Joint Data Collection on Non-Monetary Health Care Statistics

Joint Questionnaire 2020

GUIDELINES FOR COMPLETING THE OECD/EUROSTAT/WHO-EUROPE QUESTIONNAIRE 2020

QUESTIONNAIRE SENT: THURSDAY, DECEMBER 19, 2019
DEADLINE FOR RETURN: FRIDAY, FEBRUARY 28, 2020

CONTACTS:
Gaëlle Balestat, OECD (JQNMHC@oecd.org)
Ángeles Hermosa López, Eurostat (ESTAT-JQNMHCS@ec.europa.eu)
Ivo Rakovac, WHO-Europe (euJDC@who.int)

Table of Contents

****

1.   Introduction

This document provides the **Guidelines for Completing the OECD/Eurostat/WHO-Europe Joint Questionnaire on Non-Monetary Health Care Statistics (JQNMHC)**. The overall objective of this joint questionnaire is to provide internationally comparable data on key aspects of health care systems as they relate to health care resources and activities.

Content of the joint questionnaire

The joint questionnaire consists of <u>four parts</u>:
- ➢ **Health employment and education** (e.g. number of physicians, nurses, graduates, etc.);
- ➢ **Health workforce migration** (e.g. stock and flow of foreign-trained physicians and nurses);
- ➢ **Physical and technical resources** (e.g. number of beds, medical equipment, etc.);
- ➢ **Health care activities** (e.g. number of consultations, hospital discharges, surgical procedures, etc.).

Each part is associated to:
- One **Excel file with numerical data** (in Excel 97-2003 format);
- One **Word file with the documentation of** *Definitions, Sources and Methods*.

Countries in the **WHO European Region** (e.g. EU members or candidates, EFTA countries, Israel) are also asked to provide Hospital Discharge Data (HDD) in a separate **comma-delimited ASCII file** (.CSV).

<u>**Note for European countries**</u>: Some **additional variables** are requested by <u>Eurostat</u>. The additional data related to **health employment**, **physical and technical resources** and **health care activities** are collected in a separate Excel file, with the metadata collected in a separate Word file.

Deadline for returning the questionnaire

The deadline for returning the questionnaire is **FRIDAY, FEBRUARY 28, 2020**.

Contacts

For the 28 EU Member States, the four EFTA countries (Iceland, Liechtenstein, Norway and Switzerland), the five EU candidate countries (Albania, Montenegro, North Macedonia, Serbia and Turkey) and one potential candidate country (Bosnia and Herzegovina), the questionnaire (updated Excel and comma-delimited ASCII files as well as *Sources and Methods* in Word format) should be returned **via eDAMIS** (see the eDAMIS information provided separately).

For countries which are only members of OECD and/or WHO-Europe, the questionnaire should be sent directly to the corresponding contact(s) at OECD and/or WHO-Europe (see below).

Any additional correspondence with OECD and/or Eurostat and/or WHO-Europe (with other aims than submitting the questionnaire) should be sent to the following contacts in each organisation:

OECD
JQNMHC@oecd.org
Ms. Gaëlle BALESTAT
OECD Health Division
2, rue André-Pascal
75775 Paris CEDEX 16, France
Phone: (+33) 1 45 24 17 45

EUROSTAT
ESTAT-JQNMHCS@ec.europa.eu
Ms. Ángeles HERMOSA LÓPEZ
EUROPEAN COMMISSION
EUROSTAT - Unit F5 - Education, health and social protection
L-2920 Luxembourg
Phone: (+352) 4301-34462

WHO-Regional Office for Europe
euJDC@who.int
Mr. Ivo RAKOVAC
WHO Regional Office for Europe
UN City, Marmorvej 51, DK-2100 Copenhagen

2. Structure of the Joint Questionnaire on Non-Monetary Health Care Statistics

The joint questionnaire on non-monetary health care statistics is divided into four main parts: 1) Health employment and education; 2) Health workforce migration; 3) Physical and technical resources; and 4) Health care activities.

The joint questionnaire covers two types of data requests: 1) data requested by the three organisations for the commonly agreed variables (see Annex 1 for the complete list of variables); and 2) additional data requested only by Eurostat (see Annex 2 for the list of variables). Countries that are not reporting their data to Eurostat do not have to complete this additional request (see the list of all countries involved in the joint data collection in Annex 3).

Common OECD/Eurostat/WHO-Europe data request

The following files are jointly collected by the three organisations:

- **Health employment and education**
  - ➢ **CCC_HealthEmployment_Year.xls**: **Excel workbook** requesting *time series*.
  - ➢ **CCC_HealthEmployment_Year.doc**: **Word document** containing the **definitions** of all variables and requesting information on *Sources and Methods*.

- **Health workforce migration**
  - ➢ **CCC_WorkforceMigration_Year.xls: Excel workbook** requesting *time series*.
  - ➢ **CCC_WorkforceMigration_Year.doc: Word document** containing the **definitions** of all variables and requesting information on *Sources and Methods*.

- **Physical and technical resources**
  - ➢ **CCC_PhysicalResources_Year.xls**: **Excel workbook** requesting *time series*.
  - ➢ **CCC_PhysicalResources_Year.doc**: **Word document** containing the **definitions** of all variables and requesting information on *Sources and Methods*.

- **Health care activities**
  - ➢ **CCC_HealthActivities_Year.xls**: **Excel workbook** requesting *time series*.
  - ➢ **HDD_CC_Year.csv**: **comma-delimited ASCII file** for hospital discharge data (inpatient cases, day cases and bed-days) by diagnosis, by age group and by gender. These CSV files are requested only from countries belonging to the **WHO European region**.
  - ➢ **CCC_HealthActivities_Year.doc**: **Word document** containing the **definitions** of all variables and requesting information on *Sources and Methods*.

- **Pilot data collection in the 2020 joint questionnaire**
  - ➢ **New data sheets** (**PILOT_HospitalBeds**, **PILOT_Screening**, **PILOT_HospitalAggregates**) **included in CCC_PhysicalResources_Year.xls and CCC_HealthActivities_Year.xls excel workbooks**.
  - ➢ **CCC_PilotIndicators_Year.doc**: **Word document** containing the **definitions** of all the pilot variables and requesting information on *Sources and Methods*.

Note: **CCC** in the names of the files corresponds to the three-character country code (ISO 3166).

Eurostat additional data request

This additional data request needs to be completed by the 28 EU Member States, the four EFTA countries (Iceland, Liechtenstein, Norway and Switzerland), the five EU candidate countries (Albania, Montenegro, North Macedonia, Serbia and Turkey) and one potential candidate country (Bosnia and Herzegovina).

- **Health employment and education**, **Physical and technical resources**, **Health care activities**
  - ➢ **CCC_EurostatModule_Year.xls**: **Excel workbook** requesting *time series* (the additional data related to health employment, physical and technical resources, and health care activities are collected by Eurostat in a single Excel file).
  - ➢ **CCC_EurostatModule_Year.doc**: **Word document** containing the **definitions** of all variables and requesting information on *Sources and Methods*.
  - ➢ **HDD_CC_Year.csv**: the common **CSV file** with hospital discharge data by diagnosis code, by age group and by gender should also contain *data by region* (NUTS 2).

The following table summarises the files that should be completed by the countries, depending on the organisations to which they belong.

Summary table for data (XLS files) and metadata (DOC files) submission, by group of countries

	CCC_Health employment		CCC_Workforce migration		CCC_Physical resources		CCC_Health activities			CCC_Eurostat module	
	.XLS	.DOC	.XLS	.DOC	.XLS	.DOC	.XLS	.CSV (HDD)	.DOC	.XLS	.DOC
Countries reporting to OECD only (i.e. Australia, Canada, Chile, Japan, Korea, Mexico, New Zealand and the United States)	✓	✓	✓	✓	✓	✓	✓		✓		
Country reporting to WHO-Europe and OECD (i.e. Israel)	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Countries reporting to Eurostat, WHO-Europe and OECD (EU member states, EU candidate countries, EFTA countries)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: In addition, all countries are invited to complete the metadata file "CCC_PilotIndicators.doc" for the pilot variables.

Features of all Excel workbooks

All Excel workbooks contain worksheets with the following titles and functions:
- "**Country**" – the first worksheet is designed to collect information on the country respondent/national focal point and allows the respondent to provide any general or specific comment on the data collection.
- "**VariablesList**" – the second worksheet serves as a table of contents, summarising the data requested in the workbook.
- "**Definitions**" – the third worksheet provides the definitions for all variables. (The definitions are also available in the Word document requesting information on *Sources and Methods*).
- **Data worksheets** – one data worksheet is provided by main topic.

The Excel worksheets have been **protected** so that only cells of the time series can be filled/updated. This protection has been set to prevent any accidental changes in the format of the files (i.e. changing the structure of tables, by adding or deleting rows or columns for example, is not allowed). If really necessary, the protection can be removed easily as no password has been assigned (open the file, select the worksheet to unprotect, click the "Review" tab on the ribbon, and click on "Unprotect Sheet").

3.   General Instructions for Updating Data and Metadata

Updating the general information in Excel workbooks

National focal points should complete/update the respondent information in the "Country" worksheet. They are also allowed to provide in this worksheet any **comments** they might have regarding their data and metadata submission.

Correspondents are encouraged, but not requested, to update the '**Check-list**' column in the "VariablesList" worksheet. This check-list may help them follow their progress in updating the questionnaire. It may also provide useful information for the three organisations about the availability of some variables not completed in the initial submission.

Update of data in the Excel worksheets

All tables are pre-filled with the data provided to the three organisations in previous years. National focal points are asked to **check the data currently available and update them**. Please fill/update only the cells of the time series, i.e. send back the Excel worksheets in exactly the same structure in which you receive them.

If long time series are not readily available, national correspondents are invited to report the most recent data since 2000. If, in addition, data could be supplied in five year intervals (1980, 1985, 1990, 1995), this would enable the construction of internationally comparative tables at fixed years, spanning a wider time range.

All correspondents are asked to mark updates of data in **BOLD** or **COLOUR** in the Excel worksheets. This extra step is extremely useful for reviewing and processing the country's submissions more efficiently and communicating changes in the data.

Please do not write any comments into the Excel data sheets. All comments should be supplied separately (in the "Country" worksheet or by e-mail) and/or inserted into the documentation of *Sources and Methods* (see below). Although we are grateful for any additional supporting documentation on paper, the data update will rely **exclusively** on the electronic files of the questionnaire received from countries.

Years included in the 2020 Joint Questionnaire

Time series in the 2020 Joint Questionnaire should cover the period at least up to **2018**, and if possible also **2019**. The questionnaire includes a code "p" that can be used to report provisional estimates for 2019 if needed (see data codification below). National correspondents are encouraged to improve as much as possible the timeliness of their data by providing final or provisional data for 2019.

Data checks

The Joint Questionnaire includes a set of automatic checks that are designed to improve data quality and minimise errors. The main purpose of the data checks is to verify the internal consistency in the data reported in the questionnaire. Some checks verify that data correctly add up in one spreadsheet (e.g. for hospital beds), or that data are consistent from one sheet to another (e.g. for physicians). They calculate the difference between the data concerned; theoretically, the result should be equal to 0. Some other checks verify the data consistency/reliability (e.g. the number of practising physicians should be smaller than the number of physicians licensed to practice; percentages should not be higher than 100%; etc.).

All **data checks** are calculated in the **grey columns** labelled '**Data errors**' at the right of data tables. When there are errors in the data, they appear in ***bold italic and red font*** in the cells below the label 'Data errors' (see an example in Figure 1 below).

Figure 1. Example of a table with data checks

	Diagnostic exams																Data errors		
	CT exams						MRI exams						PET exams							
	Total	Code*	Hospital	Code*	Ambulatory care	Code*	Total	Code*	Hospital	Code*	Ambulatory care	Code*	Total	Code*	Hospital	Code*	Ambulatory care	Code*	CT MRI PET	
YEARS	Number		Number		Number		Number		Number		Number		Number		Number		Number			
1990																			.. .. ..	
1991																			.. .. ..	
1992																			.. .. ..	
1993																			.. .. ..	
1994																			.. .. ..	
1995	21000000		18400000		2600000		9100000		5100000		4000000								- - ..	
1996	22600000		19700000		2900000		9800000		5600000		4200000								- - ..	
1997	25100000		21600000		3500000		10900000		6300000		4600000								- - ..	
1998	26300000		22600000		3700000		11900000		6900000		5000000								- - ..	
1999	30600000		25800000		4800000		13800000		8000000		5800000								- - ..	
2000	34900000		29000000		5900000		15800000		9100000		6700000								- - ..	
2001	39600001		33100000		6500000		18000000		10200000		7800000								1 - ..	
2002	45400000		37900000		7500000		21900000		12400000		9500000								- - ..	
2003	50100000		41400000		8700000		24200000		12600000		11600001								- -1 ..	
2004	53900000		44300000		9600000		24700000		13100000		11600000								- - ..	
2005	57600000		47200000		10400000		25300000		13600000		11700000								- - ..	
2006	62000000		51000000		11000000		26600000		14600000		12000000								- - ..	
2007	68700000		55700000		13000000		27500000		14900000		12600000								- - ..	
2008	73100000		58400000		14700000		28400000		14200000		14210000								- *-10000* ..	
2009	77500000		61100000		16400000		29400000		13600000		15780000								- *20000* ..	
2010	81900000		63900000		18000000		30200000		14300000		15900000								- - ..	
2011																			.. .. ..	
2012																			.. .. ..	

Legend: - : no error (data do add up).   .. : missing data.   -1/1: small rounding error only.
Figure/text in bold, italic and red: there is an error (here, data do not add up for MRI exams in 2008 and 2009).

In the sheets "***Physicians by categories***" and "***Physicians by age and gender***", the sum of data should theoretically add up to the total number of ***practising*** physicians reported in the sheet "***Physicians***" (see below [Guidelines by Topic and Variable](#)). However, in some countries it is not possible to provide the number of physicians disaggregated by categories or by age and gender according to the "practising" concept. In this case, the sum of data by categories or by age/gender should add up to the total number of professionally active physicians or to the total number of physicians licensed to practice – depending on the concept used to report physicians by categories and by age and gender. Three check columns are then provided to compare the sum of data from the sheets "***Physicians by categories***" and "***Physicians by age and gender***" with the three variables from the sheet "***Physicians***", but only the appropriate data check should be considered.

National correspondents are strongly encouraged to use these data checking tools and to **correct any data inconsistencies** before sending back the questionnaire, or **provide explanation in the** *Sources and Methods* when the data do not add up or are not consistent.

Data codification

A tool to flag important information related to the data is included in the Joint Questionnaire. Correspondents will find in the Excel questionnaire additional columns (named "Code"), following each single data column, with the possibility of including four letters if necessary (or to revise them if they already exist). These four letters (b, d, e, p) should be used to reflect any of the four following data issues:
- **b** = **break in time series;**
- **d** = **deviation from definition;**
- **e** = **estimated data;**
- **p** = **provisional data** – ***to be used for Year-1 only*** (i.e. for 2019 in the 2020 Joint Questionnaire).

These four letters should be used accurately and only when necessary. For example, the "b" letter should be used when a change in the source or in the methodology has involved a significant break in the series. The letters intend to warn data users about important issues in the data. These issues should be briefly described in the *Sources and Methods*.

The "p" flags are not intended to stay for more than one year. They should be used **only for provisional data that are expected to be revised and replaced by final data** in the near future (i.e. during the current data collection or in the next round).

Breaks in time series

Breaks in time series occur with changes in sources or methodologies. While these changes in sources or methodologies may improve the quality of the data, they seriously limit the possibility to do any trend analysis. When adopting any new source or methodology, national correspondents are encouraged to revise the data for previous years to maintain some reasonably long time series. Some methods that can be used to overcome the breaks in time series are provided in Annex 4, but correspondents are welcome to use any other appropriate methods depending on the information available in their country. The adjusted data can be highlighted with an "e" code associated with the data (as described in Data codification above).

If it is not possible to adjust the time series, any significant breaks should be highlighted with a "b" code associated with the data (with some explanation for the breaks provided in the *Sources and Methods)*, or else the data before the break should be deleted.

Updating the documentation of *Sources and Methods*

National focal points are invited to document the *Sources and Methods* underlying the data for each variable based on a common structure agreed-upon by the three organisations (see below). The *Sources and Methods* have been pre-filled with information that countries have already provided to the international organisations. The electronic files to update the *Sources and Methods* are provided in DOC format, and should be **opened, updated and then returned using WORD** (or any text-processing software). Correspondents are asked to **use TRACK CHANGES MODE in Word to highlight all changes to the *Sources and Methods***.

Structure to follow for the collection of information on *Sources and Methods*

Source of data	1. Indicate the data source, i.e. the name of the agency and/or the complete citation of the publication. 2. Indicate the full title of the original survey collection, administrative source, database or publication. 3. Indicate if different sources were used for different years. 4. Indicate the reference period (e.g. annual average, data as of December 31, etc.). 5. Add URL for website where more information can be found.
Coverage	Indicate the data coverage if it is less than complete (geographical, population, institutions, etc). - Do the data cover the entire country or only some part(s)? (Please specify if the geographic coverage is partial.) - Do the data cover both the public and private sectors? (If not, please specify the limitation in coverage.) - Are there any other limitations in the data coverage (e.g. military services, prisons, social services)?
Deviation from definition	Indicate if the data supplied do not match the proposed definition (please specify).
Estimation method	Explain if data are an estimation, interpolation or any other relevant information.
Break in time series	Indicate if there is a break in the time series, due to changing definition, source or calculation method (please specify).

4.   Guidelines by Topic and Variable

The **definitions** for all variables are provided either in the Excel workbooks (see the sheet "Definitions") or in the Word document requesting information on *Sources and Methods*.

The definitions rely as much as possible on the following existing **international classifications**:

- International Classification of Diseases (ICD), WHO (http://www.who.int/classifications/icd/en/);
- International Classification for Health Accounts (ICHA), *A System of Health Accounts*, OECD, WHO and Eurostat (2011) (http://www.oecd.org/health/a-system-of-health-accounts-9789264116016-en.htm);
- International Standard Classification of Occupations (ISCO-08), ILO (http://www.ilo.org/public/english/bureau/stat/isco/index.htm).

Health employment and education

Most variables related to health employment (e.g. physicians, midwives and nurses, dentists, pharmacists, etc.) are requested according to three concepts:

- "**practising**" (i.e. health care professionals directly providing services to patients);
- "**professionally active**" (i.e. the "practising" category plus other health professionals working in administration and research who are not directly providing services to patients but for whom their medical/paramedical education is a prerequisite for the execution of the job);
- "**licensed to practice**" (i.e. entitled to practice as health professionals).

National correspondents are strongly encouraged to identify suitable data sources or new estimation methods in order to fill any persisting data gaps for the "**practising**" concept. This request concerns especially countries which have only submitted data for the "licensed to practice" concept. Priority should be given to practising **physicians** and **nurses**.

Physicians by age and gender

Since 2016, the data on physicians aged "65+" have been split into two age groups: "**65-74**" and "**75+**". This additional breakdown was designed to obtain more precise information about the growing number of physicians that continue to practice after age 65 and even in some cases after age 75. The age group "65-74" was pre-filled with the data previously reported in the group "65+". If not already done, national correspondents are invited to revise their data in order to reflect more accurately the new more specific age groups.

Data should be provided for **practising** physicians by age and gender, where possible. If the data are not available according to the "practising" category, they should be reported for professionally active physicians or physicians licensed to practise. Please clearly indicate in the *Sources and Methods* if your data refer to a different concept than "practising".

Physicians by categories

The common data collection includes three broad categories of doctors (corresponding to the ISCO-08 codes at the 4-digit level) and eight sub-categories:

1) **Generalist medical practitioners (ISCO-08 code: 2211)**
   1. General practitioners
   2. Other generalist (non-specialist) medical practitioners

2) Specialist medical practitioners (ISCO-08 code: 2212)
    3.   General paediatricians
    4.   Obstetricians and gynaecologists
    5.   Psychiatrists
    6.   Medical group of specialists
    7.   Surgical group of specialists
    8.   Other specialists not elsewhere classified

3) Medical doctors not further defined (ISCO-08 code: 2210)

A split was introduced in 2012 under the broad category "General medical practitioners" to distinguish more clearly the number of "general practitioners" ("family doctors"), providing **continuous medical care** to individuals and families most often in primary care sector, from other generalists/non-specialists who may be working in hospitals.

Theoretically, the two first broad categories ("Generalist medical practitioners" and "Specialist medical practitioners") should cover all physicians. In accordance with ISCO-08 definition, the last category "**Medical doctors not further defined**" should be used only if some doctors cannot be classified under "Generalist medical practitioners" or "Specialist medical practitioners". Correspondents are strongly encouraged to verify that physicians have been properly allocated to the specific categories, and to make any adjustments if necessary.

The table on physicians by categories should cover the whole physician workforce, including **medical interns and residents** as well as physicians without specialisation. The sum of the three broad categories should add up to the total number of (practising) physicians. Medical interns and residents who are pursuing postgraduate training should be allocated in their area of training. If they have not chosen a specialisation yet, they should be reported in the category "Other generalist (non-specialist) medical practitioners". Physicians in training who cannot be split according to the specialty in which they are training should be reported in the category "Medical doctors not further defined".

Some guidelines are provided to help classify different specialties, especially between the "**medical group**" and the "**surgical group**". Please indicate in the *Sources and Methods* which specialties have been allocated to these two broad categories, particularly if the allocation does not correspond to the proposed guidelines.

Data for different categories of doctors should be provided for **practising** physicians, where possible. If the data are not available according to this concept, they should be reported for professionally active physicians or physicians licensed to practise. Please clearly indicate in the *Sources and Methods* if your data refer to a different concept than "practising".

To avoid double counting doctors with more than one specialty, the following criteria are proposed to report each doctor only once:

    1)   the predominant (main) area of practice of doctors; or
    2)   the last specialty for which they have received registration.

Midwives, Nurses and Caring personnel (personal care workers)

Data are collected separately for midwives and nurses. If data are not available separately for midwives, the total number of midwives and nurses should be reported as **nurses**. This should be clearly mentioned in the *Sources and Methods*.

The main priority is to report data on the **total number of nurses**, regardless of their category or level. When relevant and possible, national correspondents are also invited to provide the breakdown between "**Professional nurses**" and "**Associate professional nurses**" in those countries where two levels of nurses exist. The ISCO-08 definitions are provided to guide this data collection. If this breakdown is provided, please note in the *Sources and Methods* which categories of nurses have been assigned to each of these two categories.

The ISCO-08 definitions are also provided to guide the data collection of "**Caring personnel (personal care workers)**", who may also be referred as "nursing aides" or given other titles in different countries. These caregivers are not nurses, but they do provide personal care to patients in institutions or at home. Please indicate in the *Sources and Methods* the categories of workers reported under this item.

Hospital employment

The main priority is to collect data on the **total** number of people working in hospitals on a **head count** basis. Data for six categories of hospital workers are also requested. In addition to head counts, data are also collected on a **FTE basis**. Three methods to convert head counts into FTE data are proposed below, and national correspondents may choose one of them to do the conversion, depending on the availability of detailed data on actual/usual or contractual hours of work.

1) **For countries which have detailed data on actual or usual working hours**

Full-time equivalent (FTE) employment should be measured by the number of hours actually or usually worked divided by the average number of hours worked in full-time jobs.

For example: if the standard working hours for a full-time job in the country is 40 hours per week, and the actual or usual working hours of a doctor or a nurse in hospital is 30 hours, s(he) should be counted as 0.75 FTE. If s(he) works 50 hours, s(he) should be counted as 1.25 FTE.

2) **For countries which only have detailed data on contractual working hours**

A worker with a full-time employment contract should be counted as 1 FTE. Concerning workers who do not have a full-time employment contract, full-time equivalent should be measured by the number of hours of work mentioned in each contract divided by the normal number of hours worked in full-time jobs.

For example: if the standard working hours for a full-time job in the country is 40 hours per week, and if the contract of a nurse is 30 hours per week, s(he) should be counted as 0.75 FTE.

3) **For countries which do not have any detailed information on working hours**

A worker with a full-time employment contract should be counted as 1 FTE. Concerning workers with part-time contracts, the practice in many countries is simply to consider that 2 part-time workers = 1 FTE.

Graduates

This part of the data collection covers medical graduates, dentist graduates, pharmacist graduates, midwife graduates and nursing graduates.

Regarding **nursing graduates**, the main priority is to collect their **total number**, regardless of their category or level. When relevant and possible, national correspondents are also invited to provide the breakdown between graduates from "professional nurse" programmes and "associate professional nurse" programmes.

Health workforce migration

The main purpose of this part of the Joint Questionnaire (which was added in 2015) is to improve the monitoring of international health workforce migration through the collection of a minimum dataset that is relevant to both source and destination countries.

The main features of the data collection are that it:

- focuses on **doctors** and **nurses** only;
- focuses mainly on the **place of training** (defined as the place of first qualification);
- collects **immigration** data from destination countries by all countries of origin, based on available national sources (e.g., professional registries, specific surveys of health personnel);
- collects data based on measures of total **stock** and annual **inflows**.

In 2020, the list of countries/places of training has been slightly updated (e.g. renaming of a few countries such as North Macedonia, resulting in the reordering of some countries, etc.). National correspondents are therefore invited to be careful when completing this table. ISO codes are available and may be used for automated data filling.

Countries which have not submitted data yet are strongly encouraged to look at possible data sources that might be used to fill this gap and to provide as much as possible data by countries of origin. The data collection also allows a reporting of the overall number of "domestic-born but foreign-trained" doctors and nurses (people born in a country who went to study in another country but have come back afterwards to practice in their home country) separately from the "foreign-born and foreign-trained". All national correspondents are invited to provide this useful additional information when possible.

Physical and technical resources

Hospitals

The total number of hospitals should include all types of hospitals, following the International Classification for Health Accounts 2011 (including HP1.1 General hospitals, HP1.2 Mental health hospitals, HP1.3 Other specialised hospitals). Although the priority is to collect data on the **total number of hospitals**, more specific information is also requested for general hospitals (HP1.1) and for the breakdown between public, not-for-profit and for-profit private hospitals, where possible.

Hospital beds

The data collection on hospital beds by categories was revised in 2016 to make it more consistent with the *System of Health Accounts* (SHA) classification that is used internationally to collect data on hospital expenditure. The two main new features of the revised data collection specifications were:

1) the addition of a column to allow countries to report data on **rehabilitative care beds** separately (rather than reporting them either in curative care beds, long-term care beds or other beds as was done before);
2) the inclusion of **psychiatric care beds** in the curative care, rehabilitative care and long-term care categories (to avoid making any distinction in the availability of beds for somatic or psychiatric reasons), while at the same time keeping a separate data collection for all psychiatric care beds.

For those countries that have difficulties in separating curative care beds from rehabilitative care beds, the proposed guideline is to report these rehabilitative care beds in the "curative care" category.

For countries that may not have sufficient information for a detailed breakdown of psychiatric care beds into curative, rehabilitative and long-term care beds categories, the currently proposed main guideline is also to allocate these psychiatric care beds in the "curative care" category. However, if this allocation is considered to be inappropriate, countries can also report these psychiatric care beds in the "other beds" category. (In any case, correspondents should ensure that "Curative care beds" reported is consistent in scope with the indicators reported in the "health care activities" module under the hospital aggregates for "Curative (acute) care": discharges, bed-days, ALOS and occupancy rate).

The data collection on hospital beds aims to cover all types of hospitals (covering HP1.1 General hospitals, HP1.2 Mental health hospitals, HP1.3 Other specialised hospitals). In addition to the breakdown by function of care (curative care, rehabilitative care, long-term care and other), a breakdown between public, not-for-profit and for-profit private hospitals is also requested, where possible. National correspondents are invited to note in the *Sources and Methods* any limitation in hospital coverage.

Hospital beds – PILOT data collection

The breakdown of psychiatric care beds into curative, rehabilitative and long-term care beds categories represents a challenge for many countries. Different methods used to allocate psychiatric care beds (about half of countries are able to distinguish them between two or three functional categories while another half cannot separate them and include all of them either under "curative care beds" or "other hospital beds") leads to significant data comparability issues.

The aim of this pilot data collection is to increase the transparency and the comparability of the beds numbers, while maintaining the initial objective of consistency with the SHA classification. To do so, data are requested at a finer level, with a breakdown between **somatic care beds** and **psychiatric care beds**, for the total number of hospital beds but also for each functional category (curative, rehabilitative and long-term care beds).

Based on previous data collections, all (or most) countries should be able to provide the total number of somatic care beds and the total number of psychiatric care beds (respectively columns D and F in the Excel worksheet "PILOT_HospitalBeds"). All (or most) countries should also be able to report the numbers of somatic care beds by function (columns J, P, V, AB). Countries which are able to distinguish psychiatric care beds by function are invited to fill in the columns L, R, X and AD, as appropriate. Countries which cannot do this breakdown are invited to report the total number of psychiatric care beds in the column AD.

When data is not available for a specific cell, this cell should remain empty. When there are no beds for a specific category, "0" should be reported. It is important to follow this recommendation, in order to distinguish real null values from missing data and interpret correctly the results.

Beds in residential long-term care facilities

The data collection on beds in residential long-term care facilities should include all types of nursing and residential care facilities as defined in the HP.2.1 and HP.2.9 categories of the International Classification for Health Accounts 2011.

Medical technology

The data collection on medical technology includes six types of diagnostic and therapeutic equipment (the data requests for angiography units and lithotriptors were discontinued in 2018 because of more limited data availability for these two technologies). The aim is to collect data on the total number of equipment in all health care facilities, including hospitals and providers of ambulatory health care. If the data in your country are only available for hospitals or ambulatory care providers, these data should be reported in the corresponding column.

Health care activities

Consultations

The aim is to collect the number of consultations with doctors and dentists *in all settings*, including in outpatient departments in hospitals. The data may either come from administrative sources or surveys. It is requested as a **rate per capita**.

Immunisation against influenza

The objective is to collect the percentage of elderly people who have been vaccinated **against seasonal influenza** during the last twelve months.

Screening

The data collection on breast cancer (mammography) screening and cervical cancer screening includes a breakdown between survey data and programme-based data. Correspondents are invited to update the data series and to note any deviation from the proposed definitions in the *Sources and Methods*.

Screening – PILOT data collection

Following breast and cervical cancer screening, many countries have introduced a **colorectal cancer screening** programme as a cost-effective public health intervention. In the past, colorectal cancer screening policies varied widely across countries but recent research show some convergence and countries follow similar policies:

- Most countries have an organised population-based colorectal cancer screening programme.
- In most countries, target people are in their 50s and 60s, and the target age range of 50-74 is used most frequently.
- Most countries use faecal immunochemical test (FIT) as the initial screening test for colorectal cancer and some countries use another stool-based test, guaiac faecal occult blood test (gFOBT).
- Several countries use endoscopic cancer screening such as (flexible) sigmoidoscopy and colonoscopy as the initial screening test for colorectal cancer, and other methods including computed tomography colonography and multitarget stool DNA are still uncommon.
- While the majority of countries provide only one screening method, a few countries offer multiple methods as the initial screening tests for colorectal cancer.
- In many countries, the screening interval of stool-based tests such as FIT and gFOBT is every year or every two years, and the interval for endoscopic colorectal screening such as sigmoidoscopy and colonoscopy is usually every 10 years or once in the lifetime. In some countries, the screening interval differs by age group.

As is the case for breast and cervical cancer screening, many countries monitor and report the coverage of colorectal cancer screening programmes by using programme data. Many countries also collect data on colorectal cancer screening through surveys.

This pilot data collection on colorectal cancer screening aims to collect the following two indicators:

1. Colorectal cancer screening rate based on <u>programme</u> data
2. Colorectal cancer screening rate based on <u>survey</u> data

Data will be requested for women, men and the total population. Basic information on colorectal cancer screening policies used in each country (e.g. **target age range**, **screening method**, etc.) and data sources and methods will also be collected. The goal of this pilot project is to assess the availability and comparability of colorectal cancer screening rates across countries. Results will be reported and the regular collection of colorectal cancer screening rates may be proposed after the 2020 data collection and validation process.

Hospital aggregates

This request is designed to collect data on total (aggregate) hospital activities, including a specific focus on curative care for some indicators. It complements the more disaggregated data collection on hospital discharges and bed-days (or length of stays) by diagnostic groups (see the following section). Besides providing an opportunity for countries to report such aggregated data if they are not able to report the more disaggregated data, this data request allows an assessment of the data coverage of the more disaggregated data supplied by diagnostic groups. This data collection also serves to collect a few aggregate indicators that can be used in analysis of hospital efficiency or utilisation rate (e.g., occupancy rates of hospital beds).

The data should cover all hospitals, and the definition of curative care is based on the standard SHA definition.

National correspondents are invited to review the consistency between the inpatient aggregated data and the disaggregated data on hospital discharges and bed-days (or length of stays) by diagnostic groups. When discrepancies exist between both data sets, some explanation should be provided in the *Sources and Methods* (e.g. differences in data sources, in settings/services coverage, etc.).

Hospital aggregates – PILOT data collection

The data reported under the curative care activities indicators ("health care activities" module) should have the same coverage as data reported under curative care beds ("physical and technical resources" module). Following the proposal to collect hospital beds (*of which* curative care beds) broken down between somatic care and psychiatric care, this pilot data collection aims to distinguish **somatic curative care activities** from **psychiatric curative care activities**, for four indicators: hospital discharges, bed-days, ALOS (directly calculated in the spreadsheet "PILOT_HospitalAggregates") and occupancy rates. National correspondents are invited to briefly describe in the *Sources and Methods* the methodology for differentiating between somatic and psychiatric activities (in CCC_PilotIndicators_2020.doc).

Discharges, bed-days, average length of stay by diagnostic categories

Data on hospital discharges, bed-days and average length of stay (ALOS) by diagnostic categories are collected in two different ways, depending on countries.

1) **Countries in the WHO European Region** (e.g. EU members or candidates, EFTA countries, Israel)

For these countries, the joint data collection on hospital discharge data (HDD) follows the approach formerly adopted by WHO-Europe and Eurostat. The data are requested for discharges (inpatient cases and day cases) and bed-days by ICD-9, ICD-10 or ISHMT code[1], by age group and by gender. The data collection on discharges and bed-days allows the calculation of average length of stay.

All countries in the WHO European Region are encouraged to supply their data using a separate **comma-delimited ASCII file** (.CSV). Annex 5 provides all information regarding the submission of the HDD files.

---

[1]   ICD: International Classification of Diseases. ISHMT: International Shortlist for Hospital Morbidity Tabulation.

2) **OECD non-European countries** (Australia, Canada, Chile, Japan, Korea, Mexico, New Zealand and the United States)

The previous OECD data collection approach is still used for non-European countries. Data are requested for hospital discharges (inpatient cases only) and ALOS by ISHMT code, for the total population only. They are collected in the common Excel workbook, in the sheets "Discharges" and "ALOS".

Hospital discharges should cover all hospitals (HP.1). They should include deaths in hospital and transfers to another hospital, but exclude transfers to other care units within a same hospital. Same day separations of inpatient cases (e.g. inpatients who die or are transferred to another hospital on the day of their admission) should be included in the number of discharges. The corresponding number of bed-days should be set to 1 day.

Countries submitting data according to ISHMT should exclude the "external causes of morbidity and mortality" (V, W, X and Y codes, chapter 20 in ICD-10), in order to avoid any double-counting of injuries which are already recorded in "injury, poisoning and certain other consequences of external causes" (S and T codes, chapter 19 in ICD-10). Discharges with unknown diagnosis should be allocated to ISHMT code 1803 (or ICD-10 code R69).

The mapping between the International Shortlist for Hospital Morbidity Tabulation (ISHMT) with ICD-10 and ICD-9 codes is available at: http://stats.oecd.org/HEALTH_QUESTIONNAIRE/ISHMT/JQNMHC_ISHMT.pdf. National correspondents are invited to update the *Sources and Methods* for these variables in the common Word file CCC_HealthActivities_Year.doc. Guidelines for completing these metadata are provided in Table 6 of Annex 5.

Diagnostic exams

The chapter includes three variables: the number of **Computed Tomography (CT) exams**, **Magnetic Resonance Imaging (MRI) exams** and **Positron Emission Tomography (PET) exams**. The aim is to collect data on all diagnostic exams, including those carried out in hospitals and outside of hospitals. To allow for a more accurate reporting in those countries which may only have partial data coverage, the data collection includes a breakdown between exams in hospitals and in ambulatory care settings.

Surgical procedures

A common shortlist of 17 surgical procedures plus 5 subgroups was initially adopted by the three organisations. However, the data requests for transplantation of kidney and secondary hip replacement have been discontinued in 2018 (because information on kidney transplants is already collected by other international organisations; data availability on secondary hip replacement is somewhat more limited than for other procedures and the total number of hip replacement seems to provide more relevant information). Hence, the current common shortlist contains 16 surgical procedures plus 4 subgroups.

In addition to the total number of procedures, a split between **inpatient cases** and **day cases** is requested for each type of surgical procedure. For cataract surgery and tonsillectomy, the number of **outpatient cases** in hospital and outside hospital is also requested, where possible, in order to provide more complete coverage of same-day surgery. Following the SHA definitions, day cases are defined as admitted patients, while outpatient cases are defined as non-admitted patients.

In reporting data on the shortlist of procedures, correspondents are invited to follow as much as possible the proposed counting method which is designed to improve data comparability by avoiding double-counting procedures when more than one code may be used depending on each national classification system. Correspondents are invited to report either **a count of the number of patients** who have received a given procedure or to **only report one code per procedure category for each patient**. For example, if a percutaneous coronary intervention including a coronary stenting is recorded as two separate codes as is the case under ICD-9-CM, only one code/procedure should be reported; if a cataract surgery is recorded as two procedures (removal of the lens and insertion of the artificial lens), only one procedure should be counted. It is particularly important for correspondents to follow this proposed counting method for percutaneous coronary interventions, coronary bypasses and cataract surgery. Correspondents are invited to briefly describe in the *Sources and Methods* the counting method used.

The **common shortlist** of surgical procedures is provided in Table 1 of the document available at the link: http://stats.oecd.org/HEALTH_QUESTIONNAIRE/Surgical%20procedures/JQNMHC_MAPPING%20ICD-9-CM.pdf.

Eurostat additional data collection

Health workers at regional level

The data collection on health workers at regional level should include **practising** physicians. If data are not available according to this concept, they should be reported for professionally active workers or workers licensed to practise. Please clearly indicate in the *Sources and Methods* if your data refer to a different concept than "practising".

The health care staff should be allocated to the place (region) where the health care service is provided.

For the regional level the **Nomenclature of Units for Territorial Statistics (NUTS)**[2] should be used for referencing the subdivisions of countries for statistical purposes. The NUTS level 2 is required for this data collection (see http://ec.europa.eu/eurostat/web/regions-and-cities/overview).

The NUTS classification was revised in 2016. Please note that following this revision, the whole time series should be revised for Ireland, France, Lithuania, Hungary, Poland and the United Kingdom in order to take into account the modifications due to this revision. These four countries are invited to re-calculate back to 1993.

Hospital beds at regional level

The data collection on hospital beds aims to collect data for all functions of care, in all types of hospitals as listed in the International Classification of Health Accounts 2011 (including HP1.1 General hospitals, HP1.2 Mental health hospitals, HP1.3 Other specialised hospitals). National correspondents are invited to note in the *Sources and Methods* any limitation in hospital coverage.

The hospital beds should be allocated to the place (region) where the health care service is provided.

Hospital technical resources

The data collection of hospital technical resources includes six types of resources available in hospitals. The data should cover both public and private hospitals. Any limitation in the data coverage should be clearly explained in the *Sources and Methods*.

---

[2] See http://ec.europa.eu/eurostat/web/nuts/background for general information.

Discharges and bed-days by diagnostic categories at regional level

The data are requested for discharges (inpatient cases and day cases) and bed-days by diagnosis code, by age group, by gender and by region.

Correspondents are encouraged to supply their regional data using the same **comma-delimited ASCII file** (.CSV) as for the national data (corresponding to the common part of the data collection). Annex 5 provides all information regarding the submission of the hospital discharge data (HDD) files.

Hospital discharges of non-resident patients

Correspondents are requested to provide any data available on non-residents discharged in their country according to the requested record structure. Data disaggregated at NUTS2 level is not requested anymore. The table has been revised and contains the data provided at country level since 2005.

ANNEX 1. LIST OF COMMON VARIABLES INCLUDED IN THE JOINT QUESTIONNAIRE ON NON-MONETARY HEALTH CARE STATISTICS

Health employment and education

Physicians (head counts)
Practising physicians
Professionally active physicians
Physicians licensed to practice

Physicians by age and gender
Physicians by age group and by gender

Physicians by categories
Generalist medical practitioners
  - General practitioners
  - Other generalist (non-specialist) medical practitioners
Specialist medical practitioners
  - General paediatricians
  - Obstetricians and gynaecologists
  - Psychiatrists
  - Medical group of specialists
  - Surgical group of specialists
  - Other specialists n.e.c.
Medical doctors not further defined

Midwives (head counts)
Practising midwives
Professionally active midwives
Midwives licensed to practice

Nurses (head counts)
Practising nurses
  - Professional nurses, practising
  - Associate professional nurses, practising
Professionally active nurses
  - Professional nurses, professionally active
  - Associate professional nurses, professionally active
Nurses licensed to practice
  - Professional nurses, licensed to practice
  - Associate professional nurses, licensed to practice

Caring personnel (personal care workers) (head counts)
Practising caring personnel (personal care workers)
Professionally active caring personnel (personal care workers)

Dentists (head counts)
Practising dentists
Professionally active dentists
Dentists licensed to practice

Pharmacists (head counts)
Practising pharmacists
Professionally active pharmacists
Pharmacists licensed to practice

Physiotherapists (head counts)
Practising physiotherapists

Hospital employment (head counts and FTE)
Total hospital employment
  - Physicians employed by hospital
  - Professional nurses and midwives employed by hospital
  - Associate professional nurses employed by hospital
  - Health care assistants employed by hospital
  - Other health service providers employed by hospital
  - Other staff employed by hospital

Graduates
Medical graduates
Dentists graduates
Pharmacists graduates
Midwives graduates
Nursing graduates
  - Professional nursing graduates
  - Associate professional nursing graduates

Health workforce migration

Foreign-trained doctors
  - Foreign-trained doctors, stock
  - Foreign-trained doctors, annual flow

Foreign-trained nurses
  - Foreign-trained nurses, stock
  - Foreign-trained nurses, annual flow

Physical and Technical Resources

Hospitals (HP.1)
Hospitals
  - Publically owned hospitals
  - Not-for-profit privately owned hospitals
  - For-profit privately owned hospitals
General hospitals

Hospital beds (HP.1)
Total hospital beds
  - Curative (acute) care beds
  - Rehabilitative care beds
  - Long-term care beds
  - Other hospital beds

  - Psychiatric care beds

  - Beds in publically owned hospitals
  - Beds in not-for-profit privately owned hospitals
  - Beds in for-profit privately owned hospitals

Hospital beds (HP.1) – PILOT
Total hospital beds
  - Somatic care beds
  - Psychiatric care beds

  - Curative (acute) care beds (total and breakdown between somatic care and psychiatric care)
  - Rehabilitative care beds (total and breakdown between somatic care and psychiatric care)
  - Long-term care beds (total and breakdown between somatic care and psychiatric care)
  - Other hospital beds (total and breakdown between somatic care and psychiatric care)

Beds in residential long-term care facilities (HP.2)
Beds in residential long-term care facilities

Medical technology (HP.1, HP.3 and HP.1+HP.3)
Computed Tomography Scanners
Magnetic Resonance Imaging Units
Positron Emission Tomography scanners
Gamma cameras
Mammographs
Radiation therapy equipment

Health care activities

AMBULATORY CARE
Consultations
Doctor consultations (in all settings)
Dentist consultations (in all settings)

Immunisation
Immunisation against influenza (population aged 65+)

Screening (survey and programme data)
Breast cancer screening (% of females aged 50-69)
Cervical cancer screening (% of females aged 20-69)

Screening (survey and programme data) - PILOT
Colorectal cancer screening (% of females, males and both)

HOSPITAL CARE
Hospital aggregates
Inpatient care discharges (all hospitals)
Inpatient care ALOS (all hospitals)

Curative care discharges
Curative care bed-days
Curative care ALOS
Curative care occupancy rates

Hospital aggregates - PILOT
Somatic curative (acute) care
  - Discharges
  - Bed-days
  - ALOS
  - Occupancy rates
Psychiatric curative (acute) care
  - Discharges
  - Bed-days
  - ALOS
  - Occupancy rates

Hospital discharges by diagnostic categories
Inpatient cases (*requested from all countries*)
Day cases (*requested from countries in WHO European Region*)

Hospital bed-days by diagnostic categories
Hospital bed-days (inpatient cases) (*requested from countries in WHO European Region*)

Hospital ALOS by diagnostic categories
Hospital average length of stay (inpatient cases) (*requested from OECD non-European countries*)

PROCEDURES
Diagnostic exams (HP.1, HP.3 and HP.1+HP.3)
Computed Tomography (CT) exams
Magnetic Resonance Imaging (MRI) exams
Positron Emission Tomography (PET) exams

Surgical procedures (shortlist of 16 procedures + 4 subgroups)
Total
  - Inpatient cases
  - Day cases
  - Outpatient cases (collected for 2 procedures only)

ANNEX 2. LIST OF EUROSTAT ADDITIONAL VARIABLES INCLUDED IN THE JOINT QUESTIONNAIRE ON NON-MONETARY HEALTH CARE STATISTICS

Health workers at regional level
Physicians

Hospital beds at regional level
Total hospital beds

Hospital technical resources
Operation theatres in hospital
Day care places altogether
Surgical day care places
Oncological day care places
Psychiatric day care places
Geriatric day care places

Hospital discharges by diagnostic categories at regional level
Inpatient cases
Day cases
Hospital bed-days (inpatient cases)

Hospital discharges of non-residents patients
In-patient cases
Day cases
Hospital bed-days (inpatient cases)

Note: These additional variables should be completed by EU Member States, EFTA countries and EU candidate and potential candidate countries (see list of countries in Annex 3).

ANNEX 3. LIST OF MEMBER STATES OF OECD, EUROSTAT AND WHO-EUROPE

Country	OECD	EUROSTAT	WHO-EUROPE
Albania		(1)	✓
Andorra			✓
Armenia			✓
Australia	✓		
Austria	✓	✓	✓
Azerbaijan			✓
Belarus			✓
Belgium	✓	✓	✓
Bosnia and Herzegovina		(1)	✓
Bulgaria		✓	✓
Canada	✓		
Chile	✓		
Croatia		✓	✓
Cyprus		✓	✓
Czech Republic	✓	✓	✓
Denmark	✓	✓	✓
Estonia	✓	✓	✓
Finland	✓	✓	✓
France	✓	✓	✓
Georgia			✓
Germany	✓	✓	✓
Greece	✓	✓	✓
Hungary	✓	✓	✓
Iceland	✓	(1) (2)	✓
Ireland	✓	✓	✓
Israel	✓		✓
Italy	✓	✓	✓
Japan	✓		
Kazakhstan			✓
Kyrgyzstan			✓
Korea	✓		
Latvia	✓	✓	✓
Liechtenstein		(2)	
Lithuania	✓	✓	✓
Luxembourg	✓	✓	✓
Malta		✓	✓
Mexico	✓		
Monaco			✓
Montenegro		(1)	✓
Netherlands	✓	✓	✓
New Zealand	✓		
North Macedonia		(1)	✓
Norway	✓	(2)	✓
Poland	✓	✓	✓
Portugal	✓	✓	✓
Republic of Moldova			✓
Romania		✓	✓
Russian Federation			✓
San Marino			✓
Serbia		(1)	✓
Slovak Republic	✓	✓	✓
Slovenia	✓	✓	✓
Spain	✓	✓	✓
Sweden	✓	✓	✓
Switzerland	✓	(2)	✓
Tajikistan			✓
Turkey	✓	(1)	✓
Turkmenistan			✓
Ukraine			✓
United Kingdom	✓	✓	✓
United States	✓		
Uzbekistan			✓

1) EU candidate and potential candidate countries covered in the Eurostat data collection.   2) EFTA countries covered in the Eurostat data collection.

ANNEX 4. APPROACHES TO OVERCOME BREAKS IN TIME SERIES

To minimise the frequency of series breaks due to measurement changes, national statistical agencies can use a number of approaches to reconstruct the series.

A first method is to estimate a conversion coefficient for the year when the break occurred by using the average growth rate for the preceding years. This method to revise the series before the break involves the following steps:

a)   Calculate the average of the annual growth rates for x years before the break in time series.
b)   Use this average growth rate to estimate a new theoretical figure for the year when the break occurred (consistent with the old series).
c)   Calculate a conversion coefficient between the actual figure and the estimate for the year when the break occurred (i.e. divide the actual figure by the estimate).
d)   Use this conversion coefficient to estimate backwards all data before the break (i.e. multiply the original data by the conversion coefficient).

Example to overcome a break in series between 2001 and 2002

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Original time series	11670	11880	12000	12100	12460	12830	13200	14600	14960	15820	16750	17680	18370	19100
Annual growth rates	2.0	1.8	1.0	0.8	3.0	3.0	2.9	10.6						
Average of growth rates over 1995-2001								2.1						
Estimate for 2002								13473						
Conversion coefficient								1.084						
Revised time series	12646e	12874e	13004e	13112e	13502e	13903e	14304e	14600	14960	15820	16750	17680	18370	19100
Annual growth rates	2.0	1.8	1.0	0.8	3.0	3.0	2.9	2.1						

[Note: A slightly different option would be to calculate the average of the annual growth rates before the break and after the break (for one or more years), and then to follow the steps b to d described above.]

A second approach involves the compilation of the series using both the old and new methods for a certain period of time around the implementation of the new method, from which a concordance coefficient can be calculated. This concordance coefficient can then be used to revise the part of the series before the break.

ANNEX 5. SPECIFICATIONS AND STANDARD DATA FORMAT FOR SUBMISSION OF NATIONAL HOSPITAL DISCHARGE DATA (HDD)

This annex concerns only countries in the <u>WHO European Region</u> (EU members or candidates, EFTA countries, Israel, and others). Table 1 below displays the complete list of countries involved in the preparation of Hospital Discharge Data (HDD) files.

Table 1. Two-character country codes for countries in the WHO European Region

AL	Albania	GE	Georgia	MT	Malta
AD	Andorra	GR	Greece	NL	Netherlands
AM	Armenia	HR	Croatia	NO	Norway
AT	Austria	HU	Hungary	PL	Poland
AZ	Azerbaijan	IE	Ireland	PT	Portugal
BE	Belgium	IL	Israel	RO	Romania
BG	Bulgaria	IS	Iceland	RU	Russian Federation
BA	Bosnia and Herzegovina	IT	Italy	SE	Sweden
BY	Belarus	KZ	Kazakhstan	SI	Slovenia
CH	Switzerland	KG	Kyrgyzstan	RS	Serbia
CZ	Czech Republic	LI	Liechtenstein	SK	Slovakia
CY	Cyprus	LT	Lithuania	SM	San Marino
DE	Germany	LU	Luxembourg	TJ	Tajikistan
DK	Denmark	LV	Latvia	TM	Turkmenistan
EE	Estonia	MC	Monaco	TR	Turkey
ES	Spain	ME	Montenegro	UA	Ukraine
FI	Finland	MD	Republic of Moldova	UK	United Kingdom
FR	France	MK	North Macedonia	UZ	Uzbekistan

Source: ISO 3166 (except UK).

Data coverage and definition

Data should include discharges from all public and private hospitals (HP.1 according to the International Classification of Health Accounts) during a given calendar year. Data from residential long-term care facilities (HP.2) should not be included.

Countries shall provide three indicators:
1. Inpatient discharges
2. Bed-days
3. Day case discharges

broken down by diagnosis, sex, age and regions at NUTS 2 and according to the definitions of the joint questionnaire on non-monetary health care statistics:

- An **inpatient discharge** is the release of a patient who was formally admitted into a hospital for treatment and/or care and who stayed for a minimum of one night.
- A **bed-day** (or inpatient day) is a day during which a person admitted as an inpatient is confined to a bed and in which the patient stays overnight in a hospital.
- A **day-care discharge** is the release of a patient who was formally admitted in a hospital for receiving planned medical and paramedical services, and who was discharged on the same day.

Further down in this document the different classifications for the diagnosis and age categories are listed. However, for the sake of harmonization, it is recommended to use ICD10 3-character for the diagnosis category and the 23 age groups list (option one) for the age category.

In case the hospital discharge records do not include the special identification of day cases, any record showing that admission and discharge dates were identical should be assigned to the day case category if the patient was alive when discharged.

Data should include discharges to home and to other inpatient institutions, and deaths in hospital.

Countries using hospital treatment episodes as individual record units (i.e. when patients are transferred among departments in the same hospital, each transfer is recorded as a new admission/discharge) should make an effort to merge related episodes into one discharge record in order to make the data internationally comparable.

In accordance with the ICD, it is recommended that the main diagnosis be defined as the condition diagnosed at the end of the hospitalisation period, primarily responsible for the patient's need for treatment or examination at the hospital. If there is more than one such condition, the one held responsible for the greatest use of resources should be selected. If no diagnosis was made, the main symptom, abnormal finding or problem should be selected as the main diagnosis.

Data files

The hospital discharge data (HDD) file should contain the number of annual inpatient discharges, bed-days and day case discharges by main diagnosis or external cause of hospitalisation, by age and sex[3].

The HDD file should be a **comma-delimited ASCII file (.CSV)** with the fixed field length of variables according to the record lay-out (see Table 2).

Table 2. Record lay-out of the HDD file

Variable no. and name	Variable description	Type	Size in case of fixed field length
1. Country	Two-character country code (ISO 3166, see Table 1 or www.iso.org). This field can be extended by adding a code for sub-national entities, thus allowing reporting at sub-national level[4].	Character	2[5]
2. Year	Year of discharge	Integer	4
3. Age	Code of age grouping (see Table 3)	Integer	1
4. Gender	Code of gender (see Table 4)	Integer	1
5. ICD version	Code of ICD version used to code diagnosis (see Table 5)	Character	3
6. Diagnosis	Code of diagnosis or external cause (ICD-9 or ICD-10 code corresponding to above code of ICD version)	Character	5
7+3n. Discharges	Number of inpatient discharges (excluding day cases) with above diagnosis, by defined age groups	Integer	10
8+3n. Bed-days	Number of bed-days used for above diagnosis, by defined age groups	Integer	10
9+3n. Day cases	Number of day case discharges with above diagnosis, by defined age groups	Integer	10

---

[3] Eurostat also requires these data by region (NUTS 2 level).
[4] Hospital discharge data are collected at the NUTS level 2 by Eurostat. See NUTS2 codes (4 characters) at: http://ec.europa.eu/eurostat/web/nuts/nuts-maps-.pdf-.

The file name should be constructed as follows: **HDD_CountryCode_Year.csv**. A 2-character ISO 3166 country code (see Table 1) should be used, e.g. a data file from Austria with data for 2012 will be named HDD_AT_2012.csv.

Each record in this data file should contain data on one diagnosis for one gender by defined age groups. The total length of the record depends on the age grouping used. Normally, 23 age groups should be used (option 1, see Table 3) and in this case the total number of variables per record is 75. In case of four age groups, or if no age disaggregation is used, there should be respectively 18 or 9 variables per record.

Examples of four records in HDD file for country CC

```
CC,2004,1,1,103,TOT,332583,3475024,7453,16877,113395,372,18958,111404,667,15427,95337,709,16512,125509
,543,15127,108670,543,12357,110752,451,11156,115751,446,13005,147047,461,14676,162940,416,18698,220397
,447,20094,241176,419,22010,265966,395,23844,278469,345,25534,304968,325,26276,313408,343,25924,305147
,286,20188,241388,158,10384,133496,77,3128,42016,24,1290,20403,9,641,10376,1,477,7009,16
CC,2004,1,2,103,TOT,444550,4369903,13560,13560,88882,284,14748,87913,447,13115,85400,579,15235,124855,
441,21119,144205,755,27092,157676,1198,28703,174108,1242,23620,159912,1280,19761,153625,1235,21137,189
059,1279,24186,228055,1169,26751,262826,953,25711,262413,660,26482,280801,568,32019,357613,512,37678,4
66531,416,36987,506571,300,24029,386811,159,8322,161754,53,3511,75761,20,634,13496,4,150,1636,6
CC,2004,1,1,103,A02,682,4316,3,42,276,0,168,1043,3,88,552,0,36,221,0,29,151,0,51,284,0,37,245,0,27,144
,0,26,175,0,29,169,0,32,188,0,26,166,0,24,157,0,28,205,0,11,96,0,8,59,0,15,144,0,4,27,0,1,14,0,0,0,0,0
,0,0,0,0,0
CC,2004,1,2,103,W54,53,309,0,0,0,0,7,28,0,6,26,0,7,24,0,2,8,0,2,9,0,2,13,0,0,0,0,3,19,0,2,7,0,1,6,0,1,
8,0,5,55,0,2,9,0,3,12,0,4,36,0,5,38,0,1,11,0,0,0,0,0,0,0,0,0,0,0,0
```

Variable definitions

3. Age: Several options are available. The preferred option is to have the data tabulated by 23 age groups (22 groups plus total population, option 1 in Table 3). This will allow age-standardisation of hospitalisation rates and a detailed age-specific analysis of the data. However, some countries may not be able to provide data in such details, due to technical reasons or others. In this case, two other options are provided, i.e. using three age groups or all ages only.

Table 3. Options for age groupings

Age group No (n=)	Option 1 (Age=1)	Option 2 (Age=2)	Option 3 (Age=3)	Option 4 (Age=4)
0	All ages	All ages	All ages	All ages
1	< 1 year	0-14 years		< 1 year
2	1-4	15-64		1-4
3	5-9	65+		5-9
4	10-14			10-14
5	15-19			15-19
6	20-24			20-24
7	25-29			25-29
8	30-34			30-34
9	35-39			35-39
10	40-44			40-44
11	45-49			45-49
12	50-54			50-54
13	55-59			55-59
14	60-64			60-64
15	65-69			65-69
16	70-74			70-74
17	75-79			75-79
18	80-84			80-84
19	85-89			85+
20	90-94			Age unknown
21	95+			
22	Age unknown			

4. Gender: see Table 4. Normally, all hospital discharge records should have valid gender codes. In cases where gender has not been recorded, the additional record with the gender code = 0 may be generated to ensure that all discharges are included in the total after adding figures for male, female and unknown gender.

Table 4. Code table for gender

Gender	Code
Male	1
Female	2
Unknown/not recorded	0

5. ICD version: Indicates which version of the International Classification of Diseases is used to code the diagnosis (see Table 5).

Table 5. Code table for ICD version used to code main diagnosis

Code	Classification system
094	ICD-9 4-character list
103	ICD-10 3-character list (**most preferable)**
104	ICD-10 4-character list
HMT	International Shortlist for Hospital Morbidity Tabulation (ISHMT) *

* Can be downloaded from:
http://www.who.int/classifications/icd/implementation/morbidity/ishmt/en/index.html.

ISHMT was adopted at the meeting of WHO Collaborating Centres for the Family of International Classifications (WHO-FIC), Tokyo, Japan, 16-22 October 2005. It contains 149 groups of diagnoses including the total. It originates from the list developed under the Hospital Data Project (funded by DG Sanco of European Commission).

Note: It is highly preferable that the original **3-character ICD-10 codes** be used whenever possible, as this will allow for much more flexibility in subsequent data analyses. When the original ICD codes are aggregated into a short list, analysis of the data by detailed individual ICD codes (diagnoses) becomes impossible.

Aggregation of the ICD codes into shorter tabulation lists can be done by international agencies depending on needs and data dissemination practices.

6. Diagnosis:
- Where ICD-10 is used, three-character codes are sufficient.
- In the case of ICD-9, four-character codes are necessary. Do not use a dot to separate the fourth character in the code. For external causes (Supplementary Classification of External Causes), use prefix E. For example, a motor vehicle traffic accident involving collision with a pedestrian who has been injured should be coded as E8147.
- The four-character codes of the International Shortlist for Hospital Morbidity Tabulation (ISHMT) can also be used.

In the case of fixed record length, codes should be aligned to the left.

The HDD file should include an additional record for each gender containing the total of all discharges with the diagnosis code '**TOT**' when "healthy newborn babies" are included; otherwise the code must be '**TOE**'.

Hospitalisations due to external causes will contribute twice: once with the code of the injury (Chapter 17 of ICD-9 or Chapter 19 of ICD-10), and once with the code corresponding to external causes

(Supplementary Classification of External Causes with prefix "E" of ICD-9 or Chapter 20 of ICD-10). <u>Care should be taken not to count such cases twice in the summary record with code "TOT" (or "TOE")</u>. If external causes are not recorded, corresponding records will not appear in the HDD file.

7+3n. Discharges: The total number of inpatient discharges (excluding day cases) with the above diagnosis for a given gender and age group. The total figure for all ages is included in position 7, age below 1 year in position 10, age group 1-4 years in position 13, etc. See Table 3 for **n** value corresponding to each age group.

8+3n. Bed-days: The total number of bed-days used by inpatients (excluding day cases) with the above diagnosis for given gender and age group. The total figure for all ages is included in position 8, age below 1 year in position 11, age group 1-4 years in position 14, etc.

9+3n. Day cases: The total number of day case discharges with the above diagnosis for a given gender and age group. The total figure for all ages is included in position 9, age below 1 year in position 12, age group 1-4 years in position 15, etc.

In case of **missing data** in any of the three categories listed above, the field must be empty. Only when there are zero cases the field must be filled with '0'. For example, if data about 'Healthy newborn babies' are not available, it shall not be included with '0' cases, or if bed-days data are not available then the column must be empty.

Metadata

The data should be accompanied by the "metadata", i.e. the description of national sources, methods, definitions, degree of data completeness and coverage, explanations of any peculiarities in national statistical and coding practices, etc. (see Table 6 for recommended content). The metadata should be entered in the common Word file (CCC_HealthActivities_Year.doc).

Table 6. Recommended content of metadata related to the sources and methods of hospital discharge data

1) List the type, name, location and owner or operator of the national hospital patient registers or discharge database(s) (NHDDB), which were used to produce the HDD file.
2) Does the NHDDB cover all inpatient institutions in the country, which are classifiable as HP.1 providers according to the "System of Health Accounts"? List types of hospitals, which are covered and not covered, e.g. private hospitals, military or prison hospitals, etc., and, if possible, estimate their total capacity as compared to those that are covered by the NHDDB.
3) Does the NHDDB include all inpatient discharges and day cases in hospitals covered? List cases, which may not be included in the NHDDB (e.g. uninsured patients, foreigners, military staff, etc). If possible, estimate the proportion of missing discharge records.
4) If the discharge records were based on treatment episodes (consultant episodes, department discharges), have such multi-episode cases been combined into one discharge record? If possible, estimate the proportion of multi-episode in-patient cases.
5) If the HDD file includes day cases, how were they defined? Was there a special index flagging planned day cases in the NHDDB or were they identified by the same admission and discharge dates?
6) Describe any other known or suspected peculiarities in the coverage of the data.
7) Explain principles involved in defining the main diagnosis (condition) in the hospital discharge record.
8) Describe any known or suspected peculiarities related to the national diagnostic and recording practices and related to how the main condition is selected.

User instructions for HDD_Converter (a computer programme for the conversion of individual records into an aggregated HDD file)

Purpose

The programme HDD_Converter has been built to help countries aggregate individual hospital discharge records into standard tabular format for further submission to international organisations, thus avoiding the burden of developing the same computer programme in each country.

The HDD_converter.exe programme is available from https://euro.sharefile.com/d/sdbbb74e47f44f12a.

Input data file

Data in hospital discharge databases in different countries may be stored in different ways. Therefore, the standard minimum set of variables has first to be retrieved from these databases and saved as a temporary data file. This temporary file then is used by the HDD_converter programme as an input for further aggregation of data into the standard format.

When pooling together hospital discharge records into a temporary file make sure that:

- they include available discharge records from all hospitals in the country during the reference year;
- discharge records are only from hospitals that are classified as HP.1, according to the International Classification of Health Accounts (i.e. general hospitals and specialised hospitals, including mental health hospitals);
- if individual records correspond to treatment episodes in different departments, these records are merged into one discharge record, if possible.

(See also "Data coverage and definitions" for further details.)

Variables in the temporary input file record must appear in the following sequence and should be separated by commas:

1) Year of discharge, e.g. 2012 (Integer).
2) Age of discharged patient at the moment of admission, in full years, e.g. 50 (Integer).
3) Gender code: 1-male; 2-female (Integer).
4) Main diagnosis ICD code, e.g. for diagnosis "Atherosclerosis" it should be I70 if ICD-10 is used or 440 in case of ICD-9 codes (Character).
5) External cause code, if hospitalisation was due to accident or other external cause and this cause was recorded (Character). If such a code is missing or not applicable leave it blank or, preferably, enter missing data code "-1".
6) Day case index: 1 - if the patient was admitted with the intention of discharging him/her on the same day, and if the patient was actually discharged on the same day; 0 - in all other cases (Integer).
7) Length of stay in a hospital in full days (Integer). If the day case index = 1, length of stay should be = 0.

Example of records in temporary input file:
2011,50,1,I70,-1,0,5
2011,33,2,L02,-1,1,0
2011,20,1,S01,V43,1,0
…….

All of the above variables have to be created and their values calculated during the data retrieval from national hospital discharge databases, e.g. age = admission date minus birth date; valid values between 0-120 years, etc.

If data are missing for some variables, leave them blank or enter missing value code "-1".

Retrieve the above variables for all discharge records for the required year(s) and save them in .CSV (comma separated value) format under the name **HDDinput.csv**.

How to use the HDD_converter:

Make sure that the above input data file HDDinput.csv and HDD_converter.exe are in the same directory.

- Start HDD_converter.exe.

- Enter the required parameters in the dialogue box or select from the list:

    o the 2-character country code;
    o the data reference year;
    o the ICD codes used (103 is default);

- Click on "proceed".

The programme reads all the individual discharge records from the HDDinput.csv file and aggregates them into the standard format in accordance with specifications. It also counts missing values and displays summary information on how many records with missing or incorrect data values were found. The summary information will be saved in the file HDD_LOG_country code_year.txt.

Click on "exit" to close the programme. The resulting file **HDD_country code_year.csv** will be created and saved in the same directory.

Contact

For assistance in case of problems or queries, please contact:
Ivo Rakovac
WHO Regional Office for Europe
Email: euJDC@who.int.