Chapters III and IV are published with the updated Eurostat/OECD/WHO International Shortlist, dated 19 January 2008; while as of Hospital Morbidity Survey 2008, these same tables include the final International Shortlist, dated 11 October 2008. Amendments made in the 7th and 8th Editions of the ICD-9-MC are included in the 2010 and 2012 Morbidity Surveys, respectively, and the modifications made in the 9th edition are included since the 2014 survey.

8 Publication of results

Publication is annual, within year $t+1$ of the reference year of the survey. National and provincial tables are published with absolute figures of hospital discharges and stays, as well as the distributions and indices of the most relevant characteristics (sex, age, main diagnosis, etc.).

Beginning with the information corresponding to 1997, an electronic publication is available and the survey information may be accessed on the INE website.

The duly anonymous microdata file is also available and via express request, specific tables can be compiled for two, three or more variables, given that the high sampling volume (99%) for hospital discharges collected in the Survey in the last few years has allowed for minimising sampling errors.

Beginning with the information from the year 2005, the microdata are distributed with a 4-digit development level for the main diagnosis field (ICD-9-MC), permitting the possibility of making personalised requests with this detailed information.