eurostat

Morbidity statistics in the EU

Report on pilot studies – 2023 edition

STATISTICAL REPORTS

Morbidity statistics in the EU

Report on pilot studies – 2023 edition

Luxembourg: Publications Office of the European Union, 2024

Collection: Statistical reports
Theme: Population and social conditions

PDF ISBN 978-92-68-08959-0 KS-FT-23-003-EN-N doi: 10.2785/511997 ISSN: 2529-3222

Table of contents

Abbreviations

AMI	Acute Myocardial Infarctus
ATC	Anatomical Therapeutic Chemical Classification System
COD	Causes of Death
DDD	Defined Daily Doses
ECDC	European Centre for Disease Control
ECHI	European Core Health Indicators
EHIS	European Health Interview Survey
EPIMS	European Project on Inventories of Morbidity Statistics (not available in the public domain)
ESS	The European statistical system
GP	General Practitioner
Guidelines	When 'Guidelines' are referred to, this means the Methodological guidelines of the MORB project used by countries for the pilot studies (not available in the public domain).
ICD	International classification of diseases
IPCP	International classification of primary care
JRC	The European Commission Joint Research Centre
MORB	The Morbidity Statistics (MORB) project which is the topic of this Statistical report
MSDG	The European Commission Morbidity Statistics Development Group
pp	Percentage points
RAG	RAG stands for red, amber, and green. These traffic light colors are used in this report to classify indicator status.

In section 2 of this report, abbreviations pertaining to a particular country's data sources are used and spelled out at the first use in the relevant section. Abbreviations concerning only one country are not referred to in detail in the above list.

Summary

Introduction

Eurostat and EU Member States have been working towards producing statistics on morbidity since 1998. Pilot studies were carried out in 2005-11 ([1]), followed by an inventory project with participation of 15 Member States. National pilot studies were again carried out in 2019 to 2021 as part of the Morbidity Statistics (MORB) project, co-funded by the European Commission (Eurostat). This document reports the results of these national pilot studies.

The purpose of the MORB project was to verify the feasibility of a regular data collection on morbidity statistics. The collection started from a shortlist of clearly defined indicators developed through earlier projects. For an overview of the shortlist indicators, see Table A following this summary. The overall aim was to develop a regular data compilation for a selected set of diseases within the European Statistical System (ESS) to provide a general picture of diagnosis-specific morbidity at population level.

The countries reporting their results were Belgium (BE), Croatia (HR), Finland (FI), France (FR), Hungary (HU), Lithuania (LT), Malta (MT), Netherlands (NL) and Poland (PL).

Feasibility of a morbidity statistics data collection

The pilot studies show that the systematic and meaningful collection of morbidity statistics is feasible, but still faces serious difficulties. Compared to earlier projects, the implementation of standard definitions and methods reduced the scale of differences between countries and produced more plausible findings. The most important difficulty is that despite the standardization of methods, unavoidable differences between national healthcare systems and their data collections continue to be a problem for comparability for many indicators. Determining whether the observed differences in incidence and prevalence are due to compatibility of the data sources or reflect epidemiological patterns is not simple and would require further research.

The need for multiple data sources

The pilot studies confirm the necessity of multiple, linked data sources to provide plausible findings. Where one or more data sources were unavailable this tended to produce unreasonable low estimates for the relevant countries. Not all countries were able to access primary care data, which is essential for most chronic diseases and those conditions which do not necessarily involve hospitalisation. Equally, for many more severe diseases, cause of death data are essential as many people die without a hospital admission.

Accessibility of the data sources

In most cases, the multiple datasets needed were available through collaboration with various institutes and were sufficiently timely for regular use by T+3 years. However, a few countries experienced severe difficulties with obtaining or linking the data which made reliable multi-source morbidity statistics impossible. Any future data collections (involving new countries) should be organised with sufficient time for data-sharing agreements to be negotiated, possibly taking about two years of preparation. One country terminated the pilot study due to lack of access to necessary data sources.

([1]) Morbidity statistics in the EU - Report on pilot studies - 2014 edition https://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/ks-tc-14-003

Institutionalised populations

Institutionalised populations could, with some small exceptions, be included in the main data sources for morbidity statistics by most countries. In a few cases, specifically serious psychiatric conditions and dementia, a minority of countries do not have complete coverage.

Private sector healthcare

While most of the healthcare is provided or paid for by the public national systems, in most countries there is a private sector whose importance varies by health condition, socially and geographically. The extent of the private sector does not seem to be understood very precisely. We recommend that for a future data collection, all countries should thoroughly collect data on private sector healthcare either directly or using proxy approaches such as comparison with the European health interview survey (EHIS) and adjust the estimates accordingly.

Inclusion of non-residents

All countries had difficulties with the counting and production of estimates for non-residents. In some cases, the data sources were themselves restricted to residents only, such as persons registered with a General Practitioner or eligible under the national health insurance. Where non-residents were included and could be separately identified, the lack of a personal identifier or inability to link to a population register meant that cases could not be counted accurately, e.g. incident cases could not be distinguished from previous cases. It is therefore recommended to exclude non-residents from the calculations to achieve the best consistency.

Recommendations on the indicators

Out of the 43 indicators on the shortlist, 13 indicators were rated as good enough to be included in a future data collection with no change, or a very minor change ('Green'). A further 13 were considered good enough to be included in a future data collection with a more significant change, and/or the comparability of the results would benefit from more research ('Amber'). Finally, 17 indicators do not appear feasible, or another indicator would be preferred ('Red').

For an overview of the shortlist indicators investigated in this project, see Table A following this summary. A fuller version is available in Annex A Table 7: 'Summary of recommendations on the indicators'. Recommended definitions for indicators for a future data collection are shown in Annex A Table 4 'Recommendations on the future shortlist: indicator definitions'.

Principles for future data collections

The results of these pilot studies suggest that harmonisation of the definitions and methods is feasible, and to a lesser extent so is the comprehensive and linked use of data sources. The value of future data collections depends partly on the ability of all participating countries to base their estimates on a full set of relevant data sources. However, the interpretation of findings will still be difficult without also taking account of the healthcare system and its effects on the data. Therefore, the preparation for any future data collections should also include research by the participating countries on the size and role of private sector healthcare; the proportion of uninsured persons, non-residents and people not registered with a General Practitioner; and likely bias caused by healthcare system financing and organisation, leading to adjustment of the estimates for each indicator on a well-documented basis.

TABLE A

List of pilot indicators investigated in the MORB project by Red-Amber-Green (RAG) rating

Nr	Label	ICD-10 code	Definition
Green: Indicator appears feasible with no or very minor change			
P5	Schizophrenia, schizotypal and delusional disorders	F20-F29	Period prevalence
P6	Mood (affective) disorders	F30-F39	Period prevalence
P9	Multiple sclerosis	G35	Period prevalence
P10	Epilepsy	G40-G41	Period prevalence
P11	Hypertensive diseases	I10-I13, I15	Incidence by person
P12	Hypertensive diseases	I10-I13, I15	Period prevalence
P15	Acute myocardial infarction	I21, I22	Incidence by person
P16	Heart failure	I50	Period prevalence
P17	Stroke	I60-I64	Incidence by person
P21	Asthma	J45, J46	Period prevalence
P23	Chronic obstructive pulmonary disease (COPD)	J44	Period prevalence
P26	Diseases of liver	K70-K77	Period prevalence
P27	Rheumatoid arthritis	M05, M06	Period prevalence
Amber: Indicator appears feasible but with significant change, or more research needed			
P1	Diabetes mellitus	E10-E14	Incidence by person
P2	Diabetes mellitus	E10-E14	Period prevalence
P3	Dementia (incl. Alzheimer disease)	F00-F03, G30	Period prevalence
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	F10	Period prevalence
P8	Parkinson disease	G20	Period prevalence
P13	Ischaemic heart diseases	I20-I25	Period prevalence
P18	Cerebrovascular diseases	I60-I69	Period prevalence
P19	Pneumonia	J12-J18	Incidence by episode
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	J40-J44, J47	Period prevalence
P28	Arthrosis	M15-M19	Period prevalence
P31	Urolithiasis	N20-N23	Incidence by person
P33	Intracranial injury	S06	Incidence by person
P35	Fracture of femur	S72	Incidence by person
Red: Indicator does not appear feasible, or another indicator is preferred			
P7	Anxiety disorders	F40-F41	Period prevalence
P14	Acute myocardial infarction	I21, I22	Incidence by episode
P20	Asthma	J45, J46	Incidence by person
P24	Alcoholic liver disease	K70	Period prevalence

Nr	Label	ICD-10 code	Definition
P25	Diseases of liver other than alcoholic	K71-K77	Period prevalence
P29	Osteoporosis	M80-M82	Period prevalence
P30	Renal failure	N17-N19	Period prevalence
P32	Intracranial injury	S06	Incidence by episode
P34	Fracture of femur	S72	Incidence by episode
PB36	Land transport accidents	V01-V89	Incidence by episode
PB37	Land transport accidents	V01-V89	Incidence by person
PB38	Accidental falls	W00-W19	Incidence by episode
PB39	Accidental falls	W00-W19	Incidence by person
PB40	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by episode
PB41	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by person
PB42	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by episode
PB43	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by person

1 Introduction

1.1. Background

This report presents the findings of national pilot studies co-funded by the European Commission (Eurostat), conducted between 2019 and 2021 as part of the Morbidity Statistics (MORB) project. The morbidity of a disease is how many people have it in a particular population. Diagnosis-specific morbidity statistics at population level would be very useful for policy-makers, researchers, industry, media and others, which is why the European Commission and Member States are working on them.

The first steps towards the development of morbidity statistics within the European Statistical System (ESS) began in 1998. Several national pilot studies were conducted between 2005 and 2011 and published in a Eurostat statistical report called 'Morbidity statistics in the EU – Report on pilot studies – 2014 edition' ([2]). It was assumed that the best possible (single) data source could be identified for each diagnosis, but this turned out to be unrealistic. Instead, the preferred methodological approach is now assumed to require information from several different, and in the best case linked, data sources to produce national estimates. The main emphasis is on a common output on EU level, based on a shortlist of agreed indicators, irrespective of national data sources.

In 2015, the European Commission (DG SANTE) funded the European Project on Inventories of Morbidity Statistics (EPIMS). This project did not collect data directly and the project report is not in the public domain. The project aimed to systematically gather information on the importance and feasibility of indicators, as well as the availability of potential data sources for morbidity statistics in each of the participating 15 EU countries. The evidence reported by countries highlighted that data sources are generally complementary but overlapping, and that personal identifiers (also anonymised) are needed to avoid double-counting.

Under the current MORB project, a further set of pilot studies, described in this document, were carried out from 2019 to 2021. Ten EU countries participated in these studies. However, due to operational difficulties, one country did not manage to submit their results in time for the project's report. The main problem was access to the necessary data sources. The countries that provided their findings were Belgium (BE), Croatia (HR), Finland (FI), France (FR), Hungary (HU), Lithuania (LT), Malta (MT), Netherlands (NL), and Poland (PL).

1.2. Aim of the project

The objective of the MORB project was to assess the practicality of establishing a regular data collection system for morbidity statistics, utilizing a concise list of well-defined indicators derived from the previous pilot studies and projects. The overarching goal was to create a systematic data compilation for a specific set of diseases within the European Statistical System (ESS) to offer a comprehensive overview of diagnosis-specific morbidity at population level.

([2]) Morbidity statistics in the EU - Report on pilot studies - 2014 edition https://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/ks-tc-14-003

1.3. Selection of indicators and methodology

Indicators were shortlisted based on the work previously carried out regarding morbidity statistics. The starting point was the document produced in 2006-2007 by the European Commission Morbidity Statistics Development Group (MSDG) on 'Principles and guidelines for diagnosis-specific morbidity statistics' (Annex 4 of the statistical report 'Morbidity statistics in the EU' from 2014). Further work by Eurostat and EU Member States resulted in a shortlist of proposed indicators, health conditions, which were considered feasible to collect.

The inclusion of health conditions in the list of proposed indicators considered previous evidence relating to:

- Public health and policy relevance, social and economic burden at EU and national levels, and importance for the costs and workload of health systems.
- Clinical and epidemiological meaningfulness of the disease or group of diseases and the ability to define the indicator without ambiguity.
- Likelihood of good comparability of the indicator between countries, taking account of differences in health systems, national practices, and the evidence of the previous pilot studies.
- Feasibility in view of the possible data sources and expected difficulty of methodological issues, based on analysis of the EPIMS national inventories.

Shortlist indicators considered in the present report were split in two lists (Annex A Table 1). List A contains the indicators required to be included for pilot data collection in the studies summarised in this report, while indicators in List B are indicators countries were asked to considered for pilot data collection.

The indicators selected were measured in incidence or prevalence. Incidence can be subdivided into incidence by person and incidence by episode. Prevalence can be subdivided into period prevalence and point prevalence. For a more detailed discussion on these measurements, please refer to the report 'Morbidity statistics in the EU' from 2014.

To avoid duplication of European data collections on morbidity, it was decided to exclude infectious diseases and cancers from the national pilot studies. Data collections on these subjects already exist under the auspices of the European Centre for Disease Control (ECDC) and Joint Research Centre (JRC) respectively. Format of this report

This report summarises the findings of the pilot studies in three main sections. There is necessarily some overlap or duplication of information between the sections. For ease of reference, all the descriptive tables and lists (which can each be referred to in several parts of the report) are gathered in an annex at the end.

The sections are arranged as follows:

Section 2. Analysis by country

This section describes the data sources, estimation methods and quality issues for each of the countries submitting data in the pilot study. The information is based on the Eurostat standard metadata sheet 'ESMS: National Reference Metadata in Euro SDMX Metadata Structure ([3])', the metadata sheets per individual indicator, and the final technical reports of the grant-supported action, for each country.

It is notable that the countries described data sources in different ways in the metadata, based on national terminology and organisational structures. For example, if both primary and secondary care data were collected in one register or database, this was sometimes described as a single data source and sometimes as multiple sources. Data from an insurance organisation on reimbursed prescriptions can be categorized as either 'insurance' or 'prescription' data. Generally, the description here is based on the terminology and concepts supplied by each country and no harmonisation has been attempted.

Section 3. Analysis by data source

This section contains summary observations and conclusions regarding use of the different types of data sources for producing morbidity estimates, based on considerations from the EPIMS report and the guidelines for the grants financing the pilot studies, and the detailed information provided in section 2.

Conclusions are provided with reference to the European Statistical System dimensions of quality ([4]) with respect to:

Relevance

- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability

([3]) European Statistical System (ESS) Handbook for Quality and Metadata Reports — re-edition 2021. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-021

([4]) European Statistical System (ESS) Handbook for Quality and Metadata Reports — re-edition 2021. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-021

(The dimension 'Accessibility' was excluded, as it is not directly relevant)

Section 4. Analysis by indicator

This section reports on the suitability of the data sources and definitions for each indicator. Detailed information is given about the codes, data limitations and other relevant matters as reported by the pilot studies. For a number of indicators, countries were requested to consider certain questions or options identified in the EPIMS project in the pilot studies. Where appropriate in this section, the entry for each indicator sets out the issues to be addressed, a summary of the findings and the conclusion.

In a few cases, countries made additional suggestions for changes to the indicator definitions, based on experience from the pilot studies. These proposals are documented in the entry for each indicator. Not every point made by the countries on each indicator is recorded here, as the focus is on issues which might affect the feasibility, definition, or interpretation. For example, implementation of the International classification of diseases version 10 (ICD-10) codes in major classifications (for example the 'International classification of primary care', ICPC) is described where appropriate, but classifications used in only one country are not referred to in detail unless needed to illustrate wider issues.

2 Analysis by country

2.1. Belgium (BE)

The pilot study in BE was carried out in 2019 by the unit for Lifestyle and Chronic Diseases within the Epidemiology and Public Health department of the institute 'Sciensano'. The institute is responsible for a wide variety of public health, scientific and environmental functions.

Estimates were reported for all the main list of indicators P1-P35 (list A) and also the optional indicators PB36-PB43 relating to external causes (list B).

The data sources used by BE were as follows:

1. *Primary care*: General Practitioners (GP) practices of the Intego Network
2. *Causes of death*: Causes of death register of Belgium
3. *Hospital inpatients*: National Hospital Stay Database of the National Institute for Health and Disability Insurance
4. *Insurance*: Intermutualistic Agency (IMA) insurance data EPS cohort – covering prescription reimbursements

2.1.1. Data sources used

2.1.1.1. Intego Network of General Practitioners

2.1.1.1.1. General description

Intego is an integrated network of General Practitioners (GPs) in Flanders, which is the biggest region in Belgium population-wise and accounts for 60% of the Belgian population. Led by the Academic Centre for General Medicine from University of Leuven, the network aims to create a large database to centralize morbidity data in primary care. An automated data collection is based on the Electronic Medical Record (EMR) of the patient as registered in their GP's practice. The collection of such EMR data allows the network to convey the incidence and prevalence of a multitude of (primarily GP, non-specialist care requiring) diseases, as well as data concerning diagnostic tests or applied therapies in the registered GP practices.

2.1.1.1.2. Population coverage

The data from the Intego Network is geographically limited, as only GP practices in the region of Flanders are included in the network. It is also unknown how representative the Intego Network of GPs is of the total GP practices in Flanders. Any regional differences between the three regions of Belgium cannot be extrapolated and should be determined from a different source. The network only covers about 2% of the Flemish population, therefore, to produce a national estimate, it is necessary to extrapolate to the whole population while also accounting for representativeness and regional differences.

Only people with a social security number are registered with a GP, the population covered is thus by definition only residents. The estimated population of non-residents, treated by the GPs attached to the network, is considered to be less than 1% of the total treated population.

2.1.1.1.3. Data availability, linkage

The data are under strict data protection controls but can be accessed through special agreements between Intego and Sciensano. The primary care-based data were accessed using a remote access system. Data within the EMR are linked per patient, but linkage to other sources was not feasible.

2.1.1.1.4. Classifications used and coding issues

The diagnoses are coded in the EMR using the ICPC-2 classification and are, in parallel, mapped using a software thesaurus to ICD-10. However, this mapping is not yet completely reliable and validated for accuracy. Therefore, the ICPC coding system was used to obtain the matching diagnoses for each indicator, even though the ICPC-2 and ICD-10 definitions do not match completely. This has an important effect on a few indicators.

2.1.1.1.5. Other problems of recording or definition

The registration in the EMR is not episode based, making it difficult to calculate the prevalence for some indicators, depending on the condition and the need to visit the GP during the episode and/or follow-up. It is unknown how exhaustive the GP's information on the patient is when reliance on other healthcare sources is necessary, for instance after hospital admission or medical imaging examinations. This may cause some underestimations when the GP is not aware of a diagnosis made by another service. For instance, in Belgium, it is not uncommon for a patient to go directly to a dermatologist or gynaecologist, without referral from a GP.

2.1.1.1.6. Relevance to indicators

Used for indicators P4, P6-P8, P10-P13, P16, and P19-P30.

2.1.1.2. Causes of death register of Belgium

2.1.1.2.1. General description

Causes of death data are derived from the mandatory medical certificate of cause of death and produced by the Belgian national statistical office (Statbel) after processing by two federal entities.

2.1.1.2.2. Population coverage

Deaths of residents in Belgium are recorded. Deaths of Belgian residents occurring abroad are also registered, but without a specific cause of death.

2.1.1.2.3. Data availability, linkage

Causes of death data for 2016 were available in mid-2019. The data were not linked to any other dataset.

2.1.1.2.4. Classifications used and coding issues

Multiple causes of death are recorded using ICD-10. Some differences in coding practice between the federal entities are known.

2.1.1.2.5. Relevance to indicators

Used for indicators P40-P43.

2.1.1.3. National Hospital Stay Database

2.1.1.3.1. General description

The National Hospital Stay Database (NHSD) is a merged dataset, which is based on the Minimum Hospital Data collected by the Ministry of Health and containing information concerning the diagnoses and procedures for each admission, and the Hospital Billing Data (HBD) from the National Health Insurance companies which contains information regarding the billing data for hospitalized patients. This merged dataset is held by the National Institute for Health and Disability Insurance (RIZIV-INAMI).

2.1.1.3.2. Population coverage

The coverage of the data is complete for all hospitalized patients in Belgium.

2.1.1.3.3. Data availability, linkage

The data allow linkage between the patient and the received care over multiple years and between multiple hospitals. However, linkage to other datasets was not feasible.

In 2019, the National Hospital Stay Database information of 2017 was available.

2.1.1.3.4. Classifications used and coding issues

ICD-9 was used for the classification in the National Hospital Stay Database prior to 2015, and ICD-10 from 2015 onwards. Therefore, some of the data used for 3-year prevalence and to identify new cases in 2016 could possibly be affected by small differences between ICD-9 and ICD-10.

2.1.1.3.5. Relevance to indicators

Used for indicators P14, P15, P17, P18, P31-P35, and PB36-PB39.

2.1.1.4. InterMutualistic Agency (IMA) insurance data EPS cohort

2.1.1.4.1. General description

The InterMutualistic Agency (IMA) receives and analyses an exhaustive collection of health care data, as processed by the health insurance services in Belgium. It also facilitates the Permanent Sample – Echantillon Permanente Steekproef (EPS) – which contains a 1/40 randomly sampled cohort of health insured individuals in Belgium for whom their data regarding health insurance, which is compulsory in BE, is collected over time, allowing to perform observations on an individual level. This cohort is considered to be a representative sample of the whole of the Belgian population (including all three main regions) and was thus identified as the preferred data source to collect the insurance-based morbidity data. The database consists of three types of data, one regarding the population, one with reimbursed health care procedures, and one with reimbursed medications.

2.1.1.4.2. Population coverage

The health insurance in BE is compulsory, meaning that the data of the IMA are complete. The EPS cohort is representative of the population.

In the EPS dataset, the residence of the patient is determined based on registered domicile. It is therefore possible to see regional differences within BE for a certain disease, which would enable the EPS data to serve as a regional correction for other data sources. Non-residents can be identified by having a foreign domicile. However, people identified as non-residents could include people living in Belgium for more than 12 months, but still having their official address in a foreign country; while patients having their official address in BE but living for more than 12 months abroad would still be seen as residents.

2.1.1.4.3. Data availability, linkage

An agreed collaboration between IMA and Sciensano allowed the use of the EPS data. The insurance-based data were accessed using a remote access system. However, linkage to other data sources was not feasible.

In 2019, the insurance-based information of 2017 was available.

2.1.1.4.4. Classifications used and coding issues

Medications were coded according to the Anatomical Therapeutic Chemical (ATC) classification. As far as possible, each indicator definition was based on a predefined pseudo-diagnosis based on a collection of ATC codes with a Defined Daily Doses (DDD) threshold, as determined by an expert group.

2.1.1.4.5. Other problems of recording or definition

The operational case definition was based on the reimbursed drug prescriptions as a proxy for the disease. This method however was handled with caution, as using such proxy diagnosis can easily lead to a wrong conclusion. For example, patients not taking the drugs but who do have the disease nonetheless are not identified as a patient when using this approach, which could lead to an underestimation of the indicator. Equally, patients taking the drugs for other reasons than the disease of interest (e.g. in the context of drug repositioning) should be excluded as these would lead to an overestimation. To exclude these non-cases, a threshold value to determine a real case is based on the predefined number of Defined Daily Doses (DDDs) per year for each disease.

2.1.1.4.6. Relevance to indicators

Used for indicators P1-P3, P5, and P9.

2.1.2. Estimation methods

2.1.2.1. Overview

The data sources available to the project in BE separately covered primary care, hospital inpatients, causes of death and insurance-based prescription data, but could not be linked. As a result, one 'best' source was chosen for each indicator. The hospital inpatient and causes of death data have comprehensive national coverage. The insurance-based data are based on a representative sample, therefore national extrapolation is feasible. The primary care data are from a sample of GP practices in the Flanders region, therefore complex estimation is required, and the accuracy of these estimates is more doubtful.

2.1.2.2. Data linkage

Each data source is believed to have accurate internal linkage (where appropriate) preventing duplication and assigning episodes of care correctly. However, the data

protection agreements and secure methods of access required to use the data meant that no linkage between the sources was possible.

2.1.2.3. Weighting and extrapolation

The method of weighting and extrapolation to the national estimate was explained in detail in the report. For the analysis, the 'mgcv' package in R (Wood 2017) was used to generate a Generalised Additive Model (GAM) of the prevalence/incidence data. Using such a model allows to approach the binary state of an individual (case vs non-case) through a binomial probability distribution and smooth the data, while accounting for the given variables as gender, residence, age. The results were then upscaled to the Belgian population, using the mid-year population data of the reference year (2016) as the reference population, by each corresponding category, broken down by age, gender and residence.

2.1.2.4. Conclusion

The BE data are based on data sources of variable quality. Although a well-designed estimation method was applied, it is unknown how far this was able to account for lack of representativeness in the primary care date. More importantly, the inability to link the different data sources meant that in the end only one data source was used as the 'best fit' per indicator, contrary to the multi-source approach which is preferred to ensure full ascertainment of cases in morbidity statistics. Therefore, in most cases the BE data cannot be considered of sufficient quality. The main weaknesses are:

- Only one data source per indicator severely restricts the ability to ascertain cases with sufficient completeness.
- Several issues with the quality of the primary care data are uncertain, including representativeness and completeness.
- The data sources use different classifications: this means that for some indicator's coverage based on primary care data using ICPC-2 are not completely comparable with ICD-10 definitions.

2.2. Croatia (HR)

Participation for HR was by the Division of Public Health in the Croatian Institute of Public Health (CIPH).

Estimates were reported for all the main list of indicators P1-P35 (list A) and also the optional indicators PB36-PB43 relating to external causes (list B).

The CIPH has collected the data from the primary healthcare system as individual records instead of aggregated data since 2015. To allow full use of this data source, the estimates for all indicators in HR were calculated with a reference year of 2017 and, in the case of period prevalence of 3 years, the period was 2015-2017.

The data sources used by HR were as follows:

Primary care: Primary care (GP´s) database containing information on all medical services provided by primary health care providers in Croatia

1. *Causes of death*: The register of causes of death in Croatia
2. *Hospital inpatients*: The database containing information on all patients having undergone hospital treatment in Croatia
3. *Disease-specific registers*: Registry of communicable diseases
4. *Other*: The Disabled Persons Registry

As all the data sources are collected in similar linked registers, the general information is covered in the first sub-section (Primary care) and only key points and differences are given for each different dataset.

2.2.1. Data sources used

2.2.1.1. Primary care database

2.2.1.1.1. General description

The CIPH is the central holder of health statistics under the legislation on official statistics. Data for all the registers are collected at individual record level, with a personal identifier which allows exclusion of duplicates and linkage between the data sources. Both primary and secondary care are covered by the data.

2.2.1.1.2. Population coverage

All patients of primary healthcare providers in the public sector were included. Primary healthcare providers include GPs, primary paediatricians, primary gynaecologists and dentists, who have contracts with the Croatian Insurance Fund. Since all GPs are in the public sector this is the great majority of the population. The institutionalised population is included.

2.2.1.1.3. Data availability, linkage

The CIPH collects data for all the registers at individual record level, with a personal identifier which allows

exclusion of duplicates and linkage between the data sources.

According to national legislation, the time between the end of the reference period and the release of data is less than one year.

2.2.1.1.4. Classifications used and coding issues

ICD-10 is used to record diagnoses in all data sources.

2.2.1.1.5. Other problems of recording or definition

In primary healthcare, there may be overestimation because of unconfirmed or 'working' diagnoses, or diagnoses based only on drug prescriptions.

2.2.1.1.6. Relevance to indicators

Used for all indicators.

2.2.1.2. Causes of death

2.2.1.2.1. General description

The Croatian Bureau of Statistics (CBS) collects death registration forms for statistical purposes on a monthly basis from all county registration offices and keeps their records. Pursuant to an agreement with the CBS, the CIPH is responsible for the quality of data on the causes of death and performs tasks of querying the country offices when needed, determining and coding the underlying causes of death.

2.2.1.2.2. Population coverage

Data on deaths refer to all deceased persons who were permanent residents of Croatia and had not been absent for longer than a year, as well as all deceased persons who were not permanent residents but have been present for a year or longer.

2.2.1.2.3. Data availability, linkage

See 'primary care' above.

2.2.1.2.4. Classifications used and coding issues

The ICD-10 is used to record underlying cause of death. Not all WHO updates are applied in HR and external causes of death are not coded.

2.2.1.2.5. Relevance to indicators

Used for all indicators.

2.2.1.3. Hospital inpatients

2.2.1.3.1. General description

The register of hospital inpatients covers episodes of care in all public and private sector hospitals.

2.2.1.3.2. Population coverage

All patients of secondary healthcare providers in both public and private sectors were included. The institutionalised population is included.

2.2.1.3.3. Data availability, linkage

See 'primary care' above.

2.2.1.3.4. Classifications used and coding issues

See 'primary care' above.

2.2.1.3.5. Relevance to indicators

Used for all indicators.

2.2.1.4. Registry of communicable diseases

2.2.1.4.1. General description

The Department of Infectious Diseases and Epidemiology of the CIPH acts as the centre for information for reporting and monitoring communicable diseases.

2.2.1.4.2. Population coverage

Individuals are included on a national basis, according to the scope of the registry. The completeness of reporting is not stated in the report.

2.2.1.4.3. Data availability, linkage

See 'primary care' above.

2.2.1.4.4. Classifications used and coding issues

See 'primary care' above.

2.2.1.4.5. Relevance to indicators

Used for indicator P19 only.

2.2.1.5. Disabled Persons Registry

2.2.1.5.1. General description

The Croatian Registry of Persons with Disabilities collects data to enable the planning of appropriate preventive measures and programs for people with disabilities.

2.2.1.5.2. Population coverage

Individuals are included on a national basis, according to the scope of the registry. The completeness of reporting is not stated in the report. According to one study, there are records on 11.9% of disabled people (Puntaric et al, 2015).

2.2.1.5.3. Data availability, linkage

See 'primary care' above.

2.2.1.5.4. Classifications used and coding issues

See 'primary care' above.

2.2.1.5.5. Relevance to indicators

Used for indicators P9, P27, P28.

2.2.2. Estimation methods

2.2.2.1. Overview

The data for HR is based on linked registers held by the CIPH for various healthcare management and monitoring purposes. Some of these are long-established, while the data on primary healthcare were available in individual form for the first time only in 2015.

The report commented that since this was the first time that data from the primary health care providers was analysed, various errors in that database for the years 2015 and 2016 were identified so that corrections and adjustments caused delay in the timetable. Although workshops with relevant stakeholders for improving the quality of morbidity statistics were planned, because of the COVID-19 pandemic they were not held.

2.2.2.2. Data linkage

The data are held in registers managed by the CIPH, linked using a personal identifier. Duplicates are therefore able to be excluded. It appears that the identification of non-residents in the data is not possible.

2.2.2.3. Weighting and extrapolation

As the data are from registers with national coverage, no weighting or extrapolation was carried out.

2.2.2.4. Conclusion

The system of registers allows reliable identification of persons across data sources and episodes using a personal identifier, and the data sources mostly have good coverage of the population. The private sector primary healthcare services including some paediatricians, gynaecologists and dentists are excluded, but these are unlikely to be important for the indicators. Therefore, it is likely that the estimates are of enough quality for most indicators, however with some reservations. The weaknesses are as follows:

- The primary healthcare data are being used for the first time and may have unknown quality issues; they are less regulated than the other registers and are thought to contain some proportion of uncertain or unconfirmed diagnoses.
- The cause of death data are based on underlying cause, while the ICD-10 classification is not updated to the most recent version, and external causes of death are not counted in cause of death statistics.
- The completeness of coverage of the register of communicable diseases and the register of persons with disabilities is not stated.

2.3. Finland (FI)

Participation for FI was by the Unit for statistics and registers in the Information services department of the Finnish Institute for Health and Welfare (THL).

Estimates were reported for all the main list of indicators P1-P35 (list A) and the optional indicators PB36-PB43 relating to external causes (list B).

The data sources used by FI were as follows:

1. *Primary care*: Hospital Discharge Register
2. *Causes of death*: The national Cause of Death Register kept by Statistics Finland
3. *Hospital inpatients*: Hospital Discharge Register
4. *Hospital outpatients*: Hospital Discharge Register
5. *Other*: Diagnoses recorded from social care institutions

As the data sources 1, 3, 4 and 5 are collected in similar linked registers, the general information is covered in the first sub-section (Primary care) and only key points and differences are given for each different dataset.

2.3.1. Data sources used

2.3.1.1. Primary care institutions

2.3.1.1.1. General description

The Hospital Discharge Register in FI includes discharges from specialised health care in hospitals, including both inpatients and outpatients; from health care centres in primary care; and from social care institutions.

Finland has a comprehensive health information system, which is based on individual-level register data. A system of unique identification numbers covers all Finnish citizens and permanent residents and enables linkage of the registers. The data protection legislation allows the collection and storage of sensitive health data, as well as their secondary use for statistical purposes and scientific research. Studies have shown both a good internal validity (e.g. Gissler & Shelley 2002) and external validity of the registers (e.g. Sund 2012).

2.3.1.1.2. Population coverage

The Hospital Discharge Register includes three parts: care in the health care institutions for specialised health care, care in the social institutions, and care in the health care institutions for primary health care. Healthcare personnel and service providers are obliged to report the requested data for the statistical authorities.

The institutionalised population are included in the estimates. All data sources cover the total population, both citizens and residents. Non-residents are identified in the registers by using the municipality code number 200, which is given for those patients and clients who are permanently living in other countries than Finland.

2.3.1.1.3. Data availability, linkage

Data are available in the registers for statistical purposes regularly. The data linkages are technically easy since all registers include unique personal identification numbers. The unique identification number enable exclusion of duplicate reports within a single data source as well as in linkages between multiple data sources.

2.3.1.1.4. Classifications used and coding issues

ICD-10 codes are in used in all health registers, but some health primary care centres provide data on diagnoses using the ICPC-2 classification. The shortlist of indicators was mapped to ICPC-2 (Finnish version).

2.3.1.1.5. Relevance to indicators

Primary care – used in all indicators P1-P35 (list A) but not PB36-PB43 (list B). External causes (such as accidents) are not reported on outpatient visits.

2.3.1.2. Causes of death

2.3.1.2.1. General description

The national Cause of Death Register is kept by Statistics Finland. All underlying and multiple causes of death are recorded.

2.3.1.2.2. Population coverage

The CoD register covers all deaths occurring in FI and of Finnish citizens dying abroad whose details are registered.

2.3.1.2.3. Data availability, linkage

The data are collected by Statistics Finland. The unique personal identifier is used for linkage to the other data sources.

2.3.1.2.4. Classifications used and coding issues

Causes of death are recorded using ICD-10.

2.3.1.2.5. Relevance to indicators

Used for all indicators.

2.3.1.3. Hospital inpatients

2.3.1.3.1. General description

See under 'primary care' above.

2.3.1.3.2. Classifications used and coding issues

Diagnoses in secondary care are recorded using ICD-10.

2.3.1.3.3. Relevance to indicators

Hospital (secondary care) inpatients – used for all indicators.

2.3.1.4. Hospital outpatients

2.3.1.4.1. General description

See under 'primary care' above.

2.3.1.4.2. Classifications used and coding issues

Diagnoses in secondary care are recorded using ICD-10.

2.3.1.4.3. Relevance to indicators

Hospital (secondary care) outpatients – used in all indicators P1-P35 (list A) but not PB36-PB43 (list B). External causes (such as accidents) are not reported on outpatient visits.

2.3.1.5. Social care institutions

2.3.1.5.1. General description

See under 'primary care' above.

2.3.1.5.2. Classifications used and coding issues

ICD-10 codes are in used in all health registers, but some health primary care centres provide data on diagnoses using the ICPC-2 classification. The shortlist of indicators was mapped to ICPC-2 (Finnish version). Although the use of diagnosis codes is obligatory for the health care registers, social care institutions provide diagnoses only if they are given by a physician.

2.3.1.5.3. Other problems of recording or definition

It is known that only a minority of patients of social care institutions have a recorded diagnosis. The coverage of diagnoses has declined over time: in 2006, any diagnosis was given for 11.0 percent of reported social care episodes, while the percentage was 6.4 percent in 2016.

2.3.1.5.4. Relevance to indicators

Hospital (secondary care) outpatients – used in all indicators P1-P35 (list A) but not PB36-PB43 (list B).

2.3.2. Estimation methods

2.3.2.1. Overview

Morbidity statistics in FI are based on long-established linked registers, which have been used for epidemiological research for many years. There is therefore likely to be good confidence in the methodology.

The population coverage of the data is comprehensive and includes non-residents (who can be clearly identified) and institutionalised people. All inpatient care in secondary health care and all outpatient care in public primary and secondary health care are covered in the national estimates.

2.3.2.2. Data linkage

Linkage is based on a single unique personal identifier for all data sources.

2.3.2.3. Weighting and extrapolation

The data are based on 100% registers, so no weighting or extrapolation was done.

2.3.2.4. Conclusion

The estimates from FI are likely to be of sufficient quality, especially because of the very comprehensive coverage, with a few exceptions which have been identified in the study report. The weaknesses are as follows:

- Indicators related to injuries and external causes are known to be underestimates, since external causes are missing from part of the data, covering outpatient visits in primary and secondary care.
- Some primary care services report using ICPC-2, which is known to differ from the ICD-10 definitions for some conditions. This particularly affects indicators P24-P26 on liver disease.

2.4. France (FR)

Participation for France was by the Office of Population Health Status within the Research, Studies, Evaluation and Statistics Department (DREES) of the Ministry of Social Security and Health.

Estimates were reported for all the main list of indicators P1-P35 (list A) and also the optional indicators PB36-PB43 relating to external causes (list B).

The data sources used by FR were as follows:

1. *Hospital inpatients*: Programme National de Médicalisation des Systèmes d'Information (PMSI) database, covering all hospital inpatient activities.
2. *Insurance*: The national health insurance fund information system (SNIIRAM database).
3. *Prescriptions*: Also collected within the SNIIRAM database, based on reimbursed drug prescriptions.

2.4.1. Data sources used

2.4.1.1. Programme National de Médicalisation des Systèmes d'Information (PMSI) database

2.4.1.1.1. General description

The main objective of the PMSI is to analyse the medical activity of hospitals for budget allocation purposes. All inpatient activity is included. As part of the PMSI, any stay in a health facility, public or private, is the subject of a systematic and minimal collection of administrative and medical information that is used mainly for the financing of health facilities (activity-based payment) and for the organization of the offer of care (planning).

More information on the hospital data (PMSI database) is published at https://www.atih.sante.fr/notice-technique-pmsi-2020-0 and in Boudemaghe & Belhadj (2017).

The centrally collected Anonymous Release Summary (RSA) contains medical information (related diagnoses, performed medical acts, etc.) as well as administrative information (identification of the facility, length of stay, mode of entry and exit of which, possibly, the death) and the patient (gender, age, geographic code of residence based on the postal code of residence). In addition to the RSA, institutions must provide information on the number of consultations and external acts performed, on the use of certain drugs and implantable medical devices (prostheses, implants) used. The data are subject to quality and accuracy controls led by the Social Security administration, potentially resulting in fines for those hospitals not respecting certain coding criteria.

2.4.1.1.2. Population coverage

The systematic collection covers all hospitalizations in curative care, rehabilitative care, psychiatric care and home hospitalizations, in the public and private sectors. For this project, hospitalizations of non-residents could not be included as they could not be linked.

2.4.1.1.3. Data availability, linkage

The PMSI is managed by the Technical Agency for Information on Hospitalization (ATIH). Information collected at the facility level is then centralized at the national level on the form of an Anonymous Release Summary (RSA). Within the PMSI database a deterministic anonymous linkage identifier is implemented allowing individuals to be accurately connected between episodes of hospital care.

The availability for statistical use is at approximately T+18 months.

The PMSI data are linked within the National Health Data System (SNDS) to other data sources, and in turn linked to the Permanent Demographic Sample (EDP) which is a large representative sample of residents of FR based on linked administrative datasets. Data linkage is ensured by means of an anonymous and unique secure identification number for each person.

2.4.1.1.4. Classifications used and coding issues

The classifications used are ICD-10 for all diagnoses recorded in hospitals.

2.4.1.1.5. Relevance to indicators

The PMSI inpatient data were used for all indicators. For 17 indicators consisting mainly of acute and serious conditions this was the only source used (P14, P15, P17, P19, P31-P35, PB36-PB43).

2.4.1.2. The national health insurance fund information system (SNIIRAM database)

2.4.1.2.1. General description

The SNIIRAM database was created to contribute to the knowledge of the expenses of all the health insurance schemes, the definition, the implementation and the evaluation of the health policies, the improvement of the quality of the care and transmission to health professionals of information relating to their activity, income and, where appropriate, their prescriptions. SNIIRAM is managed by the National Health Insurance Fund for Salaried Workers (CNAMTS). It is an anonymous data warehouse gathering information from reimbursements made by all health insurance, containing approximately 1.2 billion care records for the entire population living in France. More information is published at *https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/sniiram/description-des-donnees-et-wiki-sniiram.php* and in descriptive articles such as Bezin (2017).

2.4.1.2.2. Population coverage

The coverage of the SNIIRAM data extends to all health insurance schemes and represents 98.8% of the resident population of France.

2.4.1.2.3. Data availability, linkage

The availability for statistical use is at approximately T+18 months.

The SNIIRAM data are linked within the National Health Data System (SNDS) to other data sources, and in turn linked to the Permanent Demographic Sample (EDP) which is a large representative sample of residents of FR based on linked administrative datasets. Data linkage is ensured by means of an anonymous and unique secure identification number for each person.

2.4.1.2.4. Classifications used and coding issues

The classifications used are ICD-10 for diagnoses.

2.4.1.2.5. Relevance to indicators

The SNIIRAM data on insurance reimbursements were used for the majority of indicators (P1-P13, P16, P18, P20-P30).

2.4.1.3. Prescriptions database

2.4.1.3.1. General description

The drug prescriptions reimbursement data are contained in the SNIIRAM insurance database, and the description is the same except where shown below.

2.4.1.3.2. Population coverage

As above.

2.4.1.3.3. Data availability, linkage

As above.

2.4.1.3.4. Classifications used and coding issues

Medicines are recorded according to the ATC classification.

2.4.1.3.5. Relevance to indicators

Used for indicators P1, P5, P6, P8, P9, P12, P22, P23, P29.

2.4.2. Estimation methods

2.4.2.1. Overview

The estimates for FR were based on combined data from insurance records, hospital inpatients and prescriptions, which are all linked within the National Health Data System (SNDS). The figures are extrapolations from a sample as explained below.

Although there are plans for the addition of cause of death data and certain other data sources in the SNDS, the use of these sources was not achievable in this project. The report from FR noted that:

Data on causes of death for 2016 are available as such but are not yet integrated into the SNDS, which means that they cannot be linked with the data of the other sources. It was also noted that the availability of CoD data is T+4 years.

- There is a lack of data from emergency units. The database exists and Individuals are well identified but the possibility of linkage is up to now still to study.
- There is no information about diagnosis in the primary care database.

2.4.2.2. Data linkage

Data linkage is ensured by means of an anonymous and unique secure identification number for each person. Non-residents having health care in France are not taken into account, but persons affiliated to the French health system and living abroad may be included.

2.4.2.3. Weighting and extrapolation

The data extracted from the National Health Data System (SNDS) are linked to the Permanent Demographic Sample (EDP) which is a large representative sample of residents of FR based on linked administrative datasets. The EDP is a longitudinal socio-demographic panel, created in 1968 by INSEE, to study the career, residential or family backgrounds of people residing in France. It brings together information from different administrative sources and investigations for individuals born on certain days of the year (4 days a year until the early 2000s, then 16 days a year, or 4.4% of the population).

Estimates are computed based on weights from the related FIDELI socio-fiscal data. One weight is computed for each pair (patient:year) to obtain representativity of the French population. Some patients found in the insurance data are excluded from the analysis if they are not present in this reference data source.

2.4.2.4. Conclusion

According to the report from FR, only five indicators were described as satisfactory (P5, P6, P30, P34, P35). All others were expected to have various degrees of underestimation

because of the absence of data from primary care, emergency care and causes of death. Six indicators were stated as definitely invalid based on the available data (P7, P20, P21, P31, P36, P37).

It is therefore to be expected that, because of the lack of important data sources, the estimates from FR overall cannot be considered satisfactory. The production of morbidity statistics will become feasible after at least the addition of CoD data, and preferably the additional sources. The practical production of regular morbidity statistics will also depend on the timeliness of the CoD data which has the longest delay of the sources considered.

2.5. Hungary (HU)

Participation for Hungary was by the Health Statistics Section of the Hungarian Central Statistical Office (HCSO).

Estimates were reported for all the main list of indicators P1-P35 (list A) and the optional indicators PB36-PB43 relating to external causes (list B).

The data sources used by HU were as follows:

1. *Causes of death*: The causes of death statistics of the Hungarian Central Statistical Office (HCSO).
2. *Insurance*: The Hungarian National Health Insurance Fund (NHIF) special care episode report databases, consisting of:
   - Outpatient care episode report database
   - Inpatient care episode report database
   - Drug prescriptions database.
3. *Other*: The European Health Interview Survey (EHIS) was used for some indicators to make estimates for the private sector.

2.5.1. Data sources used

2.5.1.1. Causes of Death Register

2.5.1.1.1. General description

Causes of death data are collected by the Hungarian Central Statistical Office (HCSO) under the relevant laws. All causes of death are recorded for statistical purposes.

2.5.1.1.2. Population coverage

The CoD data cover all deaths in the territory of Hungary, and in addition the deaths of persons having a permanent address in Hungary that occurred abroad and were registered in Hungary. Causes of death data for all indicators of the pilot data collection were queried from the CoD database of HCSO. All diagnoses in death certificates were included in the queries regardless of the type of the cause of death: whether it was the underlying cause, consequence of the underlying cause or the direct cause of death.

2.5.1.1.3. Data availability, linkage

CoD data refer to the previous calendar year.

The owner of causes of death data is the HCSO. Strict rules that apply to the management of CoD data made necessary to find special ways to be able to use these data in the indicator calculations. After processing CoD data, Health Insurance Numbers are deleted from the database according to data protection rules. Since the CoD data are then available without a unique identifier, linkage was carried out based on sex, date of birth and date of death.

2.5.1.1.4. Classifications used and coding issues

Causes of death were recorded according to the International Classification of Diseases, 10th revision (ICD-10). All recorded causes of death were used.

2.5.1.1.5. Relevance to indicators

The CoD data was used for all indicators.

2.5.1.2. Hungarian National Health Insurance Fund (NHIF) databases

2.5.1.2.1. General description

Hungarian National Health Insurance Fund (NHIF) special care episode report database includes all of the following:

1. General Practitioners' care episode report database
2. Outpatient care episode report database
3. Inpatient care episode report database
4. Dialysis care episode report database
5. Drug prescriptions database
6. Report on individually financed devices and procedures
7. Database of Health Insurance Identification Number (TAJ).

Out of these, the data on outpatients (2), inpatients (3) and prescriptions (5) were used for the estimates, while the ID number data (7) were used for identification. Prescriptions were only used for some indicators. Although these are all part of the NHIF data, each database can be separated, and for completeness the sources are described separately.

2.5.1.2.2. Population coverage

The NHIF databases cover all publicly financed health services. Some 95-98% of the population are entitled for health care services financed by NHIF. Common regulations and audited IT-tools guarantee statistical consistency within the data sets. All residents including institutionalized individuals and all age-groups are included. Data on non-residents are not available.

NHIF regularly monitors and checks health service providers sending data to the database.

2.5.1.2.3. Data availability, linkage

The primary owner of health insurance data is the NHIF. According to the legislation, the Ministry of Human Capacities and the National Healthcare Service Centre (NHSC) are entitled to use the health insurance databases of NHIF except for the General Practitioners' care episode report database. There are initiatives by NHSC to change the legal regulation to obtain access to the GP data, but they have not been successful so far.

NHIF data are collected for reimbursement purposes, and data refer to the previous calendar year. Databases are available at T+9 months.

The NHIF provides data to NHSC replacing the Health Insurance Number by a pseudo identifier, so the data are anonymised and linkage to the CoD data is by sex, date of birth and date of death.

2.5.1.2.4. Classifications used and coding issues

Diagnoses are recorded according to the International Classification of Diseases, 10th revision (ICD-10). Medications prescribed were recorded according to the ATC classification.

Inpatient diagnosis data were queried taking into account the following types of ICD codes:

disease underlying the principal diagnosis justifying care/ in the case of rehabilitation wards disease underlying the principal diagnosis justifying rehabilitation care

- principal diagnosis justifying care/in the case of rehabilitation wards principal diagnosis justifying rehabilitation care
- complication
- accompanying disease
- external causes of injuries and poisonings
- ancillary ICD code with flag '* ' to the principal diagnosis justifying care/in the case of rehabilitation wards to the principal diagnosis justifying rehabilitation care

In the case of outpatient care episode data, data of diagnostic wards were excluded from the data calculation because diagnoses at these wards are preliminary diagnoses.

2.5.1.2.5. Other problems of recording or definition

In the cases of indicators P4, P7, P9, P14-P19, P27 and P30, queries from the drug prescriptions database of NHIF were carried out by ICD codes but data were not included in the union of NHIF databases. When calculating the data source matrix, it was examined whether pseudo identifiers of cases in the union of the other NHIF databases were included in the 'filtered' drug prescriptions database, i.e. whether the person had a drug prescription with the ICD code of the indicator, and the result was included in the matrix. No unique cases were discovered by these searches.

2.5.1.2.6. Relevance to indicators

The health insurance (NHIF) data was used for all indicators.

Regarding P14, P15, P17, inpatient diagnosis data were examined using the following types of ICD codes only:

- principal diagnosis justifying care/in the case of rehabilitation wards principal diagnosis justifying rehabilitation care
- disease underlying the principal diagnosis justifying care/ in the case of rehabilitation wards disease underlying the principal diagnosis justifying rehabilitation care

For these three indicators, inpatient care data were used exclusively because the occurrence of these acute diseases always needs hospital treatment.

2.5.1.3. European Health Interview Survey (EHIS)

2.5.1.3.1. General description

The 2019 EHIS for HU was used to estimate the proportion of persons for some indicators using private sector care, and therefore not counted in the health insurance data. Because the EHIS is not itself a morbidity data source, a full description is not required here.

Diseases with any chance of exclusive private health care treatment were inserted in the list of chronic diseases of EHIS 2019 in HU and additional questions were added to the questionnaire. The health conditions covered for this purpose, the wording of the questions, and the principles and assumptions applied for use of the EHIS data are shown

in annex – B 1.1 Use of the European Health Interview Survey (EHIS) in HU to identify proportions of individuals with certain conditions using private healthcare.

2.5.1.3.2. Population coverage

EHIS data are based on a representative sample of 15+ year old population having a permanent address in Hungary.

2.5.1.3.3. Data availability, linkage

EHIS data are collected for statistical purposes, data refer to the previous year and definitions are used according to the EHIS implementing Regulation (EU) 2018/255 [5].

The EHIS data are anonymised and were only used in aggregate form.

2.5.1.3.4. Relevance to indicators

The EHIS data was used to adjust for patients in the private sector, for indicators P1, P2, P6, P7, P11-P13, P20-P22, and P27-P29.

2.5.2. Estimation methods

2.5.2.1. Overview

The health insurance databases for outpatient and inpatient care were used, linked to the CoD database. Drug prescriptions, which also form part of the health insurance data, were searched for cases matching the relevant indicators for comparison only.

EHIS data were used to determine the proportion of public and private services (see below).

Data from the Report on certain diseases of persons registered at the general practitioners' service and aggregated data of the National Registry of Myocardial Infarction were used independently for validation only.

2.5.2.2. Data linkage

Records from the separate sources within the health insurance (NHIF) data were linked to each other by the pseudo Health Insurance Number and the date of health care contact. CoD death data were then linked to NHIF linked database by sex, date of birth and date of death.

2.5.2.3. Weighting and extrapolation

The NHIF and CoD data used are complete and of national scope without any known geographical bias.

For health conditions where it is possible to carry out treatment exclusively in the private sector, the proportions of persons using private sector care were estimated based on questions included in the EHIS. Parameters were calculated based on the number of persons reporting themselves as having used exclusively private services, exclusively public services, or both, for that health condition. Numbers from the linked NHIF and CoD data were multiplied by the relevant parameter to produce the estimate, for each age-sex group separately.

The algorithm of this calculation was stated as:

Terms: N: Prevalence or incidence to be estimated (public and private sector together)

Npu: Exclusive users of public services

Npr: Exclusive users of private services

Nmix: Users of both public and private services

NR: Data from NHIF database (register) and causes of death statistics

Calculation: $N_e x$ is the estimation of Nx from EHIS where 'x' can be 'pu' or 'pr' or 'mix'.

$NR = Npu + Nmix$

$N = NR + Npr$

If $P = 1 + {}_e pr / (N_e pu + N_e mix)$, then

$N_e = NR*P$

Because of outliers caused by calculation based on small numbers, especially at the younger ages, standard deviations of the parameters by age-group were calculated for all diseases and the maximum was determined as average + 3*standard deviation, so higher values were cut back to that level.

2.5.2.4. Conclusion

The data from HU have not been previously used for morbidity statistics in this form but come from routine sources which are well validated and quality assured. The success rate of linkage using sex, date of birth and date of death was not stated, but neither is the method considered a problem. The coverage of health insurance is 95-98% of the population. HU made efforts to adjust the estimates

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R0255

for persons using private sector services where that was appropriate but expressed the opinion that this method did not fully account for the issue. Generally, it is likely that the estimates from HU are of sufficient quality.

The possible weaknesses are as follows:

Certain private sector healthcare providers and uninsured persons are excluded, despite the adjustment made.

- Counting of non-residents is not feasible from the available data in HU.
- The insurance-based data may be affected by incentives to report diagnoses attracting higher reimbursement or for similar reasons.

2.6. Lithuania (LT)

Participation for Lithuania was by the Health Statistics Department of the Institute of Hygiene.

Estimates were reported for all the main list of indicators P1-P35 (list A) and the optional indicators PB36-PB43 relating to external causes (list B).

There were only two data sources used by LT, as follows. Further details of these are given below.

1. *Causes of death*: The Causes of Death Register, including underlying and secondary causes of death.
2. *Insurance*: The Compulsory Health Insurance Fund Information System (CHIF IS) including all registered diagnoses reported by providers of primary, hospital inpatients, hospital outpatients, and emergency care.

2.6.1. Data sources used

2.6.1.1. Causes of Death Register

2.6.1.1.1. General description

Death certificates are collected by Institute of Hygiene according to the Law on Causes of Death Register. Underlying cause of death and multiple causes are all collected. Metadata for causes of death statistics in LT are published – see http://hi.lt/kodeks/Mirties_priezasciu_metainfo_en.docx.

2.6.1.1.2. Population coverage

All deaths of residents of LT are included, but not non-residents.

2.6.1.1.3. Data availability, linkage

Causes of death data are fully available to the Institute of Hygiene. However, because the CHIF IS insurance-based data are anonymised, linkage is carried out using a deterministic method (by date of birth, date of death and sex) and not by a personal identifier. Linkage is only on a one-to-one relation and is highly successful – particularly because the population of LT is quite small, so there is limited potential for duplication of individual values. Data for 2010-18 were linked and for each year the linkage rate was more than 99%.

2.6.1.1.4. Classifications used and coding issues

For CoD data, validation is done using mathematical and logical rules. Data are checked and coded by coders according to the international rules of coding of death certificates. Coders compare diagnosis on the death certificate with the diagnosis registered in CHIF IS, and if there is inconsistency the coder can contact the physician who issued the death certificate for clarification and correction.

2.6.1.1.5. Relevance to indicators

The CoD data was used for all indicators.

2.6.1.2. Compulsory Health Insurance Fund Information System (CHIF IS)

2.6.1.2.1. General description

The Compulsory Health Insurance Fund Information System (CHIF IS) is the most comprehensive database of health information in LT and has been used by the Institute of Hygiene for purposes of statistical data calculation including diagnosis-specific morbidity calculation since 2004. Therefore, it is the main source for morbidity data, excluding infectious diseases and cancer for which specific data sources (surveillance system and registers) exist. CHIF IS covers all health care institutions having contracts with CHIF to provide care under the national health insurance.

2.6.1.2.2. Population coverage

The coverage of the CHIF IS data is: primary health care visits (incl. all registered diagnoses) – 100%, hospital and emergency inpatients and outpatients episodes (incl. the main diagnosis, complications and comorbidities) – 99%, outpatient visits in hospitals and specialised outpatient health care institutions (incl. all registered diagnoses) – 90%, dentist visits (incl. all registered diagnoses) – 40-

50%. CHIF IS does not cover a few state and private care institutions, having no contracts with CHIF, most of which are quite small.

2.6.1.2.3. Data availability, linkage

The Institute of Hygiene receives CHIF IS data regularly for statistical purposes under an agreement with the CHIF, but only as an anonymised copy. Within the CHIF IS database, all records are identified by a personal ID number except for foreign non-residents. See under Causes of Death above for linkage between the two data sources.

2.6.1.2.4. Classifications used and coding issues

Diagnoses are coded according to ICD-10.

The records of CHIF IS include the diagnostic modifiers '+' meaning a new case and '0' meaning a suspected diagnosis. Cases with the '0' modifier were not considered.

Although the '+' modifier is used to identify a new case in the national methodology of LT for morbidity statistics, for these estimates it was not applied is it is used in outpatient care only and the consistency of recording the modifier is not thought to be sufficient. Therefore, LT followed the Guidelines for incidence per person and considered a new case as one with no previous contact in two years, for incidence by episode, the length of episode for particular diseases was used.

2.6.1.2.5. Other problems of recording or definition

The situation regarding residents and non-residents is complex in LT as many Lithuanian citizens, who have a personal ID, are living abroad. Non-residents are therefore divided into Lithuanian non-residents and foreign non-residents. In the national methodology for morbidity statistics, but not for these estimates, all persons having a Lithuanian personal ID are included.

In CHIF IS, for persons having a Lithuanian ID number only the date of last emigration (or death) is recorded. There is no history of movement of the person. Therefore, if the status of residence was changed during the year, only the last status for that year can be considered only. Emigration status is not known for dead persons.

For foreign non-residents, indicators could not be calculated correctly. As they have no ID code it is impossible to do record linkage and eliminate duplication of cases, to add diseases registered during 2 years before reference year (for period prevalence), or to distinguish between episodes (for incidence by episode). Therefore, every registered diagnosis of a foreign non-resident in the year 2016 was included into incidence by episode and incidence by person.

2.6.1.2.6. Relevance to indicators

The CHIF IS data was used for all indicators.

2.6.2. Estimation methods

2.6.2.1. Overview

The project team decided on linking CHIF IS and the Causes of Death Register (including multiple causes of death) as the basis for all indicators. This linkage was said to be necessary to increase the coverage for diseases with high sudden mortality rate (like myocardial infarction, injuries). Including multiple causes of death could be important for many diseases especially when the person had no recent healthcare contacts.

2.6.2.2. Data linkage

Linkage within the CHIF IS carried out reliably using a personal ID number for residents and Lithuanian non-residents, but not for foreign non-residents. This means that the calculation of episodes and identification of new cases should be accurate except for foreign non-residents, as noted above.

Because the CHIF IS data available to the Institute of Hygiene are anonymised, linkage to the CoD data was carried out using a deterministic method (by date of birth, date of death and sex) and not by the personal identifier. Linkage is only on a one-to-one relation and is highly successful – particularly because the population of LT is quite small, so there is limited potential for duplication of individual values. Data for 2010-18 were linked and for each year the linkage rate was more than 99%.

2.6.2.3. Weighting and extrapolation

The data cover the whole of LT and are considered comprehensive, except for healthcare contacts in a small number of institutions that are not funded by the CHIF. Therefore, no weighting or extrapolation was carried out.

2.6.2.4. Conclusion

The data from LT are based on well-understood insurance and CoD data which have already been used for national

morbidity statistics for many years. Although the linkage of the data sources was in an anonymised form, the method used is more than 99% successful. Generally, the estimates from LT can be assumed to be of enough quality.

The possible weaknesses are as follows:

Certain private sector healthcare providers and uninsured persons are excluded.

- Counting of non-residents is not reliable, as individuals or episodes cannot be linked.
- The insurance-based data may be affected by incentives to report diagnoses attracting higher reimbursement or for similar reasons.

2.7. Malta (MT)

Participation for Malta was by the unit for Morbidity Statistics within the Directorate for Health Information and Research of Ministry for Health.

Estimates were reported for all the main list of indicators P1-P35 (list A) except for:

P25: Diseases of liver other than alcoholic (K71-K77) – Period prevalence

- P28: Arthrosis (M15-M19) – Period prevalence
- P29: Osteoporosis (M80-M82) – Period prevalence
- P32: Intracranial injury (S06) – Incidence by episode
- P34: Fracture of femur (S72) – Incidence by episode

In addition, estimates were reported for the optional indicators PB36-PB43 relating to external causes (list B) except for:

- PB36: Land transport accidents (V01-V89) – Incidence by episode
- PB38: Accidental falls (W00-W19) – Incidence by episode

The data sources used by MT were as follows:

Causes of death: National Mortality Registry (NMR)

1. *Hospital inpatients*: National Hospital Information System (NHIS)
2. *Prescriptions*: Pharmacy of your choice (POYC) database
3. *Emergency care*: Injury database (IDB)
4. *Other*: Patient master index (PMI)

2.7.1. Data sources used

2.7.1.1. National Mortality Registry (NMR)

2.7.1.1.1. General description

The NMR is a national register, which collects information on all deaths in the Maltese Islands.

2.7.1.1.2. Population coverage

The NMR collects information on deaths of residents and non-residents occurring in the Maltese Islands, as well as deaths of Maltese citizens occurring abroad, when available.

2.7.1.1.3. Data availability, linkage

The linkage between the data sources is done via the unique personal identifier.

2.7.1.1.4. Classifications used and coding issues

The International Classification of diseases, 10th edition (ICD-10) is used for the classification of multiple causes of death.

2.7.1.1.5. Relevance to indicators

Used for indicators P2-P6, P8-P10, P12-19, P21, P24, P26-P27, P30, P33, P35, PB37, PB39-PB43.

2.7.1.2. National Hospital Information System (NHIS)

2.7.1.2.1. General description

The NHIS provides information on all patients admitted to a hospital for an overnight stay by diagnosis on discharge, and covers episodes of care in all public and private sector hospitals in Malta.

2.7.1.2.2. Population coverage

The NHIS provides data on both residents and non-residents.

2.7.1.2.3. Data availability, linkage

The linkage between the data sources is done via the unique personal identifier.

2.7.1.2.4. Classifications used and coding issues

The International Classification of diseases, 10th edition (ICD-10) is used for the classification in the NHIS.

2.7.1.2.5. Other problems of recording or definition

In case of prevalence, multiple diagnoses were checked for a particular health condition. In contrast, for incidence cases only the first primary diagnosis was used. One of the main limitations in using NHIS as data source for the morbidity project was exclusion and inclusion of specific cases. For example, conditions like renal failure or heart failure could be considered as a terminal causes in patients with cancer. Therefore, looking at chronic kidney failure or chronic heart disease in this type of data collection seems to be more accurate.

2.7.1.2.6. Relevance to indicators

Used for indicators P1-P27, P30-P31, P33, P35, PB37, PB39-PB43.

2.7.1.3. Pharmacy of your choice (POYC) information system

2.7.1.3.1. General description

Pharmacy of your choice (POYC) is a register, which collects information about all persons receiving medication from the free government scheme by disease, and by year they were registered to obtain the medication. This information is often available for many years and for many diseases, and is linked via the unique personal identifier.

2.7.1.3.2. Population coverage

The POYC collects information about residents only.

2.7.1.3.3. Data availability, linkage

The linkage between the data sources is done via the unique personal identifier.

2.7.1.3.4. Classifications used and coding issues

Not stated in the report.

2.7.1.3.5. Other problems of recording or definition

The POYC information system due to usage of 3-year period prevalence (rather than longer periods) and exclusion of dispensing history produces underestimates for some of the health conditions. As a result, person with long-term medical condition that was registered with POYC in 2012 would not feature in POYC data for 2013-2016 due to the fact that POYC has only information on the medical condition for which patient is entitled to receive medication, however, does not include dispensing history for the patient.

2.7.1.3.6. Relevance to indicators

Used for indicator P1-P13, P16, P18, P20-P24, P26-27, P30.

2.7.1.4. Injury Database (IDB)

2.7.1.4.1. General description

The injury database (IDB) collects information about persons seen at the accident and emergency department of the main public hospitals in Malta and Gozo.

2.7.1.4.2. Population coverage

The completeness of reporting is not stated in the report.

2.7.1.4.3. Data availability, linkage

The linkage between the data sources is done via the unique personal identifier.

2.7.1.4.4. Classifications used and coding issues

The coding is not related to ICD-10. Nevertheless, Malta did not experience any issues when translating the required codes into the morbidity causes needed for the project.

2.7.1.4.5. Other problems of recording or definition

The IDB together with the NHIS register required manual reviewing of multiple episodes in order to calculate the incidence by episode. Decision whether the episode is new or not raised problems.

2.7.1.4.6. Relevance to indicators

Used for indicator P33, P35, PB37 and PB39-PB41.

2.7.1.5. Patient Master Index (PMI)

2.7.1.5.1. General description

The PMI is a data file which contains case-based information by unique identifier of all persons who had

been in touch with the government hospital system over the years whether as inpatients or outpatients. The PMI is mainly a demographic file and was used as linkage to other sources of information that included date of birth.

2.7.1.5.2. Population coverage

Not stated in the report.

2.7.1.5.3. Data availability, linkage

The linkage between the data sources is done via the unique personal identifier.

2.7.1.5.4. Classifications used and coding issues

Not stated in the report.

2.7.1.5.5. Relevance to indicators

Used for indicators P1-2.

2.7.2. Estimation methods

2.7.2.1. Overview

The data sources available to the project in MT covered hospital inpatients, causes of death, prescriptions and emergency care data, and they were linked via the common personal identifier.

The report commented that the COVID-19 pandemic caused delay in the timetable and some objectives, namely development of a software for direct POYC queries, could not be undertaken. Nevertheless, great efforts were done to provide data on different morbidity indicators.

2.7.2.2. Data linkage

Linkage is based on a unique personal identifier for all data sources.

2.7.2.3. Weighting and extrapolation

As the data are from registers with national coverage, no weighting or extrapolation was carried out.

2.7.2.4. Conclusion

The system of registers allows reliable identification of persons across data sources and episodes using a personal identifier, and the data sources mostly have good coverage of the population. It is likely that the estimates are of enough quality for most indicators, however with some reservations. The weaknesses are as follows:

- Certain health conditions with 3-year prevalence may be underestimated because of missing information on dispensing history in the data related to prescriptions (POYC)
- Lack of important data sources, namely primary healthcare and hospital outpatients data, as well as dispensing history data for medicines (prescriptions).

2.8. Netherlands (NL)

Participation for the Netherlands was by the Health and Care Team of Statistics Netherlands (CBS).

Estimates were reported for all the main list of indicators P1-P35 (list A) except for:

- P24: Alcoholic liver disease (K70) – Period prevalence
- P25: Diseases of liver other than alcoholic (K71-K77) – Period prevalence

No estimates were reported for the optional indicators PB36-PB43 relating to external causes (list B).

The data sources used by NL were as follows:

1. *Primary care*: Nivel Primary Care Database (Nivel-PCD), a representative selection of general practitioner practices covering about 8% of the Dutch population.
2. *Causes of death*: National Causes of Death (CoD) data including both underlying and secondary causes of death of persons deceased in the index year
3. *Hospital inpatients*: The national Hospital discharge register (HDR) which contains both inpatients and day patients.
4. *Combined hospital data*: DTC-SSC (Diagnosis Treatment Combinations Somatic Specialist Care) data, which contains outpatients, day patients and inpatients of hospitals and independent treatment centres.
5. *Prescriptions*: Dispensed medicines data, which contains prescribed medicines provided through pharmacies and covered by the health insurance system.
6. *Other*: DTC-MHC (Diagnosis Treatment Combinations Mental Health Care) which contains data on specialized mental health care.
7. *Long term care eligibility decisions* (LTC-E CIZ) and Long-term care co-payments (LTC-C CAK) – these two sources are described together.

2.8.1. Data sources used

2.8.1.1. Nivel Primary Care Database (Nivel-PCD)

2.8.1.1.1. General description

In the Netherlands several networks of general practices exist, some in a specific region and some more widely spread across the country. One of them, Nivel Primary Care database (https://nivel.nl/en/nivel-primary-care-database) is widespread throughout NL, considered to be nationally representative, and suitable for use in Morbidity Statistics as it comprises both contacts and diagnostic information, and data can be linked on person-level to other sources. It covered in 2016 about 8% of the Dutch population (https://www.nivel.nl/nl/nivel-zorgregistraties-eerste-lijn/methode-cijfers-huisartsen).

As each GP practice has a fixed population, the epidemiological denominator is known, and the source is suitable for the determination of incidence and prevalence measures. In NL, almost all inhabitants are registered with a GP practice and it is only possible to register with one GP practice at a time.

Episode information includes date of onset and end date, first and last contact with the GP practice, date of first and last prescription. A disease episode starts at the first health contact and ends a predefined period after the last health contact, depending on the type of health problem. No end of episode is defined for chronic diseases. More information about the construction of episodes and validation of the algorithm in Nivel-PCD is published in Nielen, Spronk, Davids, et al (2019).

2.8.1.1.2. Population coverage

As Nivel-PCD covers a relatively small part (around 8%) of the population, it is required to scale up the number of cases with a particular health problem to the national population. This is done by using a weighting method to correct at the same time for any remaining lack of representativeness.

General practitioners provide general primary care for the non-institutionalised population. In 2016, 2% of the population was resident in institutions, particularly older people. A small proportion of the population, mainly elderly, is therefore missing from the Nivel-PCD data.

By extrapolation of the Nivel-PCD population to the total population, incidence and prevalent rates at all ages are assumed to be equal in both the institutionalised and the non-institutionalised population. This assumption probably is not true, as especially in recent years institutionalisation only applies to those who need care 24 hours a day and are clearly in poor health.

Since the basis of the Nivel-PCD data is persons registered with a general practitioner, by definition, the non-residents are not included. Non-residents may sometimes access care from a general practitioner on a temporary basis, but such contacts are not recorded in an equivalent way to residents and could not be used in the project.

The data sources used for Morbidity Statistics cover health provided in the public sector. In NL the private sector is small.

An additional issue is that for some chronic disorders (non-insulin-dependent diabetes mellitus, COPD, asthma) so called 'multi-disciplinary coordinated care' (also known as 'chain care') was recently introduced. Several medical disciplines provide the required care together (for example, the general practitioner in collaboration with a dietician, a physiotherapist and/ or a remedial therapist) coordinated by the general practitioner. This type of care is financed in a different way than other care provided by the general practitioner and contacts are registered in separate database. In 2016, chain care contacts were not all covered by Nivel-PCD, which may affect prevalence and incidence data for corresponding indicators.

2.8.1.1.3. Data availability, linkage

Data from the Nivel-PCD primary care network are regularly available to Statistics Netherlands and are also used for a variety of research purposes.

Timely availability of the data is good. The 2019 data are expected to be available for morbidity statistics in the second quarter of 2021.

Individuals recorded in the Nivel-PCD data are residents registered with a general practitioner and are identified by a personal identifier that can be linked to the population register. The methods used for linkage are a standard practice in Statistics Netherland.

Non-residents are not recorded in an equivalent way and linkage is not possible.

Missing contact information from 'chain care' for chronic diseases may result in the overestimation of incidence for these diseases. However, 'chain care' contacts have become more available so once integrated with Nivel-PCD data this issue will be less prominent.

2.8.1.1.4. Classifications used and coding issues

Diagnoses are coded using the International Classification for Primary Care, version 1 (ICPC-1). A mapping between ICPC-1 and ICD-10 was created by the project team based on the work of Okkes, Oskam and Lamberts (2005).

For many indicators there is a direct relationship between ICPC-1 and ICD-10, for example: - ICD-10: E10-E14 Diabetes mellitus – ICPC-1: T90 Diabetes (identical coverage to ICD-10)

For others there is a mapping that is close but imperfect, for example: - ICD-10: F10 Mental and behavioural disorders due to use of alcohol – ICPC-1: P15 Chronic alcohol abuse (corresponding to ICD-10 F10.1-F10.9, plus G31.2 Degeneration of nervous system due to alcohol) – ICPC-1: P16 Acute alcohol abuse (corresponding to ICD-10 F10.0).

ICPC-1 codes P15 and P16 together give full coverage of ICD-10 F10, but also include ICD-10 G31.2.

A small number of mappings between ICPC-1 and ICD-10 are not possible, notably on the distinction between alcohol and non-alcoholic liver disease (see below).

2.8.1.1.5. Other problems of recording or definition

The definition of 'episodes' in this source depends on the type of the health problem: for chronic diseases disease episodes are never closed, for other health problems an episode closes after a certain period without any further contacts with the GP. For some indicators of morbidity statistics, a 3-year prevalence was requested on health problems that were not considered as chronic in Nivel-PCD (mostly mental health problems). In that case, sometimes no episode for the disease was present in 2016-data. Therefore, also episodes from 2014 and 2015 had to be used.

2.8.1.1.6. Relevance to indicators

Used for all the indicators reported by NL except P14-P15 Acute myocardial infarction; P30 Renal failure; P34-P35 Fracture of femur.

The indicators P24 Alcoholic liver disease and P25 Diseases of liver other than alcoholic could not be reported, because the ICPC-1 code D97: Cirrhosis/other liver disease makes no distinction on whether alcohol is involved or not.

Other classification issues reported by the project, causing the Nivel-PCD data to be unsuitable for certain indicators, were: - P14-P15 Acute myocardial infarction: The number of cases found only in primary care was unexpectedly high. It was thought to indicate the use of a different 'concept' of the disease or health problem in primary care and hospital care. Including primary care cases would increase the number of acute myocardial infarctions by 130 percent. – P27 Rheumatoid arthritis: It was known from other studies that the diagnostic code in primary care (ICPC-1 L88) was also used for other forms of arthritis. As it was assumed that most patients with rheumatoid arthritis would also be known from sources on specialized somatic care it was decided not to use Nivel-PCD. Inclusion of primary care would result in an increase of 114 percent of the number of cases found. – P30 Renal failure: No adequate ICPC-1 code was available to cover only renal failure. It is included in ICPC-code U99 (Urinary disease, other) which would lead to the inclusion of many other health problems than only renal failure. – P34-P35 Fracture of femur: As for acute myocardial infarction, the number of cases was unexpectedly and implausibly high. Including primary care cases would increase the number of femur fractures by 34 percent.

NL reported some other issues regarding the mapping of ICPC-1 to ICD-10 and consequent variations in the scope of the indicator definitions. These are mentioned in the appropriate sections of this report under 'Analysis by indicator'.

2.8.1.2. Hospital discharge register (HDR)

2.8.1.2.1. General description

The hospital discharge register contains primary and secondary diagnosis for hospital admissions (inpatients, day care, observations) including date of admission and discharge. The owner is the Dutch Hospital Data (DHD) (https://www.dhd.nl/producten-diensten/lbz/Paginas/Dataverzameling-LBZ.aspx). The information is used by for individual hospitals and specialists for planning and benchmark purposes and research.

Discharges in 2017 with admission in 2016 were included.

2.8.1.2.2. Population coverage

National scope, relevant to the whole population. The data cover all discharges from all general and university hospitals and specialised hospitals except for epilepsy clinics and long-stay centres for rehabilitation and asthma treatment. Independent treatment centres and private clinics are not included, but the private sector is small.

Inpatient care and day cases (patients admitted to a bed for one day only) are included, except for day patient care

for childbirth, psychiatric treatment and rehabilitation treatment. However, a weakness of this source is that it does not include outpatient care provided in hospitals.

2.8.1.2.3. Data availability, linkage

These data are regularly available to Statistics Netherlands and are also used for a variety of planning and research purposes.

Timely availability of the data is good. Full data on 2019 (including discharges in 2020 of admissions in 2019) are expected to be available for morbidity statistics in the first quarter of 2022. Only relatively few hospital stays pass the turn of the year.

Patients are identified by a personal identifier and can be linked to the population register. The methods used for linkage are a standard practice in Statistics Netherland.

Non-residents are treated in hospitals and can be defined as individuals who cannot be linked to the population register. However, the available data are limited. Since linkage is impossible, it is possible for an individual to be counted multiple times, and separate episodes of care for the same individual cannot be associated correctly.

2.8.1.2.4. Classifications used and coding issues

Diagnoses are coded using the International Classification of Diseases, 10th edition (ICD-10).

2.8.1.2.5. Other problems of recording or definition

The diagnostic information was incomplete for 20% of day care admissions in 2014-16 and is known to be incomplete for about 23% in the subsequent years. For regular statistical purposes imputation is used, but this is not feasible when person-level data are needed for linkage. The direct consequences of this limitation are difficult to estimate, as missing diagnoses may be covered within the HDR during other admissions and may be found in other sources.

2.8.1.2.6. Relevance to indicators

Used for all the indicators reported by NL.

2.8.1.3. National Causes of Death (CoD) data

2.8.1.3.1. General description

Causes of death are reported by physicians to the civil register of the municipality where the person died. The information collected includes both the underlying cause of death, i.e. the disease or injury initiating the chain of morbid events leading directly to death, and the non-underlying (i.e. intermediate or contributory) causes ('multiple causes of death').

The cause-of-death certificate, which is used exclusively for statistical purposes, is subsequently sent to Statistics Netherlands. Since 2014 cause-of-death certificates are also sent digitally to Statistics Netherlands.

2.8.1.3.2. Population coverage

Population coverage of death registration is generally complete with national scope. Causes of death are available for more than 98.4% of deaths in the Dutch population. Dutch residents who died abroad are not included.

Since the Nivel-PCD sample is the basis of the estimates, only persons who could be linked via the population register are taken into consideration.

2.8.1.3.3. Data availability, linkage

Since the cause-of-death certificates are sent to Statistics Netherlands and used for regular statistics, timeliness and availability of the data are good. Linkage is carried out using personal identifiers. Non-residents who do not have a relevant identifier cannot be linked to the population registry. The methods used for linkage are a standard practice in Statistics Netherland.

2.8.1.3.4. Classifications used and coding issues

Causes of death are coded using the International Classification of Diseases, 10th edition (ICD-10). World Health Organisation guidelines are used where possible to classify and code causes of death. From year 2013 coding is partly done automatically by the IRIS software and all underlying diseases on the cause-of-death certificate are coded. Underlying cause and secondary causes of death were all considered for the morbidity statistics.

2.8.1.3.5. Relevance to indicators

Used for all the indicators reported by NL.

2.8.1.4. Dispensed medicines data

2.8.1.4.1. General description

The data on prescriptions (drug register) are collected by the National Health Care Institute and represent all medicines dispensed by a pharmacy that are reimbursed under the statutory basic medical insurance. Medication provided in hospitals and institutions, over-the-counter medicines and medicines that are not covered by basic health insurance are not included.

2.8.1.4.2. Population coverage

The scope of the data is national, and complete within its remit. Medications provided during a hospital stay or in institutions are not included.

Since the Nivel-PCD sample is the basis of the estimates, only persons who could be linked via the population register are taken into consideration.

2.8.1.4.3. Data availability, linkage

The prescriptions data collected by the National Health Care Institute are made available to Statistics Netherlands, and timeliness and availability of the data are sufficient. The 2019 data are expected to be available for morbidity statistics in the first quarter of 2021.

Linkage is carried out using personal identifiers. Non-residents who do not have a relevant identifier cannot be linked to the population registry. The methods used for linkage are a standard practice in Statistics Netherland.

2.8.1.4.4. Classifications used and coding issues

Dispensed medicines are coded using the Anatomical Therapeutic Chemical Classification System (ATC). This classifies the type of medicine and aspects of its medical indication, but no specific diagnosis is recorded. For some indicators, a selection of the ATC-groups was made based on indications for use as mentioned in the 'Dutch Pharmacotherapeutic Compass' (https://www.farmacotherapeutischkompas.nl/) and the actual use of medication by subjects having and having not the disease in the Nivel-PCD.

2.8.1.4.5. Relevance to indicators

Data on dispensed medicines were used for P1-P2 diabetes mellitus; P3 Dementia (incl. Alzheimer disease); P8 Parkinson disease; P10 Epilepsy. In the case of P3, P8, and P10 a selection of medicines was used instead of the complete ATC groups N06D (anti-dementia drugs), N04 (anti-Parkinson drugs) and N03 (anti-epileptics).

2.8.1.5. Diagnosis Treatment Combinations Somatic Specialist Care (DTC-SSC)

2.8.1.5.1. General description

The data source Diagnosis Treatment Combinations Somatic Specialist Care (DTC-SSC) was collected by the DTC Information System DIS belonging to the Dutch Healthcare Authority (the current version of the DTC-SSC used by Statistics Netherlands is provided by Vektis). The register includes inpatient, day patient and outpatient hospital care, including care provided at independent treatment centres, specialized hospitals, and centres for rehabilitative care, kidney dialysis, audiology and radio therapy. The data are therefore wide-ranging in its coverage of secondary/specialist care.

2.8.1.5.2. Population coverage

The scope of the data is national. There is coverage from most types of healthcare institution (see above). Because the system is used for payments, completeness is good.

Since the Nivel-PCD sample is the basis of the estimates, only persons who could be linked via the population register are taken into consideration.

2.8.1.5.3. Data availability, linkage

The DTC-SSC data are readily available from the Dutch Healthcare Authority to Statistics Netherlands, and timeliness and availability of the data are enough. Full 2019 data including DTC's started in 2019 and ended in 2020 are expected to be available for morbidity statistics in the fourth quarter of 2021.

Linkage is carried out using personal identifiers. Non-residents who do not have a relevant identifier cannot be linked to the population registry. The methods used for linkage are a standard practice in Statistics Netherland.

2.8.1.5.4. Classifications used and coding issues

Diagnosis Treatment Combinations (DTC) are the basis for payment in somatic specialist care (SSC). Each medical specialty has its own set of DTC-SSC codes that often indicate a recognizable description of the treatments given. In total about 4400 different DTC codes exist. From the descriptions, diagnostic information can be derived. The DTC-SSC is not a classification, but rather a payment system

whereby, information on both the health care activities and a description of the particular health problem is provided.

In 2016 a start was made to also include ICD-10 codes with these data. In the 2017 data, ICD-10 codes were missing in 5 percent of the DTCs. This may be useful for future data collections. However, for the data required in the present pilot data collection of Morbidity Statistics (2014-2016), ICD-10 codes were still not available for the bigger part. Also, the quality of this new ICD-10 information remains to be checked. However, the new ICD-10 information proved to be of great help in the selection of relevant DTC codes for each of the indicators of the indicator shortlist.

To understand the mapping between ICD-10 and DTC-SSC, all DTC-SSC's containing ICD-10 information (in 2016 and 2017) were pooled and the ICD-10 codes per indicator were used to find all corresponding DTC-SSC codes. For most indicators, many different DTC-SSC codes were found, mainly because all medical specialties have their own set of codes. More than one medical specialty may regularly be involved with the treatment of a disease and each specialty may have one or more separate codes that refer to the disease. For example, in the case of diabetes mellitus, each of the specialty's ophthalmology, surgery, orthopaedics, internal medicine, paediatrics, geriatrics, and gastroenterology have one or more DTC-SSC codes referring to diabetes, its complications or frequently occurring treatments.

Another reason for the fact that many different DTC-SSC codes are found is that some DTC-SSC's can be used for many different diseases. In the case of diabetes, this can occur within the specialties mentioned above, but also in other specialties. For example, the specialism 'rehabilitation' has a code for 'Other disorders of lower extremities'. Clearly, this DTC-SSC code also may be used in other health problems than diabetes.

Using the descriptions of the DTC-SSC codes and the frequency with which they occur in combination with a certain ICD-10 code, for each indicator the most important codes were identified. For example, Parkinson's disease (defined by the single ICD-10 code G20) corresponded to 73 different DTC-SSC-codes. However, two of those codes accounted for 96% of the cases, one in the medical specialty 'neurology' (DTC-SSC-description Parkinson Disease) and one in the specialty 'geriatrics' (DTC-SSC-description Parkinson/ Parkinsonism). Clearly, these two codes both have a strong relationship with ICD-10 code G20 (and with each other). For most indicators, two to five DTC-SSC codes together covered more than 95% of the cases selected based on the ICD-10 definition.

Subsequently, all DTC-SSC's with the best matching descriptions were selected and the corresponding ICD-10 codes listed, to check whether the selected DTC-SSC-codes referred only (or predominantly) to the requested ICD-10 codes, or whether also other ICD-10 codes were found. For DTC-SSC-codes that frequently translate to ICD-10 codes outside the definition, a choice had to be made between inclusion (and consequently to include too many cases) or exclusion (leading to missing cases). In case of doubt, different selections were used to study the impact of different choices on the final indicator calculated after the combination of sources. Based on these two actions, selections of the DTC-SSC codes have been made for each indicator.

2.8.1.5.5. Relevance to indicators

Used for all the indicators reported by NL.

2.8.1.6. Diagnosis Treatment Combinations Mental Health Care (DTC-MHC)

2.8.1.6.1. General description

The data source Diagnosis Treatment Combinations Mental Health Care (DTC-MHC) is collected by the DTC Information System DIS belonging to the Dutch Healthcare Authority. The register includes inpatient, day patient and outpatient specialized mental health care as provided by institutions, psychotherapists and psychiatrists.

Diagnosis Treatment Combinations (DTCs) are DRG-like units that form the payment system of specialist mental health care. Diagnoses and detailed data on care provided (type of treatment, number and complexity of overnight stays, start date, end date) are included.

2.8.1.6.2. Population coverage

The scope of the data is national. All ambulatory specialized mental health care is included, however residential specialized mental health care is covered only until the first year of stay is completed (for new cases from 2015, until the third year is completed). Importantly, it does not cover basic mental health care which might be provided through general practitioners. Because the system is used for payments, completeness is good.

Since the Nivel-PCD sample is the basis of the estimates, only persons who could be linked via the population register are taken into consideration.

2.8.1.6.3. Data availability, linkage

The DTC-SSC data are readily available from the Dutch Healthcare Authority to Statistics Netherlands, and timeliness and availability of the data are enough historically. Linkage is carried out using personal identifiers. Non-residents who do not have a relevant identifier cannot be linked to the population registry. The methods used for linkage are a standard practice in Statistics Netherland.

It is important to note that although the DTC-MHC data are available for recent years and was used in the pilot project, it is expected that this source will probably not be available in the future due to a change in payment system. This will influence the total number of cases found for those indicators relating to mental health.

2.8.1.6.4. Classifications used and coding issues

The DTC-MHC is the basis the basis for payment for specialised mental health care. DTC-MHC defines mental health problems using the Diagnostic and Statistical Manual of Mental Disorders version IV (DSM-IV). The coding table of DTC-MHC includes conversions to ICD-10. Using the ICD10-definitions of the shortlist the corresponding DSM-IV codes could be selected.

2.8.1.6.5. Relevance to indicators

Used for indicators P3 Dementia (incl. Alzheimer disease); P4 Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence); P5 Schizophrenia, schizotypal and delusional disorders; P6 Mood (affective) disorders; P7 Anxiety disorders.

2.8.1.7. Long term care eligibility decisions (LTC-E CIZ) and Long-term care co-payments (LTC-C CAK)

2.8.1.7.1. General description

The data sources 'Long term care eligibility decisions (LTC-E CIZ)' and 'Long term care co-payments (LTC-C CAK)' both relate only to persons receiving long-term institutional care (nursing homes) and were used together to identify the institutional population for indicator P3 Dementia (incl. Alzheimer disease).

2.8.1.7.2. Population coverage

These data sources are national and relate only to persons receiving long-term institutional care.

2.8.1.7.3. Data availability, linkage

The exact data sources are the CIZ Register of eligibility decisions to long-term care as funded by the long-term care law, held by the Centre for Care Assessment; and the record of co-payments for use of long-term care, held by the Central Administration Office (CAK). These data can be accessed by Statistics Netherlands and linkage is carried out using personal identifiers.

For future use, CIZ eligibility decisions may also contain ICD10-information. The quality and completeness however remain to be checked and the information remains limited to subjects entering long term care.

2.8.1.7.4. Classifications used and coding issues

The data LTE-E CIZ represent eligibility decisions for admission to long term care. For morbidity statistics, reasons for admission are divided into 'psychogeriatric' or 'other'.

The data source LTC-C CAK represents co-payments for long term care and indicates either 'use' or 'no use' of insured long-term care financed under Long-Term Care Act.

Therefore, the combination of these two indicators allowed the identification of the population with a psychogeriatric diagnosis and receiving long-term care in the reference period. It was assumed that a 'psychogeriatric' diagnosis refers to some form of dementia in the great majority of cases for this age group.

2.8.1.7.5. Relevance to indicators

Used only in the case of P3 Dementia (incl. Alzheimer disease).

2.8.2. Estimation methods

2.8.2.1. Overview

For the NL estimates, the primary care data was used as the baseline and other data sources linked to it. The primary care data from Nivel-PCD is a 'sample' data source based on only one of several networks of general practitioners, covering around 8% of the resident population. Therefore, the national data sources were represented by only around 8% of cases, and the estimates were produced by weighted extrapolation of the resulting figures to the whole national population.

2.8.2.2. Data linkage

All data sources were linked at the individual level using personal identifiers. The methods of linking individuals to the population register and linking data sources together via the personal identifier is a common process of Statistics Netherlands – see Bakker, van Rooijen, and van Toor (2014).

2.8.2.3. Weighting and extrapolation

As the Nivel-PCD data covers about 8% of the population, the number of cases found in this population were extrapolated to the full Dutch population. Any inconsistencies between the population characteristics of Nivel-PCD and the general Dutch population are taken into account at the same time using a weighting procedure. The method is comparable to methods used for interview surveys to control for non-response. The weighting variables that were used are based on the demographic and socio-economic variables and included age, sex, income, degree of urbanization, migration background and person years (part of the year persons were registered with a Nivel-PCD practice). Weighting was performed using a specialised software 'Bascula' – see Nieuwenbroek and Boonstra (2002).

For dementia an exception was made for extrapolation: as the nursing home population was known (defined by persons paying a co-payment for use of long term care) and it was known which persons received this care because of psychogeriatric reasons (assumed being largely dementia), persons with dementia found in the Nivel-PCD population were extrapolated to the non-institutionalised population and subsequently the psychogeriatric nursing home population was added.

The methods for construction of episodes and other technical aspects for using the Nivel-PCD data are well understood and have been previously published – see Nielen, Spronk, Davids, et al (2019). The methodological concept of combining data sources to Nivel-PCD and weighted grossing up to national totals was an innovation of the project and was discussed with experts within Statistics Netherlands and colleagues from other institutions.

2.8.2.4. Conclusion

The method for producing the estimates for NL is unusual in that the basis is a sample data source covering only 8% of the population. All data sources are linked via the population register; however, the multiple data sources are not fully utilised because only that 8% of the population can be linked to them.

In addition, the primary care data exclude residents of institutions. With one exception of the indicator for dementia, such populations (mainly in care homes) are not directly represented. Therefore, there is the possibility of a bias causing the frequency of conditions which are most prevalent in elderly institutionalised persons to be at least slightly underestimated.

The data sources used by NL generally exclude private or non-insured healthcare, but the private sector is said to be small (no exact quantification is available) and this is unlikely to cause an important bias.

In general, the documentation and references show that the representativeness of the Nivel-PCD sample is good, and the technical methods relating to its epidemiological use are well-understood. The linkage processes are commonly used in Statistics Netherlands and allow effective assembly of the data, with exceptions relating to non-residents. Therefore, overall we can conclude that the methods used by NL seem to be well justified and the estimates are likely to be of sufficient quality.

The possible weaknesses are as follows:

Certain private sector healthcare providers and uninsured persons are excluded, although as health care insurance is compulsory for almost all residence, those excluded (i.e., military personnel) comprise a very small proportion of the population.

- The representativeness of the sample primary care data could change over time or vary between different health conditions for unobserved reasons.
- Certain chronic conditions may be under-represented because of gaps in the data around shared care ('chain care').
- The institutionalised population is excluded from most indicators.

2.9. Poland (PL)

Participation for Poland was by the Centre for Health and Health Care Statistics of the Statistical Office in Krakow.

Estimates were reported for all the main list of indicators P1-P35 (list A) and also the optional indicators PB36-PB43 relating to external causes (list B).

There were three data sources used by PL, as follows:

1. *Causes of death*: The Statistical Survey of Mortality based on deaths registered in PL by registry offices. Both

underlying and secondary causes of death of persons deceased in the index year are included.
2. *Insurance*: The database of healthcare services financed from public funds, kept by the National Health Fund (NHF). All types of insured health services are covered.
3. *Prescriptions*: The database of medicine reimbursements, kept by the National Health Fund (NHF). All medications dispensed under the public insurance are included.

2.9.1. Data sources used

2.9.1.1. Statistical Survey of Mortality

2.9.1.1.1. General description

The Statistical Survey of Mortality is the cause of death (CoD) register used for official statistics for the evaluation of the public health and creating health policies in Poland. The data are collected based on Certificates of Death registered at civil registry offices.

2.9.1.1.2. Population coverage

The mortality data include all deaths of individuals residing in PL, and in addition the deaths of Polish residents occurring abroad, registered in PL by civil registry offices. It appears that deaths of non-residents (short-term visitors) may not be covered.

2.9.1.1.3. Data availability, linkage

The mortality data collected by the registry offices become available to the Statistical Office for analysis around 14 months after the reference period. Therefore, complete data on deaths occurring in the calendar year 2019 would be available in approximately March 2021.

This data source was not linked to the other sources at individual level, rather for selected indicators aggregate numbers of deaths were used.

2.9.1.1.4. Classifications used and coding issues

Causes of death are coded according to the International Classification of Diseases, 10th Edition (ICD-10) based on all conditions mentioned on the Certificate of Death.

2.9.1.1.5. Relevance to indicators

The mortality data was used as an addition to the health insurance data only for the indicators in List B. For those indicators, the NHF data on services were selected to include only deaths in hospital (including emergency care). Therefore, deaths recorded in the mortality data as occurring in places other than hospital could be added without duplication.

2.9.1.2. NHF Healthcare services financed

2.9.1.2.1. General description

The database of healthcare services financed from public funds is kept by the National Health Fund (NHF). All types of services are covered when delivered to individuals entitled under the national health insurance, by healthcare providers contracted to provide healthcare services financed from public funds.

The NHF databases are continuously validated and audited and are considered as a valuable data source for a regular, high quality morbidity statistics.

2.9.1.2.2. Population coverage

Data from the National Health Fund (NHF) database have almost complete coverage of the national population, as the great majority of people in PL are entitled to receive healthcare services financed from public funds. Individuals entitled to healthcare services financed from public funds accounted for 91.5% of the population in 2016. Publicly funded care provided by both public and private providers is included.

Services provided outside the national health insurance are not included; the scope of privately funded care is said to be very small. There is the possibility of geographical differences in accuracy because of the level of income in different regions and ability to finance out-of-pocket healthcare, or different share of non-insured people, but this should not impact the national figures.

Services provided to non-residents are recorded in the database, but in the absence of an identifier cannot be linked. Non-residents are a small part of the population (less than 2%).

2.9.1.2.3. Data availability, linkage

The NHF database of services financed is in the form of individual records with a unique identifier (see below). All statistics needed are available according to the breakdowns.

Data are reported once a month by healthcare providers to the NHF's subsidiary and then transmitted every two/three months to NHF's headquarters. The data can change after validation and correction, so that the newest data are subject to the highest number of possible changes.

Generally, for compilation of morbidity statistics, the timeliness of NHF's data for the reference period is T+6 months.

Linkage in the NHS database is carried out using a unique personal identifier (PESEL number) which is issued to all persons in PL who are eligible for the national health insurance, including foreign citizens who are resident more than 3 months or under cooperation agreements. Non-residents cannot be directly identified in the database, although analysis on residence is possible by linkage to the population register. Permanent residents make up more than 98% of the population of PL. Care provided to persons without a PESEL number is also recorded but may be subject to some risk of duplication or errors.

Certain categories of individual are (a) less likely to have a PESEL number issued, or (b) may have a PESEL number without it being recorded in the data, in particular:

non-residents, especially short-term visitors (a)

- homeless people and other marginalised groups (a & b)
- patients who are unconscious and cannot be identified (b)

Within the data used for the estimates, the percentage of individuals without the PESEL number accounted for less than 1.0% for period prevalence and up to 2.0% for incidence for most indicators. It is notable that the conditions showing higher missingness than this are likely to be differentially associated with one or more of the groups mentioned above, for example Land transport accidents, Intentional self-harm (incl. suicidal attempt). A full breakdown of this analysis is in annex B 1.2 Coverage of the health insurance data in PL: percentage of patients without a personal identifier.

2.9.1.2.4. Classifications used and coding issues

Diagnoses for the health services provided are coded according to the International Classification of Diseases, 10th Edition (ICD-10). The conditions coded are the main cause of the service provision and also any relevant coexisting diseases/comorbidities.

Because the purpose of the data is reimbursement for insured services, the codes provided relate only to the services provided on that occasion or the health conditions which justified those services. A common issue in reimbursement-based data is that codes may sometimes be chosen to maximise the level of payment, with potential systematic effects on the statistics. This tendency may be reduced by the regular audit by the NHF.

It was also noted that the data may include 'suspected' cases at the diagnostic stage.

For data originating from primary care contacts, ICD-10 diagnoses beginning with letter 'Z' (Chapter XXI 'Factors influencing health status and contact with health services') make up a significant proportion. In 2016 these amounted to 31% of the total number of services reported in primary care, increasing in 2018 to 35%. Contacts with the main diagnosis in the 'Z' block, and with the main diagnosis in the 'Z' block and no reported comorbidities, also increased. The latter made up 25% of the total in 2016 and 28% in 2018.

This means that around a quarter of primary care records do not provide a meaningful diagnosis for morbidity statistics. This finding is likely to reflect a large number of primary care contacts where no new condition was diagnosed, but instead were 'check-ups' or renewal of previous treatment plans and prescriptions. Many cases were reported with only the ICD-10 code Z76.0 Issue of repeat prescription. Because general practitioners are mainly paid per patient and not by the recorded diagnosis, they are not strongly motivated to complete the diagnostic information.

2.9.1.2.5. Relevance to indicators

The NHF database of services financed was used to estimate all indicators, alone or with one other data source.

2.9.1.3. NHF Medicine reimbursements

2.9.1.3.1. General description

The database of medicine reimbursements is kept by the National Health Fund (NHF). All medications dispensed under the public insurance are included, including those dispensed by private sector pharmacies.

The NFH databases are continuously validated and audited and are considered as a valuable data source for a regular, high quality morbidity statistics.

2.9.1.3.2. Population coverage

Population coverage is as for the data source 'NHF Healthcare services financed' (above). All medications dispensed are covered, including those prescribed by a health practitioner in the private sector. However out-of-pocket purchases are excluded.

2.9.1.3.3. Data availability, linkage

The NHF database is in the form of individual records with a unique identifier. All statistics needed are available according to the breakdowns.

In the case of data concerning reimbursed medicines, data are reported twice a month by pharmacies to the NHF's subsidiary. Validation and audit procedures are similar to those described for the insured health services (see above). Generally, for compilation of morbidity statistics, the timeliness of NHF's data for the reference period is T+6 months.

2.9.1.3.4. Classifications used and coding issues

Data on prescriptions are coded according to the Anatomical Therapeutic Chemical Classification System (ATC). In discussion with national experts, it was decided to use the ATC codes as part of the definition for certain conditions (for example, diabetes). For other conditions it was believed that medicines are in use not for a specific disease/condition but rather for the treatment of a given sort of symptoms (e.g. antidepressants, anticonvulsants) and thus are not sufficiently specific.

2.9.1.3.5. Relevance to indicators

The NHF medicine reimbursements data was used in addition to the insured health services for the indicators P1, P2, P3, P8, P11, P12, P13, P14 and P15.

2.9.2. Estimation methods

2.9.2.1. Overview

The estimates were based on only three sources, namely the databases of healthcare services and medicine reimbursements under the national health insurance, and the cases of death data. These data cover more than 90% of the healthcare services and prescriptions in PL and are considered representative. High quality is ensured by the regular validation processes, although some bias may be introduced by reporting practices related to payments. Contacts recorded in primary care are missing a diagnosis in around a quarter of cases.

2.9.2.2. Data linkage

The two databases belonging to the National Health Fund contain individual records linked by a unique identifier, which is present for the great majority of cases. There may be a small risk of duplication in the minority of cases where no identifier is recorded.

The cause of death data was not linked to the NHF database at individual level. For the relevant indicators, the health insurance data were restricted only to hospital care, while only deaths outside hospital were added, thus avoiding duplication.

2.9.2.3. Conclusion

The estimates from PL are based on well-established data sources with strong quality control, although some typical weaknesses of insurance-based data may be present. The coverage is quite comprehensive, and the estimates are likely to be of enough quality for most conditions.

The possible weaknesses are as follows:

Certain private sector healthcare providers and uninsured persons are excluded.

- Counting of non-residents is not reliable, as individuals or episodes cannot be linked.
- The insurance-based data may be affected by incentives to report diagnoses attracting higher reimbursement or for similar reasons.

3

Analysis by data source

This section contains summary observations and conclusions regarding use of the different types of data sources for producing morbidity estimates. They are based on the expectations of the EPIMS report and the MORB project guidelines provided to participating countries, and the detailed information provided in section 2 on analysis by country.

Conclusions are provided with reference to the European Statistical System dimensions of quality [6] in relation to each data source, with respect to:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability

Of the dimensions of quality, Accessibility is not addressed here as it is not directly relevant.

3.1. Primary care

3.1.1. Preparatory remarks

Primary care records may be long-term and comprehensive, and their usefulness may be enhanced by information from other services, the results of investigations, etc. On the other hand, the clinical information may be mostly symptom-based, or not clearly defined, and specific diagnostic criteria are often not applied.

The data in primary care sources may be coded according to the ICPC or some other system, and not ICD-10. While it is possible to translate between ICPC and ICD-10 for some diagnoses, for others there is no straightforward conversion. Typically, the ICPC codes are less diagnostically and anatomically specific than ICD-10.

Some countries have sample-based primary care data, sometimes based on voluntary networks of doctors or providers. It is necessary to adjust such data appropriately for use in the national estimates, having regard to both the proportion of population coverage, and the differences between the sample population and the national population in demographic characteristics.

Lack of comparability is likely to result from differences in the organisation of healthcare, especially whether or not the general practitioner acts as gatekeeper. The use of different classifications will also be important in the case of some indicators. Comparability issues which are caused by differences in population coverage should be able to be corrected to some extent by the methods of adjustment.

3.1.2. Considerations according to the ESS dimensions of quality

3.1.2.1. Relevance

It was envisaged that primary care data would be essential to achieving comprehensive estimates. In fact, primary care (mainly GP) data was used by all except two countries in the pilot studies, either directly or as part of an insurance database. The exceptions were MT and FR which were not able to acquire GP data in time for the project.

[6] European Statistical System (ESS) Handbook for Quality and Metadata Reports — re-edition 2021. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-021

Primary care data was most essential for chronic diseases and those unlikely to require frequent hospitalisation, for example diabetes mellitus or anxiety disorders. The exact contribution of the primary care dataset could be determined only in those cases where it was identified separately from insurance, as the latter usually also included hospital episodes. In HR, primary care was the most frequent source for 32 indicators and included all the cases identified for 5 indicators. In NL the primary care dataset included between 50 and 80 percent of the cases identified for most indicators.

3.1.2.2. Accuracy and Reliability

Various difficulties were reported with the use of primary care data. Regarding the accuracy of diagnoses, several countries reported that 'suspected' or 'provisional' cases are contained in the GP database, and these cannot always be identified or excluded. Therefore, cases are likely to be overestimated unless the country has a reliable way to flag these suspected cases.

As expected, some primary care sources used the ICPC-1 or ICPC-2 classification. There was little difference between the two versions of ICPC for the purposes of the indicators. The most important difference between ICPC and ICD-10 for the indicators had already been identified in the EPIMS project, namely that the ICPC code D97 Liver disease covers the ICD-10 codes for both alcoholic (K70) and non-alcoholic (K71-K77) liver disease. As a result, NL did not report the indicators P24 and P25. BE and FI did report them, but the measures cannot be considered reliable because some of the data used ICPC.

Some other indicators were affected by smaller differences between ICPC and ICD-10, such as those reported by NL around different forms of alcohol abuse.

An important finding highlighted by LT and NL likely to have general validity is that acute myocardial infarction and stroke appear with unexpected frequency in primary care records. As these are severe conditions which are often fatal and invariably require hospital care, it is likely that cases recorded in primary care (or outpatient care) are previously treated patients seen for purposes such as rehabilitation and renewal of medication and should not be included as incident cases on that basis.

In many countries, but not all, the GP has a 'gatekeeper' role to secondary care services. However, it cannot always be assumed that the GP will receive timely information from the specialists on the diagnosis of the patients referred to them. In other countries there is direct access to some specialists, for example dermatologists and gynaecologists, Thus, there are differences in the completeness of the GP data between countries. Similarly, in most countries of Europe the primary care system has largely universal coverage, but there are varying proportions of people in some countries who may use a private sector GP.

In the case of NL, the estimates were based on data linked to patients who were covered by the Nivel network. In several countries it was apparent that only persons registered with a GP or linked to a population register could be counted, thus excluding in various cases not only non-residents, but also sometimes institutional populations or some under-served groups.

3.1.2.3. Timeliness and Punctuality

The primary care data sources were mostly available within the necessary timescales for calculation of indicators at T+3, that is, using data for the 2016 reference year for calculations in 2019. It was necessary for some countries to arrange special data-sharing permissions, which can delay the data availability.

3.1.2.4. Coherence and Comparability

Issues of comparability regarding classifications are described above.

In BE and NL, the primary care data were obtained from GP practices forming a sample of the population (BE the Intego network covering 2% and located in the Flanders region; NL the Nivel network covering 8% of the national population). Such sources can be of high quality in terms of accurate and complete reporting, but have to be thoroughly understood in terms of their representativeness in multiple dimensions – representativeness of the patients within the GP practice, if inclusion is optional; representativeness of the GP practice relative to other practices not participating in the network (for example the expertise and care of the GPs); and representativeness of the geographical region covered if the sample is not from the whole country. Inadequate accounting for these dimensions of representativeness may lead to bias, for example if the participating GPs are concentrated in more affluent areas than others.

Where data from regional and national sources are mixed, there is also an issue of coherence between the different datasets contributing to the national estimates.

3.1.2.5. Conclusions

Primary care data are essential to achieve full coverage, especially of chronic diseases and conditions unlikely to require hospitalisation.

- Sources of primary care data need to be considered carefully for their population coverage and representativeness.
- Differences in the national healthcare systems (e.g. respective scope of primary care and hospital outpatient care) and use of the private sector mean that completeness and comparability are not guaranteed.
- Specific issues make primary care data unsuitable for specific measures – distinction between alcoholic and non-alcoholic liver disease, recording of acute myocardial infarction and stroke patients for rehabilitation or similar purposes.
- National practice on recording of suspected diagnoses and the definition of episodes of care should be studied carefully to optimize the implemented definitions.

3.2. Hospital inpatients

3.2.1. Preparatory remarks

Hospital inpatient records are mostly of high quality. The information is often extensive and supported by the results of diagnostic investigations, regular observations of vital signs, etc. Because of the hospital administrative procedures, demographic and residence information should be well recorded. Multiple codes may be recorded covering not only the main diagnosis of the admission but comorbidities and symptoms. ICD-10 is most often used for the coding of diagnosis.

The data may often be organised according to episodes of care rather than per illness or per individual patient, so it is essential to avoid duplication and to link or extract the data in the correct form to give a true picture of the beginning and end of the episode of illness. The episode of care is usually one continuous period from admission to discharge, but there may be complications caused by (for example) transfer of the patient between different hospitals.

Inpatient records are a good data source for the more serious and acute diseases and injuries which universally require hospitalisation, such as acute myocardial infarction, stroke, and fracture of femur. On the other hand, they are of little value for chronic conditions, which are typically managed in primary care. Even for some of the most acute diseases, it is important to supplement inpatient data with mortality data, particularly since some people may die suddenly (for example from acute myocardial infarction) without any hospital inpatient care.

Comparability between countries should be quite high, provided that inpatient care can be clearly separated from any other aspect of hospital activity (such as outpatient) in the data source. However, there may be national differences in diagnostic criteria and recording practices, and in the application of codes. The data relating to certain categories of patients or institutions, such as long-term psychiatric patients, may be collected separately. Since the records are often used for financial purposes, the national reimbursement systems may create incentives for recording cases in different ways.

3.2.2. Considerations according to the ESS dimensions of quality

3.2.2.1. Relevance

Hospital inpatient data was the only source that was used by all the countries, including those where the hospital data were part of the insurance database. Inpatient data was the main data source for severe, acute conditions and those requiring surgical intervention. For example, it is to be expected that all incident cases of acute myocardial infarction will be represented by either inpatient episodes or causes of death, or both.

3.2.2.2. Accuracy and Reliability

There is a general view that inpatient records are of high quality, however some issues are known. Several countries pointed out that, since the hospital records are the basis of financing for healthcare providers (via insurance reimbursements or national planning mechanisms) there is an incentive to record more severe or complex diagnoses, or others which represent the treatment as having a necessarily high cost.

Regarding injuries, there were differences of opinion on whether the external cause as opposed to the nature of the injury would be recorded in hospital records. In some countries this was said to be unlikely even for inpatients, while in one it was said that external causes are likely to be recorded for inpatients, but not for outpatients.

There are differences in the coverage of the hospital data depending on the national health insurance system. In some cases, the coverage is complete, but in others there is a proportion (usually small) of people who are uninsured,

which affects the completeness of hospital data when it is reported through the insurance database. The proportions of persons using private sector hospital care varies nationally and between different diagnoses but is generally low for more serious health conditions.

Since inclusion in the hospital inpatient data results from attending in hospital as a patient, records automatically exist for non-residents, institutionalised people and other groups. However, there is not always a way to identify non-residents within the data; alternatively, non-residents may be identified but may be unable to be linked due to lack of a personal identifier.

3.2.2.3. Timeliness and Punctuality

In all cases, hospital inpatient data were found to be available within feasible timescales.

3.2.2.4. Coherence and Comparability

In most countries diagnoses in inpatient data are recorded using the ICD-10 classification, so this is a factor which promotes comparability. Multiple diagnoses are recorded. NL reported using a unique specialist coding system, which nevertheless allowed for accurate mapping to ICD-10. In a few instances, national differences in coding practice may affect comparability, such as for dementia.

As regards minor and chronic conditions, there are likely to be differences between countries on which are commonly treated on an inpatient basis and which as outpatient (ambulatory) cases or in primary care.

3.2.2.5. Conclusions

Hospital inpatient data are essential to ensure coverage of more serious health conditions and those requiring surgical interventions.

- There is generally comprehensive diagnostic information using comparable classifications, except for small differences in national practice.
- The scope of private sector hospital care varies, as does the interaction between primary and secondary care according to national healthcare systems, but these factors are unlikely to affect most indicators if multiple sources are used.

3.3. Hospital outpatients

3.3.1. Preparatory remarks

Outpatient services vary somewhat between health systems, but the general defining characteristic is that the patient attends an appointment for consultation with a health professional, diagnostic investigations or treatment, without any overnight stay in the hospital. Sometimes there may be a 'walk in' system rather than a prearranged appointment.

There are many different configurations and titles of such services. For example, some countries recognise a category of 'day case' indicating that the patient occupies a hospital bed for a short time, while recovering from a minor surgical procedure, but does not stay overnight. Outpatient sessions may take place in satellite clinics or primary care premises as well as in the hospital. However, it is universally an aspect of specialised (secondary or tertiary) care and not part of primary care.

The inclusion of outpatient data may vary between health systems and may not be collected centrally or be collected in less detail than inpatient data. In insurance-based systems, however, the data should be complete.

3.3.2. Considerations according to the ESS dimensions of quality

3.3.2.1. Relevance

Hospital outpatient data was used by five countries, four of those as part of an insurance database or combined hospital register. Outpatient data was the most important source of cases for several indicators in FI, and a smaller number in HU.

Outpatient data are often seen as less important and of lower quality than inpatient records. In some health systems there is limited recording of diagnoses.

3.3.2.2. Accuracy and Reliability

The considerations on the completeness of hospital outpatient data are mostly similar to the situation of inpatient care, particularly on issues such as the private sector, uninsured individuals and non-residents.

The coding of diagnoses in outpatient care is not always as complete as for inpatients. Regarding injuries, there were differences of opinion on whether the external cause as

opposed to the nature of the injury would be recorded in hospital records. In at least one country (FI) and potentially others external causes are not recorded for outpatients.

LT mentioned that a modifier "+" is used in outpatient care only to indicate a new case, however this was considered unreliable and length of episode according to the Guidelines was used instead.

For severe, acute conditions it is likely that outpatient attendances are for rehabilitation and routine follow-up of previous patients, therefore care is needed when defining episodes. For some conditions such as acute myocardial infarction it is probably best not to identify incident cases from outpatient data, similarly to primary care.

3.3.2.3. Timeliness and Punctuality

In all cases where hospital outpatient data were used, the data were available within feasible timescales.

3.3.2.4. Coherence and Comparability

Diagnoses recorded in outpatient data are generally similar to inpatient records, using the ICD-10 classification, so this is a factor which promotes comparability. However, multiple diagnoses might not be recorded.

As regards minor and chronic conditions, there are likely to be differences between countries on which are commonly treated on an inpatient basis and which as outpatient (ambulatory) cases, and between the respective scope of outpatient care and primary care.

3.3.2.5. Conclusions

Hospital outpatient data may be useful for completeness but can be regarded as secondary for most purposes (but see next point).

- The respective scope of outpatient care and primary care varies between countries, for example regarding their respective importance in the monitoring of chronic conditions, renewal of medications or disability claims.
- The diagnostic information recorded in outpatient care is sometimes less complete than for inpatients.

3.4. Causes of death

3.4.1. Preparatory remarks

Typically, cause of death coding can be described as accurate but not usually precise. The quality of the information depends in the first instance on the knowledge of the certifying physician, who may not have been closely involved in the person's care. Further, it is well known that quite broad and imprecise diagnostic terms are often used on death certificates. Despite these limitations, cause of death data are of great importance for the most serious and acute diseases and is essential to identify cases where death was sudden, and no healthcare was provided.

In the EU countries it is usually assumed that cause of death data are complete since their collection is required by law. However, there may be delays in the compilation of the data because of administrative procedures. In some countries this may especially affect the timeliness of data on deaths from external causes if a judicial enquiry into the accident or violence is required.

Depending on the national laws and procedures, deaths of usual residents which occur outside the country may not be included in the data. Similarly, deaths of non-residents or non-citizens may be processed differently and might be omitted from the regular database or contain less detail in some respects.

3.4.2. Considerations according to the ESS dimensions of quality

3.4.2.1. Relevance

Causes of death data were used by all countries except one. The exception was FR where it was impossible to obtain and link the COD data of 2016 in time for the project.

COD data was rarely the main data source for the indicators, the only cases being for some indicators on injuries in BE, HR and FI. However, causes of death contributed unique cases to most indicator values, because of the need to identify cases from death certificates especially for diseases which are acute and fatal. The recording of multiple causes also adds to the completeness of ascertainment for some chronic diseases.

3.4.2.2. Accuracy and Reliability

In most European countries the cause of death data are considered complete and reliable, and the ICD-10 classification is used. Most countries record multiple causes of death including external causes of injury. There are some small known differences in national practice; two countries commented on the codes used for dementia and Alzheimer's disease.

A specific issue identified is the use of the term heart failure (ICD-10 I50) which in some countries can be interpreted to mean a terminal event rather than a longer-term cardiac condition. Because of this ambiguity, it would be preferable for indicator P16 on heart failure to be identified using underlying cause of death only.

3.4.2.3. Timeliness and Punctuality

For most countries there were suitable arrangements in place for use of the COD data, however delays are possible where the data are provided by a different institute. Because of the time taken for complete readiness of the COD data in France followed by linkage to other data sources, FR was unable to use the 2016 mortality data.

3.4.2.4. Coherence and Comparability

All the countries used the ICD-10 classification for causes of death. Several use the Iris COD coding software, and all follow the international rules for classification, so comparability is expected to be good. In fact, the extent of possible national differences in clinical and diagnostic practices and the differences in death certification due to language are not well understood. Inconsistencies in reporting of injury-related deaths due to different medicolegal systems are also possible. These factors cannot easily be identified or accounted for but for these indicators are assumed not to be large.

In respect of the ICD-10 classification, there is also good coherence between COD data and most hospital datasets and insurance-based data.

3.4.2.5. Conclusions

Causes of death are an essential data source both to identify cases of acute and fatal conditions or injuries, but also to contribute cases from the recording of chronic diseases as contributory causes.

- Comparability is thought to be good internationally because of compliance with ICD-10 and its classification rules.
- Specific issues were identified relating to a few indicators.

3.5. Disease-specific registers

3.5.1. Preparatory remarks

Disease-specific registers exist only for a limited number of diseases: mainly tuberculosis, cancer, diabetes, cardiovascular diseases, and end stages of renal failure. Registers may exist for the purpose of managing specialist services or for research use.

Two registers for the same disease should in theory produce very comparable data. However, the many differences such as inclusion or exclusion of cases identified at death make this more questionable. In addition, registers vary in their legal basis, organisation and resources, potentially affecting their ability to collect data in a consistent way and ensure its completeness and quality.

3.5.2. Considerations according to the ESS dimensions of quality

3.5.2.1. Relevance

The most important disease-specific registers are cancer registers which exist in many European countries or regions. However, as the draft indicators specific to cancer were not part of the pilot studies, these were not relevant.

A large variety of registers do exist in different countries, often managed by academic institutes or medical organisations, but these are likely to be useful mainly for validation. For example, HU used the National Registry of Myocardial Infarction for validation of indicators P14 and P15.

Disease-specific registers were used by only one country in the pilot studies to make up the estimates, and for only one indicator. HR used the Registry of communicable diseases, for indicator P19 Pneumonia only. This data source contributed 3.5% of cases, while the majority of cases were identified in primary care. No detailed information about this registry was provided, and its scope and completeness of reporting are not stated in the report.

3.5.2.2. Conclusions

Disease-specific registers are not normally useful for these statistics but could be included if a need is identified by national experts to enable completion of a particular indicator. However, care should be taken as inconsistent use of specialised sources could reduce the comparability.

- Registers are more likely to be valuable for validation of the estimates.

3.6. Surveillance systems

3.6.1. Preparatory remarks

Surveillance systems relate mainly to infectious diseases. Although the surveillance data perform an important purpose from the point of view of public health, there are limits to the usability of the data in many cases.

3.6.2. Considerations according to the ESS dimensions of quality

3.6.2.1. Relevance

Surveillance systems relate mainly to communicable diseases. As the draft indicators on communicable diseases were not part of the pilot studies, such reporting was not found to be relevant, and no use of these systems was made by the countries.

3.6.2.2. Conclusions

Surveillance systems are not relevant to the list of pilot indicators.

3.7. Emergency care

3.7.1. Preparatory remarks

Emergency care data are the specific records of emergency rooms/accident and emergency departments or emergency ambulance/paramedic services. These data sources are often not integrated into the regular hospital databases or not collected centrally, so their statistical and research use has so far been limited. Relatively little is known about the quality of emergency care data at present. It may be expected that, if the systems are not integrated with the hospital database, they may be mainly episode based rather than person based. Diagnostic information may be limited to the immediately presenting and treated problems, while demographic information may be limited.

3.7.2. Considerations according to the ESS dimensions of quality

3.7.2.1. Relevance

Emergency care data were used by only three countries, LT, MT and PL. In case of MT, national register (Injury Database – IDB) collects information on patients seen at the main government hospital accident and emergency department with any form of injury. In contrast, in LT and PL cases the emergency care data were contained within the insurance dataset along with other types of care. Most use of this data source was made by LT, where it was the source that provided the largest number of cases for 10 indicators relating to injuries. No detailed information specific to emergency care was given, but the considerations relevant to insurance-based data in general are likely to be relevant.

3.7.2.2. Conclusions

- Emergency care data may be useful if available especially for complete reporting of sudden acute conditions and injuries.
- For countries where the use of emergency care data is not already well-established, caution is advised because of risks such as lower quality or incomplete diagnostic coding and difficulty with accurate linkage to other datasets.

3.8. Insurance

3.8.1. Preparatory remarks

The use of insurance databases is well developed in a few countries, where it was achieved successfully in the earlier pilot studies. In other countries, access to the data is a more recent development and has involved overcoming legal obstacles and long negotiations with the stakeholders.

Insurance databases often have the possibility to improve accuracy by linking together data from different sources, and through checking and quality assurance processes. On the other hand, the cases and diagnoses reported are likely to be influenced by the specific rules of the reimbursement system.

The previous pilot studies suggested that in countries which share broadly similar health insurance systems, the statistical results are mostly quite comparable. However, national differences in healthcare organisation and the rules

of each insurance or reimbursement system may cause differences.

3.8.2. Considerations according to the ESS dimensions of quality

3.8.2.1. Relevance

Data from the national health insurance systems was the main source for three countries and a less important source for another two. In PL it was the main source for all indicators and the only source for 24, while for HU and LT it was the main source for a large number. The importance of insurance data for some healthcare systems reflects both the complete or almost complete population coverage, and the fact that it typically includes within its data from multiple sources such as both primary and secondary care. In some countries, such as LT, there is already a history of using insurance data to produce morbidity statistics.

3.8.2.2. Accuracy and Reliability

It was reported that insurance data are generally complete and reliable, with detailed coding of multiple diagnoses. In most cases reporting from healthcare providers to the insurance database is mandatory and strictly regulated, with processes for validation and the querying of errors and inconsistencies. Because of the financial implications, there can be penalties for providing inaccurate information.

An important concern is that because the diagnoses reported by healthcare providers are used to determine reimbursement, there is an incentive to record more severe diagnoses or those requiring high-cost medical interventions. Clinicians may also record the diagnosis in the most favourable way to justify the reimbursement of medicines, reference to rehabilitation or for disability recognition.

Since insurance data are relevant only to persons covered by the health insurance system, there is possibility of bias, although in most cases the population covered is very broad. People excluded may be for example non-residents, some institutionalised populations whose healthcare is provided separately, or individuals seeking private sector care on an out-of-pocket basis.

PL distinguished in their insurance database between individuals with a personal identifier enabling linkage, and those without. The highest percentages without an identifier occurred for the indicators PB36-PB37 Land transport accidents (8.5%) and PB40-PB41 Intentional self-harm (incl. suicidal attempt) (7.8%) while for most conditions the figure was around 1%. The individuals without a personal identifier in the database typically include non-residents and individuals who could not be identified at the time of care (for example they were unconscious), but also homeless people and some other disadvantaged groups.

3.8.2.3. Timeliness and Punctuality

For countries using insurance data, the data were typically very timely because of their use for reimbursements. In some cases, special agreements had to be set up with the relevant health insurance institutions for access to the data.

3.8.2.4. Coherence and Comparability

In most cases coding is based on ICD-10 for diagnoses, and ATC if prescriptions are separately recorded, so that comparability is good in terms of classifications.

The health insurance system of each country is different, leading to differences in the scope of the insurance database in terms of different care services included. There are also differences in population coverage and whether the insurance is mandatory. This makes it difficult to compare the data directly on its own, but comparability is expected to be good where the relevant services are grouped together.

It was noted in the earlier pilot studies of 2005 onwards that for some conditions, the rates estimated by countries with insurance-based health systems were quite similar to each other, and higher than the rates of countries with different systems. In the current pilot studies HU, LT and PL do tend to have the highest age standardised morbidity rates for some indicators, but further research would be needed to determine whether this is for epidemiological reasons or due to data issues.

3.8.2.5. Conclusions

Insurance data vary in their scope but are generally complete and of high quality because of their importance for reimbursement and the funding of healthcare providers. For some countries they are the most essential data source.

- The role of the data in attracting reimbursement means that there is a risk of over-estimation or the transfer of patients to more severe diagnoses.
- The effects of uninsured persons and use of the private sector vary between countries and would merit further study.

3.9. Prescriptions

3.9.1. Preparatory remarks

Specialist databases of prescribed medications are likely to be accurate as regards the substances prescribed, their dosages, etc, and ATC codes are often used. The principal difficulty is that the prescription databases do not often have specific information on the diagnosis or indication for the treatment, so that the medication has to be sufficiently specific to a particular diagnosis to be used as a proxy. This could be the case for a limited number of diseases only.

For less serious conditions some patients might buy over-the-counter medicines instead of seeking a prescription. For some diseases, patients could be treated with or without medicine (e.g. diabetes could be treated with diet) making the use of prescriptions as a proxy incomplete.

3.9.2. Considerations according to the ESS dimensions of quality

3.9.2.1. Relevance

Specific ATC codes were recommended in the Guidelines for some indicators, while it was mentioned that the use of prescriptions might be feasible for some others. Decision in discussion with national experts was advised.

Prescriptions were used as a data source by six countries. In two of those, BE and HU, the prescriptions were recorded within an insurance database. The application of prescriptions to specific indicators ranged from the majority of indicators to less than half. Prescriptions were the largest data source for at least one indicator in FR and were identified as covering the majority of cases for 17 indicators in HU, and 22 indicators in MT.

3.9.2.2. Accuracy and Reliability

The use of drug prescriptions to identify cases is difficult, as many conditions do not have unique drugs used for their treatment, while some drugs are not sufficiently specific to only one diagnosis. For example, many of the same drugs are used to treat asthma, chronic obstructive pulmonary disorder (COPD) and other chronic respiratory diseases.

Other potential causes of error when using prescriptions as a proxy for a health condition include patients who have the disease of interest but do not take prescribed drugs, for example diabetes controlled by diet only, leading to underestimation of the indicator. There is also the possibility of patients taking a drug for other reasons than the disease of interest (e.g. in the context of drug repositioning); BE gave the example of a person with obesity using diabetes-related medication (categorised as ATC A10) for the purpose of weight control.

In some cases, very specific lists of ATC codes were applied for a few indicators, as in NL. BE calculated for each ATC a threshold value to determine a 'real' case based on a predefined number of Defined Daily Doses (DDDs) per year for each disease.

3.9.2.3. Timeliness and Punctuality

Where prescriptions data were used, they were regularly collected and available in a timely way, often for the administration of reimbursements.

3.9.2.4. Coherence and Comparability

Prescriptions are generally recorded using the ATC classification, and in some cases ICD-10 codes are also recorded showing the clinical indication for the prescription. The latter greatly increases the usability of the data for morbidity statistics. The ability to use prescriptions for the indicators varies between countries, as do the methods for ensuring the most precise application. Further, there is no doubt variation in prescribing practices between the countries, so that comparability is not straightforward.

3.9.2.5. Conclusions

Prescriptions data can be useful to ensure complete recording of some indicators, particularly well-controlled chronic conditions where there may be no hospitalisation or contact with any physician for some time.

- However, considerable caution is needed to ensure that only drugs with very specific application to the relevant condition are selected, and national differences in prescribing practice need to be considered.
- Prescriptions should preferably not be used as the only source for any indicator if possible.

3.10. Combined hospital data

3.10.1. Preparatory remarks

Considerations relating to combined sources of hospital data (for example, covering both inpatients and

outpatients) are like those mentioned for the individual sources above. Provided effective linkage can be achieved, the completeness and value of combined hospital data are likely to be high and has the potential to overcome limitations that exist in relation to sources such as outpatient or emergency care records when these are available only in isolation.

3.10.2. Considerations according to the ESS dimensions of quality

3.10.2.1. Relevance

Combined hospital data was only reported in that form by one country, NL. Other countries reported hospital inpatient and outpatient data together within an insurance database. The considerations on the quality of these data sources are the same as for hospital inpatient and outpatient data.

3.10.2.2. Conclusions

- Combined hospital data are likely to be of similar quality to hospital inpatient or outpatient data separately, or insurance data based on hospital records, and the same issues apply.

3.11. Other data sources

3.11.1. Relevance

A small number of other date sources, not falling into the categories above, were used by four countries:

HR: The Disabled Persons Registry

FI: Diagnoses recorded from social care institutions

MT: Patient Master Index (PMI)

NL: Diagnosis Treatment Combinations Mental Health Care (DTC-MHC), Long term care eligibility decisions (LTC-E CIZ), and Long-term care co-payments (LTC-C CAK)

These were used to fill known gaps in the coverage of the data for certain conditions, for example in NL the DTC-MHC data improved the completeness of information on mental health conditions because it represents the specialist psychiatric services. In MT, the PMI was used as a demographic file, that allowed linking the health information with other data sources by using the unique personal identifier.

The features of these data sources are mentioned under the entries above for the relevant countries. Because of their small number and diversity, it is not possible to generalize much about the appropriateness, reliability, or other quality aspects of such datasets.

Where such data sources relate to persons registered for social security benefits, eligibility for social care, co-payments and so on, there are considerations similar to those for insurance-based data that may be relevant – namely the possibility of diagnostic practice being affected by the aim of proving a certain type of disability or eligibility for services.

3.11.2. Conclusions

- Other data sources may be appropriate to fill known gaps in the completeness of specific indicators. These should be used with care in discussion with national experts, with careful consideration to representativeness of the population, reliability of recording the diagnoses and classifications used.
- There is a risk that the use of some unusual data sources could reduce comparability, for example because of differences in the purpose of the data collections or their scope in terms of social instead of health care.

4

Analysis by indicator

The guidelines of the MORB project recommended that specific questions or options for certain indicators should be addressed in the pilot studies. Here, these indicators are indicated by an asterisk '*' after their number.

The overall quality of each indicator is summarised using a Red-Amber-Green, so called 'RAG rating':

Red Not suitable for use, or another indicator is preferred.

Amber Suitable for use with reservations, requiring care in interpretation or further research to understand differences.

Green Suitable for use in a future data collection, with no more than minor reservations.

It is important to recognise that this rating relates to the quality and feasibility of the indicator in general and does not consider specific problems one or more of the pilot studies may have had, such as lack of primary care data.

4.1. Indicators in List A (indicators to be included for pilot data collection)

4.1.1. P1 Diabetes mellitus (incidence by person)

RAG status: Amber

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT
- Reported with some differences: BE, NL, PL

Classification

- ICD-10: E10-E14
- ICPC-1: T90 (NL)
- ICPC-2: T89-T90 (FI)
- ATC: A10 Drugs used in diabetes (BE, NL, PL)

Codes of the DTC-SSC clinical classification used for specialist care in NL are not shown as they are restricted to one country. However, the codes used to identify diabetes are of interest as an illustration of the range of medical specialties, reported conditions and treatments required to cover all likely contacts relating to this condition - see annex B section 1.3 Indicator P1 Diabetes: relevant codes in the DTC-SSC classification used in NL.

4.1.1.1. Notes on data sources and issues

FR: Significant underestimation. We know that diabetes is difficult to find in French insurance claims as well as in hospitalization data. This is mainly due to the under-diagnosis of this pathology.

MT: The data does not represent all the incidence cases, due to the fact that patients accessing primary care and outpatient care in the public and private settings were not included in the estimates.

NL: For incidence of diabetes, there was an effect of missing health contacts in primary care leading to over-counting of incident cases (that is, some will be identified as new cases by mistake) thus leading to an over-estimation which could be as great as 25%. This was partly caused by the

introduction of a new method of care organisation, 'chain care', for which only limited data on contacts was available.

4.1.1.2. Triangulation and plausibility

FR: A difference of prevalence of a factor two has been found when using the Constances cohort information or only the Insurance and Hospitalization data (SNDS): 5.89% vs 3.75%. This difference is roughly the same for diabetes incidence cases (0.43% vs 0.26% with only the SNDS).

It is generally accepted that cases of diabetes mellitus can go undiagnosed for a long time, so any measure based on service use is likely to underestimate the true epidemiological situation.

See also under P2.

4.1.1.3. Conclusions

A reasonable estimate of the incidence of diabetes mellitus can be made, but there are known weaknesses (both overestimates and underestimates) in several countries.

- The main data source is likely to be primary care. In hospital data and causes of death, use of secondary diagnoses and multiple causes of death data are essential.
- The codes for diagnosis and medications are quite specific and easy to implement.
- Because of its public health importance, the indicator should be included in any future morbidity data collection.
- However, it is likely that healthcare system and data collection differences make comparability questionable.
- Further research is needed to understand the observed differences.

4.1.2. P2 Diabetes mellitus (period prevalence)

Rag status: Amber

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT
- Reported with some differences: BE, NL, PL

Classification

- ICD-10: E10-E14
- ICPC-1: T90 (NL)
- ICPC-2: T89-T90 (FI)
- ATC: A10 Drugs used in diabetes (BE, NL, PL)

4.1.2.1. Notes on data sources and issues

FR: Significant underestimation. See under P2 above.

LT: By increasing the reference period from 3 years to 4, 5 or 6 years, the reported period prevalence increases by 4.0%, 6.7% and 8.6% in turn.

MT: Underestimation of the prevalence of diabetes mellitus, as a result of missing information on primary care and outpatient data from public and private sector.

NL: In 2016, 94 percent of all people with an ongoing diabetes episode recorded in primary care did have a registered GP contact in the past three years. After integration with other sources, this percentage increased to 96 percent. Therefore, implementation of the indicator was very effective for prevalence.

PL: By increasing the reference period from 3 years to 4, 5, 6, or 7 years, the reported period prevalence increases by about 5% for each extra year.

4.1.2.2. Triangulation and plausibility

FR: See under P2 above.

LT: Prevalence of diabetes mellitus according to the 2014 EHIS was quite similar.

NL: The rates were comparable to the national health interview survey (HIS) at younger ages, but substantially higher than the HIS at older ages. For all persons aged 65+ the indicator rate was 6.6 percentage points (pp) higher than the HIS. Possible explanations: difference between current experience and 3-year prevalence; a larger proportion at older ages might have Type 2 diabetes controlled by diet, and therefore be less aware of their condition; older people with diabetes might be more frail and therefore in care homes or less likely to respond to the HIS.

It is known that diabetes mellitus can go undiagnosed for a long time, while once diagnosed, some individuals may have their condition controlled by diet only and have no regular need for medication or other contacts.

The ECHI [7] indicator HSIND027001: Proportion of people reporting diabetes in the past 12 months, with the reference year 2014, was highest in FR followed by HU and FI. In the experimental rates from the pilot studies, PL and HU were highest while FR and FI were quite low. This is

(7) https://ec.europa.eu/health/indicators_data/indicators_en

consistent with the estimates from FR being a significant underestimate, but the reasons why lower rates are observed in FI and high rates in PL are unknown.

4.1.2.3. Conclusions

A reasonable estimate of the period prevalence of diabetes mellitus can be made, but while some countries report plausible comparisons with alternative data sources, others have known weaknesses (both overestimates and underestimates).

- The main data source is likely to be primary care. In hospital data and causes of death, use of secondary diagnoses and multiple causes of death data are essential.
- The codes for diagnosis and medications are quite specific and easy to implement.
- A question was raised on the possibility of a longer reference period than three years, but to keep consistency between indicators a change is not recommended.
- Because of its public health importance, the indicator should be included in any future morbidity data collection.
- However, it is likely that healthcare system and data collection differences make comparability questionable.
- Further research is needed to understand the observed differences.

4.1.3. P3 Dementia (incl. Alzheimer disease) (period prevalence)

RAG status: Amber

- Reported exactly according to the Guidelines: FI, FR, HU, LT, MT
- Reported with some differences: BE, HR, NL, PL

Classification

- ICD-10: F00-F03, G30 (HR: add F05.1) (NL: add G31)
- ICPC-1: P70 (NL)
- ICPC-2: P70 (FI)
- ATC: N06DA, N06DX01 (BE) N06DA02, N06DA03, N06DA04, N06DX01 (or: N06D Anti-dementia drugs excluding N06DX02 Ginkgo folium because of its frequent use for other indications) (NL)

4.1.3.1. Notes on data sources and issues

FI: Compared to the pilot study of 2005, the prevalence of Alzheimer disease is 27% lower despite the addition of primary care data; this may be because reporting from social care institutions has fallen and access to the data on disability benefits was not available.

FR: Considered satisfactory, with, however, the possibility of underestimating. The number of patients treated may be underestimated, since some patients may not register a long-term sickness under this pathology and not having been hospitalized in the reference period (2014-2016). We do not have any information about diagnosis in the primary care database.

MT: The number of patients treated for dementia is very likely to be underestimated. This is because the main source of information (POYC) collects information only about the patients receiving Donepezil from the government formulary. Patients who are receiving other anti-dementia medications are not included in this source of information. In addition, there is lacking information on dementia patients from the community and elderly homes.

4.1.3.2. Triangulation and plausibility

Clinical thinking and diagnostic practice regarding different forms of dementia vary between countries, as illustrated by the two proposed changes to ICD-10 codes. Many people with dementia are resident in care homes, so the extent to which institutional populations are included in the estimates is important for comparability.

According to ECHI indicator HSIND025015: EuroCoDe - As percentage of total population the highest rate of dementia is likely to be in BE, FR and FI (but figures are dated from 2006). The observed rates in FI are consistent with this, while low rates in FR are consistent with significant underestimation due to the lack of primary care data. The very low observed rates in BE suggest that the insurance-based source used for this indicator is inadequate for the purpose. Given the age of the ECHI data, effects of demographic and epidemiological change over time in some countries are also quite possible.

4.1.3.3. Changes proposed by the countries

HR proposed to add ICD-10 code F05.1 'Delirium superimposed on dementia' as this is often used for dementia in hospitals according to national practice. We recommend accepting this change to allow for national differences in practice, noting that this diagnosis is specific to dementia.

NL proposed to add ICD-10 code G31 'Other degenerative diseases of nervous system, not elsewhere classified' at the

request of national specialists; this code covers diseases such as Pick disease and Lewy bodies disease which were said to belong to a more 'modern' interpretation of dementia. When comparing total numbers using both variations, NL found that adding G31 resulted in a 0.6% increase of the total number of prevalent cases reported. Note that when mapping ICPC-1 to indicator P3, G31.2 is included. We recommend not accepting this change as the scope of G31 includes conditions different from dementia (including nervus degeneration due to alcohol) while specifically excluding Alzheimer, Lewy body disease and senility NOS.

4.1.3.4. Conclusions

A reasonable estimate of the period prevalence of dementia incl. Alzheimer disease can be made, depending heavily on the availability of suitable data sources including primary care, and covering institutional populations. Not all countries could achieve plausible estimates.

- A combination of data sources is likely to be needed, including primary care, hospital care (including secondary diagnoses) and causes of death (including multiple causes).
- There were different opinions on the codes for diagnosis. It is recommended that ICD-10 F05.1 should be added to the definition but G31 should not. Codes for medications are not completely specific but may be useful to identify cases.
- Because of its public health importance, the indicator should be included in any future morbidity data collection.
- However, it is likely that healthcare system and data collection differences make comparability questionable.
- Further research is needed to understand the observed differences.

4.1.4. P4* Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) (period prevalence)

RAG status: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: F10
- ICPC-1: P15 Chronic alcohol abuse (corresponding to ICD10 F10.1-F10.9, +G31.2), P16 Acute alcohol abuse (ICD10 F10.0). P15+P16 give full coverage of ICD10 F10, but also include G31.2: Degeneration of nervous system due to alcohol (NL)
- ICPC-2: P15-P16 (BE, FI)

4.1.4.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Considered satisfactory, with, however, the possibility of underestimating.

LT: Data for prevalence of mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) (F10) could not be fully complete: data from alcohol and drug abuse hospitals are not included in CHIF IS (but since 2018 such data are already included in CHIF IS), anonymously treated patients are not included in CHIF IS. Due to the nature of the diagnosis only cases when the person is seeking care could be included as well as the cases when person is treated from other diseases or injuries and F10 appears as a comorbidity. Almost for half of persons, diagnosis F10 is registered as comorbidity (on hospital discharge cards, death certificates, in emergency care episodes).

4.1.4.2. Triangulation and plausibility

There are few data sources to compare the observed rates against, for example the ECHI indicators relating to alcohol consumption are available for only a few countries. The indicator DHIND025010: Recorded adult (15+ years) per capita consumption (in litres of pure alcohol) shows LT as highest followed by HU for most years, which is consistent with the findings.

4.1.4.3. Issues from the EPIMS project

The completeness of reporting and comparability between the countries are both doubtful, therefore the pilot studies should review the quality of their data carefully. Depending on the pilot results, this indicator may not be feasible as a regular indicator.

4.1.4.4. Conclusions

Despite the reservations based on the EPIMS report, it seems that a reasonable estimate of the period prevalence of mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) is feasible, though some countries thought underestimation to be likely.

- The most suitable data sources vary between countries. Primary care, hospital care and causes of death are all helpful, but also emergency care if available.
- It was noted that the ICD-10 code G31.2 is specific to consequences of alcohol and this should be added to the definition.
- Potentially large epidemiological differences between the countries are possible, however the effects of cultural factors (willingness to seek care, attitudes to alcoholism) and healthcare system and data collection differences also need to be considered.
- Further research is needed to understand the observed differences and accuracy in the different countries.

4.1.5. P5 Schizophrenia, schizotypal and delusional disorders (period prevalence)

RAG status: Green

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: BE, NL

Classification

- ICD-10: F20-F29
- ICPC-1: P72 Schizophrenia ((ICD10 F20-F22, F24-F28), P98 Psychosis NOS/other ((ICD10 F23, F29, F53.1) (too many ICD10-codes included: ICD10- F53.1: Severe mental and behavioural disorders associated with the puerperium, not elsewhere classified) (NL)
- ICPC-2: P72 (FI)
- ATC: N05A (except N05AN) (BE)

4.1.5.1. Notes on data sources and issues

MT: The reported estimates for the period prevalence of schizophrenia, schizotypal and delusional disorders are very likely to be underestimated, as the important data sources such as primary care and outpatient mental health data were not available at the time of running the project.

4.1.5.2. Triangulation and plausibility

No comments were made, and the availability of existing indicators for easy reference is limited. It is unknown whether the observed differences are epidemiological or result from underestimation or overestimation in some countries. National differences in diagnostic practice are also likely to exist. Some of the people with serious psychiatric illness will be in institutional settings, so the coverage of these populations is important for comparability although overall population prevalence is quite low.

4.1.5.3. Changes proposed by the countries

In the Guidelines ICPC-1 code P72 was suggested for indicator P5. NL found that to better cover the scope of this indicator ICPC-1 code P98 (Psychosis NOS/other) should also be included. Adding this code increased the number of prevalent cases reported by 40%.

4.1.5.4. Conclusions

A reasonable estimate of the period prevalence of schizophrenia, schizotypal and delusional disorders is feasible. Although there are marked differences between the countries, no specific reservations were expressed about accuracy of the estimates.

- The most suitable data sources vary between countries and no one source is likely to be sufficient on its own. Coverage of institutional populations is important.
- When using ICPC-1, code P98 should be added to the definition.

4.1.6. P6 Mood (affective) disorders (period prevalence)

RAG status: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT PL
- Reported with some differences: NL
- Classification
- ICD-10: F30-F39
- ICPC-1: P73 Affective psychosis (ICD10 F30, F31, F34.0), P76 Depressive disorder (ICD10 F32-F39 (except F34.0), F41.2, F53.0). Too many ICD-codes included: F41.2: Mixed anxiety and depressive disorder F53.0: Mild mental and behavioural disorders associated with the puerperium, not elsewhere classified (NL)
- ICPC-2: P73, P76 (BE, FI)

4.1.6.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

4.1.6.2. Triangulation and plausibility

NL: Compared to the national HIS response 'Had a depression in last 12 months', the survey rates were 2.8pp higher than the indicator overall, dividing into 3.3pp higher for males and 2.3pp higher for females. In both sources, female rates were higher than male. Possible explanations: not every person who feels depressed will receive a medical diagnosis of affective disorder; the sex difference reflects a commonly found epidemiological pattern and may also show greater willingness among women to seek medical care for such conditions.

The pattern of higher rates in females than males, in some countries by a large amount, is consistent with findings from various studies of mental health. ECHI indicator HSIND028001: Proportion of people reporting depression in the past 12 months for 2014 showed the highest rate in FI, which is not found in the estimates. However, this might be caused by the difference between self-reported depression and a diagnosed affective disorder. In some countries, cultural and diagnostic differences might be relevant.

4.1.6.3. Conclusions

A reasonable estimate of the period prevalence of mood (affective) disorders is feasible. Although there are differences between the countries, few reservations were expressed about accuracy of the estimates. However, healthcare system and other differences may affect comparability.

- The most important data source is likely to be primary care, especially as most people with these disorders will not require hospitalisation.
- There are some inconsistencies between ICD-10 and ICPC in the scope of the available codes.

4.1.7. P7* Anxiety disorders (period prevalence)

RAG rating: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: F40-F41
- ICPC-1: P74 Anxiety disorder/anxiety state (ICD10 F41 (excl F41.2)) and P79 (corresponds to F40, F42). Not included: ICD10 F41.2: Mixed anxiety and depressive disorder (NL)
- ICPC-2: P74, P79 (BE, FI)

4.1.7.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Uncertainty about the validity of this indicator. The number of patients treated may be underestimated, since some patients may not register a long-term sickness for their anxiety disorders and not having been hospitalized in the reference period (2014-2016). We do not have any information about diagnosis in the primary care database.

4.1.7.2. Triangulation and plausibility

The pattern of higher rates in females than males, in some countries by a large amount, is consistent with findings from various studies of mental health. In some countries, cultural and diagnostic differences might be relevant. The extreme differences between countries, especially the high prevalence in HR and HU and suggest a strong influence of healthcare system or diagnostic differences. Identification of cases is unlikely to be successful without primary care data, causing the very low rate observed for FR.

4.1.7.3. Issues from the EPIMS project

The feasibility including differences in severity measured and identification of the appropriate medications has to be reviewed.

4.1.7.4. Changes proposed by the countries

In the Guidelines ICPC-2 codes P74 and P79 were suggested for indicator P7. For ICPC-1 only P74 was proposed. NL proposed that as the description of P79 in ICPC-1 (Other neurosis, Phobia, Compulsive neurosis) is almost equal to ICPC-2; this code should also be included when using ICPC-1. Adding this code increased the number of prevalent cases reported by 9%. We recommend accepting this change.

4.1.7.5. Conclusions

Compared to P6, there were greater differences between the countries and more reservations on the data quality were expressed. The data sources, healthcare system differences, diagnostic and cultural factors may all affect comparability. The results are not easily validated.

- No information on severity or on the use of medications was collected.
- The most important data source is likely to be primary care, especially as most people with these disorders will not require hospitalisation.
- There are some inconsistencies between ICD-10 and ICPC in the scope of the available codes. For ICPC-1, code P79 should be added to the definition.
- This indicator is not recommended for inclusion in future data collections.

4.1.8. P8 Parkinson disease (period prevalence)

RAG rating: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: NL, PL

Classification

- ICD-10: G20
- ICPC-1: N87 Parkinsonism (ICD10 G20, G21, G22); too many ICD-codes included (NL)
- ICPC-2: N87 (BE, FI)
- ATC: N04BA02, N04BA03, N04BD, N04BX01, N04BX02. Other N04-codes are not considered specific enough (NL) N04 (PL)

4.1.8.1. Notes on data sources and issues

FR: Considered satisfactory with, however, the possibility of underestimating.

BE: The results, obtained by analysing the prevalence/incidence of Parkinson's disease, should be analysed with caution as a difference exists between Parkinsonism and Parkinson's disease. Treatments exist to reduce the Parkinson-like symptoms that are not part of Parkinson's disease, but are initiated due to a different mechanism of action. These Parkinson-like symptoms are defined as Parkinsonism and can be mistakenly diagnosed as Parkinson's disease, especially using proxy diagnoses.

Also BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

MT: The number of period prevalence cases may be underestimated, since there is no information about diagnosis in the primary care database.

4.1.8.2. Triangulation and plausibility

There were few comments suggesting weaknesses in this indicator. The comparison between countries was reasonable, except that for males in HU there was an extremely high observed rate; the reason for this is unknown. These findings suggest that there are important differences of diagnostic practice or coding between some countries.

4.1.8.3. Changes proposed by the countries

In the Guidelines medications of ATC-group N04 (anti-Parkinson drugs) were suggested to identify cases. However, NL found that the number of cases found increased by 157% using the full group N04; 70% of cases identified using N04 medication were not known to have Parkinson disease from any other source. Some of the medicines in this group are known to be used for other health problems. Therefore, it is suggested to remove this ATC group from the Guidelines or use it with caution. It is recommended to change the suggested ATC codes accordingly.

4.1.8.4. Conclusions

- For most countries, estimates of the prevalence of Parkinson disease are feasible. However, there are likely to be significant differences in diagnostic or coding practice

which must be understood before reliable comparison is possible.
- The most important data source is likely to be primary care, especially as most people with these disorders will not require hospitalisation.
- The coding is mostly simple, however with some questions about the possible inclusion of cases of Parkinsonism which are not Parkinson diseases.
- Further research is needed on the observed differences between countries.

4.1.9. P9 Multiple sclerosis (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: BE, NL

Classification

- ICD-10: G35
- ICPC-1: N86 (NL)
- ICPC-2: N86 (FI)
- ATC: L03AB07, L03AB08, L03AX13, L05AA23, L05AA27, L05AA31, L05AA34, N07XX09 (BE)

4.1.9.1. Notes on data sources and issues

FR: Considered satisfactory with, however, the possibility of underestimating.

MT: The number of period prevalence cases may be underestimated, since only the 3-year prevalence was used. POYC (prescription register) represents cardholders eligible for medication, who had their card renewed or received a new card that year. As a result, those patients who are receiving medication and do not have to renew/receive a new card in those 3 years, will be excluded.

4.1.9.2. Triangulation and plausibility

No comments were made by the countries. The range of observed rates across the countries was plausible, but somewhat low for BE suggesting inadequate identification of cases. There was a notably higher prevalence shown in women which is consistent with the epidemiology of MS.

4.1.9.3. Conclusions

- Reasonable estimates of the prevalence of multiple sclerosis are feasible, provided the data sources have suitable coverage.
- The most important data sources differ between countries and are likely to include primary care, hospitals and to a lesser extent causes of death. Disability insurance or eligibility data, if available, will also be relevant.
- There were no difficult issues regarding coding, and a feasible list of ATC codes was used by BE.

4.1.10. P10 Epilepsy (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: G40-G41
- ICPC-1: N88 (NL)
- ICPC-2: N88 (BE, FI)
- ATC: N03AD01, N03AF03, N03AG04, N03AX10, N03AX14, N03AX15, N03AX17, N03AX18, N03AX22, N03AX23. Other N03-codes are not considered specific enough (NL)

4.1.10.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Considered satisfactory with, however, the possibility of underestimating.

4.1.10.2. Triangulation and plausibility

The variation in observed rates between the countries was quite wide, but not implausible considering the differences in data sources used and the wide variations found by studies (see e.g. Beghi, Giussani, Nichols, et al, 2019). The greater prevalence in males than females is to be expected.

4.1.10.3. Changes proposed by the countries

In the Guidelines medications of ATC-group N03 (anti-epileptics) were suggested to identify cases. However, NL found that the number of cases found increased by 270& using the full group N03, while 80% of cases were not known from any other data source. Some of the medicines in the group are known to be used for other health problems. Therefore, it is suggested to remove this ATC group from the Guidelines or use it with particular care. We recommend limiting the group of ATC codes accordingly.

4.1.10.4. Conclusions

- Reasonable estimates of the prevalence of multiple sclerosis are feasible, provided the data sources have suitable coverage, though some caution on comparability is needed.
- The most important data sources differ between countries and are likely to include primary care, hospitals and to a lesser extent causes of death.
- A more specific list of ATC codes was suggested by NL and should be implemented.

4.1.11. P11 Hypertensive diseases (incidence by person)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: NL, PL

Classification

- ICD-10: I10-I13, I15
- ICPC-1: K86 Essential hypertension without organ damage (ICD10 I10), K87 Hypertension with organ damage / secondary hypertension (I11-I13, I15 + I67.4). too many ICD-codes included: ICD10 I67.4: Hypertensive encephalopathy (NL)
- ICPC-2: K86-K87 (BE, FI)
- ATC: C03AA, C03AB, C03AH, C03AX01, C02CA04, C03BA, C03DB, C03EA, C09BA02-9, C09BB, C09DB, C09DA02-4, C09DA06-7, C09DA01, C02AB01-2, C02AC01-2, C02AC04-5, C02DB02-4, C02DC01, C02DD01, C02DG01, C02KA01, C02KB01, C02KC01, C02KD01, C02KX01, C09XA (PL)

4.1.11.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

4.1.11.2. Triangulation and plausibility

FR: Significant underestimation. A difference of prevalence of a factor two has been found when using the cohort information or only the Insurance and Hospitalization data (SNDS): 32.40 % vs 13.75 %. This difference is roughly the same for hypertension incidence cases (0.95 % vs 0.46 % with only the SNDS).

Hypertension is likely to be widely treated in primary care. Mild cases may be undiagnosed for a long time or be revealed only during examination for other health conditions. Therefore, the available data sources are important and national differences in diagnostic practice could also have an effect. Comparisons made by the countries on prevalence suggest that the findings are plausible.

4.1.11.3. Conclusions

- Reasonable estimates of the incidence of hypertensive diseases are likely to be feasible for most countries.
- Plausible figures depend mostly on the availability of primary care data.
- There are some minor inconsistencies with ICPC coding.

4.1.12. P12 Hypertensive diseases (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: NL, PL

Classification

- ICD-10: I10-I13, I15
- ICPC-1: K86 Essential hypertension without organ damage (ICD10 I10), K87 Hypertension with organ damage / secondary hypertension (I11-I13, I15 + I67.4). too many ICD-codes included: ICD10 I67.4: Hypertensive encephalopathy (NL)
- ICPC-2: K86-K87 (BE, FI)
- ATC: C03AA, C03AB, C03AH, C03AX01, C02CA04, C03BA, C03DB, C03EA, C09BA02-9, C09BB, C09DB, C09DA02-4,

C09DA06-7, C09DA01, C02AB01-2, C02AC01-2, C02AC04-5, C02DB02-4, C02DC01, C02DD01, C02DG01, C02KA01, C02KB01, C02KC01, C02KD01, C02KX01, C09XA (PL)

4.1.12.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

MT: The number of period prevalence cases for hypertensive diseases may be underestimated, due to application of 3-year period prevalence only. POYC (prescription register) represents cardholders eligible for medication, who had their card renewed or received a new card that year. As a result, those patients who are receiving medication and do not have to renew/receive a new card in those 3 years, will be excluded. In addition, important data source, namely public primary care was not available at the time of running the project.

4.1.12.2. Triangulation and plausibility

FR: Significant underestimation. See under P11 above.

LT: Prevalence of hypertension according to the 2014 EHIS was quite similar.

NL: Compared to 'Hypertension in last 12 months' in the national HIS, results were plausible, but the indicator was especially higher for older ages. Possible explanations: difference between current experience and 3-year prevalence/selection of healthier persons into the survey data; lack of awareness of the diagnosis.

4.1.12.3. Conclusions

- Reasonable estimates of the prevalence of hypertensive diseases are likely to be feasible for most countries.
- Plausible figures depend mostly on the availability of primary care data.
- There are some minor inconsistencies with ICPC coding.

4.1.13. P13 Ischaemic heart diseases (period prevalence)

RAG rating: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: NL, PL

Classification

- ICD-10: I20-I25
- ICPC-1: K74-K76 (NL)
- ICPC-2: K74-K76 (BE, FI)
- ATC: C01DA (PL)

4.1.13.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Possible underestimation because lack of diagnostic in the primary care database.

MT: Possible underestimation, as an important data source, namely primary care data could not be used in this pilot data collection.

4.1.13.2. Triangulation and plausibility

NL: Compared to 'Serious heart condition in the last 12 months, such as heart failure or angina pectoris' in the HIS, the indicator was much higher (12.9 pp) in the oldest age group. Possible explanations: difference between current experience and 3-year prevalence/selection of healthier persons into the survey data.

The ECHI indicator HSIND012150: Standardised death rate per 100,000 inhabitants, ischaemic heart disease for 2016 showed the highest mortality in LT, followed by HU and then HR. This is partly consistent with the observed estimates, though by contrast, mortality in P was much lower. The observed higher prevalence in males than females is to be expected.

4.1.13.3. Conclusions

- Reasonable estimates of the prevalence of ischaemic heart disease are likely to be possible, but there are some reservations. Comparison with mortality rates shows only partially consistent patterns.
- Mixed sources including primary care, hospital inpatients and causes of death should all be relevant.
- There are no difficulties regarding the classifications.

- Further research on the differences, and their relationship to mortality rates, would be beneficial.

4.1.14. P14* Acute myocardial infarction (incidence by episode)

RAG rating: Red

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT
- Reported with some differences: BE, NL, PL

Classification

- ICD-10: I21, I22 (BE: ICD-9 401, ICD-10 I21-I24)
- ICPC-2: K75 (FI)

4.1.14.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence.

FR: Possible underestimation. The number of episodes may be underestimated, because the causes of death are missing in this pilot exercise.

HU: Concerning P14, P15 and P17 AMI and strokes incidences, out of the NHIF databases, inpatient care data were used exclusively because the occurrence of these acute diseases always needs hospital treatment. Cases with respective ICD codes in the outpatient or drug prescriptions databases regard the treatment of consequences of earlier AMI or stroke cases.

LT: It was noted that significant part of myocardial infarction and stroke cases (27% for myocardial infarction and 40% for stroke) were registered in outpatient care and had no hospital stay (and it was not death case). It was noted by specialists that stroke or myocardial infarction could be treated in outpatient care (without admission to the hospital) in very rare cases. Sometimes diagnosis of stroke or myocardial infarction could be used in outpatient care to justify reimbursement of medicine, reference to rehabilitation care or for disability recognition. Therefore, specialists suggested using hospital discharge data and death certificates (when person died suddenly and not reach health care institution) only for the calculation of stroke and myocardial infarction data.

NL: After linkage of data sources, cases found in primary care acute for P14 and P15 were not found in hospital care or causes of death, which seems unlikely. It turned out that in primary care health contacts that take place a long time after the acute phase of the actual event were registered using the diagnosis code of the actual event. Possibly the ICPC classification has no good alternative. For this reason, NL decided to calculate indicators P14 and P15 without data from primary care. This led to a reduction of 60% in the number of cases reported.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.14.2. Triangulation and plausibility

HU: In the case of AMI indicators (P14, P15), aggregated data from the National Registry of Myocardial Infarction were also used for validation.

NL: Compared to 'Myocardial infarction in last 12 months' in the HIS, the rates found were almost identical at all ages.

4.1.14.3. Issues from the EPIMS project

The intension is to include only one of the indicators for Acute myocardial infarction in the shortlist: either incidence by episode or incidence by person. A decision will be made after the pilot studies based on the feasibility and comparability.

4.1.14.4. Conclusions

- The indicator on incidence by episode is likely to be feasible, however, a choice should be made between indicators P14 and P15.
- Indicator P15 Incidence by person is preferred as it avoids difficulties with the definition of episodes, which may vary between countries. Incidence by episode shows greater variability between the countries.
- This indicator should not be included in the future data collections.

4.1.15. P15* Acute myocardial infarction (incidence by person)

RAG rating: Green

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT
- Reported with some differences: BE, NL, PL

Classification

- ICD-10: I21, I22 (BE: ICD-9 401, ICD-10 I21-I24)

- ICPC-2: K75 (FI)

4.1.15.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence.

FR, HU, LT, NL: See P14 above.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.15.2. Triangulation and plausibility

HU: In the case of AMI indicators (P14, P15), aggregated data from the National Registry of Myocardial Infarction were also used for validation.

4.1.15.3. Issues from the EPIMS project

See P14 above.

4.1.15.4. Conclusions

- Estimates of the incidence by person from acute myocardial infarction appear to be possible. Variations between the countries are plausible, as is the greater incidence in males than females.
- The main data sources are hospital inpatients and causes of death. The counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation.

4.1.16. P16* Heart failure (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: I50
- ICPC-1: K77 (NL)
- ICPC-2: K77 (BE, FI)

4.1.16.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Possible underestimation because lack of diagnostic in the primary care database.

4.1.16.2. Triangulation and plausibility

The differences between countries in observed prevalence are likely to reflect issues noted below on diagnostic and death certification practice.

4.1.16.3. Issues from the EPIMS project

Because of expected variations in the coding and recording practice, the feasibility and comparability is uncertain. A decision will be made after the pilot studies.

FI proposed an option to expand the ICD-10 definition to include I11.0 Hypertensive heart disease with (congestive) heart failure), I13.0 Hypertensive heart and renal disease with (congestive) heart failure), and I13.2 Hypertensive heart and renal disease with both (congestive) heart failure and renal failure). We do not recommend adding these codes because of conflict with the preferred proposal below.

LT: Heart failure (I50) is often written on death certificates as the condition or complication of other diseases appearing during the last days or hours before death. Heart failure diagnosis was written on 51.8% of death certificates as one of the multiple causes of death (21,301 out of 41,106 deaths in 2016). It was the main cause of death in only 0.05% of all deaths. In 2016, 8,493 persons died having had no heart failure diagnosis during the years 2014-2016. Although such cases constituted only 4.8% of all prevalence cases of heart failure, the question is if we need to include heart failure diagnosis from death certificate when it is not the main cause of death.

NL: By using the broader definition proposed, the total number of cases was increased by only 0.04 percent.

4.1.16.4. Conclusions

- Reasonable estimates of the prevalence of heart failure are likely to be feasible, but care must be taken to exclude the diagnosis of 'heart failure' in some countries as a

symptom or mode of dying rather than an ongoing cardiac condition.
- Therefore, for COD data, heart failure should be counted only where it is the underlying cause of death.
- Care should be taken with hospital inpatient data for the same reason, considering distinguishing between main diagnosis and secondary diagnoses in discussion with national experts.
- Management of chronic heart failure is likely to involve primary care and possibly outpatient care, which are therefore necessary for adequate coverage.

4.1.17. P17 Stroke (incidence by person)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: I60-I64 (BE: ICD-9 430-434 ICD-10 I60-I64)
- ICPC-1: K90 Cerebrovascular accident (ICD10 I60-I64 and G46 Vascular syndromes of brain in cerebrovascular diseases) (NL)
- ICPC-2: K90 (FI)

4.1.17.1. Notes on data sources and issues

BE: Only taking into account hospitalizations might lead to an underestimation of the actual prevalence.

FR: The number of episodes may be underestimated, because the causes of death are missing in this pilot exercise.

HU: Concerning P14, P15 and P17 AMI and strokes incidences, out of the NHIF databases, inpatient care data were used exclusively because the occurrence of these acute diseases always needs hospital treatment. Cases with respective ICD codes in the outpatient or drug prescriptions databases regard the treatment of consequences of earlier AMI or stroke cases.

LT: It was noted that a significant part of myocardial infarction and stroke cases (27% for myocardial infarction and 40% for stroke) were registered in outpatient care and had no hospital stay (and it was not cause of death). It was noted by specialists that stroke or myocardial infarction could be treated in outpatient care (without admission to the hospital) in very rare cases. Sometimes diagnosis of stroke or myocardial infarction could be used in outpatient care to justify reimbursement of medicine, reference to rehabilitation care or for disability recognition. Therefore, specialists suggested using only hospital discharge data and death certificates (when a person died suddenly and had not reached a health care institution) for the calculation of stroke and myocardial infarction data.

4.1.17.2. Triangulation and plausibility

NL: Compared to 'Stroke, cerebral haemorrhage/infarction in last 12 months' in the HIS, rates were quite similar.

The pattern across countries generally seems plausible, however the very high rate for males in HU stands out and possible reasons should be investigated.

4.1.17.3. Changes proposed by the countries

LT: For calculating the incidence of stroke, checking two years back for incidence case could be insufficient, as the first stroke could happen earlier. However, for reasons of consistency with other indicators, we do not recommend any change in the methodology.

4.1.17.4. Conclusions

- Estimates of the incidence by person for stroke appear to be possible. Variations between the countries are plausible, as is the greater incidence in males than females.
- The main data sources are hospital inpatients and causes of death. As with AMI, the counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation.

4.1.18. P18 Cerebrovascular diseases (period prevalence)

RAG rating: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: I60-I69 (BE: ICD-19 430-438 ICD-10 I60-I69)
- ICPC-1: K90 Cerebrovascular accident (ICD10 I60-I64 and G46 (Vascular syndromes of brain in cerebrovascular diseases) (NL)

- ICPC-2: K90-K91 (FI)

4.1.18.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual prevalence.

FR: Possible underestimation because of lack of diagnostic in the primary care database.

4.1.18.2. Triangulation and plausibility

NL: Compared to 'Ever had a stroke, cerebral haemorrhage or cerebral infarction' in the HIS, rates were quite similar.

Prevalence of cerebrovascular diseases is quite complex and includes acute as well as longer-term illness. For example, it is influences by the survival from stroke.

ECHI indicator HSIND012130: Standardised death rate per 100,000 inhabitants, cerebrovascular disease for 2016 shows the highest mortality in LT followed by HR. In the observed morbidity HU and PL also had quite high rates, so the pattern is not quite consistent.

4.1.18.3. Conclusions

- Reasonable estimates of the prevalence of cerebrovascular diseases are likely to be feasible. Observed differences might be because of healthcare system features and diagnostic practice, for example relating to patients having care for rehabilitation.
- Primary care, hospital care and causes of death are all likely to make some contribution.
- There are minor possible inconsistencies with the ICPC coding.
- Research may be needed on differences in diagnostic practice and the organisation of care.

4.1.19. P19 Pneumonia (incidence by episode)

RAG rating: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: J12-J18
- ICPC-1: R81 Pneumonia (ICD10 J12-J18, but also: J10.0: Influenza with pneumonia, other influenza virus identified, J11.0: Influenza with pneumonia, virus not identified, A48.1: Legionnaire's disease) (NL)
- ICPC-2: R81 (BE, FI)

BE: The ICPC-2 code (R81) is not completely matched to the original proposed ICD-10 definition (J12-J18).

4.1.19.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Possible underestimation because of lack of diagnostic in the primary care database.

MT: Possible underestimation because of lack of data on pneumonia in the primary care database.

4.1.19.2. Triangulation and plausibility

NL: Comparing an 'incidence per person' version of indicator to 'Bronchitis or pneumonia during the last 2 months, (excluding. chronic bronchitis)' in the HIS, rates were partly similar with the indicator being lower at younger ages and higher at 65+. Possible explanations: a higher figure would be expected because the time horizon of the HIS question is shorter, but this might be balanced by a lower likelihood of younger people accessing medical care for the symptoms.

The reasons for differences observed are unclear, and do not appear to correspond to differences in mortality (Marshall et al, 2018). Therefore, it is likely that diagnostic practice or data collection differences are involved.

4.1.19.3. Conclusions

- No important problems were expressed on this indicator. However, large differences were observed which are likely to relate to the data sources used. Further research on comparability would be beneficial and comparisons should be made with caution.
- Hospital inpatient care, primary care and causes of death are all likely to contribute cases. Pneumonia often occurs among hospital inpatients and institutionalised elderly people. Coverage of the institutional populations is therefore important.

4.1.20. P20 Asthma (incidence by person)

RAG rating: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: J45, J46
- ICPC-1: R96 (NL)
- ICPC-2: R96 (BE, FI)

4.1.20.1. Notes on data sources and issues

BE: The differentiation between asthma and COPD is not always straightforward, even for the physician diagnosing the patient. Therefore, even when using diagnosis-based morbidity data, caution should be exercised when interpreting these results.

Also BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: The data source used cannot give a valid estimate.

MT: Lack of primary care data – settings where asthma is often diagnosed and managed.

4.1.20.2. Triangulation and plausibility

FR: Large underestimation. A difference of prevalence of a factor 10 has been found when using the Constance cohort information or only the Insurance and Hospitalization data (SNDS): 12.45% vs 0.14 %. This difference is of a factor 10 for asthma incidence cases (1.08 % vs 0.09 % with only the SNDS).

Comparing the difference between countries for incidence and prevalence of asthma, the pattern for prevalence is much less varied. This suggests that reliable measurement of prevalence is easier than the identification of new cases. See also under P21.

4.1.20.3. Conclusions

- The estimation of incidence of asthma seems to be feasible, but it is likely that measurement of prevalence is more reliable. It is less likely that the first point of diagnosis can make an accurate distinction between asthma and other respiratory conditions (COPD, bronchitis). This affects the counting of incidence.
- Prevalence is more suitable as asthma is a very long-term chronic disease.
- This indicator should not be included in the future data collections.

4.1.21. P21 Asthma (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: J45, J46
- ICPC-1: R96 (NL)
- ICPC-2: R96 (BE, FI)

4.1.21.1. Notes on data sources and issues

BE: The differentiation between asthma and COPD is not always straightforward, even for the physician diagnosing the patient. Therefore, even when using diagnosis-based morbidity data, caution should be exercised when interpreting these results.

Also BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: The data source used cannot give a valid estimate.

MT: Probably underestimated, as asthma is often a condition seen in primary care. Primary care database is not included in this pilot project.

4.1.21.2. Triangulation and plausibility

FR: Large underestimation. See under P20 above.

NL: Compared to 'Asthma in last 12 months' in the HIS, the indicator was slightly higher for most age-sex combinations. Possible explanations: difference between current experience and 3-year prevalence.

ECHI indicator HSIND029001: Proportion of people reporting asthma in the past 12 months for 2014 showed very high rates of asthma in FI and FR and quite uniform rates in the other countries. Other studies have also suggested higher prevalence in western and northern Europe, however many countries in eastern and central Europe are known to have higher mortality from respiratory diseases generally.

4.1.21.3. Conclusions

- Reasonable estimation of prevalence of asthma seems to be feasible. The countries with low observed rates are likely to be due to inadequacy of the data sources available for the pilot study.
- As a long-term disease, asthma is most likely to be covered by primary care data. However, hospital care and causes of death may add extra cases especially when secondary diagnoses are included.

4.1.22. P22* Chronic lower respiratory diseases other than asthma (incl. COPD) (period prevalence)

RAG rating: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: J40-J44, J47
- ICPC-1: R91 Chronic bronchitis / bronchiectasis, R95 Emphysema / COPD (NL)
- ICPC-2: R78, R79, R95, R99 (BE, FI)

BE: Identical ICPC-2 codes were used for P22 and P23. Therefore, identical results will be produced based on this definition, while the original ICD-10 proposed definition would imply (slightly) dissimilar results.

4.1.22.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Likely under-estimate because of lack of diagnostic in the primary care database. Must be considered as a lower bound of the prevalence.

MT: Likely under-estimate because of lack of diagnostic in the primary care database.

4.1.22.2. Triangulation and plausibility

LT: Prevalence of chronic lower respiratory diseases other than asthma according to the 2014 EHIS was quite similar.

NL: Compared to 'COPD, chronic bronchitis or pulmonary emphysema in last 12 months' in the HIS, the indicator was lower at younger ages but higher at 65+. Possible explanations: difference between current experience and 3-year prevalence/selection of healthier persons into the survey data; awareness of the correct diagnosis at older ages.

Differences observed were mainly plausible after accounting for the data sources used. The observed rate for BE was very high, which may reflect differences in the data sources and ICPC codes.

4.1.22.3. Issues from the EPIMS project

The inclusion of the code J47 bronchiectasis is uncertain and observations on its validity are invited.

NL: Exclusion of J47 resulted in a 2% reduction of cases found in the data sources using ICD-10, and only 0.2% in the total number of cases. This difference is likely because ICPC-1 code R91, along with ICPC-1 code R95 used for primary care, also includes bronchiectasis (along with chronic bronchitis), so that subjects with bronchiectasis remain to be included based on primary care data. In ICPC-2, codes are slightly different and bronchiectasis seems part of ICPC-2 R99 (respiratory disease, other). NL did not recommend using that ICPC-2 code to complete ICD10 J47, as it also includes many other respiratory diseases.

NL further comment: Indicator P22 Chronic lower respiratory diseases other than asthma (incl. COPD) and P23 (Chronic obstructive pulmonary disease (COPD) differ in only a few ICD-10 codes, with the result that P22 has 13% more cases.

4.1.22.4. Conclusions

- Reasonable estimates for this group of diseases seem to be feasible, but a very high rate for BE compared to other countries should be investigated. Since differences in diagnostic practice can affect the allocation of cases between respiratory diseases, some care with comparability is needed.
- Only NL commented on the use of ICD-10 code J47. It seems that there are contradictory issues concerning comparability with ICPC-1 and ICPC-2. We recommend that J47 should be included for completeness.

4.1.23. P23 Chronic obstructive pulmonary disease (COPD) (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, MT PL
- Reported with some differences: BE, NL

Classification

- ICD-10: J44
- ICPC-1: R95 Emphysema/COPD (NL)
- ICPC-2: R95 (FI) R78, R79, R95, R99 (BE)

BE: Identical ICPC-2 codes were used for P22 and P23. Therefore, identical results will be produced based on this definition, while the original ICD-10 proposed definition would imply (slightly) dissimilar results.

4.1.23.1. Notes on data sources and issues

BE: The differentiation between asthma and COPD is not always straightforward, even for the physician diagnosing the patient. Therefore, even when using diagnosis-based morbidity data, caution should be exercised when interpreting these results.

Also BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Likely under-estimate because of lack of diagnostic in the primary care database. Must be considered as a lower bound of the prevalence.

MT: Likely under-estimate because of lack of diagnostic in the primary care database.

4.1.23.2. Triangulation and plausibility

FR: Work done for a separate study with more specific definitions estimates the prevalence of COPD between 4% and 8% in the population aged more than 40 depending on the definition of the population of interest. Here we have a prevalence of 2.1% for this population.

Figures were relatively consistent across the countries, except for a very high rate in BE which is likely to be a result of differences in the data sources and ICPC codes.

4.1.23.3. Conclusions

Reasonable estimates for COPD seem to be feasible, while a very high rate for BE compared to other countries is likely to be due to methodology.

- Hospital inpatients and causes of death are likely to contribute cases in addition to those in primary care.

4.1.24. P24* Alcoholic liver disease (period prevalence)

RAG rating: Red

- Reported exactly according to the Guidelines: HR, FR, HU, LT, MT, PL
- Reported with some differences: BE, FI
- Not reported: NL

Classification

- ICD-10: K70
- ICPC-2: D97 (BE, FI)

4.1.24.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

MT: P24 is not considered a reliable indicator, because of lack of primary care data.

4.1.24.2. Triangulation and plausibility

The variation between countries was plausible, taking account of the data sources used in the pilot studies, except that the high rate for BE may be due to specific issues with recording of this diagnosis.

4.1.24.3. Issues from the EPIMS project

There are different views among the countries on the feasibility of separating alcoholic from other liver diseases. A decision will be made after the pilot studies.

BE: There is no possibility to differentiate the alcoholic and non-alcoholic liver diseases in primary care (both are D97 coded). Therefore, only the pooled results can be precisely determined.

FI reported all the indicators P24-P26 using the appropriate ICD-10 codes, but only the identical ICPC-2 code D97 for cases reported by some primary care institutions. Therefore, there could be a small overestimate from patients reported only in primary care.

NL could not report indicator P24 or P25 but were able to report P26. The main reason is that the ICPC-1 code D97 (cirrhosis/other liver disease) used in primary care data cannot distinguish according to whether the liver disease was connected to alcohol. On other data sources, NL also reported that their DTC-SSC codes used in hospital care cannot make this distinction.

It was noted in the pilot studies that the number of cases reported for indicator P26 is not necessarily the exact sum of those for P24 and P25, because it is possible for the same individual to be recorded with both alcoholic and non-alcoholic types of liver disease, either within one data source or in different data sources. The original validation of the estimates was changed to allow this possibility.

4.1.24.4. Conclusions

- The estimation of prevalence of alcoholic liver disease is feasible for some countries. However, because of the known inconsistency between ICD-10 and ICPC, estimates based on primary care data using ICPC are not possible. No alternative approach for these primary care data was suggested.
- This indicator should not be included in the future data collections.
- If there is a high demand for this indicator for policy reasons, an alternative might be to include it as an 'optional' indicator.

4.1.25. P25* Diseases of liver other than alcoholic (period prevalence)

RAG status: Red

- Reported exactly according to the Guidelines: HR, FR, HU, LT, PL
- Reported with some differences: BE, FI
- Not reported: MT, NL

Classification

- ICD-10: K71-K77
- ICPC-2: D97 (BE, FI)

4.1.25.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

4.1.25.2. Triangulation and plausibility

The variation between countries was plausible, taking account of the different data sources used in the pilot studies.

4.1.25.3. Issues from the EPIMS project

See P24 above.

4.1.25.4. Conclusions

- The estimation of prevalence of non-alcoholic liver disease is feasible for some countries. However, because of the known inconsistency between ICD-10 and ICPC, estimates based on primary care data using ICPC are not possible. No alternative approach for these primary care data was suggested.
- This indicator should not be included in the future data collections.
- If there is a high demand for this indicator for policy reasons, an alternative might be to include it as an 'optional' indicator.

4.1.26. P26* Diseases of liver (period prevalence)

RAG status: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: K70-K77
- ICPC-1: D97 (NL)
- ICPC-2: D97 (BE, FI)

4.1.26.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: The number of patients treated may be underestimated, since some patients may not register a long-term sickness for their disease of liver and not having been hospitalized in the reference period (2014-2016). We do not have any information about diagnosis in the primary care database.

MT: The number of patients treated may be underestimated, as there is no information about the diagnosis in the primary care settings.

4.1.26.2. Triangulation and plausibility

NL: Compared to 'Cirrhosis of the liver in last 12 months' in the HIS, the indicator was higher at all ages, with the difference increasing with age. Possible explanations: this is to be expected as cirrhosis is only a subset of diseases of the liver.

The variation between countries was not unreasonable. The low rate in FR is potentially due to the lack of primary care data. However, the low rate in FI is unexplained and should be investigated.

4.1.26.3. Issues from the EPIMS project

See P24 above.

4.1.26.4. Conclusions

- Estimation of the prevalence of liver disease (including both alcoholic and non-alcoholic liver diseases) seems to be feasible. A few national differences would benefit from investigation.
- Perhaps because this is quite a mixed group of diseases, no one data source can be guaranteed to provide the majority of cases.
- For this indicator rather than P24 and P25, there are no difficulties with classification.

4.1.27. P27* Rheumatoid arthritis (period prevalence)

RAG rating: Green

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, PL
- Reported with some differences: MT, NL

Classification

- ICD-10: M05, M06
- ICPC-2: L88 (BE, FI)

4.1.27.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: The number of patients treated may be underestimated, since some patients may not register a long-term sickness for their rheumatoid arthritis disease and not having been hospitalized in the reference period (2014-2016). We do not have any information about diagnosis in the primary care database.

NL: For indicator P27 it was decided to exclude primary care data. A validation study revealed that about 30% of patients registered with ICPC-1 code L88 ('Rheumatoid arthritis/related condition') could not be confirmed with inflammatory arthritis based on additional diagnostic information. L88 appears to cover in practice much more than rheumatic arthritis alone (apart from the inclusion of Bechterev and juvenile arthritis) - see Nielen et al (2013).

MT: No information about diagnosis in the primary care and outpatient settings. In addition, the 3-year period prevalence could not be used.

4.1.27.2. Triangulation and plausibility

NL: Compared with 'Chronic arthritis (chronic rheumatism, rheumatoid arthritis) in last 12 months' in the HIS, the indicator was much lower at older ages. Differences were large as 8.1 pp for males and 18.9 pp for females at 65+. Possible explanations: rheumatoid arthritis is only a subset of 'chronic rheumatism'; individuals may not be familiar with the symptoms or diagnosis (see Simons et al, 2017); chronically affected people especially at older ages might not seek medical care within a 3-year period; the criteria for medical diagnosis might be much more restrictive than understood by survey respondents. However, this discrepancy suggests that the indicator could be unreliable as a measure of the real population prevalence.

The variation between countries is small for males, but much larger for females. Rheumatoid arthritis is more prevalent in females partly because of their greater average longevity. The fact that greater variation is restricted to one sex suggests that national methodologies or diagnostic practice are not responsible.

4.1.27.3. Issues from the EPIMS project

The inclusion of the code M06.4 is uncertain and observations from the pilot studies are invited.

A suggestion had been made (HU) that ICD10-code M06.4 Inflammatory polyarthropathy should be excluded.

NL: The effect of exclusion of M06.4 was small, resulting in a reduction of the total number of cases by only 0.01 %. We do not recommend any change.

4.1.27.4. Conclusions

- Estimation of the prevalence of rheumatoid arthritis seems to be feasible. National differences can plausibly be attributed to demographic and epidemiological factors.
- It is not clear how well the indicator reflects the level of epidemiological prevalence, taking account of mild and undiagnosed cases.
- There are some difficulties with classification, but these were not universally found. No change is suggested.

4.1.28. P28* Arthrosis (period prevalence)

RAG status: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, PL
- Reported with some differences: NL
- Not reported: MT

Classification

- ICD-10: M15-M19
- ICPC-1: L89 Osteoarthrosis of hip, L90 Osteoarthrosis of knee, L91 Osteoarthrosis other (M15-M19; M13) also includes ICD10-M13 (other arthritis) (NL)
- ICPC-2: L89-L91 (BE, FI)

4.1.28.1. Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Possible underestimation because of lack of diagnostic in the primary care database.

4.1.28.2. Triangulation and plausibility

LT: Prevalence of arthrosis according to the 2014 EHIS was quite similar.

NL: Compared to 'Arthrosis of hips or knees in last 12 months' in the HIS, the indicator was substantially lower especially at 65+. As for P27, possible explanations include a difference between the actual medical diagnosis and the understanding of survey respondents. However, this discrepancy also suggests that the indicator should be treated with caution.

4.1.28.3. Issues from the EPIMS project

There is some doubt on the comparability of data obtained from primary care (using ICPC codes) and data using the ICD codes. Observations from the pilot studies on the feasibility and comparability are invited.

4.1.28.4. Conclusions

- The estimation of prevalence of arthrosis seems reasonably possible, with some reservations. There was quite wide variation between the countries for unknown reasons. Because of this, further investigation would be beneficial.
- No comments were made on the comparability of ICD-10 and ICPC and the national results did not suggest a major difference.
- Primary care is likely to be the main source of cases, but hospital care may also be important, as well as disability benefits and eligibility if available.

4.1.29. P29* Osteoporosis (period prevalence)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, PL
- Reported with some differences: NL
- Not reported: MT

Classification

- ICD-10: M80-M82
- ICPC-1: L95 (NL)
- ICPC-2: L95 (BE, FI)

4.1.29.1 Notes on data sources and issues

BE: It is not known to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

FR: Knowing the prevalence of osteoporosis is difficult because osteoporosis is not a reason for hospitalization but can appear as an associated diagnosis during hospitalizations for fracture treatment. Therefore, there is strong underestimation of people with osteoporosis. Additionally, osteoporosis evolves with age and what is measured by bone densitometry is the level of osteoporosis in relation to a given age: the prescription of a bone densitometry follows strict recommendations.

4.1.29.2. Triangulation and plausibility

FR commented on the difficulty of consistent diagnosis (above). Certain countries (HR, HU, NL) showed much higher rates than others, and HU showed a substantial rate for males as well as females. The findings confirm the difficult of comparable measurement of this condition.

4.1.29.3. Issues from the EPIMS project

There is known to be wide variation in the reported diagnosis of osteoporosis. The pilot studies are invited to comment on the adequacy of diagnosis and on influencing factors such as screening.

4.1.29.4. Conclusions

Making reasonable estimates of prevalence of osteoporosis seemed feasible for some countries, however implausibly wide variation confirmed the difficulty of consistent definition. Differences are also likely to be caused by national practices such as screening programmes.

- This indicator should not be included in the future data collections.

4.1.30. P30* Renal failure (period prevalence)

RAG status: Red

- Reported exactly according to the Guidelines: HR, FI, FR, HU, LT, PL
- Reported with some differences: BE, MT, NL

Classification

- ICD-10: N17-N19
- ICPC-2: U99 (BE, FI)

4.1.30.1. Notes on data sources and issues

BE: It is unknown to what extent external diagnoses (e.g. from specialist care or lab results) are recorded by the GP. Bypass of the GP when consulting direct specialised care will lead to an underestimation of the actual prevalence. As the data are case management-based, it is not always possible to identify a visit to the GP as a follow-up visit for a given condition or for a different disease.

MT: The 3-year period prevalence could not be used due to limitations of available data sources.

NL: For indicator P30 no adequate ICPC-1 code was available to cover only renal failure. It is included in ICPC-code U99 (Urinary disease, other) which would lead to the inclusion of many other health problems than only renal failure.

4.1.30.2. Triangulation and plausibility

Considering the non-specific nature of the ICPC code, as mentioned by NL, it is surprising that the rates observed are not more divergent. There was information in the EPIMS report on the variability of diagnosis in different countries; however, the same ICD-10 codes were used for all secondary care sources. Since it is not clear what the scope of coverage is in each country and how the primary and secondary care definitions overlap in practice, it is doubtful how meaningful this indicator will be.

4.1.30.3. Issues from the EPIMS project

Because of the variety of diagnostic codes used in the countries, observations are invited from the pilot studies on the most feasible and comparable definition.

4.1.30.4. Conclusions

- Although the rates observed are not very divergent, there is doubt as to how meaningful the indicator will be in some countries and whether the cases identified are comparable.
- This indicator should not be included in the future data collections.

4.1.31. P31* Urolithiasis (incidence by person)

RAG status: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT, PL
- Reported with some differences: NL

Classification

- ICD-10: N20-N23 (BE: ICD-9 592, 594 ICD-10 N20-N23)
- ICPC-1: U14, U95 (NL)
- ICPC-2: U14, U95 (FI)

4.1.31.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence.

FR: Considered unsatisfactory.

MT: Only includes patients admitted to hospital with this condition.

4.1.31.2. Triangulation and plausibility

Few specific comments were made. The observed rates varied widely without clear explanation. Some of the difference is no doubt caused by the data sources and their relevance and completeness, including differences in clinical practice and management of the condition. Epidemiological studies have found quite widely varying prevalence internationally (see e.g. Romero et al, 2010) so the picture is complicated, and it might be difficult to determine the meaning of the indicator with sufficient clarity.

4.1.31.3. Issues from the EPIMS project

There were doubts about the feasibility of this indicator, therefore observations from the pilot studies are invited.

4.1.31.4. Conclusions

- The rates observed are quite different, but at the same time within the wide range of prevalence rates identified in different epidemiological studies. In some cases, the low rates will be because of the lack of primary care data. The indicator may be feasible but further research is needed to enable meaningful interpretation.
- Multiple data sources are needed as management may be in primary care or hospital.

4.1.32. P32* Intracranial injury (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, NL
- Reported with some differences: PL
- Not reported: MT

Classification

- ICD-10: S06 (BE: ICD-9 592, 594 ICD-10 S06)
- ICPC-1: N79 Concussion (ICD10 S06.0 Concussion) (NL)
- ICPC-2: N79-N80 (FI)

4.1.32.1. Notes on data sources and issues

BE: Only taking into account hospitalizations might lead to an underestimation of the actual incidence.

FR: Possible underestimation because lack of linkage with emergency care database.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.32.2. Triangulation and plausibility

Difficulties were reported on the identification of cases using ICPC. There were very low rates observed in most countries, with a very high rate reported by FI which is likely to be a result of the methodology. Countries which used primary care or outpatients are likely to have detected less severe injuries, and possibly a wider range of head injuries. For incidence by person there was more variability between countries but also a smaller overall range, which seems more plausible. Therefore, incidence per person is preferred.

4.1.32.3. Issues from the EPIMS project

Several countries proposed variations of the coding; it was also suggested that it is feasible only to record the more severe cases requiring hospitalisation. Observations on these issues from the pilot studies are invited. The relative feasibility and value of incidence by person and incidence by episode is to be considered.

NL: In the Guidelines, ICPC-1 codes N79 and N80 are suggested for use in primary care data. NL found that use of N80 led to the inclusion many other head injuries, not necessarily intracranial, with an increase in total numbers of 134%.

4.1.32.4. Conclusions

There is a need to decide between incidence per episode and per person for the intracranial injury's indicator.
The range of observed rates seemed more plausible for incidence per person, although both indicators have problems of definition in the different data sources.

This indicator should not be included in the future data collections.

4.1.33. P33* Intracranial injury (incidence by person)

RAG status: Amber

- Reported exactly according to the Guidelines: BE, HR , FI, FR, HU, LT, MT, NL
- Reported with some differences: PL

Classification

- ICD-10: S06 (BE: ICD-9 850-854 ICD-10 S06)
- ICPC-1: N79 Concussion (ICD10 S06.0 Concussion) (NL)
- ICPC-2: N79-N80 (FI)

4.1.33.1. Notes on data sources and issues

BE: Only taking into account hospitalizations might lead to an underestimation of the actual incidence.

FR: Possible underestimation because lack of linkage with emergency care database.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.33.2. Triangulation and plausibility

Difficulties were reported on the identification of cases using ICPC. There were very low rates observed in most countries, with a very high rate reported by FI which is likely to be a result of the methodology. Countries which used primary care or outpatients are likely to have detected less severe injuries, and possibly a wider range of head injuries. For incidence by person there was more variability between countries but also a smaller overall range, which seems more plausible. Therefore, incidence per person is preferred.

If the full range of data sources is used, there will be different ranges of injuries covered by primary and secondary care, and perhaps between inpatient and outpatient services. There is also the possibility that following severe injuries, attendance at primary care or outpatient settings is for follow-up and rehabilitation, so the cases are not incident. Therefore, it is recommended that only data on causes of death, hospital inpatients, and emergency care if available, should be used.

4.1.33.3. Issues from the EPIMS project

See P32 above.

4.1.33.4. Conclusions

- There is a need to decide between incidence per episode and per person for the intracranial injuries' indicator. The range of observed rates seemed more plausible for incidence per person, although both indicators have problems of definition in the different data sources.
- To record only incident cases of severe intracranial injury, excluding other head injuries and excluding persons attending for rehabilitation or other non-acute reasons, we recommend that only data on causes of death, hospital inpatients, and emergency care if available, should be used.

4.1.34. P34* Fracture of femur (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT
- Reported with some differences: NL, PL
- Not reported: MT

Classification

- ICD-10: S72 (BE: ICD-19 820-821 ICD-10 S72)
- ICPC-2: L75 (FI)

4.1.34.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence.

FI: Compared to the pilot study of 2005, the prevalence of fracture of femur is 27% lower despite the addition of primary care data; this may be because reporting from social care institutions has fallen but the reason is not clear.

NL: After linkage of data sources, cases found in P34 and P35 were not found in hospital care or causes of death, which seems unlikely. It turned out that in primary care health contacts that take place a long time after the acute phase of the actual event were registered using the diagnosis code of the actual event. Possibly the ICPC classification has no good alternative. For this reason, NL decided to calculate indicators P34 and P35 without data from primary care. This led to a reduction of 25% in the number of cases reported.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.34.2. Triangulation and plausibility

There was much greater uniformity in the observed rates for incidence by person than for incidence by episode. The pattern is more plausible and incidence by person is preferred; this is also consistent with previous decisions.

4.1.34.3. Issues from the EPIMS project

Observations on the definition, especially the inclusion of surgical procedures, from the pilot studies are invited. The relative feasibility and value of incidence by person and incidence by episode is to be considered.

4.1.34.4. Conclusions

- No comments were made on the recording of surgical procedures. NL reported issues making the use of primary care data coded with ICPC difficult.
- Incidence by person seemed to produce more feasible comparative results and is also consistent with previous decisions, so is preferred.
- This indicator should not be included in the future data collections.

4.1.35. P35* Fracture of femur (incidence by person)

RAG status: Amber

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: NL, PL

Classification

- ICD-10: S72 (BE: ICD-19 820-821 ICD-10 S72)
- ICPC-2: L75 (FI)

4.1.35.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence.

NL: After linkage of data sources, cases found in P34 and P35 were not found in hospital care or causes of death, which seems unlikely. It turned out that in primary care health contacts that take place a long time after the acute phase of the actual event were registered using the diagnosis code of the actual event. Possibly the ICPC classification has no good alternative. For this reason, NL decided to calculate indicators P34 and P35 without data

from primary care. This led to a reduction of 25% in the number of cases reported.

PL: Restriction of health insurance cases to hospital care, avoiding overlap with causes of death which were restricted to non-hospital.

4.1.35.2. Triangulation and plausibility

There was much greater uniformity in the observed rates for incidence by person than for incidence by episode. The pattern is more plausible and incidence by person is preferred; this is also consistent with previous decisions.

4.1.35.3. Issues from the EPIMS project

See P34 above.

4.1.35.4. Conclusions

- No comments were made on the recording of surgical procedures. NL reported issues making the use of primary care data coded with ICPC difficult.
- Incidence by person seemed to produce more feasible comparative results and is also consistent with previous decisions, so is preferred.
- To record only incident cases of fracture of femur, excluding other cases such as person attending for rehabilitation, we recommend that only data on causes of death, hospital inpatients, and emergency care if available, should be used.

4.2. Indicators in List B (indicators to be considered for pilot data collection)

4.2.1. PB36* Land transport accidents (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT
- Reported with some differences: PL
- Not reported: MT, NL

Classification

- ICD-10: V01-V89 (BE: ICD-9 E800-E829 ICD-10 V01-V89)

4.2.1.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence. Although there are ICD-10 codes to indicate a land transport accident, due to the nature of the data (which are administrative, primarily focused on the financial aspects of secondary care), a primary focus on the given care might exist. This leads to the inclusion of the codes regarding fractured bones or other traumas rather than indicating an accident took place.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: We know that ICD10 codes for land transport accidents are hardly ever used by practitioners. Thus, these figures are not a good estimation.

LT: For calculation of incidence of injuries and external causes only curative cases are taken into account, rehabilitation and long-term treatment cases are excluded. Coding of external causes is not very accurate. External causes are coded mostly in emergency care departments and hospitals. In outpatient settings (incl. primary health care) the coding of external causes is not obligatory. Therefore, some mild injuries could be missed.

MT: Did not report this indicator.

NL: Did not report this (or any of the subsequent) indicators.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.1.2. Triangulation and plausibility

The plausibility of the rates observed is difficult to judge, since there are at the same time quite large differences in the rates, differences in the data sources and their completeness, and in the epidemiological incidence of road traffic accidents and injuries. There was no similarity to the pattern of ECHI indicator HSIND034001: Proportion of individuals aged 15+ reporting to have had an accident in road traffic during the past 12 months. Therefore, no conclusion could be drawn from the findings.

Several countries reported that details of the external cause of an injury are not recorded, or are incomplete, especially for outpatient and emergency care. In primary care, only very mild injuries or cases for rehabilitation are likely to be recorded.

4.2.1.3. Issues from the EPIMS project

This indicator is optional in the pilot studies. Observations on the feasibility, validity and comparability are invited. The relative feasibility and value of incidence by person and incidence by episode is also to be considered.

4.2.1.4. Changes proposed by the countries

FR: Concerning these two indicators, the basic information necessary to inform them is collected for each accident and processed by a specialized organization: The French Road Safety Observatory (ONISR). These data are sent to the Directorate-General MOVE. Our correspondent in the observatory informs us that it is agreed that the transmission of these data to the European Commission must serve all the necessary uses. Therefore, the question is how to work on the data that is available to the DG MOVE.

4.2.1.5. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) and the statistics from national road safety observatories and similar institutes are a more useful source of information.
- This indicator should not be included in the future data collections.

4.2.2. PB37* Land transport accidents (incidence by person)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: PL
- Not reported: NL

Classification

- ICD-10: V01-V89 (BE: ICD-9 E800-E829 ICD-10 V01-V89)

4.2.2.1. Notes on data sources and issues

BE: Only taking into account hospitalizations might lead to an underestimation of the actual incidence. Although there are ICD-10 codes to indicate a land transport accident, due to the nature of the data (which are administrative, primarily focused on the financial aspects of secondary care), a primary focus on the given care might exist. This leads to the inclusion of the codes regarding fractured bones or other traumas rather than indicating an accident took place.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: We know that ICD10 codes for land transport accidents are hardly ever used by practitioners. Thus, these figures are not a good estimation.

MT: Transport accidents seen in the primary care setting were not included due to unavailability of the data.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.2.2. Triangulation and plausibility

See under PB36.

4.2.2.3. Issues from the EPIMS project

See PB36 above.

4.2.2.4. Changes proposed by the countries

FR: See PB36 above.

4.2.2.5. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) and the statistics from national road safety observatories and similar institutes are a more useful source of information.
- This indicator should not be included in the future data collections.

4.2.3. PB38* Accidental falls (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT
- Reported with some differences: PL
- Not reported: MT, NL

Classification

- ICD-10: W00-W19 (BE: ICD-9 E880-E888 ICD-10 W00-W19)

4.2.3.1. Notes on data sources and issues

BE: Only considering hospitalizations might lead to an underestimation of the actual incidence. Similar concerns as for PB36 and PB37 are expressed; a fall might cause a diagnosis of trauma, which would most likely be included in the diagnosis, rather than the fall itself.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Possible underestimation because of coding defects and lack of linkage with the emergency care database.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.3.2. Triangulation and plausibility

It is difficult to assess the plausibility of the rates observed due to quite large differences in the rates, and differences in the data sources and their completeness. Therefore, no conclusion could be drawn from the findings. No alternative source for ready comparison was found.

Several countries reported that details of the external cause of an injury are not recorded, or are incomplete, especially for outpatient and emergency care. In primary care, only very mild injuries or cases for rehabilitation are likely to be recorded.

4.2.3.3. Issues from the EPIMS project

See PB36 above.

4.2.3.4. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, though representing only the most severe cases.
- This indicator should not be included in the future data collections.

4.2.4. PB39* Accidental falls (incidence by person)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: PL
- Not reported: NL

Classification

- ICD-10: W00-W19 (BE: ICD-9 E880-E888 ICD-10 W00-W19)

4.2.4.1. Notes on data sources and issues

BE: Only taking into account hospitalizations might lead to an underestimation of the actual incidence. Similar concerns as for PB36 and PB37 are expressed; a fall might cause a diagnosis of trauma, which would most likely be included in the diagnosis, rather than the fall itself.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Possible underestimation because of coding defects and lack of linkage with the emergency care database.

MT: Accidental falls seen in the primary care setting were not included.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.4.2. Triangulation and plausibility

See under PB38.

4.2.4.3. Issues from the EPIMS project

See PB36 above.

4.2.4.4. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, though representing only the most severe cases.
- This indicator should not be included in the future data collections.

4.2.5. PB40* Intentional self-harm (incl. suicidal attempt) (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: PL
- Not reported: NL

Classification

- ICD-10: X60-X84

4.2.5.1. Notes on data sources and issues

BE: Only considering fatal cases leads to an underestimation of the actual incidence. The indicator could experience the same problems as the previous indicators in list B. Additionally, the definition of the indicator is rather broad, as the indicator includes anything between suicidal attempts, intentionally inflicting cuts, or intentional self-poisoning. This could lead to a significantly varying case definition depending on the sources available to investigate the indicator, leading to vastly different results among the Member States.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Considered unsatisfactory because of lack of linkage with the emergency care database and lack of diagnostic in the primary care database.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.5.2. Triangulation and plausibility

The plausibility of the rates observed is difficult to judge, but for some countries seemed disproportionately high. Although the ECHI indicator HSIND012270: Standardised death rate per 100,000 inhabitants, suicide and intentional self-harm for 2016 showed LT having a much higher rate than most of the countries, the overall comparison was not plausible.

Several countries reported that details of the external cause of an injury are not recorded, or are incomplete, especially for outpatient and emergency care.

4.2.5.3. Issues from the EPIMS project

See PB36 above.

4.2.5.4. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, though representing only a subset of intentional self-harm (that is, successful suicide attempts).
- This indicator should not be included in the future data collections.

4.2.6. PB41* Intentional self-harm (incl. suicidal attempt) (incidence by person)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: PL
- Not reported: NL

Classification

- ICD-10: X60-X84

4.2.6.1. Notes on data sources and issues

BE: Only considering fatal cases leads to an underestimation of the actual incidence. The indicator could experience the same problems as the previous indicators in list B. Additionally, the definition of the indicator is rather broad, as the indicator includes anything between suicidal attempts, intentionally inflicting cuts or intentional self-poisoning. This could lead to a significantly varying case definition depending on the sources available to investigate the indicator, leading to vastly different results among the member states.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Likely a significant underestimation due to lack of linkage with emergency care database and lack of diagnostic in the primary care database.

MT: No data from primary care.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.6.2. Triangulation and plausibility

See under PB40.

4.2.6.3. Issues from the EPIMS project

See PB36 above.

4.2.6.4. Conclusions

- The pilot studies showed that the external cause of an injury is not recorded or incomplete in many healthcare settings and is unlikely to provide meaningful information. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, though representing only a subset of intentional self-harm (that is, successful suicide attempts).
- This indicator should not be included in the future data collections.

4.2.7. PB42* Complications of medical and surgical care (incidence by episode)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT
- Reported with some differences: MT, PL
- Not reported: NL

Classification

- ICD-10: Y40-Y66, Y69-Y84

4.2.7.1. Notes on data sources and issues

BE: Only considering fatal cases leads to an underestimation of the actual incidence. The same concerns as above regarding the broad spectrum of possible case definitions could be made, as the indicator could include anything between a minimal allergic reaction to antibiotics to major complications during surgery leading to the death of the patient.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Likely underestimate because of coding defects.

MT: Likely underestimate, as only codes Y40-Y59 were collected.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.7.2. Triangulation and plausibility

FR: Note that there are related statistics by ECDC on healthcare-associated infections (https://www.ecdc.europa.eu/en/healthcare-associated-infections).

The plausibility of the rates observed is difficult to judge but varied widely for unexplained reasons and in some cases seemed disproportionately high. Likely there are national differences in the recording and coding of complications of medical and surgical care. The relevant external cause code may not always be used.

4.2.7.3. Issues from the EPIMS project

See PB36 above.

4.2.7.4. Conclusions

- The pilot studies reported several doubts about the reliability and consistency of recording, and the range of observed variation is unlikely. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, representing severe and well-recognised cases).
- This indicator should not be included in the future data collections.

4.2.8. PB43* Complications of medical and surgical care (incidence by person)

RAG status: Red

- Reported exactly according to the Guidelines: BE, HR, FI, FR, HU, LT, MT
- Reported with some differences: PL
- Not reported: NL

Classification

- ICD-10: Y40-Y66, Y69-Y84

4.2.8.1. Notes on data sources and issues

BE: Only taking into account fatal cases leads to an underestimation of the actual incidence. The same concerns as above regarding the broad spectrum of possible case definitions could be made, as the indicator could include anything between a minimal allergic reaction to antibiotics to major complications during surgery leading to the death of the patient.

FI: Poor validity, data covers only inpatient hospitalizations. No data from hospital outpatient care or primary care available.

FR: Likely under-estimate because of coding defects.

MT: Likely underestimate, as only codes Y40-Y59 were collected.

PL: Able to report this indicator, using the combination of health insurance data (restricted to hospital cases only) and causes of death (for cases outside hospital).

4.2.8.2. Triangulation and plausibility

FR: Note that there are related statistics by ECDC on healthcare-associated infections (https://www.ecdc.europa.eu/en/healthcare-associated-infections).

See under PB42.

4.2.8.3. Issues from the EPIMS project

See PB36 above.

4.2.8.4. Conclusions

- The pilot studies reported several doubts about the reliability and consistency of recording, and the range of observed variation is unlikely. It is probable that the most accurate information is obtained from death certification, but only for the most severe cases.
- Therefore, we recommend that COD (single source data) is a more useful source of information, representing severe and well-recognised cases).
- This indicator should not be included in the future data collections.

5

Summary and conclusions

5.1. Morbidity statistics; points for consideration

5.1.1. Accessibility of the data sources

The EPIMS project produced an inventory of morbidity statistics' sources in 15 Member States. The results made it clear that, despite the existence of relevant EU regulations covering the statistical collections of data, the legal situation in each country is different. Factors such as ownership of a data source, the legal basis for its collection and use, or frameworks for data sharing between institutions influence the feasibility of accessing and linking data, and the timescales for doing so. National differences in data confidentiality practices may affect not only the availability of data sources, but also the level of detail that can be accessed within a data source (for example, exact date of birth or only completed years of age).

For some of the countries in the present pilot studies, availability of data was not a problem. Agreements were usually in place between the institutions collecting or owning the data, where appropriate. Most datasets were feasible to use by year T+3, for example 2016 data were available in time to use in 2019, and often the availability was much faster – in one case within a few months. This was especially the case where the data were collected regularly for administrative purposes such as insurance reimbursements.

However, some of the pilot studies experienced significant problems. Therefore, data were generally not adequate to produce reliable estimates according to the MORB project guidelines in several countries.

BE reported that acquiring access to microdata from various sources was a difficult and a cumbersome process under the national data protection regime. In a specific research project, the use of pseudonymised individual (personal) data can only be authorized after a complex procedure of investigating the security risks of each dataset and receiving positive advice from the Information Security Committee. Additionally, to comply with the privacy regulations, the patients should be aware of the information that is being processed (by the means of a clause or personal contact) or a legal framework should allow such processing.

Furthermore, most of the data sources used by BE did not provide the dataset itself to the researchers. Instead, a remote access system was used to grant access to the data. As each data source uses their own (proprietary) system, creating a linkage between the data from one source to another was not possible. Therefore, the intended concept of multi-source linked data for morbidity statistics could not be achieved.

FR had intended to use causes of death data, but the delivery and processing of this dataset was not sufficiently timely. While the COD data for 2016 had been made available, it was not yet linked to the SNDS data source used within the timescale of the project.

HU was unable to include primary care diagnoses in their estimates. The team's efforts to reach an agreement with the institution responsible for the GP database (the National Health Insurance Fund) were unsuccessful.

MT reported that the primary care and outpatient data (both private and public), as well as dispensing data for medicines were not accessible at the time of the project. They may be available for future data collections.

Summary

In most cases, the multiple datasets needed were available through collaboration with various institutes and were sufficiently timely for regular use by T+3 years. However, a few countries experienced difficulties with obtaining or linking the data which made reliable multi-source morbidity statistics impossible. Future data collections (involving new countries) should be organised with sufficient time for data-sharing agreements to be negotiated, possibly taking about two years of preparation.

5.1.2. Issues affecting institutionalised populations

5.1.2.1. Care of the elderly

The share of the population who are resident in institutions, most often care homes/nursing homes, increases with age. For example, NL noted that in 2016, 2% of the entire population resided in an institution of some type. The share living in an institution was 4% at age 80 years, 11% at age 85, 23% at 90, and 43% at 95 and over. Around 4% of the population was 80 years or older. Elderly women more often live in institutions than men of the same age.

In NL, general practitioners do not provide care for the institutionalised population, who instead receive primary care from a specialised elderly care physician. Since the main source of data was the Nivel-PCD network of GPs, this meant that the institutionalised proportion of the population could not be specifically included for most indicators. The assumption was made that the incidence or prevalence of conditions is identical in the institutionalised population as in people in private residences. NL was able to make estimates for the elderly institutionalised population only for indicator P3 Dementia (incl. Alzheimer disease) for which social care data provided an acceptable proxy.

In the other countries, the inclusion of institutional populations varied according to the countries' healthcare systems and insurance laws. Primary care data are the most likely to be affected, as in some countries the ongoing care of institutionalised people is not carried out by GPs. For other data sources there should not be significant problems or differences in this regard.

5.1.2.2. Psychiatric institutions

The inclusion of long-term residents of psychiatric institutions may be important for the certain indicators on mental health.

In some data sources, psychiatric hospitals are only partly covered. For example, in NL the specialist data DTC-MHC includes health records up to one year of stay in a psychiatric hospital. In NL in 2016 a total of 8650 persons received institutionalized mental health care for longer than one year, but no information was available on these diagnoses. It was estimated that these individuals could potentially increase the indicators relevant to mental illnesses (P4-P8) by 0.5% each.

In most countries the number of individuals in psychiatric hospitals or other mental health institutions is quite small, and most people with psychiatric conditions live in community settings. The measures specific to psychiatric diagnoses are most likely to be affected if this group is omitted from the available data sources.

5.1.2.3. Other institutional populations

Other groups which may not be covered by the major data sources include prisoners and military personnel. Generally, these are small groups relative to the total population. Military personnel tend to be younger and in good health, so they could be expected to add relatively few cases for most indicators. Some countries had no accessible health records for these groups (NL).

5.1.2.4. Summary

For most countries, institutionalised populations (with some small exceptions) are included in the main data sources for morbidity statistics. In a few cases, specifically in case of serious psychiatric conditions and dementia, a minority of countries do not have complete coverage.

5.1.3. Private sector healthcare

In some countries (NL) data from private sector health services or services not covered by the national health insurance are not included. In others (PL), treatment by private sector providers is mostly included because it is provided to most people under the national health insurance. For example, PL reported that 90.1 % of the population are eligible for publicly funded healthcare and prescriptions.

Both private sector healthcare provision and the number of people not covered by health insurance are said to be small in comparison to public services, typically less than 10% of the total. Precise figures are often not available. The nature of the private sector varies between healthcare systems: private providers sometimes concentrate on certain kinds of health conditions and treatments, such as minor elective surgeries, so their relevance varies depending on the indicator. The presence and use of private healthcare also vary geographically within some countries, depending on the socioeconomic composition of areas.

Private sector pharmacies are often the source of prescription data, which is included in the data sources used when the prescriptions are under the national health insurance.

HU used the national EHIS of 2019 to adjust the indicators for the proportion of people using private healthcare, for a number of health conditions which can be identified from the EHIS and where private sector care is feasible. The EHIS question CD1 on 'chronic diseases in the past 12 months' was complemented by additional questions on diagnosis by a medical doctor and medicine use, which are usually parts of the national EHIS questionnaire, and a newly added question concerning the type of health services used. The list of conditions covered and question wording can be found in annex B 1.1 Use of the European Health Interview Survey (EHIS) in HU to identify proportions of individuals with certain conditions using private healthcare.

Summary

While it is generally understood that the majority of healthcare is provided or paid for by the public national systems, in most countries there is a private sector whose importance varies by health condition, socially and geographically. The extent of the private sector does not seem to be understood very precisely. We recommend that for a future data collection, all countries should thoroughly collect data on private sector healthcare either directly or using proxy approaches such as comparison with the EHIS and adjust the estimates accordingly.

5.1.4. Inclusion of non-residents

The Guidelines recommended that the estimates should be provided divided into residents of the country and non-residents where possible. By 'residents' is meant all persons having their usual residence in the Member State on the reference date ([8]), while a non-resident means any person not meeting that definition of residence. An additional category of 'Unknown' was allowed in case of missing data, or if the data sources used could not distinguish between residents and non-residents. The 'Total' category is then the sum of all these categories.

The countries mostly had little success in accurately identifying non-residents or including them separately in the estimates.

For BE, specific issues were mentioned regarding the proposed definitions of residence. Some of the data sources determined the residence of the patient based on their domicile address, not on their length of stay in Belgium. Some sources had no case definition for non-residents and were thus not capable of registering non-residents by default. As a result of these issues, meaningful estimates by residence were not possible.

For LT, non-residents could be identified, and indicators could be calculated in the same way as for residents. However, for foreign non-residents (as opposed to Lithuanian citizens resident abroad) linkage of records was not possible due to the absence of a personal identifier, so that elimination of duplicates, linkage between the insurance data and deaths, and the identification of incident cases were all impossible. Similarly, FI was able to identify non-residents using a specific area code, but cases could not be accurately linked over time or between data sources.

In NL the main source of data was the Nivel-PCD database from a network of GP practices. Because GPs provide regular care only to registered individuals living in a defined area, this source includes only residents. While people who are not registered with the GP can access care on an emergency or casual basis, the data available for such exceptional cases is not recorded in a form feasible to use for statistical purposes, and there is no personal identifier. Similarly, in FR, some registers used were restricted to residents who are eligible for certain services and benefits.

In PL data on publicly funded services for non-residents are included in the National Health Fund database, but in the case of short-term visitors there is no personal identifier (PESEL number). Non-residents make up less than 2% of the PL population. In the findings from PL it was noted that people without an available PESEL number include not only non-residents, but also homeless people and perhaps other marginalised groups.

Within the data used for the estimates, in PL, the percentage of individuals without the PESEL number

([8]) Regulation (EU) 1260/2013 on European demographic statistics (http://eur-lex.europa.eu/eli/reg/2013/1260/oj)

accounted for less than 1.0% for period prevalence and up to 2.0% for incidence for most indicators. The conditions showing the highest missingness were incidence of Land transport accidents (8.5%), incidence of Intentional self-harm (incl. suicidal attempt) (7.8%), and prevalence of Epilepsy (4.4%). The first of these could plausibly affect non-residents at a higher rate, and the second might more affect homeless people or others in disadvantaged situations; however, the reason for the finding on epilepsy is not apparent. A full breakdown of this analysis is in annex - B.1.2 Coverage of the health insurance data in PL: percentage of patients without a personal identifier.

In cases where non-residents could be identified, an interesting finding is that the incidence and prevalence for some conditions differs from non-residents, presumably for reasons including demographic differences, receipt of healthcare in other countries, and reasons for and modes of travel.

In a similar analysis by LT, non-residents were identified and the indicator with the highest proportion of non-residents was incidence of Land transport accidents (1.36%), the only indicator with a proportion greater than 1%.

Summary

The counting and production of estimates for non-residents was difficult in all countries. In various cases, the data sources were themselves restricted to residents only, such as persons registered with a GP or eligible under the national health insurance. Where non-residents were included and could be separately identified, the lack of a personal identifier or inability to link to a population register meant that cases could not be counted accurately, e.g. incident cases could not be distinguished from previous cases. Therefore, to achieve the greatest consistency, we recommend that non-residents should be excluded as far as possible from the calculations.

5.1.5. Additional matters for consideration

Several further items were raised by the countries, as follows.

The possibility of calculation of longer period prevalence should be discussed (comment by LT and PL). This would make the reported prevalence closer to the epidemiological prevalence, which can be much greater for chronic conditions. The annex - B.1.4 Differences in period prevalence of diabetes mellitus (E10-E14) in LT according to the length of reference period. This can be considered in the future, but we recommend that for consistency a uniform three-year period prevalence should be applied for all indicators. The use of different reference periods might cause confusion to users. A longer period would require access to additional years of data, which could be difficult for some countries or introduce complications due to changes over time in service provision or diagnostic practice.

For some conditions like AMI and stroke, decide whether 'first ever' or 'first in reference period' is more important (comment from LT). We recommend that generally the definition of incidence for all relevant indicators is 'incidence per person, first occasion in reference period'. This is more reliable than 'first ever' as there is no guarantee in many data sources that the record will make clear whether the individual had any previous occurrence, before the first year of the reference period. If a variation in the counting of incidence is essential, it will be specified in the indicator definitions.

Differences between versions of ICD-10 and the future introduction of ICD-11 should be considered (comment from NL). The implementation by the countries may vary in terms of the annual updates of ICD-10, and also some national modifications can be used. The versions implemented in healthcare and mortality may be different. The complexity of different versions and national implementations would require additional research and is potentially too large a subject to cover, therefore we recommend that these issues be left to the responsibility of national experts.

The NL raised a question about the effect of variations over time due to changes in the data sources. This is a valid concern but is the responsibility of the national experts to consider.

LT and NL commented that a new term such as 'contact prevalence' could be created, to distinguish between the estimates based on healthcare contact reported in the pilot studies and the 'actual' prevalence measured by epidemiological studies. Such an unfamiliar term might be confusing to readers; however, we recommend that in descriptions the words 'based on healthcare contact' should be appended where appropriate.

Several countries raised that the practical resources and time required for the work should be made clear. All countries devoted substantial resources to the pilot studies. Even after the initial development of methods and programs, in some cases the regular production of morbidity statistics would require the full-time work of perhaps two people. Most countries described thorough

efforts to develop reproducible dataset linkage and analysis methods, and it would be valuable for examples of these methods (even including program code) to be published in detail.

5.2. Quality of the data sources and estimates

5.2.1. Considerations from the EPIMS project and the Guidelines

Based on the findings of the EPIMS project as well as the Quality Assurance Framework of the ESS [9], the Guidelines set out the need to establish the quality of the available morbidity estimates in the following terms:

- **Relevance:** The data source is suitable in terms of (a) its coverage or representation of the national population, and (b) its appropriateness for ascertainment of cases of the relevant health conditions.
- **Accuracy:** The data are (a) subject to regular processes of validation and quality assurance, (b) believed to be sufficiently free of errors and omissions, and (c) able to be delivered from the original source to the national project team without loss of integrity.
- **Timeliness and punctuality:** The data are (a) available for all of the reference period, (b) able to be delivered to the national project team within the timescales needed, and (c) unlikely to be subject to any excessive delay.
- **Accessibility:** The national project team is confident of being able to obtain the data, including (a) completion of any regulatory procedures, and (b) meeting any necessary costs.
- **Clarity and interpretability:** There is sufficient knowledge on the data and data source available to the national project team, based on (a) detailed metadata, and (b) access to technical experts.
- **Coherence and consistency:** The data contains the variables, codes and classifications which are needed for production of the indicators, generally following international practice in the use of the classifications. Variations within the data (e.g. changes in codes used between different years) are understood.
- **Comparability:** Data processing and recording (a) follows generally accepted practices and conventions, and (b) is consistent over time and place (e.g. between regions). Relevant national differences in medical practice, diagnostic recording, etc are able to be taken into account.

5.2.2. Quality of the underlying data sources

In most cases the underlying data used for the morbidity estimates are from administrative sources recording healthcare activity, health insurance payments, decisions on eligibility for care, prescription or dispensing of medicines, or the certification of deaths. These sources can typically be considered fairly reliable in terms of accuracy because the data are used regularly for operational purposes (e.g. by clinicians), and in many cases are already well-established as sources for published medical research or statistical reporting.

BE described a logical sequence of questions for choosing the best sources of data:

1. Is the source based on medical diagnoses or on proxy diagnoses (e.g. the use of prescription medication)?
2. Is the source exhaustive or sample-based?
3. If the source is sample based, is the sample population well defined and representative?
4. Does the source capture all cases? Or in the case of a sample, is the source exhaustive for the covered population?
5. Is the coverage regional or national?
6. (If the source is regional, how can the regional differences be accounted for (if existing)?
7. Does the source allow the computation of annual figures?

Routine procedures to assure the quality of administrative data were described. For example, in PL the health insurance data submitted to the National Health Fund are validated with more than 900 rules and audited continuously. Errors are reported back to the pharmacies and health care providers. If necessary, the NHF can issue a penalty for incorrect reporting of data, and report issues to different authorities including the police.

However, it was reported that in some data sources it is known from previous experience that certain diagnostic codes can be used in ways that are different from the formal description.

The **completeness** of data seems to be most at risk when key actors do not consider the data important, or when data collection takes place in situations where recording is less routine or convenient. For example, NL reported that in

[9] https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646

hospital discharge data, diagnostic information is missing from around 20% of day care admissions.

The **accuracy** is at risk where the data are used for payments to hospitals, performance management or similar purposes. Here, there may be an incentive to record diagnoses in certain ways, such as to assign the patient to a casemix group attracting higher reimbursement. For example, LT commented that in insurance data, a more severe diagnosis could be chosen for better reimbursement, to justify reimbursement of medicine, reference to rehabilitation care or for disability recognition. There is consequently potential for misleading comparisons between countries if some data sources are subject to such issues.

In terms of demographics, most countries reported little or no missing data for age and sex. In some cases, this information was available through linkage to the population registry. The situation regarding data on residence or non-residence was much more difficult - this is described in section 5.1.4 Inclusion of non-residents.

Known quality issues of the underlying data are reported for each country and data source in section 2 on analysis by country and section 3 on analysis by data source.

5.2.3. Quality assurance processes in the pilot studies

The pilot studies reported various methods of reaching valid decisions on methodological questions and subsequently assuring the quality of their estimates. The quality assurance was generally quite extensive and used a combination of approaches as discussed below.

5.2.3.1. Reference to existing documentation and scientific literature

References were given to official publications and published academic literature which variously described the data sources and their quality, the statistical methods (e.g. for data linkage), or previous research findings based on those data sources. Where websites were mentioned for general information these links are given in the text of this report, while citations of specific reports and articles can be found in the annex Annex C References.

5.2.3.2. Exploration of the data in multiple ways

The accuracy of translation between ICD-10 definitions and other classifications used in the data sources could sometimes be checked by exploring the data in alternative ways. For example, NL examined the mapping of the indicators to their DTC-SSC codes used in hospital data using data from 2016 and 2017 which had been partially coded with both ICD-10 and DTC-SSC, first mapping the frequency of association between each DTC-SSC and ICD-10 code and selecting the most strongly associated DTC-SSC codes for each definition, and then mapping the DTC-SSC codes back to ICD-10 to assess the sensitivity and specificity of the resulting selections of cases.

5.2.3.3. Comparison with the previous morbidity pilot studies

FI compared the results with the morbidity pilot exercise of 2005. They found that the number of registered cases was substantially higher for most medical conditions: the number doubled for alcoholic liver disease, diabetes mellitus, rheumatoid arthritis, multiple sclerosis, accidental falls, arthrosis, mental and behavioural disorders due to use of alcohol including alcohol dependence, heart failure, and osteoporosis. The increase was even higher for diseases of liver other than alcoholic and pneumonia (three-fold), chronic lower respiratory diseases other than asthma including COPD (four-fold), depression and other affective disorders and anxiety disorders, and renal failure (six-fold or more). The increase was mostly because of the inclusion of data from primary care.

For two medical conditions there was a reduction from the 2005 pilot study of 27 percent. No obvious reason was found for this. For fracture of femur, the same data sources were used to find the cases. For dementia (incl. Alzheimer's disease), FI included data on primary care visits as a new data source in 2016, but unlike 2005 had no access to data on disability benefits. There was also a decline in the coverage of diagnoses in the Hospital Discharge Register for social care institutions.

5.2.3.4. Triangulation between data sources

The quality of the estimates was evaluated by self-assessment of the project teams and colleagues, comparison with results obtained during previous actions, and the results of other data sources e.g. psychiatric morbidity study, general hospital morbidity study and European Health Interview Survey (description from PL).

For many indicators, comparison could be made between the estimates and other available health statistics, for example a health interview survey. Such comparisons are not always possible for all the indicators, and there is not always an exact comparison, for example because:

Not all the indicators are feasible for self-reporting in a survey, such as mental illness and conditions with high fatality.

- Self-reported illness is not expected to match exactly with medical diagnosis.
- Categories or codes used in different data sources may not match sufficiently.

However, general patterns seen in the data may be expected to be consistent across the different data sources. Where there is inconsistency, it should be possible to explain this either by known differences (e.g. of non-matching categories) or by plausible hypotheses drawing on known issues such as types of bias in surveys, including the omission of institutionalised people from most national surveys.

FR compared the estimates against statistics from the Constances cohort, a general-purpose population-based prospective epidemiological cohort under the responsibility of the Inserm-University of Versailles-Saint Quentin en Yvelines Joint Service Unit (UVSQ) 'Epidemiological Cohort in Population'. The cohort is representative of the French population affiliated to the national workers social security system (salaried workers, professionally active or retired and their family; about 85% of the French population). The main data which are collected by various means (self-administered questionnaires, medical questionnaires administered by physicians, health examinations, extraction of national databases) at inclusion and during follow-up concern the following areas: sociodemographic characteristics, examination data on health (including collection of biological material for the biobank), and health data collected by questionnaires. There is a longitudinal follow-up, without limitation of duration. Over the 2014-2016 inclusion period, the sample will consist of 88,109 individuals, of whom 39,420 are aged 50 and over. More information on the Constances cohort is available at https://www.constances.fr/.

HU used two specific sources for validation of certain indicators: (a) a healthcare-based full coverage survey on certain diseases of persons registered at the general practitioners' service, and (b) the National Registry of Myocardial Infarction.

LT compared the period prevalence for a number of indicators to the European Health Interview Survey (EHIS) done in 2014 by Statistics Lithuania. For some diseases (diabetes mellitus, hypertension, chronic lower respiratory diseases other than asthma, arthrosis) the results were quite similar, while for others the differences were more significant. LT noted that for many conditions' prevalence was higher in the calculated indicators than in the EHIS. As well as noting the difference of two years from 2014 to 2016, health selection bias and exclusion of the institutional population in the survey could be important.

NL compared a few indicators with their national health interview survey for the years 2016 and 2017 combined ([10]).

For most of the 13 indicators they compared, the patterns by age and sex were largely consistent between the two sources. In some cases, there were relatively small differences in the rates which can plausible be explained by reporting and methodological issues. In others there were marked differences which can nevertheless be explained.

Indicator P6 Mood (affective) disorders (prevalence) was compared to 'Had a depression in last 12 months'. The survey rates were 2.8 percentage points (pp) higher than the indicator overall, dividing into 3.3 pp higher for males and 2.3 pp higher for females. In both sources, female rates were higher than male. This finding is plausible, since not every person who feels depressed will receive a medical diagnosis of affective disorder. The sex difference reflects a commonly found epidemiological pattern and may also show greater willingness among women to seek medical care for such conditions (see e.g. Cox (2014); Mackenzie, Gekoski and Knox (2006); Piccinelli and Wilkinson (2000)).

In PL in 2020, new survey questions will provide a comparison to the estimates and test the assumption that the results are applicable to the whole population, including those accessing healthcare outside the national insurance. These questions asking about medical diagnosis of specified conditions will be added to the regular household sample survey "Health care in households" which is conducted by Statistics Poland periodically every 3-5 years.

The most informative findings based on triangulation exercises by the countries are reported in section 4 on analysis by indicator.

([10]) http://opendata.cbs.nl/statline/#/CBS/nl/dataset/83384NED/table?dl=73759

5.2.3.5. Consultation with national experts

All countries mentioned extensive consultation with national experts, both in preparation for the pilot studies (for example, to confirm the most suitable sources or to refine the choice of codes) and to interpret and comment on the findings. For example, NL described repeated discussions with expert colleagues which checked all the following information points:

- Do the total number of cases and the calculated rates comply with results that were expected?
- Is the distribution of cases by age and sex as expected?
- Is the number of cases found in each source as expected?
- Is the number of unique cases provided by one single source as expected?
- Does the overlap between primary care and hospital care fit with assumptions?
- Is the overlap between different hospital-based services (e.g. inpatient discharges and casemix measures) as expected?
- In case of doubt on the selection of diagnostic codes: compare results on different selections, how large is the impact of choices after integration with other data sources? What is the right choice?

HR mentioned that although workshops with relevant stakeholders for improving the quality of morbidity statistics were planned, because of COVID-19 pandemic they were not organized.

5.3. General conclusions on future data collections

5.3.1. Feasibility of a morbidity statistics data collection

The pilot studies show that the systematic and meaningful collection of morbidity statistics is feasible, but still faces serious difficulties. Compared to earlier projects, the implementation of standard definitions and methods reduced the scale of differences between countries and produced more plausible findings. The most important difficulty is that despite the standardization of methods, unavoidable differences between national healthcare systems and their data collections continue to be a problem for comparability for many indicators. Determining whether the observed differences in incidence and prevalence are due to compatibility of the data sources or reflect epidemiological patterns is not simple.

Some healthcare systems are likely to incorporate systematic biases, such as incentives to record a more severe diagnosis in insurance-based data to attract higher reimbursement. These issues can be identified only by the national experts, by detailed studies of diagnostic and coding practice.

Another difficult issue is the use of different classifications, particularly ICPC in primary care. Because the scope of ICPC codes does not always correspond to ICD-10 codes in a way that allows precise mapping, certain indicators cannot be feasible if some countries rely on ICPC-based records for primary care, unless the relevant conditions can be fully covered using other data sources.

5.3.2. The need for multiple data sources

The results of the EPIMS project identified the importance of linking data in almost all the possible scenarios. The individual data sources have different forms of bias or omissions and cannot produce a reliable or comparable estimate if they are not linked with other databases. The evidence reported from the countries in the EPIMS project highlighted that in most cases the sources are complementary to each other with various degrees of overlapping. Algorithms should be developed across sources to obtain best estimates and identification of individuals (including by anonymised methods) is generally necessary to avoid double counting.

The pilot studies confirmed the necessity of multiple, linked data sources to provide plausible findings. Where one or more data sources were unavailable this tended to produce unreasonably low estimates for the relevant countries. Not all countries were able to access primary care data, which is essential for most chronic diseases and those conditions which do not necessarily involve hospitalisation. Equally, for several more severe diseases, cause of death data are essential as many people die without a hospital admission or long course of illness.

It is essential for data sources to be linked, whether by a personal identifier or (if necessary) by probabilistic methods, since otherwise duplicates across datasets cannot be identified and new cases may be impossible to distinguish from recurring cases. Of concern is the experience in BE of gaining access to the data only through separate secure mechanisms, making linkage impossible. Avoiding such problems, like gaining access to data overall, depends on both the national legislation on privacy and

data protection and the willingness of data owners to reach suitable agreements.

5.3.3. Plausibility and interpretation of the findings

The main purpose of the project was to establish the feasibility of the data collections, not to measure or make substantive comparisons between the rates of incidence and prevalence shown by the indicators. At the same time, the plausibility of the patterns and differences observed was one of the factors considered when evaluating the likely meaningfulness of each indicator. Where an indicator was considered feasible, however, this does not mean that all national estimates for that indicator were plausible.

It was noted that, compared to the earlier pilot studies of 2005 onwards, the implementation of standard methods generally reduced the range of variation between the countries. In the earlier pilot studies, it was observed that those countries with insurance-based data collections often had rates quite like each other, and higher than countries with different systems. This still appeared to be the case for some indicators, though in a less consistent way.

The observed incidence and prevalence rates are the result of a complex interaction between the data sources available and their characteristics in each country, other features specific to each country such as healthcare system organisation and diagnostic practice, and the actual epidemiology of the condition of interest. In some cases, an indicator was considered not feasible because of the difficulty of separating these factors. In others, it seemed likely that understanding could be improved with further research.

Within the scope of this project, plausibility could be checked by the consistency or expected patterns between the countries, and with reference to readily available comparators such as the ECHI indicators where suitable. In some cases, reference to specific epidemiological articles was made. However, it was not possible to undertake a systematic comparison of the findings with the epidemiological literature, national reports, or other data collected by international organisations. Such additional research would be very valuable to aid interpretation of the indicators in future but would constitute a separate project.

5.3.4. Proposed shortlist of indicators for future data collections

The overall quality of each indicator is summarised section 4 on analysis by indicator and in Table 7: Summary of recommendations on the indicators using a 'RAG rating' as follows:

Green Suitable for use in a future data collection, with no more than minor reservations.

Amber Suitable for use with reservations, requiring care in interpretation or further research to understand differences.

Red Not suitable for use, or another indicator is preferred.

It is important to recognise that this rating relates to the quality and feasibility of the indicator in general and does not take into account specific problems one or more of the pilot studies may have had, such as lack of primary care data.

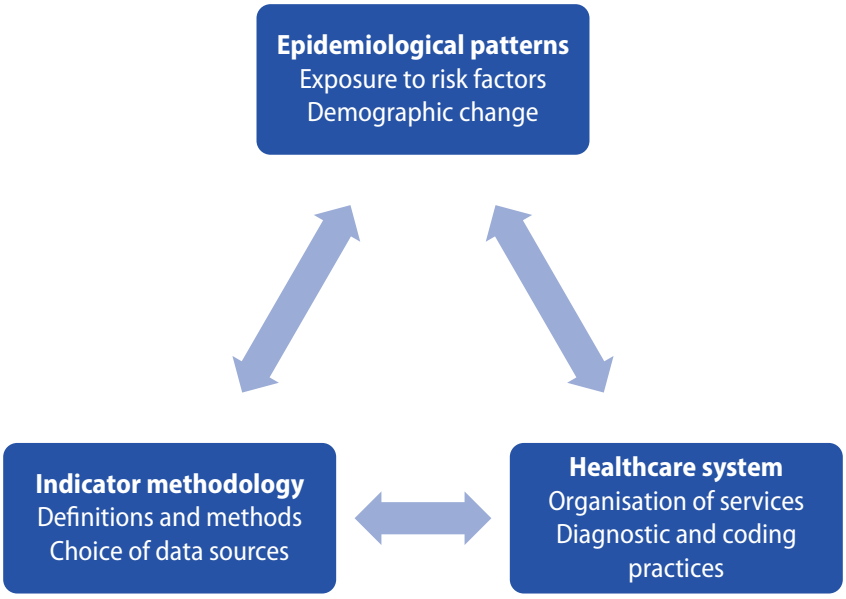
Recommended definitions for indicators for a future data collection are shown in Annex D Recommendations on the future shortlist: indicator definitions.

5.3.5. Principles for future data collections

The factors influencing the observed incidence and prevalence based on healthcare contact, and the national differences and other patterns, can be said to be determined by a 'triangle' of three types of factors – the epidemiological patterns, indicator methodology, and healthcare system features - as shown in Figure 1. The purpose of the indicators is to identify the actual epidemiological patterns, but the ability to do so depends on sufficient uniformity in the other two aspects, without which differences in the observed patterns cannot be meaningfully interpreted.

FIGURE 1

Factors influencing the indicator data collection



The pilot studies have so far suggested that harmonisation of the definitions and methods is feasible. To a lesser extent so is the comprehensive and linked use of data sources. The most obvious deficiencies were caused by some indicators where harmonisation of the definition was obstructed by technical issues, and by certain countries being unable to obtain or link important datasets. Thus, the value of any future data collections depends partly on the ability of all participating countries to base their estimates on a full set of relevant data sources.

However, the interpretation of findings will still be very difficult without also accounting for the healthcare system and its effects on the data. Therefore, the preparation for any future data collections should also include research by all the participating countries on at least the following aspects, leading to adjustment of the estimates for each indicator on a well-documented basis:

- The size and role of private sector healthcare, how it varies socially and geographically within the country and how it affects each indicator. Depending on the availability of information, this analysis might be done either directly or by proxy using sources such as the EHIS.
- The proportion of uninsured persons and non-residents, and where relevant the proportion of residents not registered with a GP. If sample sources are used, the representativeness of the sample and its suitability to be generalised to the national population must be carefully considered.
- Likely bias caused by the healthcare system financing and organisation, for example incentives to record particular diagnoses to attract higher reimbursement or show eligibility for benefits and services. Other organisational issues may also need consideration, such as record-keeping practices when care is shared between providers. These will require consultation with national experts and analysis of (for example) patterns of diagnosis over time.

Other subjects for which further research would be beneficial but is less essential at this point include specific clinical and diagnostic practices which may differ between countries and over time, such as the use of non-standard codes; cultural issues affecting diagnostic language; and changes in the setting in which treatment is carried out (such as increases in minor surgeries done in primary or outpatient care),

The time needed for these activities, and for practical preparations and the arrangements to obtain the data, will have to be considered in the timescales for future development.

Annex A. Tables

TABLE 1

List of pilot indicators

No.	Disease or group title	ICD-10 codes	Measure	Reference period	ICD-10 Chapter	Pilot questions (*)
List A: indicators to be included for pilot data collection						
P1	Diabetes mellitus	E10-E14	Incidence by person	1 year	IV	N
P2	Diabetes mellitus	E10-E14	Period prevalence	3 years	IV	N
P3	Dementia (incl. Alzheimer disease)	F00-F03, G30	Period prevalence	3 years	V	N
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	F10	Period prevalence	3 years	V	Y
P5	Schizophrenia, schizotypal and delusional disorders	F20-F29	Period prevalence	3 years	V	N
P6	Mood (affective) disorders	F30-F39	Period prevalence	3 years	V	N
P7	Anxiety disorders	F40-F41	Period prevalence	3 years	V	Y
P8	Parkinson disease	G20	Period prevalence	3 years	VI	N
P9	Multiple sclerosis	G35	Period prevalence	3 years	VI	N
P10	Epilepsy	G40-G41	Period prevalence	3 years	VI	N
P11	Hypertensive diseases	I10-I13, I15	Incidence by person	1 year	IX	N
P12	Hypertensive diseases	I10-I13, I15	Period prevalence	3 years	IX	N
P13	Ischaemic heart diseases	I20-I25	Period prevalence	3 years	IX	N
P14	Acute myocardial infarction	I21, I22	Incidence by episode	1 year	IX	Y
P15	Acute myocardial infarction	I21, I22	Incidence by person	1 year	IX	Y
P16	Heart failure	I50	Period prevalence	3 years	IX	Y
P17	Stroke	I60-I64	Incidence by person	1 year	IX	N
P18	Cerebrovascular diseases	I60-I69	Period prevalence	3 years	IX	N
P19	Pneumonia	J12-J18	Incidence by episode	1 year	X	N
P20	Asthma	J45, J46	Incidence by person	1 year	X	N
P21	Asthma	J45, J46	Period prevalence	3 years	X	N

No.	Disease or group title	ICD-10 codes	Measure	Reference period	ICD-10 Chapter	Pilot questions (*)
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	J40-J44, J47	Period prevalence	3 years	X	Y
P23	Chronic obstructive pulmonary disease (COPD)	J44	Period prevalence	3 years	X	N
P24	Alcoholic liver disease	K70	Period prevalence	3 years	XI	Y
P25	Diseases of liver other than alcoholic	K71-K77	Period prevalence	3 years	XI	Y
P26	Diseases of liver	K70-K77	Period prevalence	3 years	XI	Y
P27	Rheumatoid arthritis	M05, M06	Period prevalence	3 years	XIII	Y
P28	Arthrosis	M15-M19	Period prevalence	3 years	XIII	Y
P29	Osteoporosis	M80-M82	Period prevalence	3 years	XIII	Y
P30	Renal failure	N17-N19	Period prevalence	3 years	XIV	Y
P31	Urolithiasis	N20-N23	Incidence by person	1 year	XIV	Y
P32	Intracranial injury	S06	Incidence by episode	1 year	XIX	Y
P33	Intracranial injury	S06	Incidence by person	1 year	XIX	Y
P34	Fracture of femur	S72	Incidence by episode	1 year	XIX	Y
P35	Fracture of femur	S72	Incidence by person	1 year	XIX	Y
List B: indicators to be considered for pilot data collection						
PB36	Land transport accidents	V01-V89	Incidence by episode	1 year	XX	Y
PB37	Land transport accidents	V01-V89	Incidence by person	1 year	XX	Y
PB38	Accidental falls	W00-W19	Incidence by episode	1 year	XX	Y
PB39	Accidental falls	W00-W19	Incidence by person	1 year	XX	Y
PB40	Intentional self-harm (incl. suicidal attempt)	X60-X84	Incidence by episode	1 year	XX	Y
PB41	Intentional self-harm(incl. suicidal attempt)	X60-X84	Incidence by person	1 year	XX	Y
PB42	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by episode	1 year	XX	Y
PB43	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by person	1 year	XX	Y

(*) The column 'Pilot questions' indicates whether participating countries were asked to address specific questions (e.g. about inclusion of certain codes) - see Table 4

TABLE 2

Data sources used per country

Type of data source	BE	HR	FI	FR	HU	LT	MT	NL	PL
Primary care	Y	Y	Y		Y	Z		Y	Z
Hospital inpatients	Y	Y	Y	Y	Z	Z	Y	Y	Z
Hospital outpatients			Y		Z	Z		Z	Z
Causes of death	Y	Y	Y		Y	Y	Y	Y	Y
Disease-specific registers		Y							
Surveillance systems									
Emergency care						Z	Y		Z
Insurance	Y			Y	Y	Y			Y
Prescriptions	Z			Y	Z		Y	Y	Y
Combined hospital data								Y	
Other		Y	Y		Y			Y (3)	

Key
Y = This type of source is used for at least one indicator
Z = Not used directly but included in another source (e.g. outpatients included in combined hospital data)
(n) = If multiple sources of the same type are used, the number of sources of this type is shown

TABLE 3

Data sources used for each indicator, per country

No.	Disease or group title	BE	HR	FI	FR	HU	LT	MT	NL	PL
P1	Diabetes mellitus	IN	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR) OT	CD IN	HI PR	PC HI CD PR CH	IN PR
P2	Diabetes mellitus	IN	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN	HI CD PR	PC HI CD PR CH	IN PR
P3	Dementia (incl. Alzheimer disease)	IN	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD PR CH OT	IN PR
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD	PC HI CD CH OT	IN
P5	Schizophrenia, schizotypal and delusional disorders	IN	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD CH OT	IN
P6	Mood (affective) disorders	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR) OT	CD IN	HI CD PR	PC HI CD CH OT	IN
P7	Anxiety disorders	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI H, PR) OT	CD IN	HI PR	PC HI CD CH OT	IN
P8	Parkinson disease	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD PR CH	IN PR
P9	Multiple sclerosis	IN	PC CD HI OT	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD CH	IN
P10	Epilepsy	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD PR CH	IN
P11	Hypertensive diseases	PC	PC CD HI	PC CD HI HO OT	HI NI	CD IN (HI HO PR) OT	CD IN	HI PR	PC HI CD CH	IN PR
P12	Hypertensive diseases	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR) OT	CD IN	HI CD PR	PC HI CD CH	IN PR
P13	Ischaemic heart diseases	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN	HI CD PR	PC HI CD CH	IN PR
P14	Acute myocardial infarction	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD	HI CD CH	IN PR

No.	Disease or group title	BE	HR	FI	FR	HU	LT	MT	NL	PL
P15	Acute myocardial infarction	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD	PC HI CD CH	IN PR
P16	Heart failure	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD CH	IN
P17	Stroke	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD	PC HI CD CH	IN PR
P18	Cerebrovascular diseases	HI	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD CH	IN
P19	Pneumonia	PC	PC CD HI DR	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD	PC HI CD CH	IN
P20	Asthma	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN	HI PR	PC HI CD CH	IN
P21	Asthma	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN	HI CD PR	PC HI CD CH	IN
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR) OT	CD IN	HI PR	PC HI CD CH	IN
P23	Chronic obstructive pulmonary disease (COPD)	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR)	CD IN	HI PR	PC HI CD CH	IN
P24	Alcoholic liver disease	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	—	IN
P25	Diseases of liver other than alcoholic	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN		—	IN
P26	Diseases of liver	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	PC HI CD CH	IN
P27	Rheumatoid arthritis	PC	PC CD HI OT	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN	HI CD PR	HI CD CH	IN
P28	Arthrosis	PC	PC CD HI OT	PC CD HI HO OT	HI IN	CD IN (HI HO PR) OT	CD IN		PC HI CD CH	IN
P29	Osteoporosis	PC	PC CD HI	PC CD HI HO OT	HI IN PR	CD IN (HI HO PR) OT	CD IN		PC HI CD CH	IN
P30	Renal failure	PC	PC CD HI	PC CD HI HO OT	HI IN	CD IN (HI HO PR)	CD IN	HI CD PR	HI CD CH	IN
P31	Urolithiasis	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI	PC HI CD CH	IN
P32	Intracranial injury	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN		PC HI CD CH	IN
P33	Intracranial injury	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD EC	PC HI CD CH	IN
P34	Fracture of femur	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN		HI CD CH	IN
P35	Fracture of femur	HI	PC CD HI	PC CD HI HO OT	HI	CD IN (HI HO PR)	CD IN	HI CD EC	HI CD CH	IN

No.	Disease or group title	BE	HR	FI	FR	HU	LT	MT	NL	PL
PB36	Land transport accidents	HI	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN		—	CD IN
PB37	Land transport accidents	HI	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD EC	—	CD IN
PB38	Accidental falls	HI	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN		—	CD IN
PB39	Accidental falls	HI	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD EC	—	CD IN
PB40	Intentional self harm (incl. suicidal attempt)	CD	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD EC	—	CD IN
PB41	Intentional self harm (incl. suicidal attempt)	CD	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD EC	—	CD IN
PB42	Complications of medical and surgical care	CD	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD	—	CD IN
PB43	Complications of medical and surgical care	CD	PC CD HI	CD HI	HI	CD IN (HI HO PR)	CD IN	HI CD	—	CD IN

Key
PC = Primary care
HI = Hospital inpatients
HO = Hospital outpatients
CD = Causes of death
DR = Disease-specific registers
EC = Emergency care
IN = Insurance
PR = Prescriptions
CH = Combined hospital data
OT = Other

TABLE 4

Most important data source for each indicator, per country

No.	Disease or group title	Measure	BE (1)			HR			FI			FR			HU (2)			LT (2)			MT			NL			PL		
			Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only
List A: indicators to be included for pilot data collection																													
P1	Diabetes mellitus	Incidence by person	IN	—	—	PC	98	98	PC	—	—	PR	—	—	PR	58	36	IN	83	83	PR	89	84	PC	68	37	IN	72	28
P2	Diabetes mellitus	Period prevalence	IN	—	—	PC	99	91	PC	—	—	IN	—	—	PR	83	20	IN	96	63	PR	93	93	PC	79	13	IN	87	12
P3	Dementia (incl. Alzheimer disease)	Period prevalence	IN	—	—	PC	95	86	HI	—	—	IN	—	—	HO	75	34	IN	88	67	PR	103	84	PC	29	12	IN	95	74
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	Period prevalence	PC	—	—	PC	91	67	HO	—	—	HI	88	80	HO	85	50	EC	51	34	HI	101	100	PC	67	48	IN	100	100
P5	Schizophrenia, schizotypal and delusional disorders	Period prevalence	IN	—	—	PC	98	76	HO	—	—	IN	—	—	PR	84	17	IN	86	49	PR	58	53	PC	61	33	IN	100	100
P6	Mood (affective) disorders	Period prevalence	PC	—	—	PC	100	95	HO	—	—	PR	—	—	PR	70	46	IN	94	81	PR	100	96	PC	72	59	IN	100	100
P7	Anxiety disorders	Period prevalence	PC	—	—	PC	100	99	HO	—	—	HI	71	69	PR	72	54	IN	86	81	PR	60	59	PC	76	70	IN	100	100
P8	Parkinson disease	Period prevalence	PC	—	—	PC	98	90	HO	—	—	IN	—	—	PR	85	17	IN	93	68	PR	118	95	PC	67	11	IN	61	13
P9	Multiple sclerosis	Period prevalence	IN	—	—	PC	97	43	HO	—	—	IN	—	—	PR	95	56	IN	49	92	PR	115	91	PC	70	12	IN	100	100
P10	Epilepsy	Period prevalence	PC	—	—	PC	98	88	HO	—	—	IN	62	49	PR	75	54	IN	77	49	PR	113	94	PC	61	21	IN	100	100

No.	Disease or group title	Measure	BE [1]			HR			FI			FR			HU [2]			LT [2]			MT			NL			PL		
			Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only
P11	Hypertensive diseases	Incidence by person	PC	—	—	PC	99	99	PC	—	—	PR	91	91	HO	54	35	IN	81	81	PR	98	98	PC	69	64	IN	95	86
P12	Hypertensive diseases	Period prevalence	PC	—	—	PC	100	99	PC	—	—	PR	—	—	PR	70	27	IN	94	70	PR	104	99	PC	78	65	IN	97	74
P13	Ischaemic heart diseases	Period prevalence	PC	—	—	PC	93	76	HI	—	—	IN	79	53	HO	72	34	IN	79	51	PR	154	93	PC	66	28	IN	93	71
P14	Acute myocardial infarction	Incidence by episode	HI	—	—	HI	45	45	HI	—	—	HI	100	100	HI	80	36	HI	92	92	HI	74	67	HI	71	17	IN	95	95
P15	Acute myocardial infarction	Incidence by person	HI	—	—	HI	47	47	HI	—	—	HI	100	100	PR	79	34	HI	92	92	HI	72	65	HI	72	17	IN	95	95
P16	Heart failure	Period prevalence	PC	—	—	PC	86	75	HO	—	—	HI	65	56	PR	60	38	IN	69	47	PR	124	96	PC	53	25	IN	100	100
P17	Stroke	Incidence by person	HI	—	—	HI	61	61	HI	—	—	HI	100	100	HO	88	57	HI	95	95	HI	101	83	PC	58	28	IN	97	97
P18	Cerebrovascular diseases	Period prevalence	HI	—	—	PC	87	63	HI	—	—	IN	65	46	PR	84	48	IN	75	51	PR	102	78	PC	64	37	IN	100	100
P19	Pneumonia	Incidence by episode	PC	—	—	PC	77	77	HO	—	—	HI	100	100	HO	64	32	IN	54	54	HI	101	93	PC	73	63	IN	100	100
P20	Asthma	Incidence by person	PC	—	—	PC	97	97	HO	—	—	PR	93	93	PR	50	28	IN	90	90	PR	103	90	PC	75	70	IN	100	100
P21	Asthma	Period prevalence	PC	—	—	PC	100	97	HO	—	—	HI	59	52	HO	64	29	IN	116	98	PR	113	96	PC	78	66	IN	100	100
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	Period prevalence	PC	—	—	PC	99	95	HO	—	—	PR	—	—	PR	45	12	IN	86	75	PR	117	84	PC	73	46	IN	100	100
P23	Chronic obstructive pulmonary disease (COPD)	Period prevalence	PC	—	—	PC	98	93	HO	—	—	PR	—	—	PR	69	44	IN	79	79	PR	124	90	PC	73	45	IN	100	100

No.	Disease or group title	Measure	BE (1)			HR			FI			FR			HU (2)			LT (2)			MT			NL			PL		
			Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only
P24	Alcoholic liver disease	Period prevalence	PC	—	—	PC	89	77	HI	—	—	HI	68	51	HI	65	29	HI	73	60	PR	65	47	X	X	X	IN	100	100
P25	Diseases of liver other than alcoholic	Period prevalence	PC	—	—	PC	98	94	HO	—	—	HI	75	68	HO	77	54	IN	84	77	X	X	X	X	X	X	IN	100	100
P26	Diseases of liver	Period prevalence	PC	—	—	PC	97	92	HO	—	—	HI	73	60	HO	76	50	IN	80	72	PR	88	62	PC	41	29	IN	100	100
P27	Rheumatoid arthritis	Period prevalence	PC	—	—	PC	99	89	HO	—	—	IN	91	74	PR	91	65	IN	96	79	PR	100	99	CH	92	73	IN	100	100
P28	Arthrosis	Period prevalence	PC	—	—	PC	99	93	PC	—	—	HI	94	92	HO	91	44	IN	94	82	X	X	X	PC	62	37	IN	100	100
P29	Osteoporosis	Period prevalence	PC	—	—	PC	100	98	HO	—	—	PR	—	—	PR	78	41	IN	94	89	X	X	X	PC	55	43	IN	100	100
P30	Renal failure	Period prevalence	PC	—	—	PC	87	67	HI	—	—	HI	78	64	HI	68	32	HI	59	38	PR	89	82	HI	63	23	IN	100	100
P31	Urolithiasis	Incidence by person	HI	—	—	PC	97	97	HO	—	—	HI	100	100	HO	92	58	IN	69	69	HI	100	100	PC	74	57	IN	100	100
P32	Intracranial injury	Incidence by episode	HI	—	—	HI	53	53	HO	—	—	HI	100	100	HO	79	44	EC	52	52	X	X	X	PC	45	39	IN	100	100
P33	Intracranial injury	Incidence by person	HI	—	—	HI	58	58	HO	—	—	HI	100	100	HO	79	42	EC	52	52	EC	97	96	PC	44	37	IN	100	100
P34	Fracture of femur	Incidence by episode	HI	—	—	HI	66	66	HO	—	—	HI	100	100	HO	99	81	EC	41	41	X	X	X	CH	84	18	IN	100	100
P35	Fracture of femur	Incidence by person	HI	—	—	HI	74	74	HI	—	—	HI	100	100	HO	100	95	EC	44	44	HI	109	96	CH	84	15	IN	100	100

No.	Disease or group title	Measure	BE (1)			HR			FI			FR			HU (2)			LT (2)			MT			NL			PL		
			Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only	Main	% of all	% only
List B: indicators to be considered for pilot data collection																													
PB36	Land transport accidents	Incidence by episode	HI	—	—	PC	65	65	HO	—	—	HI	100	100	HO	89	73	EC	75	75	X	X	X	X	X	X	IN	97	97
PB37	Land transport accidents	Incidence by person	HI	—	—	PC	63	63	HO	—	—	HI	100	100	HO	89	72	EC	75	75	EC	99	91	X	X	X	IN	97	97
PB38	Accidental falls	Incidence by episode	HI	—	—	PC	91	91	HO	—	—	HI	100	100	HO	94	82	EC	84	84	X	X	X	X	X	X	IN	99	99
PB39	Accidental falls	Incidence by person	HI		—	PC	91	91	PC	—	—	HI	100	100	HO	94	81	EC	84	84	EC	96	85	X	X	X	IN	99	99
PB40	Intentional self harm (incl. suicidal attempt)	Incidence by episode	CD	—	—	PC	42	42	CD	—	—	HI	100	100	HI	74	61	EC	43	43	HI	77	54	X	X	X	IN	80	80
PB41	Intentional self harm (incl. suicidal attempt)	Incidence by person	CD	—	—	CD	43	43	CD	—	—	HI	100	100	HI	74	59	EC	42	42	HI	76	53	X	X	X	IN	79	79
PB42	Complications of medical and surgical care	Incidence by episode	CD	—	—	PC	85	85	PC	—	—	HI	100	100	HI	38	34	HI	87	87	HI	100	99	X	X	X	IN	99	99
PB43	Complications of medical and surgical care	Incidence by person	CD	—	—	PC	98	98	PC	—	—	HI	100	100	HI	35	35	HI	86	86	HI	100	99	X	X	X	IN	99	99

Key
Main =The data source providing the largest number of cases included in this estimate
% of all =The percentage of all cases included in the estimate that were counted in this data source (not exclusively)
% only = The percentage of all cases included in the estimate that were counted in this data source only, not any other source

PC = Primary care
HI = Hospital inpatients
HO = Hospital outpatients
CD = Causes of death
DR =Disease-specific registers
SU = Surveillance systems
EC = Emergency care
IN = Insurance
PR = Prescriptions
CH = Combined hospital data
OT = Other
' X ' =Indicator not reported
' - ' = Could not be calculated from the metadata due to weighting procedures or other calculation issue

Notes:
(1) For BE, all indicators were based on one source only
(2) For HU and LT, all the sources listed are collected as part of the national health insurance data

TABLE 5

Questions identified for consideration in the pilot studies, per indicator

No.	Disease or group title	ICD-10 codes	Measure	Question identified for pilot reports
List A: indicators to be included for pilot data collection				
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	F10	Period prevalence	The completeness of reporting and comparability between the countries are both doubtful, therefore the pilot studies should review the quality of their data carefully. Depending on the pilot results, this indicator may not be feasible as a regular indicator.
P7	Anxiety disorders	F40-F41	Period prevalence	The feasibility including differences in severity measured and identification of the appropriate medications has to be reviewed.
P14	Acute myocardial infarction	I21, I22	Incidence by episode	It is intended to include in the shortlist only one of the indicators for Acute myocardial infarction: either incidence by episode or incidence by person. A decision will be made after the pilot studies based on the feasibility and comparability.
P15	Acute myocardial infarction	I21, I22	Incidence by person	As P14 above.
P16	Heart failure	I50	Period prevalence	Because of expected variations in the coding and recording practice, the feasibility and comparability is uncertain. A decision will be made after the pilot studies.
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	J40-J44, J47	Period prevalence	The inclusion of the code J47 bronchiectasis is uncertain and observations on its validity are invited.
P24	Alcoholic liver disease	K70	Period prevalence	There are different views among the countries on the feasibility of separating alcoholic from other liver diseases. A decision will be made after the pilot studies.
P25	Diseases of liver other than alcoholic	K71-K77	Period prevalence	As P24 above.
P26	Diseases of liver	K70-K77	Period prevalence	As P24 above.
P27	Rheumatoid arthritis	M05, M06	Period prevalence	The inclusion of the code M06.4 is uncertain and observations from the pilot studies are invited.
P28	Arthrosis	M15-M19	Period prevalence	There is some doubt on the comparability of data obtained from primary care (using ICPC codes) and data using the ICD codes. Observations from the pilot studies on the feasibility and comparability are invited.
P29	Osteoporosis	M80-M82	Period prevalence	There is known to be wide variation in the reported diagnosis of osteoporosis. The pilot studies are invited to comment on the adequacy of diagnosis and on influencing factors such as screening.

No.	Disease or group title	ICD-10 codes	Measure	Question identified for pilot reports
P30	Renal failure	N17-N19	Period prevalence	Because of the variety of diagnostic codes used in the countries, observations are invited from the pilot studies on the most feasible and comparable definition.
P31	Urolithiasis	N20-N23	Incidence by person	There were doubts about the feasibility of this indicator, therefore observations from the pilot studies are invited.
P32	Intracranial injury	S06	Incidence by episode	Several countries proposed variations of the coding, it was also suggested that it is feasible only to record the more severe cases requiring hospitalisation. Observations on these issues from the pilot studies are invited. The relative feasibility and value of incidence by person and incidence by episode is to be considered.
P33	Intracranial injury	S06	Incidence by person	As P32 above.
P34	Fracture of femur	S72	Incidence by episode	Observations on the definition, especially the inclusion of surgical procedures, from the pilot studies are invited. The relative feasibility and value of incidence by person and incidence by episode is to be considered.
P35	Fracture of femur	S72	Incidence by person	As P34 above.
List B: indicators to be considered for pilot data collection				
PB36	Land transport accidents	V01-V89	Incidence by episode	This indicator is optional in the pilot studies. Observations on the feasibility, validity and comparability are invited. The relative feasibility and value of incidence by person and incidence by episode is also to be considered.
PB37	Land transport accidents	V01-V89	Incidence by person	As PB36 above.
PB38	Accidental falls	W00-W19	Incidence by episode	As PB36 above.
PB39	Accidental falls	W00-W19	Incidence by person	As PB36 above.
PB40	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by episode	As PB36 above.
PB41	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by person	As PB36 above.
PB42	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by episode	As PB36 above.
PB43	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by person	As PB36 above.

Note: Only those indicators where participating countries were asked to address specific questions are included in this table. The original source of the list is the Methodological Guidelines of the project, Annex A.

TABLE 6

Types of difference from the guidelines for each indicator, per country

No.	Disease or group title	Measure	BE	HR (1)	FI	FR	HU	LT	MT	NL (2)	PL
List A: indicators to be included for pilot data collection											
P1	Diabetes mellitus	Incidence by person	B	D	—	—	—	—	—	B	B
P2	Diabetes mellitus	Period prevalence	B	D	—	—	—	—	—	B	B
P3	Dementia (incl. Alzheimer disease)	Period prevalence	B	A D	—	—	—	—	—	A B	B
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	Period prevalence	—	D	—	—	—	—	—	B	—
P5	Schizophrenia, schizotypal and delusional disorders	Period prevalence	B	D	—	—	—	—	—	B	—
P6	Mood (affective) disorders	Period prevalence	—	D	—	—	—	—	—	B	—
P7	Anxiety disorders	Period prevalence	—	D	—	—	—	—	—	B	—
P8	Parkinson disease	Period prevalence	—	D	—	—	—	—	—	B	B
P9	Multiple sclerosis	Period prevalence	B	D	—	—	—	—	—	B	—
P10	Epilepsy	Period prevalence	—	D	—	—	—	—	—	B	—
P11	Hypertensive diseases	Incidence by person	—	D	—	—	—	—	—	B	B
P12	Hypertensive diseases	Period prevalence	—	D	—	—	—	—	—	B	B
P13	Ischaemic heart diseases	Period prevalence	—	D	—	—	—	—	—	B	B
P14	Acute myocardial infarction	Incidence by episode	B	D	—	—	—	—	—	B	C
P15	Acute myocardial infarction	Incidence by person	B	D	—	—	—	—	—	B	C
P16	Heart failure	Period prevalence	—	D	—	—	—	—	—	B	—
P17	Stroke	Incidence by person	—	D	—	—	—	—	—	B	—
P18	Cerebrovascular diseases	Period prevalence	—	D	—	—	—	—	—	B	—
P19	Pneumonia	Incidence by episode	—	D	—	—	—	—	—	B	—

No.	Disease or group title	Measure	BE	HR (1)	FI	FR	HU	LT	MT	NL (2)	PL
P20	Asthma	Incidence by person	—	D	—	—	—	—	—	B	—
P21	Asthma	Period prevalence	—	D	—	—	—	—	—	B	—
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	Period prevalence	—	D	—	—	—	—	—	B	—
P23	Chronic obstructive pulmonary disease (COPD)	Period prevalence	B	D	—	—	—	—	—	B	—
P24	Alcoholic liver disease	Period prevalence	B	D	B	—	—	—	—	X	—
P25	Diseases of liver other than alcoholic	Period prevalence	B	D	B	—	—	—	—	X	—
P26	Diseases of liver	Period prevalence	—	D	—	—	—	—	—	B	—
P27	Rheumatoid arthritis	Period prevalence	—	D	—	—	—	—	—	B	—
P28	Arthrosis	Period prevalence	—	D	—	—	—	—	—	B	—
P29	Osteoporosis	Period prevalence	—	D	—	—	—	—	—	B	—
P30	Renal failure	Period prevalence	B	D	—	—	—	—	—	B	—
P31	Urolithiasis	Incidence by person	—	D	—	—	—	—	—	B	—
P32	Intracranial injury	Incidence by episode	—	D	—	—	—	—	—	—	C
P33	Intracranial injury	Incidence by person	—	D	—	—	—	—	—	—	C
P34	Fracture of femur	Incidence by episode	—	D	—	—	—	—	—	B	C
P35	Fracture of femur	Incidence by person	—	D	—	—	—	—	—	B	C
List B: indicators to be considered for pilot data collection											
PB36	Land transport accidents	Incidence by episode	—	D	—	—	—	—	—	X	C
PB37	Land transport accidents	Incidence by person	—	D	—	—	—	—	—	X	C
PB38	Accidental falls	Incidence by episode	—	D	—	—	—	—	—	X	C
PB39	Accidental falls	Incidence by person	—	D	—	—	—	—	—	X	C
PB40	Intentional self harm (incl. suicidal attempt)	Incidence by episode	—	D	—	—	—	—	—	X	C

No.	Disease or group title	Measure	BE	HR (1)	FI	FR	HU	LT	MT	NL (2)	PL
PB41	Intentional self harm (incl. suicidal attempt)	Incidence by person	—	D		—	—	—	—	X	C
PB42	Complications of medical and surgical care	Incidence by episode	—	D	—	—	—	—	—	X	C
PB43	Complications of medical and surgical care	Incidence by person	—	D	—	—	—	—	—	X	C

Key
-' = No difference
X = Indicator not reported
A = Difference in the ICD-10 codes applied due to national practice
B = Difference in implementation of the ICD-10 definition due to use of other classifications
C = Special selection criteria of service or episode types applied
D = Difference in time periods

Notes
(1) For HR, the reference year was 2017, and for three-year prevalence the period was 2015-17
(2) For NL, all estimates are based on an 8% representative sample in primary care

TABLE 7

Summary of recommendations on the indicators

Green Suitable for use in a future data collection, with no more than minor reservations.

Amber Suitable for use with reservations, requiring care in interpretation or further research to understand differences.

Red Not suitable for use, or another indicator is preferred.

No.	Disease or group title	ICD-10 codes	Measure	Recommendation
		List A: indicators to be included for pilot data collection		
P1	Diabetes mellitus	E10-E14	Incidence by person	Amber: Include without change; further research on comparability needed
P2	Diabetes mellitus	E10-E14	Period prevalence	Amber: Include without change; further research on comparability needed
P3	Dementia (incl. Alzheimer disease)	F00-F03, G30	Period prevalence	Amber: Include with addition of ICD-10 code F05.1; further research on comparability needed
P4	Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	F10	Period prevalence	Amber: Include with addition of ICD-10 code G312; further research on comparability needed
P5	Schizophrenia, schizotypal and delusional disorders	F20-F29	Period prevalence	Green: Include with addition of ICPC code P98
P6	Mood (affective) disorders	F30-F39	Period prevalence	Green: Include despite minor inconsistencies with ICPC codes
P7	Anxiety disorders	F40-F41	Period prevalence	Red: Do not include; too much uncertainty about comparability
P8	Parkinson disease	G20	Period prevalence	Amber: Include, but research needed on diagnosis in some countries
P9	Multiple sclerosis	G35	Period prevalence	Green: Include without change
P10	Epilepsy	G40-G41	Period prevalence	Green: Include with refinement of the ATC codes
P11	Hypertensive diseases	I10-I13, I15	Incidence by person	Green: Include without change
P12	Hypertensive diseases	I10-I13, I15	Period prevalence	Green: Include without change

No.	Disease or group title	ICD-10 codes	Measure	Recommendation
P13	Ischaemic heart diseases	I20-I25	Period prevalence	Amber: Include, but research on differences needed
P14	Acute myocardial infarction	I21, I22	Incidence by episode	Red: Do not include; indicator P15 is preferred
P15	Acute myocardial infarction	I21, I22	Incidence by person	Green: Include; use inpatient and COD data only
P16	Heart failure	I50	Period prevalence	Green: Include; use underlying COD and possibly main inpatient diagnosis
P17	Stroke	I60-I64	Incidence by person	Green: Include; use inpatient and COD data only
P18	Cerebrovascular diseases	I60-I69	Period prevalence	Amber: Include; but research needed on diagnosis and care organisation
P19	Pneumonia	J12-J18	Incidence by episode	Amber: Include, but research needed on comparability
P20	Asthma	J45, J46	Incidence by person	Red: Do not include; indicator P21 is preferred
P21	Asthma	J45, J46	Period prevalence	Green: Include without change; primary care is main source
P22	Chronic lower respiratory diseases other than asthma (incl. COPD)	J40-J44, J47	Period prevalence	Amber: Include, but research needed on comparability
P23	Chronic obstructive pulmonary disease (COPD)	J44	Period prevalence	Green: Include without change
P24	Alcoholic liver disease	K70	Period prevalence	Red: Do not include; cannot be recorded in some countries (or optional)
P25	Diseases of liver other than alcoholic	K71-K77	Period prevalence	Red: Do not include; cannot be recorded in some countries (or optional)
P26	Diseases of liver	K70-K77	Period prevalence	Green: Include without change
P27	Rheumatoid arthritis	M05, M06	Period prevalence	Green: Include without change
P28	Arthrosis	M15-M19	Period prevalence	Amber: Include, but research needed on comparability
P29	Osteoporosis	M80-M82	Period prevalence	Red: Do not include; too much uncertainty about comparability

No.	Disease or group title	ICD-10 codes	Measure	Recommendation
P30	Renal failure	N17-N19	Period prevalence	Red: Do not include; too much uncertainty about comparability
P31	Urolithiasis	N20-N23	Incidence by person	Amber: Include, but research needed on comparability
P32	Intracranial injury	S06	Incidence by episode	Red: Do not include; indicator P33 is preferred
P33	Intracranial injury	S06	Incidence by person	Amber: Include, but restrict to COD, inpatient and emergency care
P34	Fracture of femur	S72	Incidence by episode	Red: Do not include; indicator P35 is preferred
P35	Fracture of femur	S72	Incidence by person	Amber: Include, but restrict to COD, inpatient and emergency care
		List B: indicators to be considered for pilot data collection		
PB36	Land transport accidents	V01-V89	Incidence by episode	Red: Do not include; unlikely to be meaningful
PB37	Land transport accidents	V01-V89	Incidence by person	Red: Do not include; unlikely to be meaningful
PB38	Accidental falls	W00-W19	Incidence by episode	Red: Do not include; unlikely to be meaningful
PB39	Accidental falls	W00-W19	Incidence by person	Red: Do not include; unlikely to be meaningful
PB40	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by episode	Red: Do not include; unlikely to be meaningful
PB41	Intentional self harm (incl. suicidal attempt)	X60-X84	Incidence by person	Red: Do not include; unlikely to be meaningful
PB42	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by episode	Red: Do not include; unlikely to be meaningful
PB43	Complications of medical and surgical care	Y40-Y66, Y69-Y84	Incidence by person	Red: Do not include; unlikely to be meaningful

Annex B. Detailed examples on selected methodological issues

1.1. Use of the European Health Interview Survey (EHIS) in HU to identify proportions of individuals with certain conditions using private healthcare

The 2019 EHIS questionnaire included the following diseases where estimates could be made for the private sector:

- Asthma
- Chronic lower respiratory diseases other than asthma (incl. COPD)
- Ischaemic heart diseases
- Hypertensive diseases
- Rheumatoid arthritis
- Arthrosis
- Diabetes mellitus
- Dementia (incl. Alzheimer's disease)
- Depression
- Other affective disorders
- Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)
- Anxiety disorders
- Schizophrenia, schizotypal and delusional disorders
- Osteoporosis

In the 2019 EHIS in HU, question CD1 which concerns 'chronic diseases in the past 12 months' was complemented by additional questions on diagnosis by a medical doctor and medicine use, which are usually parts of the national EHIS questionnaire, and a newly added question concerning the type of health services used. The wording of the questions is shown below.

When using the statistics from EHIS to inform the estimates, the following principles and assumptions were applied:

a. Use of private services depends on the disease, the sex and the age of the respondent. In the cases of dementia (incl. Alzheimer's disease), mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) and schizophrenia, schizotypal and delusional disorders, the number of cases in EHIS are very low, so no private service will be calculated in any age-group-sex data cell.
b. Regarding the 15+ population, the proportion of public and private services in EHIS –except outliers- represents the real situation in the population. In younger age-groups where the level of use of health services is usually low, outliers have to be corrected. In other age-groups parameters calculated from EHIS can be used in the estimation.
c. In the case of the 0-14 population, all diseases are assumed to be included in health insurance database because of the strict and regular health control of this age-group.
d. Patients with incidence data in the health insurance database who used public or private services exclusively according to EHIS, had their first contact at the health care provider indicated in EHIS.
e. If patients attended both public and private services, it is not known where they got their diagnosis first, but they appear in the health insurance database according to the definition of incidence for that database.
f. It is assumed that not only the distribution of prevalence by type of service provider but that of incidence by person is also the same as the distribution in EHIS. In the calculation of estimates, only proportions are used from EHIS regardless of the number of cases.
g. EHIS data are not appropriate for making estimates of incidence by episode.
h. Health insurance data will be complemented by estimation based on EHIS data if there are cases in the age-group-sex cell who used public and also who used private services. Otherwise, the insurance data remain unchanged.

Wording of the EHIS questions

1. During the past 12 MONTHS, have you had any of the following diseases or conditions? (CD1)

If yes:

2. Was this disease/condition diagnosed by a medical doctor?

If yes:3. For this disease/condition have you taken any medicine on medical advice?

Please include each medicine or dietary supplement (herbal medicine, vitamin) taken on medical advice, regardless of prescription drug or not.

4. During the past 12 MONTHS, what type of health services have you used in connection with this disease?

1. publicly financed services exclusively
2. privately financed services exclusively
3. publicly and privately financed services both
4. none

1.2. Coverage of the health insurance data in PL: percentage of patients without a personal identifier

In the health insurance system of PL, all persons included can be identified by a unique identifier (the PESEL number). Population coverage of the health insurance system is very wide; however, the data show small numbers of cases where no identifier was recorded. The individuals concerned typically include individuals who could not be identified at the time of care (for example they were unconscious), but also homeless people and some other groups. It is notable that the highest percentages without an identifier occur for such indicators as PB36-PB37 Land transport accidents and PB40-PB41 Intentional self-harm (incl. suicidal attempt).

Health condition	Percent
Period prevalence	
Diabetes mellitus	0.4
Dementia (incl. Alzheimer's disease)	0.2
Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence)	2.3
Schizophrenia. schizotypal and delusional disorders	1.0
Mood (affective) disorders	0.3
Anxiety disorders	0.2
Parkinson's disease	0.1
Multiple sclerosis	0.3
Epilepsy	4.4
Hypertensive diseases	0.4
Ischaemic heart diseases	0.3
Heart failure	0.3
Cerebrovascular diseases	0.6
Asthma	0.3
Chronic lower respiratory diseases other than asthma (incl. COPD)	0.3
Chronic obstructive pulmonary disease (COPD)	0.3
Alcoholic liver disease	0.7
Diseases of liver other than alcoholic	0.3
Diseases of liver	0.3
Rheumatoid arthritis	0.1
Arthrosis	0.1
Osteoporosis	0.0
Renal failure	0.4
Incidence	
Diabetes mellitus	0.8
Hypertensive diseases	1.1
Acute myocardial infarction	1.1
Stroke	1.5
Pneumonia	0.5
Asthma	0.6
Urolithiasis	0.7
Intracranial injury	2.0
Fracture of femur	1.3
Land transport accidents	8.5
Accidental falls	2.0
Intentional self-harm (incl. suicidal attempt)	7.8
Complications of medical and surgical care	1.8

1.3. Indicator P1 Diabetes: relevant codes in the DTC-SSC classification used in NL

The DTC-SSC classification used in NL specialised care performs a DRG-type function for the organisation of payments. Codes are separately formulated for each medical specialty. The codes indicate the medical specialty followed by the specific condition or treatment. As can be seen in this example, a wide range of codes can be needed to identify all the diagnoses, treatments and related clinical designations for one indicator. In some cases, the same identical clinical entity is represented by more than one code, because of the separation between specialties.

Code	Specialty	Health condition or treatment
0303/0522	Surgery	Islet transplantation
0303/0531	Surgery	Kidney and pancreas transplantation
0303/0553	Surgery	Islet transplantation process receiver
0303/0559	Surgery	Kidney and pancreatic transplantation recipient
0303/0560	Surgery	Pancreas transplantation trajectory receiver
0303/0562	Surgery	Liver and pancreatic transplantation recipient
0303/0563	Surgery	Liver, pancreas and intestinal transplantation recipient
0305/2065	Orthopaedics	Diabetic foot
0313/0072	Internal Medicine	Islet transplant pathway receiver
0313/0078	Internal Medicine	Kidney and pancreatic transplant recipient
0313/0079	Internal Medicine	Pancreas transplantation trajectory receiver
0313/0082	Internal Medicine	Liver and pancreatic transplant recipient
0313/0083	Internal Medicine	Liver, pancreatic and intestinal transplant recipient
0313/0221	Internal Medicine	Diabetes mellitus without secondary complications
0313/0222	Internal Medicine	Diabetes mellitus with secondary complications
0313/0223	Internal Medicine	Diabetes mellitus chronic pump therapy
0313/0345	Internal Medicine	Kidney and pancreas transplantation <= 365 days
0313/0347	Internal Medicine	Kidney and pancreas transplant> 365 days
0316/7104	Pediatrics	Diabetes mellitus
0316/7113	Pediatrics	Diabetes mellitus with chronic pump therapy
0316/7114	Pediatrics	Diabetes mellitus other
0316/7903	Pediatrics	Islet transplant pathway recipient
0316/7909	Pediatrics	Kidney and pancreatic transplant recipient
0316/7910	Pediatrics	Pancreas transplantation recipient
0316/7923	Pediatrics	Liver and pancreatic transplant recipient
0316/7924	Pediatrics	Liver, pancreatic and intestinal transplantation recipient
0318/0767	Gastroenterology and liver disorders	Liver and pancreatic transplant recipient
0318/0768	Gastroenterology and liver disorders	Liver, pancreatic and intestinal transplant recipient
0318/0902	Gastroenterology and liver disorders	Diabetes mellitus
0335/0222	Clinical Geriatrics	Diabetes Mellitus
0362/0400	Radiology	Islet transplant pathway receiver
0301/0754	Ophthalmology	NPDRP

Code	Specialty	Health condition or treatment
0301/0755	Ophthalmology	Preprolif. DRP
0301/0757	Ophthalmology	PDRP
0301/0759	Ophthalmology	Other pathology DRP
0303/0432	Surgery	Diabetic foot (diabetes NOS)
0303/0521	Surgery	Pancreas transplantation
0303/0522	Surgery	Islet transplantation
0303/0531	Surgery	Kidney and pancreas transplantation
0303/0553	Surgery	Islet transplantation process receiver
0303/0559	Surgery	Kidney and pancreatic transplantation recipient
0303/0560	Surgery	Pancreas transplantation trajectory receiver
0303/0562	Surgery	Liver and pancreatic transplantation recipient
0303/0563	Surgery	Liver, pancreas and intestinal transplantation recipient
0305/2065	Orthopaedics	Diabetic foot
0313/0072	Internal Medicine	Islet transplant pathway receiver

1.4. Differences in period prevalence of diabetes mellitus (E10-E14) in LT according to the length of reference period

LT calculated the period prevalence of diabetes mellitus according to different reference periods up to 6 years (2016 only up to 2011-16). The table shows, for each reference period, the number of persons identified, and the percentage increase compared to 1 year prevalence, compared to 3 year prevalence, and relative to the previously listed reference period.

Reference period	Years included	Persons	Increase (%) compared to:		
			One year	Three years	Previous
One year	2016	104,910	—	—	—
Two years	2015-16	123,784	18	—	18
Three years	2014-16	132,543	26.3	—	7.1
Four years	2013-16	137,804	31.4	4	4
Five years	2012-16	141,481	34.9	6.7	2.7
Six years	2011-16	143,979	37.2	8.6	1.8

Annex C. References

Bakker BFM, van Rooijen J, van Toor L (2014). 'The System of Social Statistical Datasets of Statistics Netherlands: An Integral Approach to the Production of Register-based Social Statistics'. *Stat J IAOS,* 30;4:411–424. https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji00803

Beghi E, Giussani G, Nichols E, et al (GBD 2016 Collaborators) (2019). Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology,* 18;4: 357-375. https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(18)30454-X

Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N (2017). The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf,* 26(8):954-962. https://onlinelibrary.wiley.com/doi/abs/10.1002/pds.4233

Boudemaghe T, Belhadj I (2017). Data Resource Profile: The French National Uniform Hospital Discharge Data Set Database (PMSI). *Int J Epidemiol,* 46, 2: 392–392d. https://doi.org/10.1093/ije/dyw359

Cox DW (2014) Gender differences in professional consultation for a mental health concern: A Canadian population study. *Canadian Psychology/Psychologie Canadienne,* 55(2), 68-74. https://doi:10.1037/a0036296

Gissler M, Shelley J (2002). Quality of data on subsequent events in a routine Medical Birth Register. *Med Inform Internet Med*, 27: 33-37. https://www.tandfonline.com/doi/abs/10.1080/14639230110119234

Mackenzie CS, Gekoski WL, Knox VJ (2006) Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. *Aging Ment Health,* 10(6):574-82. https://doi:10.1080/13607860600641200

Marshall DC, Goodson RJ, Xu Y, et al, (2018). Trends in mortality from pneumonia in the Europe union: a temporal analysis of the European detailed mortality database between 2001 and 2014. *Respir Res*, 19; 81. https://respiratory-research.biomedcentral.com/articles/10.1186/s12931-018-0781-4

Nieuwenbroek N, Boonstra HJ (2002). Bascula 4.0 for weighting sample survey data with estimation of variances. *Surv Stat,* 46:6-11. http://isi-iass.org/home/wp-content/uploads/N46.pdf

Nielen MMJ, Spronk I, Davids R, Korevaar JC, Poos R, Hoeymans N, Opstelten W, van der Sande MAB, Biermans MCJ, Schellevis FG, Verheij RA (2019). Estimating Morbidity Rates Based on Routine Electronic Health Records in Primary Care: Observational Study. *JMIR Med Inform,* 26;7(3):e11929. https://pubmed.ncbi.nlm.nih.gov/31350839

Nielen MM, Ursum J, Schellevis FG, Korevaar JC (2013) The validity of the diagnosis of inflammatory arthritis in a large population-based primary care database. *BMC Family Practice,* 14, 79. https://doi.org/10.1186/1471-2296-14-79

Okkes IM, Oskam SK, Lamberts H (2005). 'ICPC in the Amsterdam Transition Project. CD-ROM.' Amsterdam: Academic Medical Center/University of Amsterdam, Department of Family Medicine. https://www.famenet.nl/about/icpc-in-the-transition-project/

Piccinelli M, Wilkinson G (2000) Gender differences in depression: Critical review. *Br J Psych* ,177(6), 486-492. https://doi:10.1192/bjp.177.6.486

Puntaric D, Stasevic I, Ropac D, et al (2015). Some indicators of health care status in Croatia. *Acta Med Croatica*, 69, 3-14. https://hrcak.srce.hr/file/215881

Romero V, Akpinar H, Assimos DG, (2010). Kidney Stones: A Global Picture of Prevalence, Incidence, and Associated Risk Factors. *Rev Urol*, 12(2-3): e86-e96. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2931286/

Simons G, Belcher J, Morton C, Kumar K, Falahee M, Mallen CD, Stack RJ, Raza K (2017) Symptom Recognition and Perceived Urgency of Help-Seeking for Rheumatoid Arthritis

and Other Diseases in the General Public: A Mixed Method Approach. *Arthritis Care Res,* 69,5:633-641. https://doi.org/10.1002/acr.22979

Sund R (2012). Quality of the Finnish Hospital Discharge Register: a systematic review. *Scand J Public Health*, 40: 505-15. https://journals.sagepub.com/doi/10.1177/1403494812456637

Wood SN (2017). Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC, Boca Raton. https://www.taylorfrancis.com/books/generalized-additive-models-simon-wood/10.1201/9781315370279

Annex D. Recommendations on the future shortlist: indicator definitions

Title	MB01 Diabetes mellitus – incidence
Measure	**Incidence by person** based on healthcare contact. The number of individuals having a first diagnosis of diabetes mellitus in the single index year.
Classification	ICD-10: E10-E14 ICPC-1: T90 ICPC-2: T89-T90 ATC: A10
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Hospital outpatients or drug prescriptions may be a suitable alternative to primary care.
Notes	Diabetes is considered a chronic disease for which a case can be new only once ever. New diagnosis is identified by specific identification of a new case (flag, etc) or no record of previous diagnosis/contact in the two previous years. An individual should not be a new case in more than one year.
Cautions	A test result e.g. for blood glucose should not be counted without a recorded diagnosis. Compared to epidemiological estimates, figures may be lower because of undiagnosed cases and people whose diabetes is controlled by diet only. A person who had existing diabetes but no contact in the previous two years may be counted as a new case.
Title	**MB02 Diabetes mellitus - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed diabetes mellitus at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: E10-E14 ICPC-1: T90 ICPC-2: T89-T90 ATC: A10
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Hospital outpatients or drug prescriptions may be a suitable alternative to primary care.
Notes	Diabetes is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	A test result e.g. for blood glucose should not be counted without a recorded diagnosis. Compared to epidemiological estimates, figures may be lower because of undiagnosed cases and people whose diabetes is controlled by diet only. A person who had existing diabetes but no contact in the three-year period will not be counted.

Title	MB03 Dementia (incl. Alzheimer disease) - prevalence
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed dementia (including Alzheimer disease) at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: F00-F03, F05.1, G30 ICPC: P70 ATC: N06D excl. N06DX02
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Drug prescriptions may be a suitable alternative to primary care, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	Dementia (including Alzheimer disease) is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	Since many people with dementia are in nursing homes or other residential establishments for the elderly, it is important to ensure that the institutionalised population is represented in the estimates.
Title	**MB04 Mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed mental and behavioural disorders due to use of alcohol (incl. alcohol dependence) at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: F10, G31.2 ICPC: P15-P16
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Other sources such as emergency care may be helpful.
Notes	While these disorders are likely to be long-term, it is possible for the diagnosis to change over time. However, given their long-term nature, three-year period prevalence is used.
Cautions	If data are not available from specialist treatment services for alcohol dependence, or general psychiatric institutions, underestimation is likely.
Title	**MB05 Schizophrenia, schizotypal and delusional disorders - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed schizophrenia, schizotypal and delusional disorders at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: F20-F29 ICPC: P72, P98 ATC: N05A excl. N05AN
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Other sources such as emergency care may be helpful. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	While these disorders are likely to be long-term, it is possible for the diagnosis to change over time. However, given their long-term nature, three-year period prevalence is used.
Cautions	The inclusion of data from psychiatric institutions and community mental health services is essential to avoid underestimation.

Title	MB06 Mood (affective) disorders - prevalence
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed mood (affective) disorders at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: F30-F39 ICPC: P73, P76
Data sources	Primary care and hospital outpatients are likely to be the most important sources. The data from community mental health services may be useful.
Notes	While these disorders are often to be long-term, it is possible for the diagnosis to change over time. However, given their long-term nature, three-year period prevalence is used.
Cautions	Causes of death and hospital inpatients are unlikely to give meaningful results. Drug prescriptions may be relevant but are unlikely to have enough consistency to be valuable. Self-reported depression is likely to be higher since not all people with depressive mood seek or receive a medical diagnosis.
Title	**MB07 Parkinson disease - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed Parkinson disease at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: G20 ICPC: N87* ATC: N04BA02, N04BA03, N04BD, N04BX01, N04BX02
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	Parkinson disease is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	* For ICPC, possible overestimation due to counting of Parkinsonism caused by other pathologies has to be considered. Coverage of the institutionalised population is important.
Title	**MB08 Multiple sclerosis - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed multiple sclerosis at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: G35 ICPC: N86 ATC: L03AB07, L03AB08, L03AX13, L05AA23, L05AA27, L05AA31, L05AA34, N07XX09
Data sources	Primary care, hospital inpatients and causes of death are essential. For causes of death and inpatients, multiple causes and secondary diagnoses must be included. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	Multiple sclerosis is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	Coverage of the institutionalised population is important.

Title	MB09 Epilepsy - prevalence
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed epilepsy at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: G40-G41 ICPC: N88 ATC: N03AD01, N03AF03, N03AG04, N03AX10, N03AX14, N03AX15, N03AX17, N03AX18, N03AX22, N03AX23
Data sources	Primary care, hospital inpatients and outpatients are important sources. Causes of death are less likely to be useful, but if used, multiple causes are needed. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	Epilepsy is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	Coverage of the institutionalised population may be somewhat important.
Title	**MB10 Hypertensive diseases - incidence**
Measure	**Incidence by person** based on healthcare contact. The number of individuals having a first diagnosis of hypertensive diseases in the single index year.
Classification	ICD-10: I10-I13, I15 ICPC: K86, K87* ATC: C03AA, C03AB, C03AH, C03AX01, C02CA04, C03BA, C03DB, C03EA, C09BA02-9, C09BB, C09DB, C09DA02-4, C09DA06-7, C09DA01, C02AB01-2, C02AC01-2, C02AC04-5, C02DB02-4, C02DC01, C02DD01, C02DG01, C02KA01, C02KB01, C02KC01, C02KD01, C02KX01, C09XA
Data sources	Primary care, hospital inpatients and outpatients are important sources. Causes of death should also be included, and multiple causes are needed. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or no record of previous diagnosis/contact in the two previous years. While a person can have more than one separate episode of hypertensive disease, the definition of incidence by person counts an individual only once in the year. An individual should not be a new case in more than one year.
Cautions	*For ICPC, possible overestimation due to the scope of the codes has to be considered. The epidemiological incidence may be higher due to undiagnosed cases. The clinical definition of arterial hypertension varies. Some patients might be treated only by diet and could be missed if prescriptions data are used.
Title	**MB11 Hypertensive diseases - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed hypertensive diseases at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: I10-I13, I15 ICPC: K86, K87* ATC: C03AA, C03AB, C03AH, C03AX01, C02CA04, C03BA, C03DB, C03EA, C09BA02-9, C09BB, C09DB, C09DA02-4, C09DA06-7, C09DA01, C02AB01-2, C02AC01-2, C02AC04-5, C02DB02-4, C02DC01, C02DD01, C02DG01, C02KA01, C02KB01, C02KC01, C02KD01, C02KX01, C09XA

Data sources	Primary care, hospital inpatients and outpatients are important sources. Causes of death should also be included, and multiple causes are needed. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	While hypertensive diseases are likely to be long-term, it is possible for the diagnosis to change over time. However, given their long-term nature, three-year period prevalence is used.
Cautions	* For ICPC, possible overestimation due to the scope of the codes has to be considered. The epidemiological incidence may be higher due to undiagnosed cases. The clinical definition of arterial hypertension varies. Some patients might be treated only by diet and could be missed if prescriptions data are used.
Title	**MB12 Ischaemic heart diseases - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed ischaemic heart diseases at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: I20-I25 ICPC: K74-K76 ATC: C01DA
Data sources	Primary care, hospital inpatients and causes of death are all important. For causes of death, multiple causes should be counted. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	While ischaemic heart diseases are likely to be long-term, it is possible for the diagnosis to change over time. However, given their long-term nature, three-year period prevalence is used.
Cautions	Epidemiological prevalence may be higher due to undiagnosed cases or variations in diagnostic practice.
Title	**MB13 Acute myocardial infarction - incidence**
Measure	**Incidence by person by year** based on healthcare contact. The number of individuals having any diagnosis of acute myocardial infarction in the single index year.
Classification	ICD-10: I21, I22 ICPC: K75*
Data sources	Hospital inpatients and causes of death are essential. The counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or an interval of at least 28 days since the end of any previously recorded episode of illness. A person can have more than one separate episode of acute myocardial infarction: the definition of incidence by person by year counts an individual only once in the year, however an individual may be a new case in more than one year.
Cautions	* ICPC is unlikely to be used due to the unsuitability of primary care data for this indicator. The counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation.
Title	**MB14 Heart failure - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed heart failure at any time during the three-year reference period and present in the index year (year 3).

Classification	ICD-10: I50 ICPC: K77
Data sources	Primary care and possibly outpatient care are the most important sources. Inpatients should be counted with caution, while for cause of death only the underlying cause should be used.
Notes	Heart failure is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	Care must be taken to exclude the diagnosis of 'heart failure' in some countries as a symptom or mode of dying rather than an ongoing cardiac condition.
Title	**MB15 Stroke - incidence**
Measure	**Incidence by person by year** based on healthcare contact. The number of individuals having any diagnosis of stroke in the single index year.
Classification	ICD-10: I60-I64 ICPC: K90*
Data sources	Hospital inpatients and causes of death are essential. The counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or an interval of at least 28 days since the end of any previously recorded episode of illness. A person can have more than one separate episode of stroke: the definition of incidence by person by year counts an individual only once in the year, however an individual may be a new case in more than one year.
Cautions	*ICPC is unlikely to be used due to the unsuitability of primary care data for this indicator. The counting of cases from primary care and outpatient care should be avoided, as these are likely to be previous cases attending for purposes such as rehabilitation. The range of definitions and codes used for cerebrovascular diseases and stroke has been found to be inconsistent between countries.
Title	**MB16 Cerebrovascular diseases - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed cerebrovascular diseases at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: I60-I69 ICPC: K90-K91
Data sources	Primary care, hospital inpatients and causes of death are all important.
Notes	The group of cerebrovascular diseases vary in nature but generally are considered chronic diseases which continue indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	The range of definitions and codes used for cerebrovascular diseases and stroke has been found to be inconsistent between countries.
Title	**MB17 Pneumonia - incidence**
Measure	**Incidence by episode** based on healthcare contact. The number of individuals having any diagnosis of pneumonia in the single index year, counting each new episode of pneumonia beginning in the index year as a separate case.
Classification	ICD-10: J12-J18 ICPC: R81

Data sources	Primary care, hospital inpatients and causes of death are all important.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or an interval of at least 28 days since the end of any previously recorded episode of illness. A person can have more than one separate episode of pneumonia in any year. An episode which began in the previous year is not counted in the index year.
Cautions	Pneumonia often occurs among hospital inpatients and institutionalised elderly people. Coverage of the institutional populations is therefore important.
Title	**MB18 Asthma - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed asthma at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: J45, J46 ICPC: R96 ATC: R03
Data sources	Primary care is likely to be the most important source. For causes of death, multiple causes should be counted. Drug prescriptions may be helpful, but the most suitable ATC codes should be chosen carefully based on national practice.
Notes	Asthma is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	There may be variations in diagnostic practice between asthma, COPD and bronchitis. It has been suggested that asthma in childhood is over-diagnosed. Some patients may be missed due to obtaining medication in the private sector.
Title	**MB19 Chronic lower respiratory diseases other than asthma (incl. COPD) - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed chronic lower respiratory diseases other than asthma (including COPD) at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: J40-J44, J47 ICPC-1: R91, R95 ICPC-2: R79, R95, R99
Data sources	Primary care is likely to be the most important source. For causes of death, multiple causes should be counted.
Notes	This group of conditions are considered chronic diseases which continue indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	There may be variations in diagnostic practice between asthma, COPD and bronchitis. Care should be taken to exclude acute bronchitis. Some patients may be missed due to obtaining medication in the private sector.
Title	**MB20 Chronic obstructive pulmonary disease (COPD) - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed chronic obstructive pulmonary disease (COPD) at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: J44 ICPC: R95
Data sources	Primary care is likely to be the most important source. For causes of death, multiple causes should be counted.

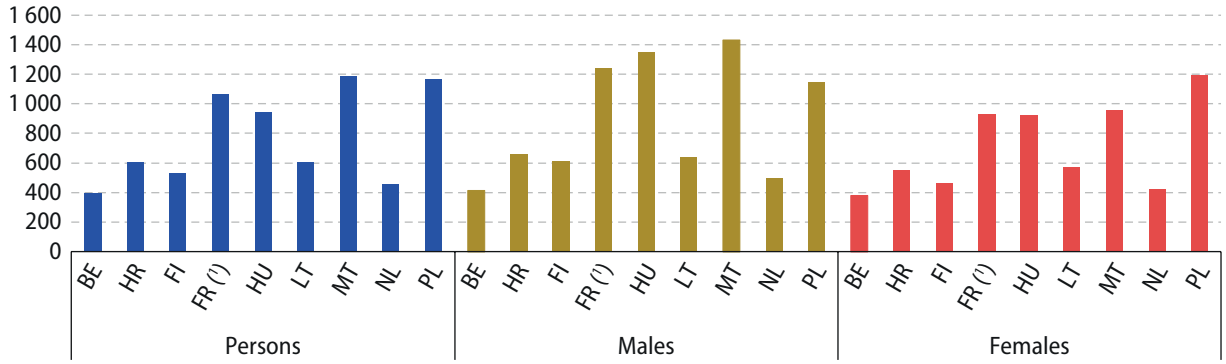
Notes	Chronic obstructive pulmonary disease (COPD) is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	There may be variations in diagnostic practice between asthma, COPD and bronchitis. Some patients may be missed due to obtaining medication in the private sector.
Title	**MB21 Diseases of liver - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed diseases of the liver at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: K70-K77 ICPC: D97 **Special note - for diagnoses based on ICD-10, but not ICPC, the following optional sub-divisions are possible:** (a) Alcoholic liver disease K70 (b) Diseases of liver other than alcoholic K71-K77
Data sources	Primary care, hospital inpatients and outpatients, and causes of death are all important.
Notes	Diseases of the liver are usually long-term and may be considered chronic, and therefore continue indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	If sub-division into alcoholic and non-alcoholic liver diseases is intended, it should be noted that as well as the inability to make this distinction in ICPC, there may be unknown variations in diagnostic practice between the categories.
Title	**MB22 Rheumatoid arthritis - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed rheumatoid arthritis at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: M05, M06 ICPC: L88*
Data sources	Primary care, hospital inpatients and outpatients are all likely to be important. Additional sources such as disability claims may be helpful. Causes of death are unlikely to be relevant.
Notes	Rheumatoid arthritis is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.
Cautions	*The ICPC code may be used for a wider range of similar conditions. National usage of the codes in primary care should be carefully assessed.
Title	**MB23 Arthrosis - prevalence**
Measure	**Three-year period prevalence** based on healthcare contact. The number of individuals having diagnosed arthrosis at any time during the three-year reference period and present in the index year (year 3).
Classification	ICD-10: M15-M19 ICPC: L89-L91
Data sources	Primary care, hospital inpatients and outpatients are all likely to be important. Additional sources such as disability claims may be helpful. Causes of death are unlikely to be relevant.
Notes	Arthrosis is considered a chronic disease which continues indefinitely after the first diagnosis. Any contact in the three-year period, including a drug prescription or claim for disability benefits, should be counted.

Cautions	The non-specific nature of joint problems in elderly patients may lead to differences in recording. Some patients may be missed due to treatment in the private sector.
Title	**MB24 Urolithiasis - incidence**
Measure	**Incidence by person** based on healthcare contact. The number of individuals having a first diagnosis of urolithiasis in the single index year.
Classification	ICD-10: N20-N23 ICPC: U14, U95
Data sources	Primary care, hospital inpatients and outpatients are all likely to be important. Causes of death are unlikely to be relevant.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or no record of previous diagnosis/contact in the two previous years. An individual should not be a new case in more than one year.
Cautions	There may be significant variations in clinical practice, such as the severity threshold for access to services. Coverage of institutionalised populations may be important. Patients treated in the private sector may be missed.
Title	**MB25 Intracranial injury - incidence**
Measure	**Incidence by person by year** based on healthcare contact. The number of individuals having any diagnosis of intracranial injury in the single index year.
Classification	ICD-10: S06 ICPC-1: N79* ICPC-2: N79-N80*
Data sources	Hospital inpatients and causes of death should be the main sources, with the addition of emergency care if available. This will restrict the estimates to severe cases and exclude patients attending only for after-care and rehabilitation. A relevant surgical procedure may be counted.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or an interval of at least 28 days since the end of any previously recorded episode of illness. A person can have more than one separate episode of intracranial injury: the definition of incidence by person by year counts an individual only once in the year, however an individual may be a new case in more than one year. Coverage of institutionalised populations is important.
Cautions	* ICPC codes are imprecise and are unlikely to be used due to the unsuitability of primary care data.
Title	**MB26 Fracture of femur - incidence**
Measure	**Incidence by person by year** based on healthcare contact. The number of individuals having any diagnosis of fracture of the femur in the single index year.
Classification	ICD-10: S72 ICPC: L75
Data sources	Hospital inpatients and causes of death should be the main sources, with the addition of emergency care if available. This will restrict the estimates to severe cases and exclude patients attending only for after-care and rehabilitation. A relevant surgical procedure may be counted.
Notes	New diagnosis is identified by specific identification of a new case (flag, etc) or an interval of at least 28 days since the end of any previously recorded episode of illness. A person can have more than one separate episode of fracture of femur: the definition of incidence by person by year counts an individual only once in the year, however an individual may be a new case in more than one year. Coverage of institutionalised populations is important.
Cautions	* ICPC codes are unlikely to be used due to the unsuitability of primary care data.

Annex E. Age-standardized morbidity rates for each indicator, by sex, per country: bar charts

P1 - DIABETES MELLITUS

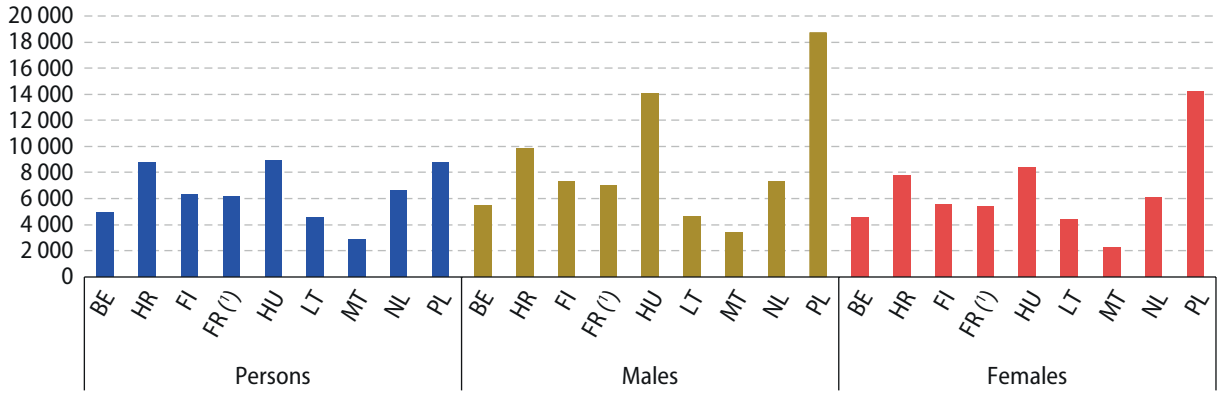
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P2 - DIABETES MELLITUS

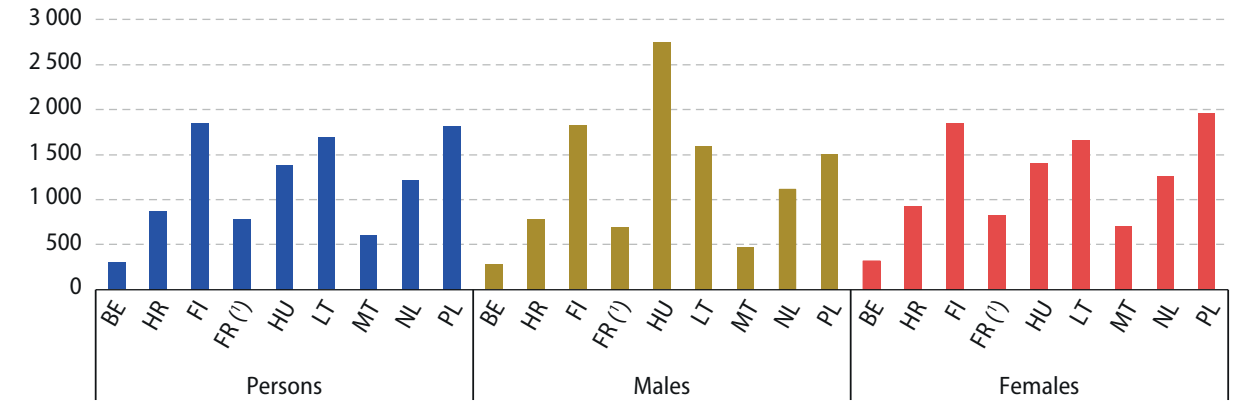
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P3 - DEMENTIA (INCL. ALZHEIMER DISEASE)

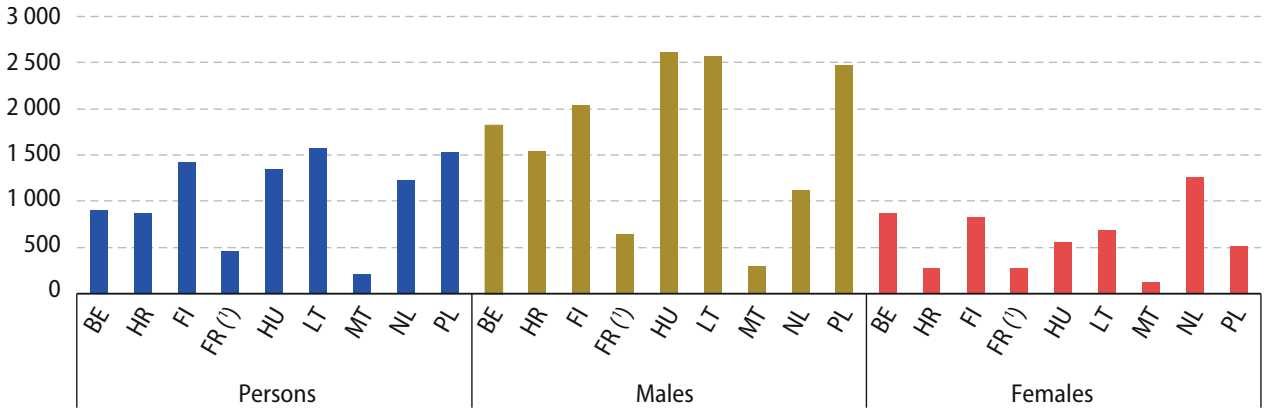
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P4 - MENTAL AND BEHAVIOURAL DISORDERS DUE TO USE OF ALCOHOL (INCL. ALCOHOL DEPENDENCE)

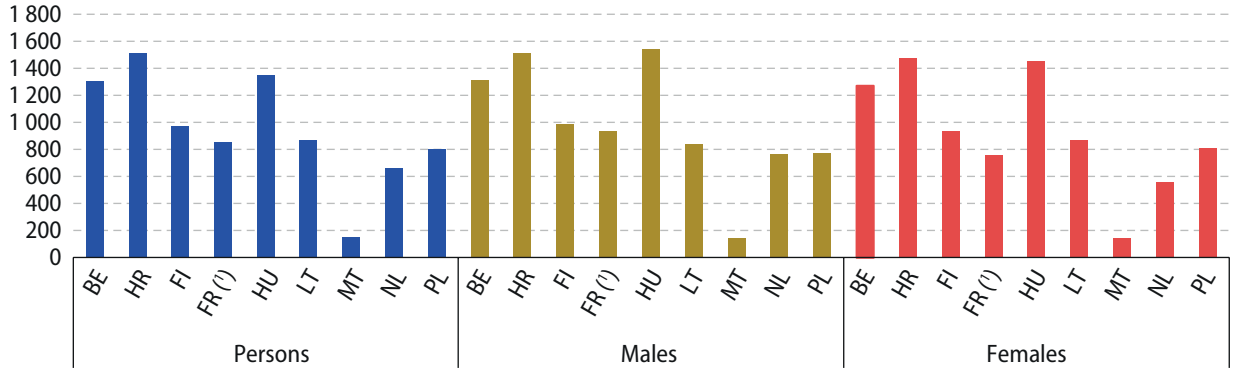
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P5 - SCHIZOPHRENIA, SCHIZOTYPAL AND DELUSIONAL DISORDERS

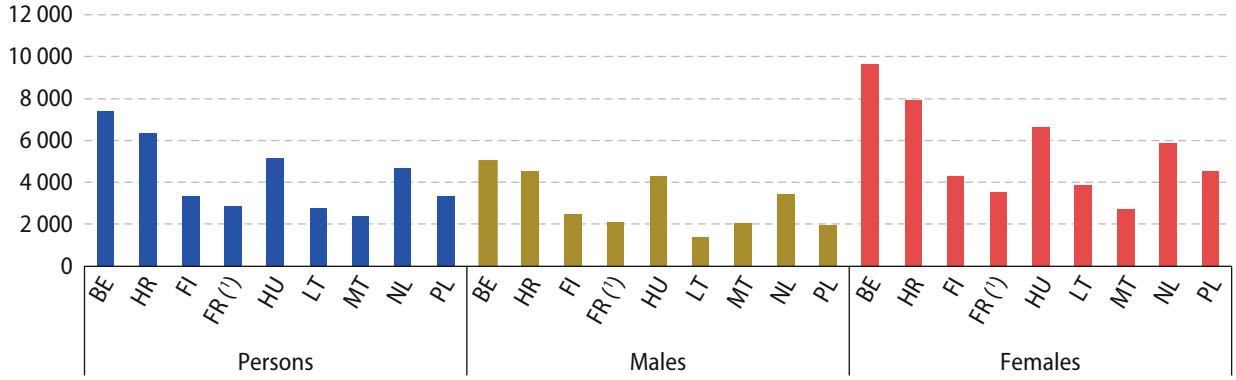
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P6 - MOOD (AFFECTIVE) DISORDERS

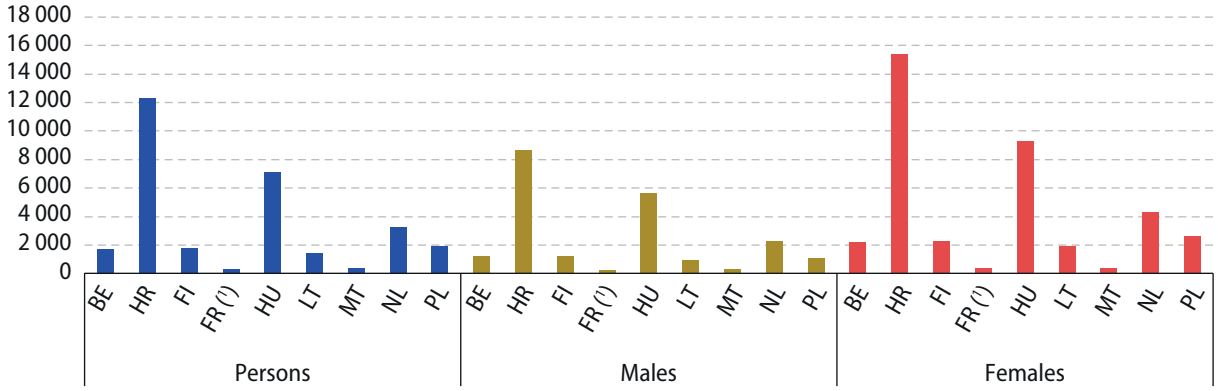
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P7 - ANXIETY DISORDERS

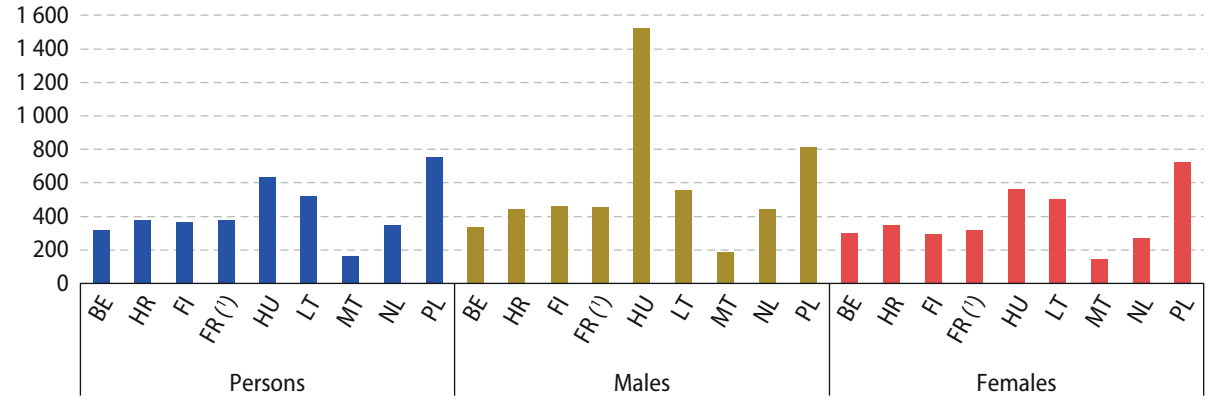
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P8 - PARKINSON DISEASE

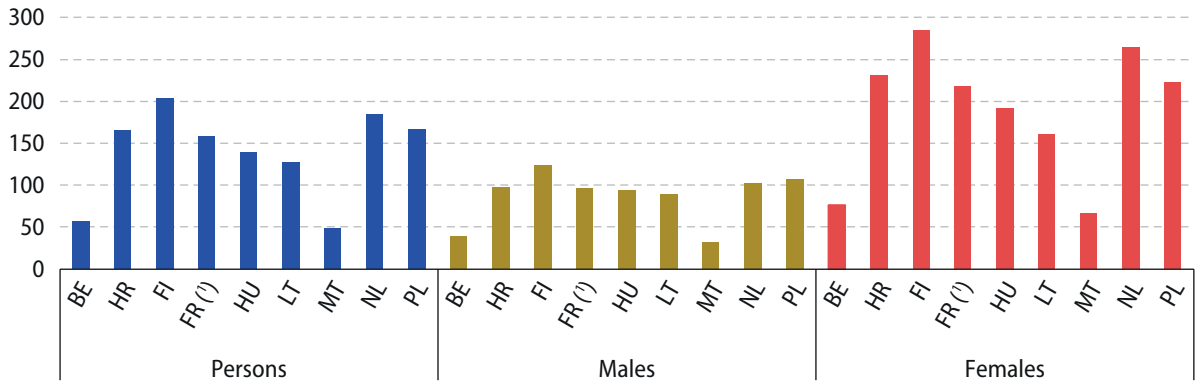
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P9 - MULTIPLE SCLEROSIS

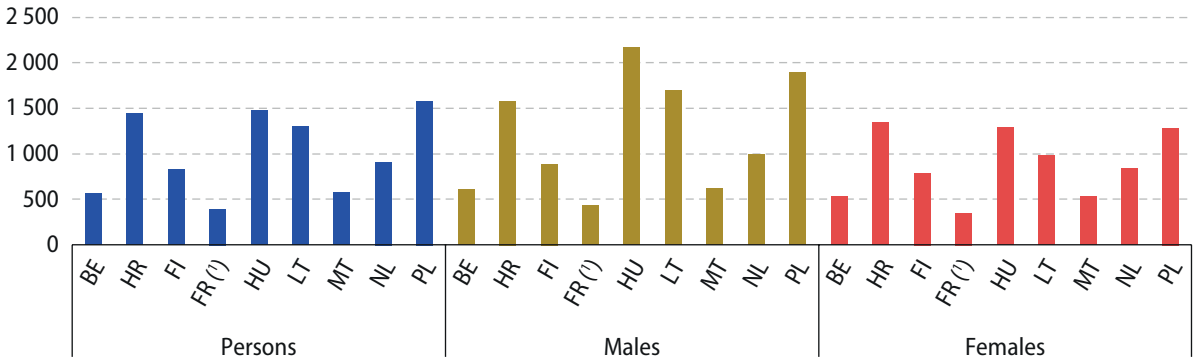
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P10 - EPILEPSY

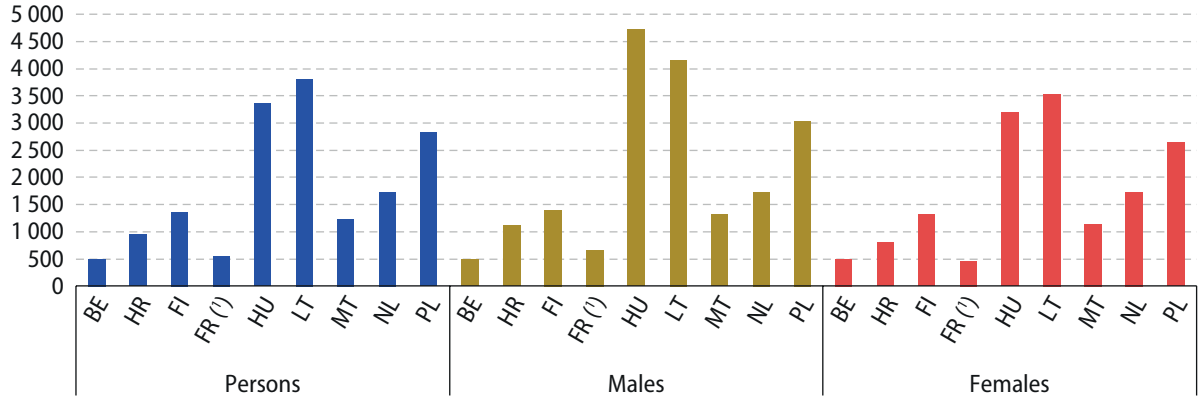
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P11 - HYPERTENSIVE DISEASES

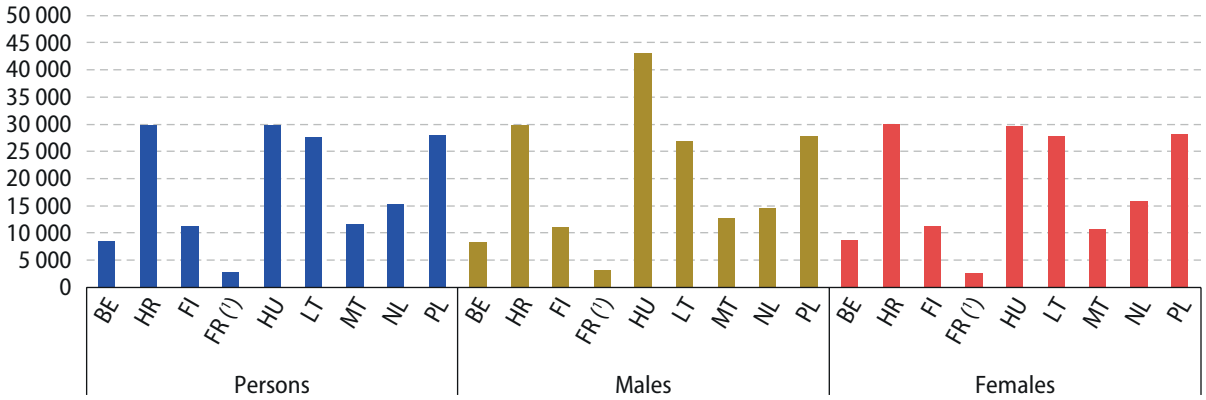
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P12 - HYPERTENSIVE DISEASES

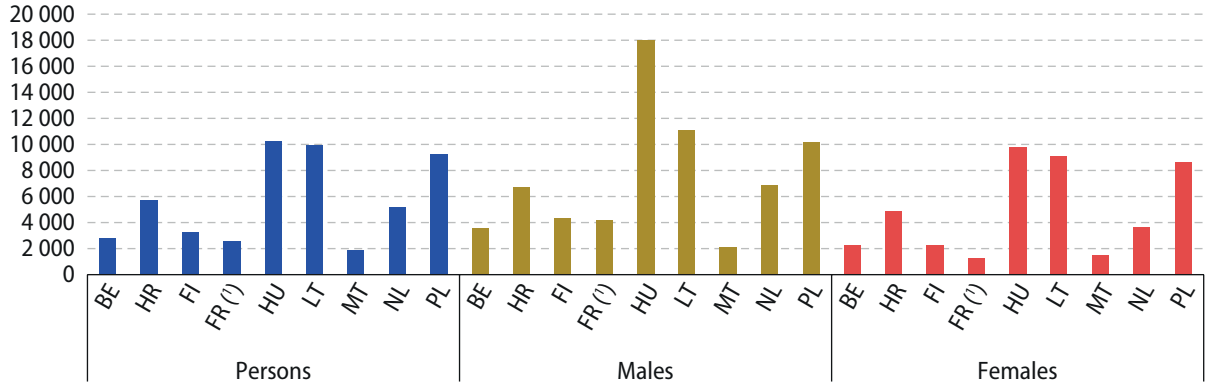
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P13 - ISCHAEMIC HEART DISEASES

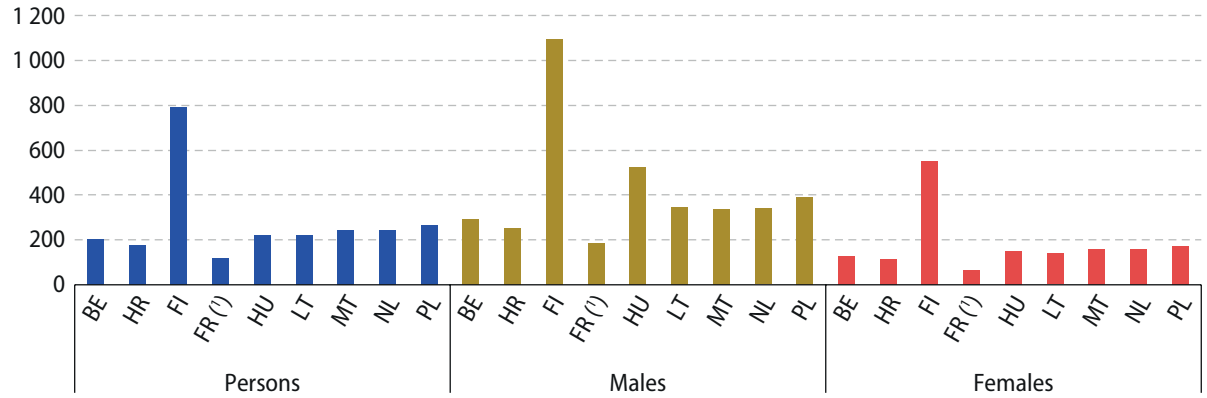
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P14 - ACUTE MYOCARDIAL INFARCTION

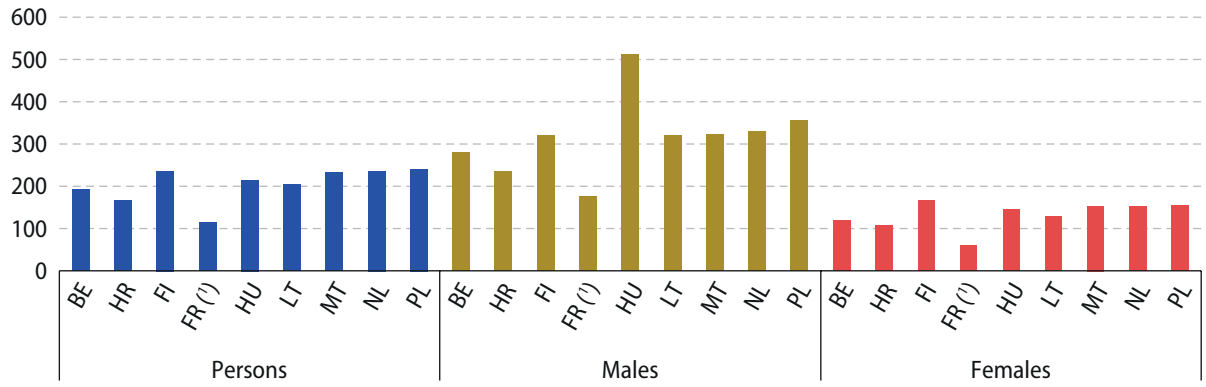
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P15 - ACUTE MYOCARDIAL INFARCTION

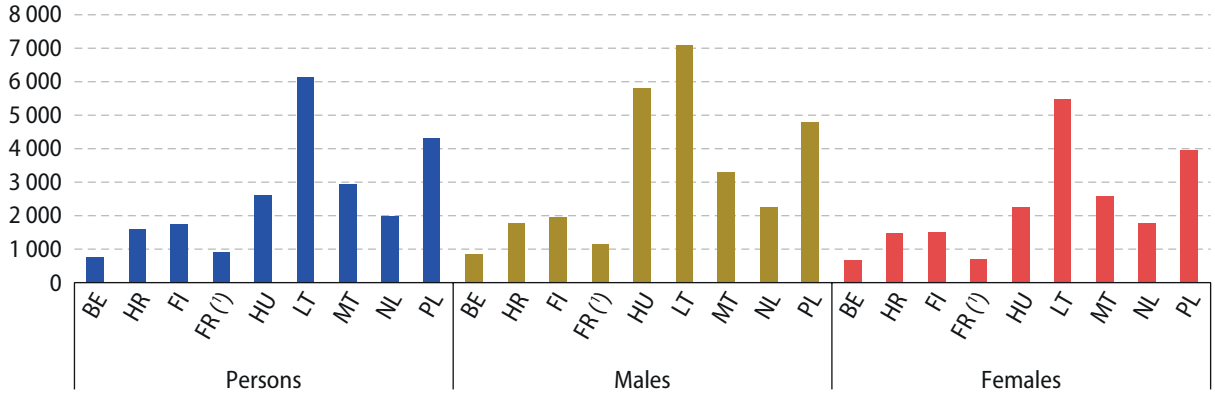
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P16 - HEART FAILURE

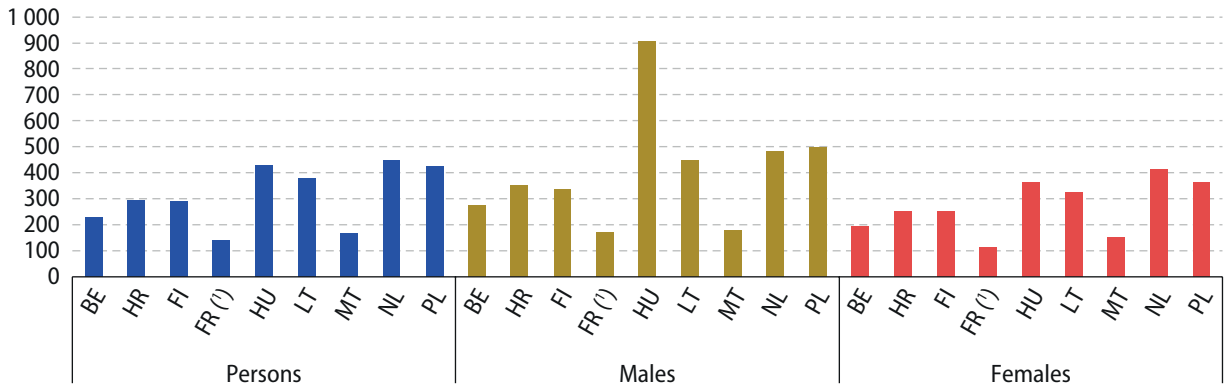
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P17 - STROKE

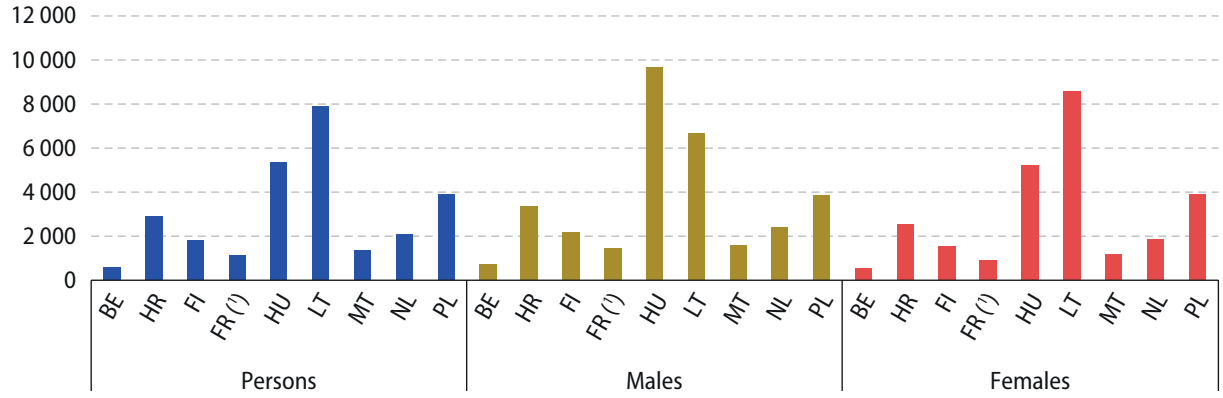
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P18 - CEREBROVASCULAR DISEASES

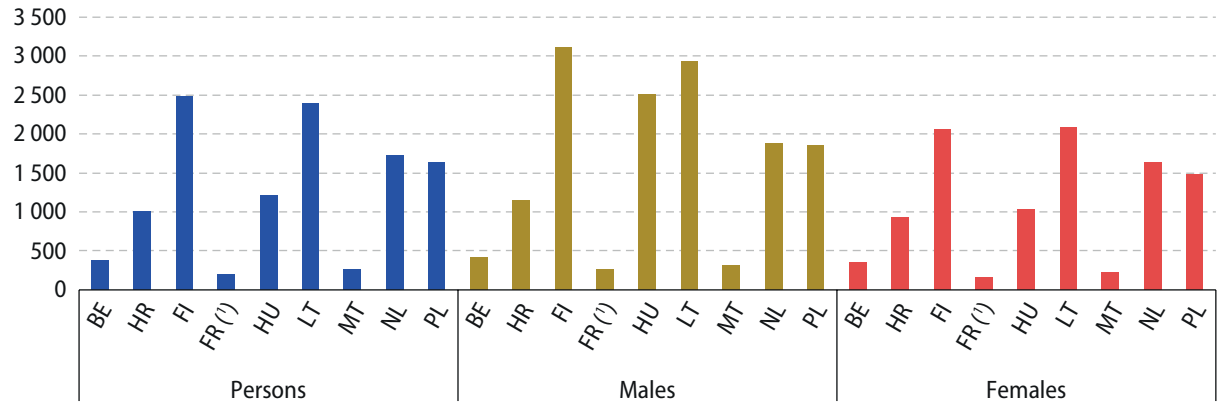
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P19 - PNEUMONIA

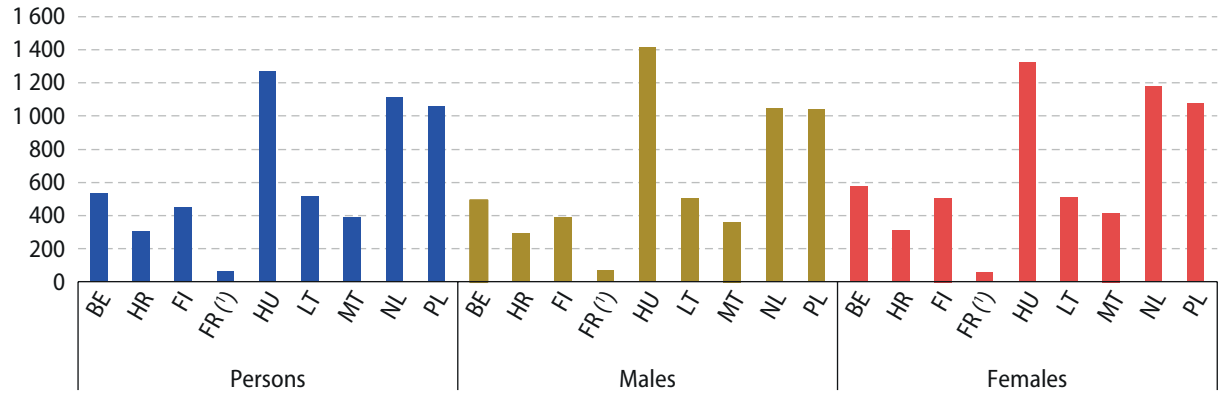
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P20 - ASTHMA

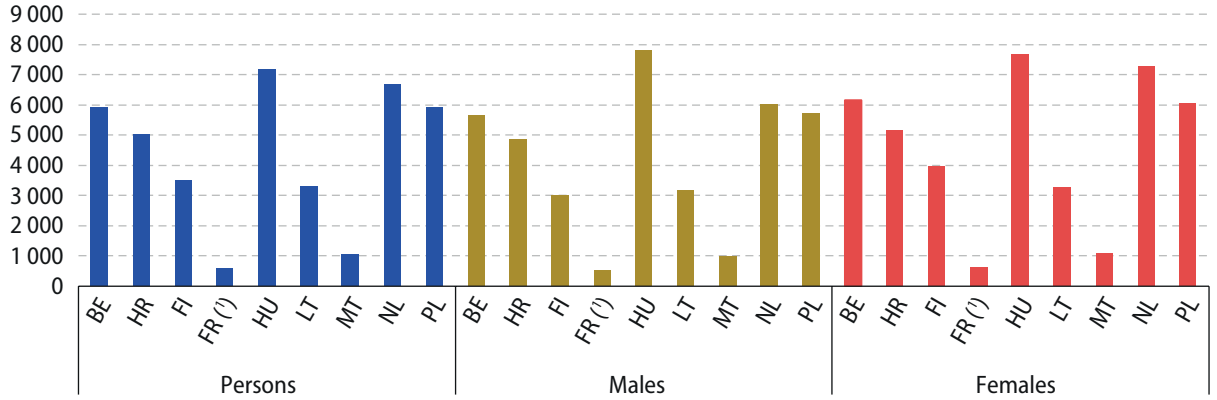
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P21 - ASTHMA

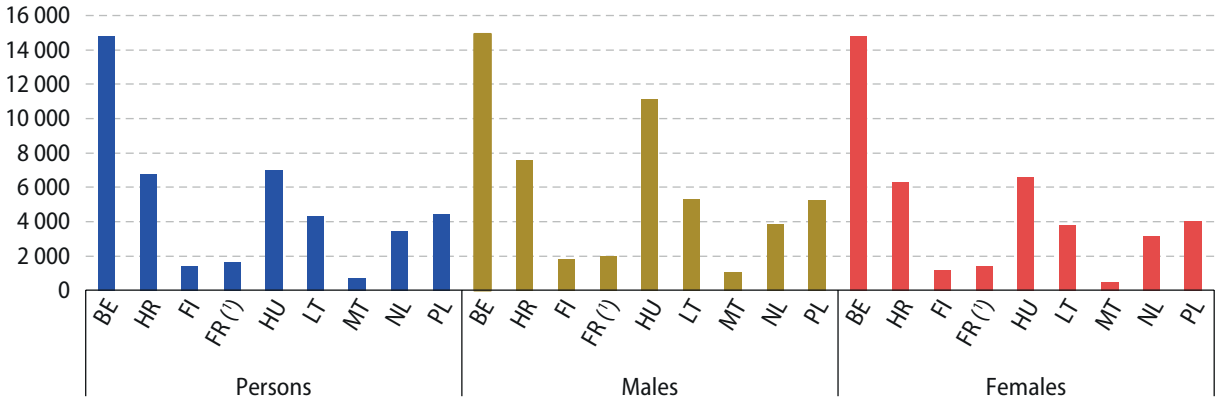
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P22 - CHRONIC LOWER RESPIRATORY DISEASES OTHER THAN ASTHMA (INCL. COPD)

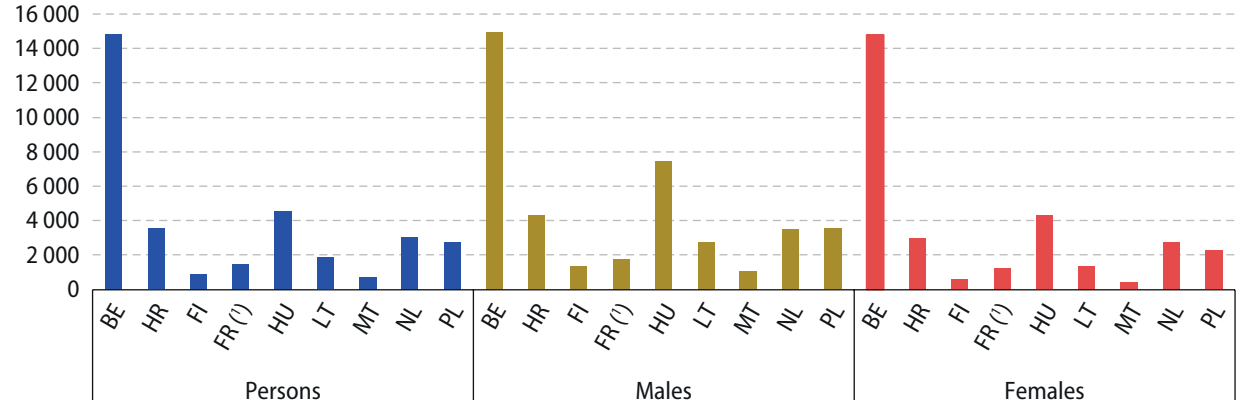
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P23 - CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

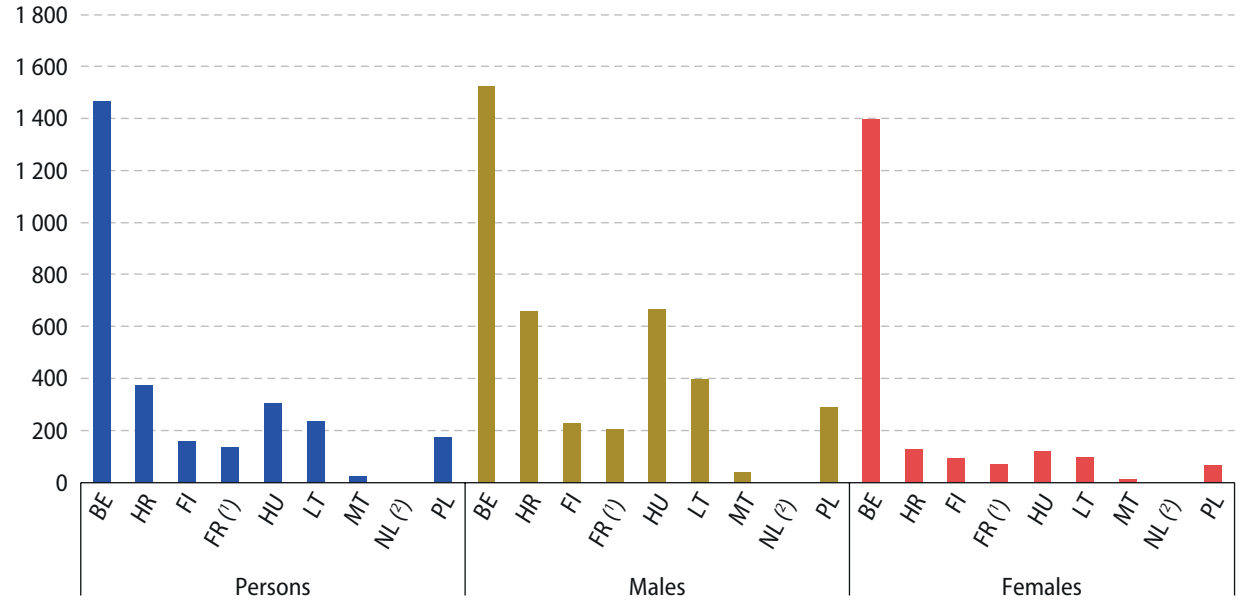
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P24 - ALCOHOLIC LIVER DISEASE

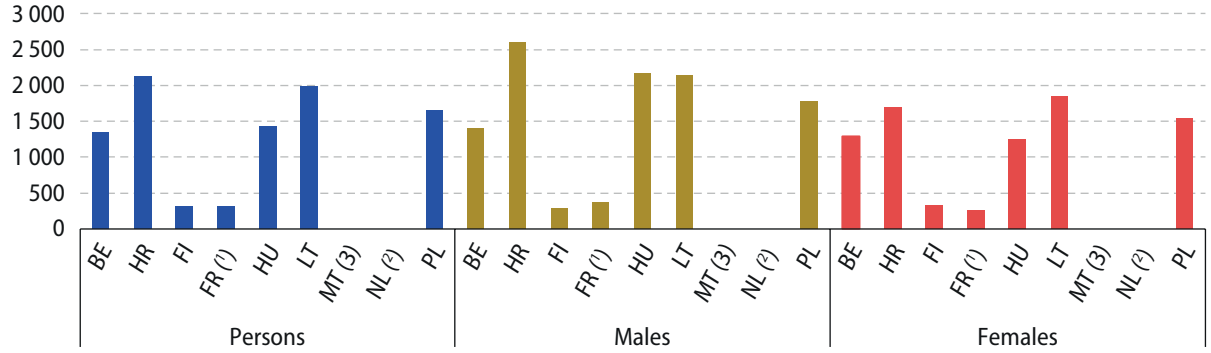
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P25 - DISEASES OF LIVER OTHER THAN ALCOHOLIC

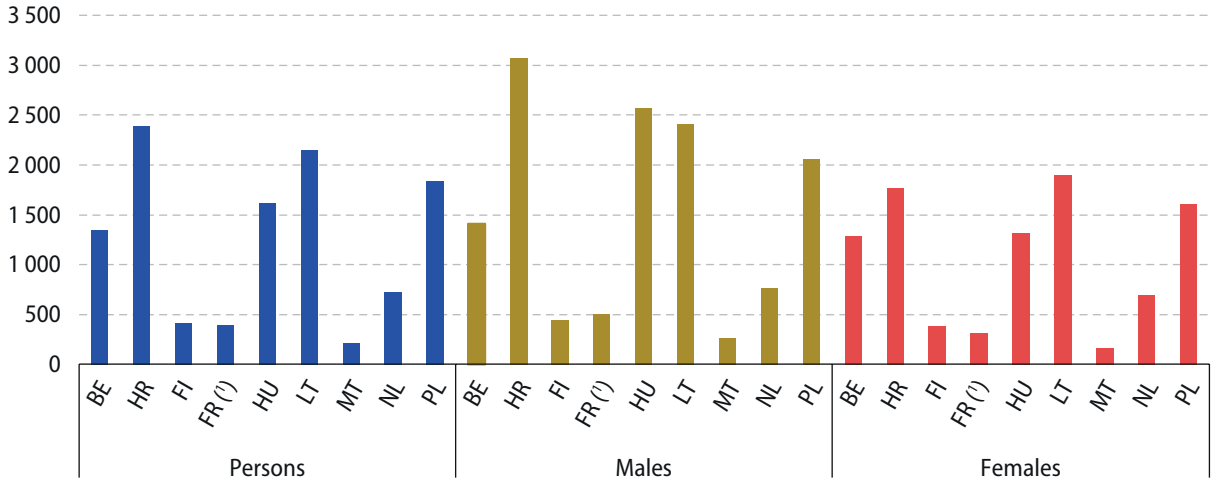
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P26 - DISEASES OF LIVER

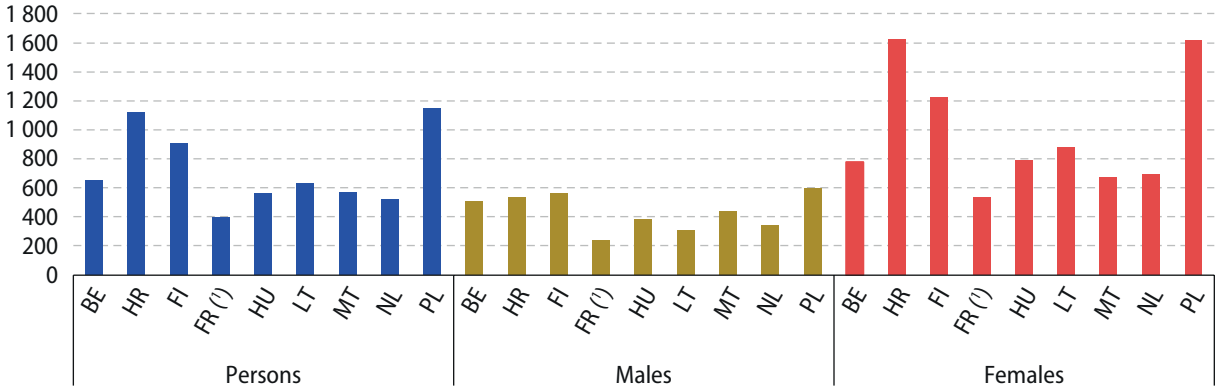
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P27 - RHEUMATOID ARTHRITIS

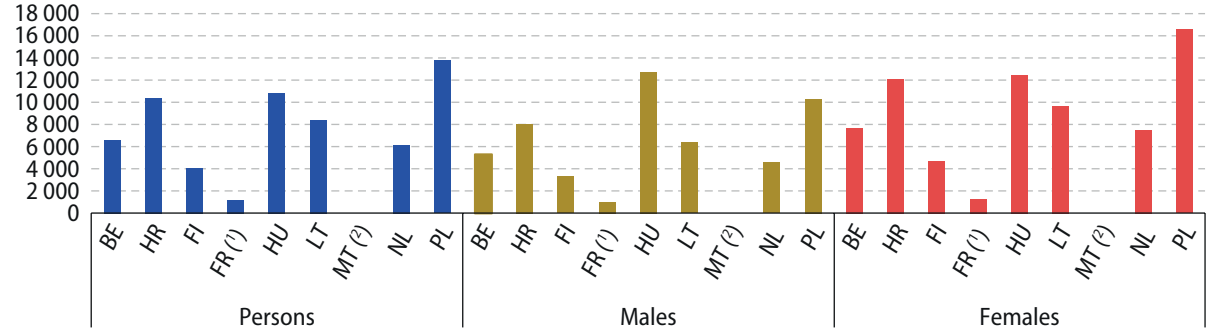
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P28 - ARTHROSIS

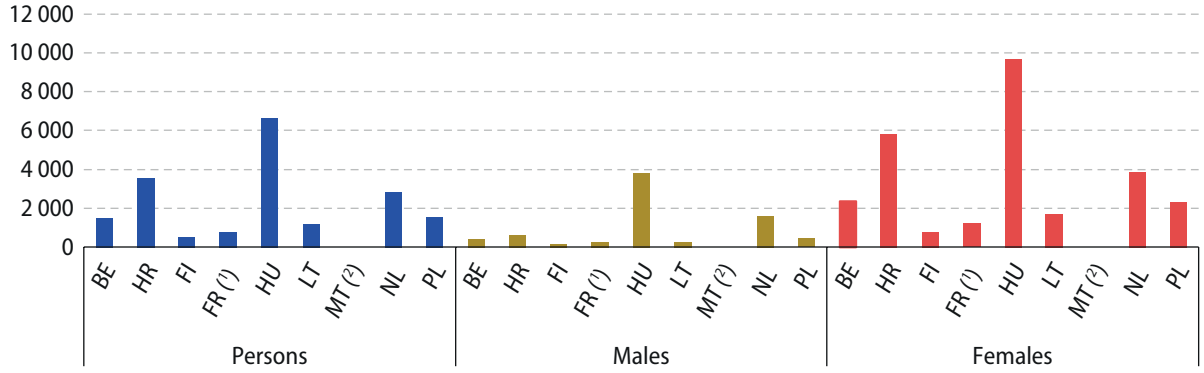
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P29 - OSTEOPOROSIS

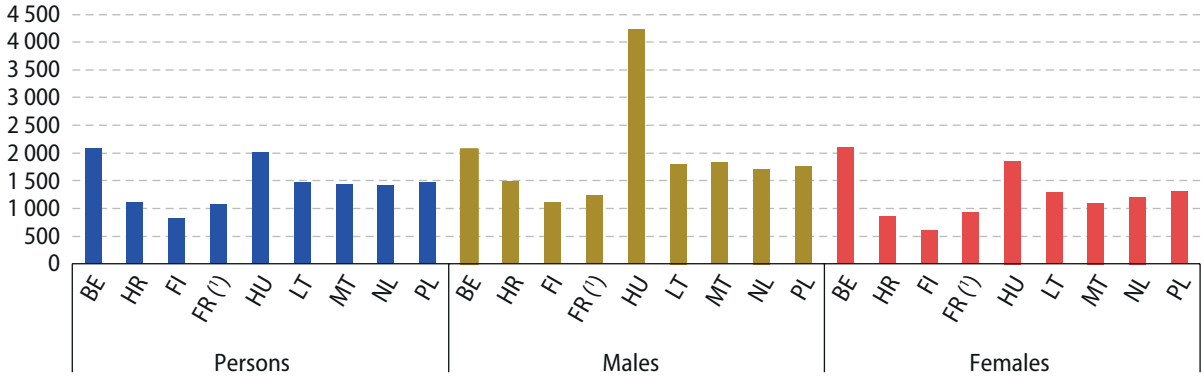
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P30 - RENAL FAILURE

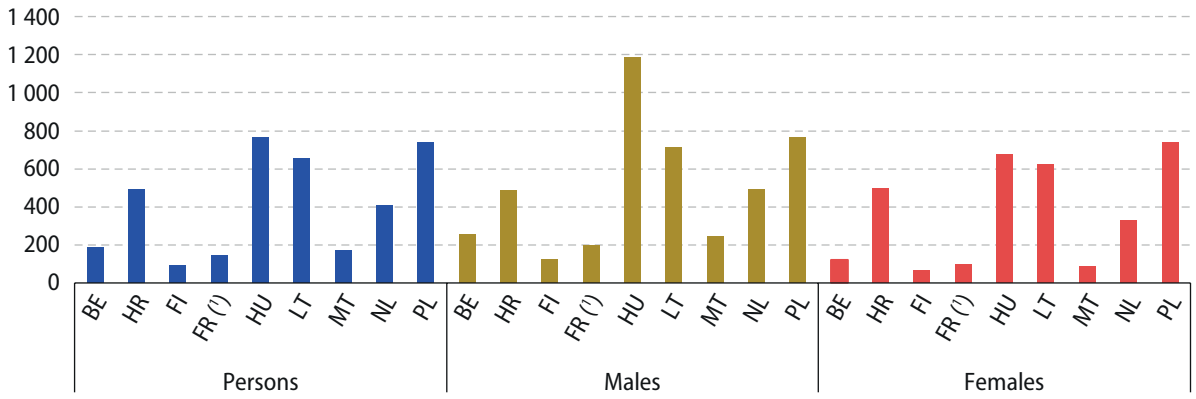
Period prevalence (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P31 - UROLITHIASIS

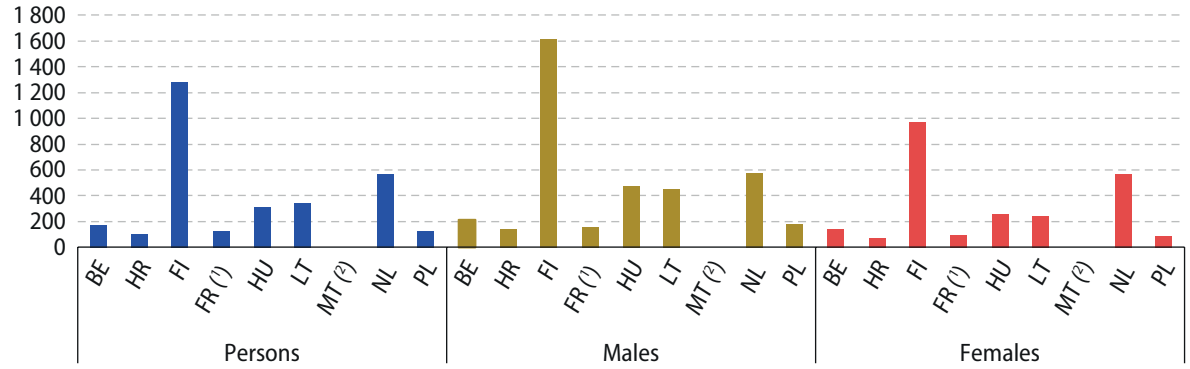
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P32 - INTRACRANIAL INJURY

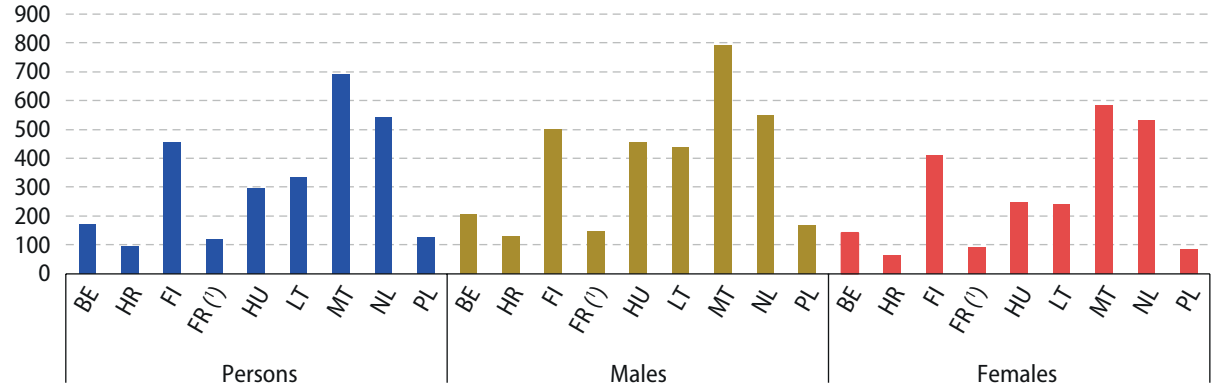
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P33 - INTRACRANIAL INJURY

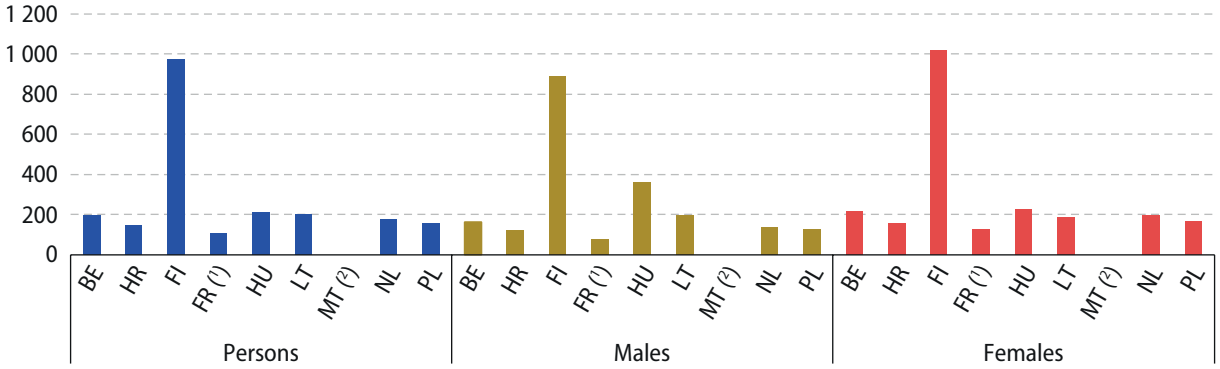
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P34 - FRACTURE OF FEMUR

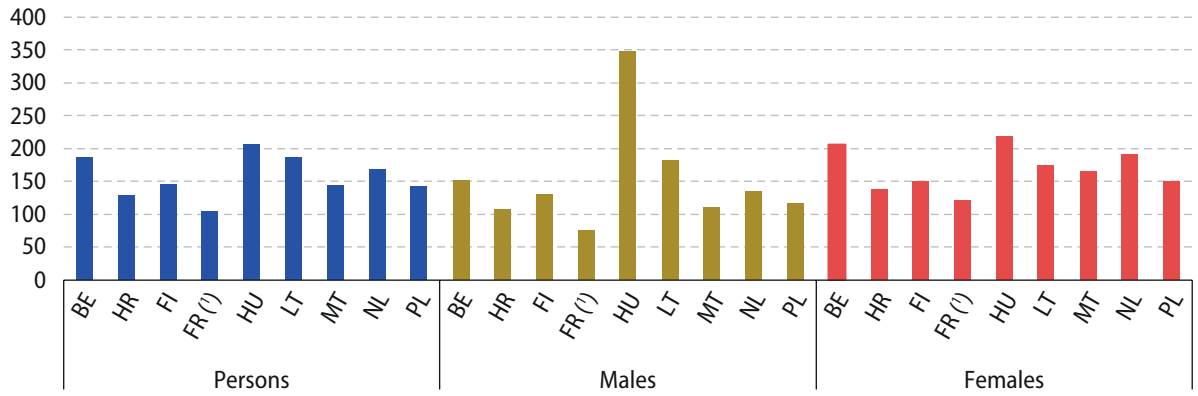
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

P35 - FRACTURE OF FEMUR

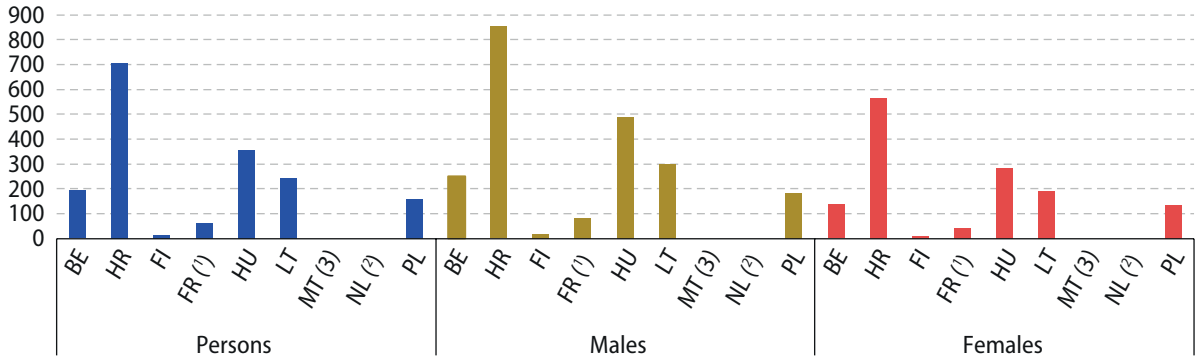
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB36 - LAND TRANSPORT ACCIDENTS

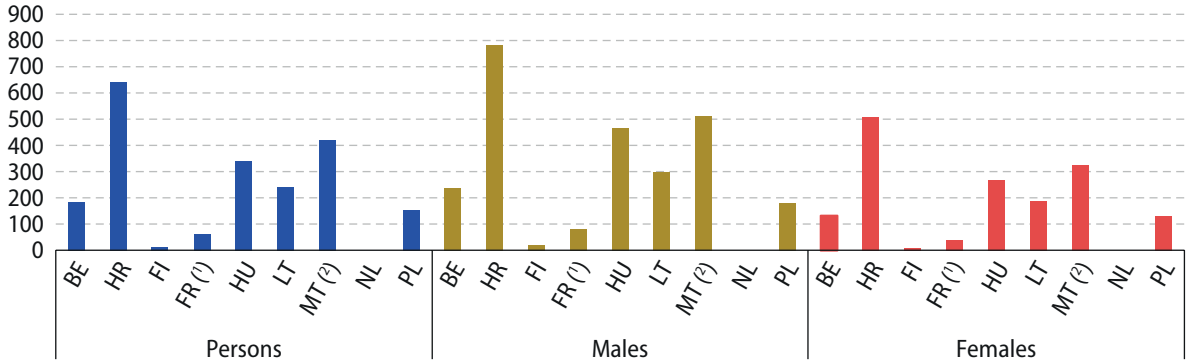
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB37 - LAND TRANSPORT ACCIDENTS

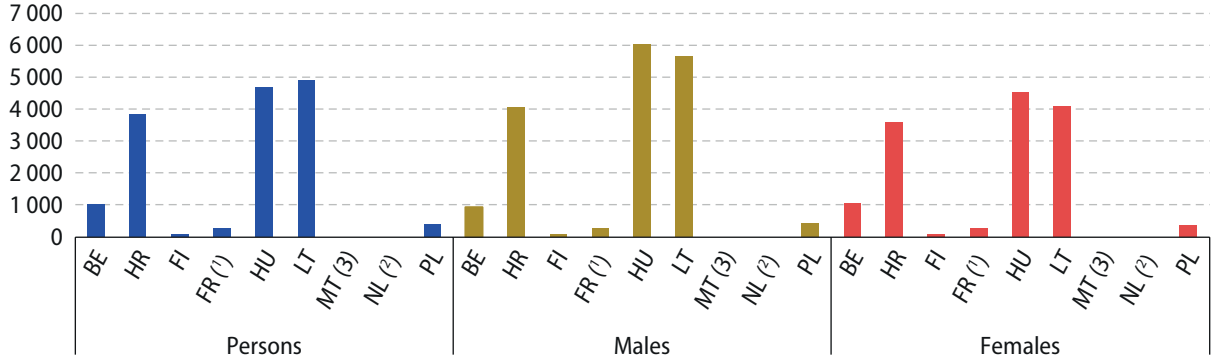
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB38 - ACCIDENTAL FALLS

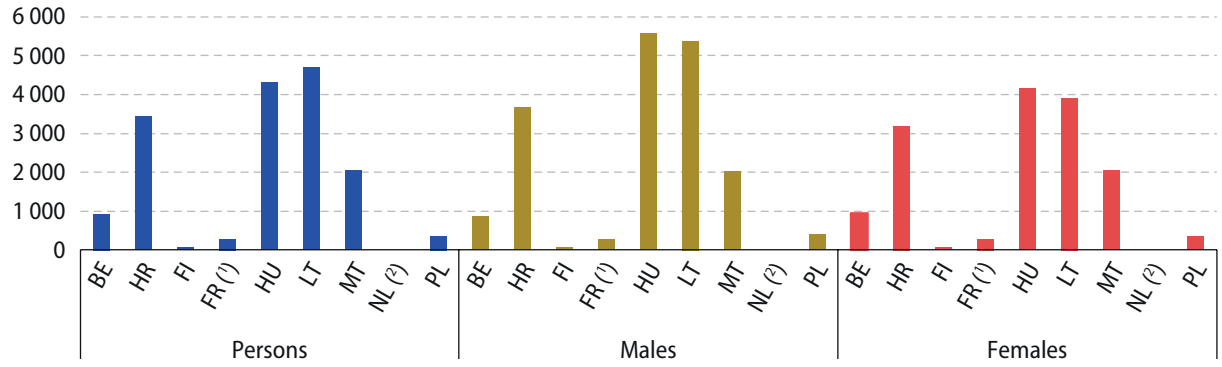
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB39 - ACCIDENTAL FALLS

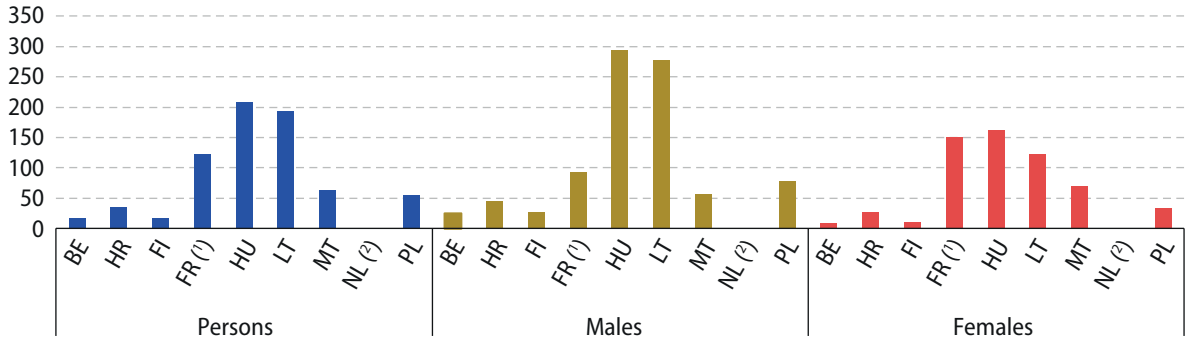
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB40 - INTENTIONAL SELF HARM (INCL. SUICIDAL ATTEMPT)

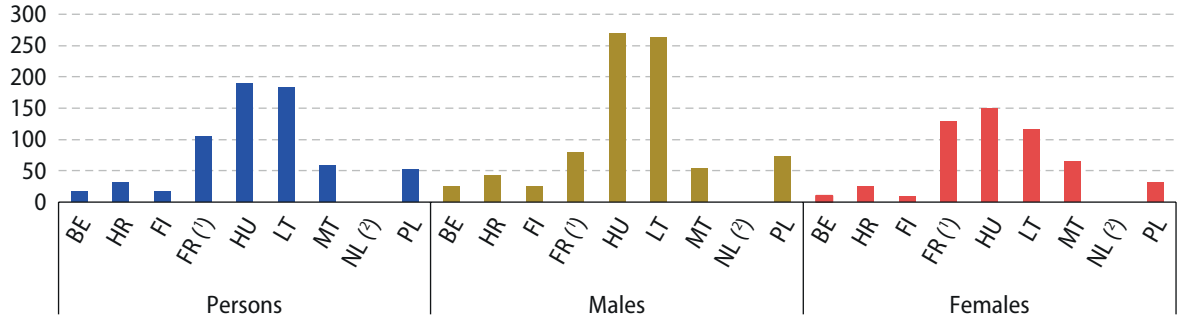
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB41 - INTENTIONAL SELF HARM (INCL. SUICIDAL ATTEMPT)

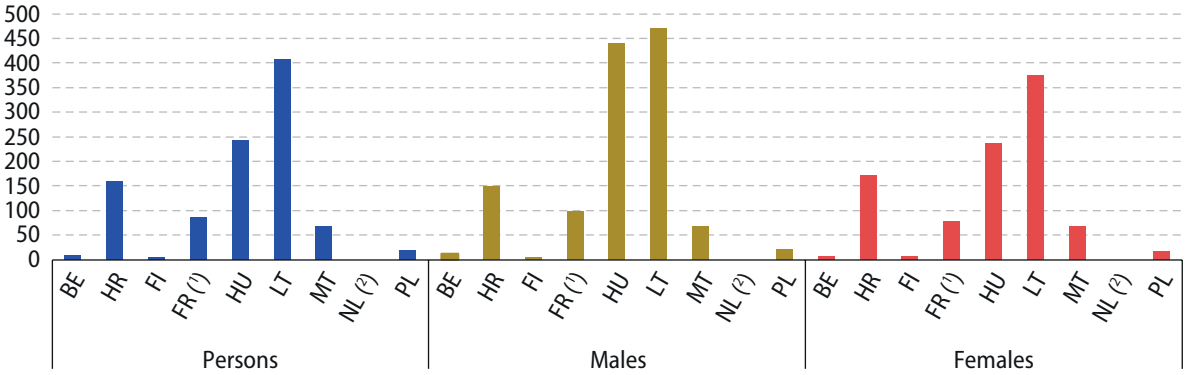
Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB42 - COMPLICATIONS OF MEDICAL AND SURGICAL CARE

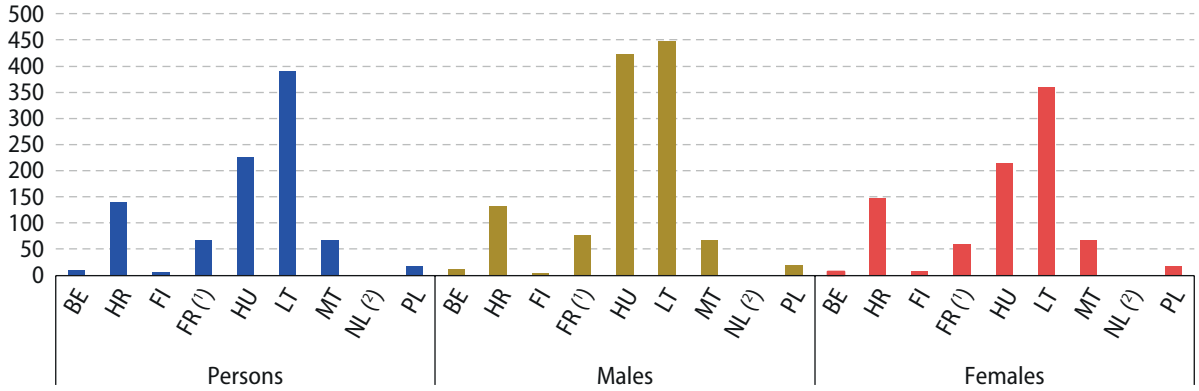
Incidence by episode (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

PB43 - COMPLICATIONS OF MEDICAL AND SURGICAL CARE

Incidence by person (age-standardised rate per 100 000)



FR (1) Age-standardised rates per 100,000 population using the 2013 European Standard Population. Highest age group is 95+ , except for France 85+
NL (2) Estimates were reported for all the main list of indicators P1-P35, except P24 and P25. Additionally, no estimates were reported for the optional indicators PB36-PB43
Note: Incidence is for the reference year 2016 and period prevalence is based on a three-year reference period 2014-16, except for Croatia reference year was 2017

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Morbidity statistics in the EU. Report on pilot studies

Between 2019 and 2021, nine Member States conducted national pilot studies focused on diagnosis-specific morbidity statistics, testing the feasibility of such data collection within the European Statistical System (ESS). These studies were centred around 43 indicators that had been developed in previous projects. This document presents their findings, shared with the Working Group on Public Health Statistics in November 2022.

For more information
https://ec.europa.eu/eurostat/

Publications Office
of the European Union