BRIDGEHEALTH
BRidging Information and Data Generation for Evidence-based Health Policy and Research

HA5 - DATA QUALITY METHODS

Draft April 2017

HA Leader Simona Giampaoli

Istituto Superiore di Sanità
Viale Regina Elena 299, 00161 ROME, Italy
e-mail: simona.giampaoli@iss.it, tel +39 0649904231

This project is funded by the Health Programme of the European Union

Table of contents

I. Introduction 6
II. Aims 8
III. Approach..... **Error! Bookmark not defined.**
IV. Definition of quality..... 9
V. Dedinition of case and indicator 10
VI. Data and data sources quality dimensions 12
A. RELEVANCE 12
B. ACCURACY (validity, reliability, precision, completeness, consistency) 16
1. Validity 17
1.1 Agreement..... 17
1.2 Missing Information 17
2. Reliability and precision 17
2.1 Assessment of reliability 17
2.2 Assessment of precision 17
3. Completeness 17
3.1 Assessment of completeness: qualitative methods..... 17
3.2 Assessment of completeness: quantitative methods 17
4. Consistency 17
C. TIMELINESS AND PUNCTUALITY 20
D. ACCESSIBILITY AND CLARITY 19
E. COMPARABILITY..... 21
F. COHERENCE 21
VII. Quality report and indicators 24
VIII. Quality methods developed in EU projects: implications and limitations..... 28
A. EURO-PERISTAT 30
B. EUROREACH PROJECT..... 33
C. EHES - European Health Examination Survey..... 34
D. EUROCISS - European Cardiovascular Indicators Surveillance Set 35
E. EUBIROD - European Best Information through Regional Outcomes in Diabetes 36
F. JAMIE - Joint Action on Injury Monitoring in Europe 38
IX. How to improve data quality 39
A. TRAINING AND TESTING..... 39
B. DATA COLLECTION..... 40

C. DATA INPUT AND DATA DELIVERY 41
D. DEFINITION OF NEW VARIABLES OR INDICATORS 42
X. Conclusions and recommendations 43
XI. References..... **Error! Bookmark not defined.**

Working Group

Simona Giampaoli (simona.giampaoli@iss.it), Istituto Superiore di Sanità, Italy

Flavia Carle (f.carle-esterno@sanita.it), Università Politecnica delle Marche, Italy

Angela Fehr (FehrA@rki.de), Robert Koch Institute, Germany

Sabrina Hense (HenseS@rki.de), Robert Koch Institute, Germany

Luigi Palmieri (luigi.palmieri@iss.it), Istituto Superiore di Sanità, Italy

Anna Di Lonardo (anna.dilonardo@iss.it), Istituto Superiore di Sanità, Italy

Serena Vannucchi (serena.vannucchi@iss.it), Istituto Superiore di Sanità, Italy

Chiara Donfrancesco (chiara.donfrancesco@iss.it), Istituto Superiore di Sanità, Italy

Hanna Tolonen (hanna.tolonen@thl.fi), National Institute for Health and Welfare, Finland

Jennifer Zeitlin (jennifer.zeitlin@inserm.fr), INSERM, France

Laura Iannucci (iannucci@istat.it), Italian Institute of Statistics, Italy

Researchers who would like to collaborate to HA5 and WP8

Executive summary

Within an health information system, pursuing and maintaining data quality is crucial to assess population health and health care performance, to monitor time trends of diseases and geographical gradients, to identify gaps and reduce inequalities. Main sources of data which contribute to health information are: 1) administrative databases (hospital diagnoses, drug prescriptions, outpatient visits, exemptions), systematically collected at national level for management of resources and health services purposes; 2) health examination surveys/health interview surveys which provide standardized data on representative samples of the general population; 3) population-based registries which provide standardized data in definite areas under surveillance. Clear definitions, harmonization and data processing procedures in computing indicators are the key issues to ensure reliability and comparability.

This horizontal activity was aimed at: identifying methods of quality assessment in data collection/data sources among previous and running European Projects, particularly in those participating to BRIDGE Health; identifying methods of quality assessment in data processing from different sources to assess indicators; creating an overview of health information areas where quality issues are faced.
Core result of the Horizontal Activity is the technical report that can contribute to build an integrate, sustainable and comprehensive EU health information system.

The work is based on experiences and good practices developed by experts in different European Projects; a questionnaire was sent to the leaders; a literature review of quality methods applied in health data, data sources and health indicators was updated. The report includes a detailed description of quality dimensions of data and data sources (relevance, accuracy, timeliness, accessibility, comparability, coherence), a description of systematic and random errors, methods to assess quality and validity of indicators, implications and limitations, including description of major difficulties encountered to ensure data quality in different European projects. Examples of quality checks for data provided by ad hoc surveys, population-based registries and administrative databases are described as well as main steps to improve quality methods.

The first step to plan and organize a quality data collection is to prepare the manual of operations, which includes a detailed description of exams/questions/data, which should follow international standardized procedures and methods in definitions of the diseases under surveillance, in data collection, and in data processing; training and testing of the personnel involved in data collection and data management ensure good quality data and reduce systematic errors; a report with detailed description of quality checks may help the harmonization of different databases to be included in an health information

system. A prompt feed back to the personnel involved in collecting, harmonizing, and processing data may improve data quality.
Without good data, quality of indicators, quality of studies, and therefore decisions on planning and evaluating preventive programs, health care delivery, resource allocation and research, are severely impaired.

Key points

Quality dimensions, methods of assessing quality in data, data sources, and indicators

I. Introduction

Within an health information system, pursuing and maintaining data quality is crucial to assess population health and health care performance to monitor time trends of diseases and geographical gradients, to identify gaps, and to reduce inequalities.

Main sources of data, which contribute to health information are surveys, registries, and administrative databases.

A survey is an investigation about the characteristics of a given population collecting standardized data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology [1]. Surveys can be distinguished in Heath Examination Survey (HES) and Health Interview Survey (HIS). HES is a population based survey in a random sample of the general population of the country; data collection is based on measures and examinations following standardised methods and procedures (e.g. systolic and diastolic blood pressure, anthropometric measures, functional activities-electrocardiograms, spirometry), biological tests based on centralised laboratory assays (e.g. lipids, glycaemia, haemachrome), standardised questionnaire(s) (e.g. chronic diseases, life styles, pharmacological treatments, family history of diseases, diet, physical performance, cognitive function, etc.).

Health Interview Survey (HIS) is a population based survey that includes interviews on health characteristics (perceived health, diseases, disability) health related behaviour (e.g. smoking habit, physical activity), use of health services. It is based on face-to-face interview and self-administered questionnaires, telephone interviews, postal surveys.

HIS are used to collect information on self-reported and perceived health status, health determinants and health care in samples of the general population; HES provide objective measurements of health related outcomes, but they are high cost and time consuming. Both HIS and HES, if conducted adopting proper standardized and harmonized approaches, may produce comparable and reliable data.

Population-based registry performs a continuous (or periodically) and systematic collection, analysis, interpretation and dissemination of information about occurrence of a disease (e.g. cancer, cardiovascular disease), use and monitoring of medical devices (e.g. joint replacement device, pace-maker) or conditions (e.g. injuries) in a defined population resident in a specific geographic area (state or region).

Population-based registries provide data concerning all cases of a specific chronic disease in a defined population, whether treated at home or in hospital, in whichever season of the year or time of the day they may occur, and also include rapidly fatal cases unable to reach the medical services.

The strength of a chronic disease registry lies in the possibility of validating each single event (or a random sample of suspected events) according to standardized

diagnostic criteria and collecting disease-specific clinical and para-clinical data. Identification of events can be obtained actively checking events in hospital by *hot pursuit* or *cold pursuit*. Hot pursuit means identifying case admissions to hospital usually within one or two days from event onset and acquiring relevant information by visiting the ward or interviewing the patient; *cold pursuit* implies the use of routine and delayed procedures, by means of hospital discharge, review of medical and death records. These procedures allow to provide very high quality data. The weakness lies in the fact that data collection is very expensive and this kind of registers can usually be maintained only for a limited period of time in a defined population of reasonable size. A pragmatic solution derives from the integration by record linkage of administrative databases, specifically mortality and hospital discharge records, for identification of potential events and validation of a sub sample of events to estimate positive predictive values (PPV). PPV, applied to overall potential events, will allow to estimate occurred events and consequently main indicators (attack rate and case fatality).

Administrative data are systematically collected in local, regional, and national databases for purposes related to management of resources, costs, and services (e.g. mortality, hospital discharge records, drug prescriptions, exemptions, etc.). They have not research aims even though, thanks to width and richness of information, administrative data are going to be more frequently used and interconnected with data from research studies (longitudinal studies, HES, HIS, registries) for research purposes.

For example, health care administrative data are generated when contacts between patient and the health care system occur, e.g. a visit to a physician's office, a diagnostic procedure, an admission to hospital, or receipt of a prescription at a community pharmacy. Administrative data are attractive for the advantages they offer in comparison to population studies: the low cost, the large number of individuals included (generally all the population), timeliness, width of periods usually not achievable through surveys. Moreover, the use of administrative data allows to reduce the number of questions or the burden of information to be collected by an ad hoc survey or to collect data on critical information that usually individuals prefer not to respond (e.g. estimating the individual mean income without asking personal income by questionnaire). Finally administrative data can also be used for secondary purposes as sources of information for registries, surveys or follow up of longitudinal studies by integration of data using record linkage procedures. Disadvantages of administrative data are lack of high quality since such databases are collected for administrative purposes and not research, lack of quality control during data collection, their availability, access for researchers, and use is highly limited for privacy and ethical issues. For all these reasons before planning the use of administrative data (e.g. for record linkage in a population-based registry), it is necessary to weigh up the advantages and disadvantages in relation to the specific

research hypothesis taking into account that a huge and capillary checking and cleaning of administrative databases is necessary.

The primary objective of population-based registries and HIS/HES is to provide information about the descriptive epidemiology of a specific health problem, such as the overall rate of the disease in the population and its change over time. This information can be also useful for evaluation of intervention programs and to address health policy decisions. A common goal of population based registries and administrative databases is to produce relevant statistics in order to manage health care, to plan health services and healthcare expenditure, and to provide data on mortality, causes of death and hospital admissions for international statistics.

Harmonization of data processing is of paramount importance to ensure comparable definitions when administrative databases are used, and when ad hoc surveys and registries are pooled.

High quality data are essential for research, for assessing health indicators used for surveillance, prevention, health care, and for supporting health policy makers in their activities and decisions.

II. Aims

The Horizontal Activity on "Data quality methods including internal and external validation of indicators" (HA5) is aimed at:

- identifying methods of quality assessment in data sources/data collection among previous and running European Projects, particularly in those participating in the BRIDGE Health;
- identifying methods of quality assessment in processing of data and data sources used to assess indicators;
- creating an overview of health information areas where quality issues are faced.

III. Approach

This report is based on experiences and good practices developed in previous and running European projects as well as in different Member States. A literature review of quality methods applied in health data, data sources, and health indicators was implemented; reports of EUROSTAT and National Institute of Statistics, and Manuals of Operations were reviewed [2-5].

A questionnaire with the following questions was filled in by work packages leaders of BRIDGE Health: 1) What kind of health information did/do you collect using standardized procedures or methods? 2) Did/do you assess ECHIM indicators following the recommended procedures/methods? In case you did/do not, please specify why. 3) How did/do you assess the completeness of events/information? 4)

How did/do you perform internal validity? 5) How did/do you assess external validity? 6) Have you received training for standardization and data quality? How many training sections and what type of training have you received? 7) What are the major difficulties that you have encountered to assure data quality?

Lessons learnt from EUROCISS [6] and EHES Projects [7] from Non Communicable Diseases Risk Factor Collaboration and World Health Organization (WHO) [8,9] on global burden of disease, and from joint experiences in the fieldwork of researchers involved at national level to collect data, to use different data sources, to process data, and to assess indicators were reported.

The report includes a detailed description of quality dimensions of data, data sources, and quality methods for indicators; quality methods developed in European Projects with implications and limits, including description of major difficulties encountered to assure data quality; main steps to improve quality methods. Final recommendations for the implementation of quality in data, data sources and indicators, including training and testing as key issues to guarantee standardization and data quality, are suggested. Examples of quality methods applied in ad hoc surveys, registries, administrative databases, are described.

IV. Definition of quality

Starting from the definition of quality reported by the Working Group Assessment of Quality in Statistics of EUROSTAT in ISO 8402-1986 as "*the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs*" [10], subsequently this definition was updated and improved and ISO 9000: 2005 (3.1.1 **quality:** *degree to which a set of inherent characteristics fulfils requirements*; 3.1.2 **requirement**: *need or expectation that is stated, generally implied or obligatory*) and in ISO 8000-8: 2015 (Quality Data are data that meet **requirement**s and are **portable** -defined syntax and defined semantic encoding-)

The assessment of data quality is crucial to produce reliable and comparable data and to build health information systems at national and European level.
Without good data, quality of indicators, quality of studies, and therefore, decisions on planning and evaluation of preventive programs, health care delivery, resource allocation, and research, are severely impaired. It is difficult to explain what quality means because of when we are speaking about quality we do not always refer to the same concept. Depending on the context, different aspects of the quality can be taken into consideration. In the case of data, quality can be assessed by evaluation of several dimensions and it is important to consider which aspects have to be measured to obtain a good quality of data.

Case definition

A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome, or other health condition. A case definition must be clear, simple, and concise, allowing it to be easily applied to all individuals in the population of interest. It typically includes both clinical and laboratory characteristics, which are ascertained by one or many methods that might include diagnosis by a physician, completion of a survey, or routine population screening methods. Adoption of the same case definition is important in surveillance to ensure the comparability of data.

Data are characteristics or information, usually numerical, that are collected through observations [11].

Data sources are specific dataset, metadata set, database or metadata repository where data or metadata are available [12]. According to the modality of data collection, sources can be distinguished in administrative, surveys (HIS, HES) and registries.

Indicators are summary measures related to a key issue or phenomenon and derive from a series of observed facts (data). Indicators are useful in identifying trends and in drawing attention to particular issues. They are necessary in monitoring disease and health, setting policy priorities, benchmarking [11].

Methods to assess quality should be referred to data, data sources, indicators, and to data management system. Quality is a multidimensional concept, that means that no one single measure for assessing quality is sufficient. Quality of indicators depends on quality of data and data sources used to generate them. The completeness of the description of each quality dimension of data and data sources can be indicative of the quality of the computed indicator and can help to harmonize data from different sources for improving comparability.

V. Definition of case and indicator

The *case* definition should consider the definition of the event of interest and of the data necessary to define the event (e.g. symptoms, duration, ECG, autopsy are information used for validation of coronary event).

The following characteristics should also be described in the case definition: variable type (e.g., numeric or string) and format (e.g., words, numbers, dates, times, percentages); mandatory fields, which can influence the proportion of missing data; and allowed values (range of allowed values or categories). Case definition depends on the objective of the study.

Case definition is a crucial element in the process adopted to build the indicator from data and data sources, because often the epidemiological case definition differs from country to country or in different time periods and usually is different from clinical definition of the disease.

The *indicator* is a statistical synthesis of data according to defined rules, laws, functions. The definition of the indicator should consider:

1) targeting population of interest (e.g., age, sex, other characteristics),
2) area under surveillance (e.g., geographical area, population size), and
3) period under consideration.

As an example, Table 1 shows different definitions for the *indicator* "prevalence of obesity" from different *case* definitions and data sources.

Table 1. Prevalence of obesity from different data sources

Project	Case definition	Target population	Geographical area	Period	Obesity Prevalence %	
					Men	Women
EHIS(a)	ECHI	> 18 years	Europe	2006-2009	15,5	16,1
EHES (b)	EHES	25-64 years	Europe	2008-2012	18,6	18,1
Italian HES (b)	EHES	35-74 years	Italy, 20 regions	2008-2012	24.5	24.9
HIS -PASSI (c)	self-reported weight , height	18-69 years	Italy	2012-2015	11,1	9,8
HIS- ISTAT (a)	ECHI	> 18 years	Italy	2015	10,8	9
GPs (d)	measured weight, height	35-74 years	Italy	2008	23	23

Legend of table1:
EHIS=European health interview survey [13];
EHES=European health examination survey [14];
Italian HES= national health examination survey [15];
HIS-PASSI=health interview survey "Progressi delle Aziende Sanitarie per la Salute in Italia"-Italy [16];
HIS-ISTAT=health interview survey-Italian National Statistical Institute [17];
GPs=General Practitioners database [18].

(a) ECHI indicator: Proportion of adult persons (18+) who are obese, i.e. whose body mass index (BMI) is ≥ 30 kg/m^2. Calculation: body mass index (BMI) is defined as the individual's body weight (in kilograms) divided by the square of own height (in meters).
Case definition adopted by ECHI: Weight and height derived from European Health Interview Survey (EHIS) questions BMI01: How tall are you? (cm), and BMI02: How much do you weight without clothes and shoes? (kg). EHIS data are not age standardized [13].

(b) EHES indicator: Proportion of adult persons (25-64 years) who are obese, whose body mass index (BMI) is ≥ 30 kg/m^2. Calculation: body mass index (BMI) is defined as the body weight (in kilograms) divided by the square of own height (in meters) measured by scale and wall height ruler following standardized procedures and methods.
Case definition adopted by EHES: Weight and height derived from European Health Examination Survey (EHES) [14,15].

(c) HIS-PASSI indicator: Proportion of adult persons (18-64 years) who are obese, whose body mass index (BMI) is ≥ 30 kg/m^2 [16].
Case definition adopted by HIS-PASSI: Self-reported weight and height derived by the questions submitted by telephone: Can you tell me your height without shoes? Can you tell me your weight without shoes and in underwear?

(d) GPs survey indicator: Proportion of adult patients (35-74 years) who are obese, whose body mass index (BMI) is ≥ 30.0 kg/m^2 or waist circumference is > 88 cm in women and > 102 cm in men [18]. Case definition adopted in GPs survey: weight, height, and waist circumference were measured using international standardized procedures and methods (e.g., standard electronic scale, wall height ruler) with the person in underwear [18].

VI. Data and data sources quality dimensions

Quality dimensions of data and data sources are listed below:

A. Relevance
B. Accuracy (validity, completeness, consistency)
C. Timeliness and punctuality
D. Accessibility and clarity
E. Comparability and
F. Coherence

All the dimensions which influence the quality of data are interrelated and influence each other, therefore partial overlaps are possible in the separate description of each dimension.
All quality dimensions are relevant for the use of primary data (e.g., ad hoc data collection by HIS or HES) or secondary data (already collected data such as administrative data or data collected for different aims). The list of dimensions can be different from one study to another, because different subdivisions or grouping of quality dimensions can be adopted. In this report we refer to the quality list used by Eurostat [10].
Following, the different dimensions of quality are described in detail and examples are given.

A . RELEVANCE

Data should be relevant to the purposes for which they are to be used and respond to potential users' needs. This will require a periodic review of requirements to reflect changing needs. The production of statistics which have ceased to be of

interest for objectives should be abandoned. Statistics about population health and health care performance are important if contribute significantly to assess morbidity/mortality, are associated to high rate of utilization, support the planning of health systems and economic resources.

Before starting a study investigators should give a response to important questions about the users, about the needs and about the way. About the user: 1) Who and how many are the users? 2) How important is each of them?; about the needs: 1) what are the needs that the users expect to be satisfied?; about the way: 1) how far are the needs met? (16)

B . ACCURACY (validity, reliability, precision, completeness, consistency)

Accuracy indicates the closeness of the estimated value to the true value. The accuracy is a multi-faced dimension of quality because it includes different aspects which are in some cases interrelated each other. Validity, precision and reliability represent some aspects that contribute to accuracy, as well as completeness and consistency.

B1 Validity

Validity may be applied to a method or to an instrument and indicates the extent to which method/instrument measure or perform what are designed to measure/perform. During data collection measurements errors may occur; they influence the validity because generate values different from the true ones. For example, a laboratory method is valid when obtained values are within an established range. Errors can be due to instruments (device or questionnaire), to respondents (giving a wrong answer consciously or unconsciously), to researcher/technician (laboratory operator using wrong methodologies, interviewers Influencing the answer). Instruments and researcher/technician errors can be evaluated by repetition of the measurement with a different instrument or laboratory test or interview performed by a different person. It is more difficult to assess error from the respondent, as it requires different sources for the same respondent. During the data editing, data inconsistencies can be detected; they suggest the presence of errors. The proportion of records that fail each edit is an indication of the quality of the data collection and data processing.

Measures of validity are:

B1.1 Agreement

A method to measure the validity is the agreement with the gold standard, e.g. agreement of hospital discharge records (HDR) with medical records. The accuracy can be estimated by two measures, sensitivity (proportion of true positives that are correctly identified) and specificity (proportion of true negatives that are correctly identified).

B1.2 Missing information

Missing information concerns proportion of registered cases with unknown values of variables, which indicate problems with data collection, and are due to inadequate case histories, investigation or ambiguity in the medical record. A large proportion of missing data may influence results of the study. For example, incident coronary events rates will be underestimated if a significant proportion of registered events appear in the "insufficient data" category rather than in "definite" or "possible" or "probable" category.

Missing data can be categorised as: 1) Missing Completely At Random (MCAR), 2) Missing At Random (MAR), 3) Missing Not At Random (MNAR).

Data are MCAR when the probability of missing data on a variable is unrelated to other measured variables and to the values of the variable itself. Missingness of data is completely unsystematic, i.e. the probability that an observation is missing is not related to any other patient/people characteristic. An example of a MCAR mechanism is given by the fact that a laboratory sample can be lost, so the resulting observation (laboratory test) is missing. In this case, there is no relationship between missing observations and any other value of the examined person in the data set.

Data are MAR when the missingeness is related to other measured variables, but not to the values of the variable itself. Therefore this type of missing data is confusingly called MAR, even if there is not random missingness. For example, if men are more likely to refer their weight than women, weight is MAR. Missing values of weight variable are not completely random but depend from the sex variable. Other example is the proportion of cases with missing data which tends to be greater amongst elderly population.

In a variable, data are MNAR when there is a relationship between the propensity of a value, or a range of values, to be missing and a specific value, or ranges of values. An example of a MNAR mechanism would be the collection of information on annual income. Typically, those with higher incomes may be less willing to reveal them, so the resulting observation is missing. In this case, missing values of annual income do not depend on other observed variables, but depend on a characteristic (higher income) of the incomplete variable itself. Another example of a MNAR mechanism would be the greater proportion of cases with missing data in elderly patients.

Population based Registries: the lack of accuracy (at recording) can derive from missing basic variables such as date of the event, site of cancer, sex, residence, date of birth; in the case of MAR, e.g., when the date of the event is missing for non-fatal events more than for fatal events; in the case of MNAR, e.g., when missing dates of the event are more frequent in specific seasons of the year in comparison to other periods or the overall year. The use of a low sensitive or low specific instrument, e.g. signs or symptoms typical of stroke instead of diagnosis of stroke by nuclear magnetic resonance/computed axial tomography.

Health interview/Health examination survey: Lack of accuracy can derive from using a low sensitive or low specific instrument, non-adherence to standardized protocols due to the person's unawareness of own condition, or unwillingness to share information. Deficit in the instruments or in the laboratory performance may lead to a systematic bias and thus lack in accuracy e low validity.

B2 Reliability and precision

Reliability is when the repetition of a method in the same conditions gives the same results. A manifestation of the reliability is the precision, that indicates how much close the measured values are to each other. Therefore, a measure is reliable when repeatedly applied to the same population, the same result is obtained in a high proportion of times.

B2.1 Assessment of reliability

For assessing reliability two procedures can be applied: test-retest procedure (the measuring procedure is performed twice on the same object, and the agreement between the results quantifies the reliability); inter-rate procedure (the measuring procedure is performed by different evaluators in independent measurements on the same object and the agreement between the results quantifies the reliability).

B2.2 Assessment of precision

The precision is the agreement among data collectors and could be evaluated re-extracting from the same source (e.g., at a later time or by someone else), or from two or more different sources, the events (e.g., coronary event from registry and acute myocardial infarction from hospital discharge registry) and assessing discrepancies by comparison of data sources.
In HES/HIS, in order to increase precision, the following methodologies could be used: regular monitoring of measurers performance; regular reviewing of

instruments; periodic assessment of the laboratory performance; data input with variable control for ranges.

B3 Completeness

Different definitions of completeness of data sources exist. We adopt the following: completeness describes the degree to which values are present in a data collection [20]. Completeness can be referred to the event or to the information.
In registries completeness of the events can be assessed by the 'coverage rate' and describes the extent to which all the expected events are registered. It is measured as percentage of registered events divided by the expected ones. The best condition is 100%. Checking of the coverage allows to identify missing or duplicate events. A method to control the completeness of the event is the record linkage with other source of information (Hospital Discharge Records linked with mortality). In a population-based registry of coronary events, non-fatal events occurred out of the surveillance area are rarely recorded; in that case to ensure completeness databases of GPs may be used to catch the event after the acute phase.
Completeness can be also referred to the information recorded for the case definition and validation. For example in the case of coronary event, ICD codes defined in HDR and/or mortality and date of the onset of the event are necessary, as well as for validation, according to MONICA criteria [19], symptoms, enzymes, coded ECG (according to Minnesota code), autopsy (in case of fatal events) are all necessary.
Case-finding may be problematic in patients with multi-morbidity, for example in elderly where the presence of multiple pathologies can make a single diagnosis by hospital admission/discharge record more difficult to be classified.
In HIS and HES, completeness can be invalidated by coverage errors caused by divergences between the target population and the frame population, defined as "the set of population units which can be actually accessed and the survey data that refer to this population". Ideally, the frame population coincides with the target population; this situation is difficult to obtain because the frame population is smaller or larger than the target population.

The coverage errors can be distinguished in:

1) Undercoverage: persons who are not accessible by the frame, for example resident in a given area, but temporarily out of the area;
2) Overcoverage: persons who are accessible by the frame but who are not belonging to the frame (e.g., Inclusion of dead people);
3) Multiple listings: persons who are present more than once in the frame (e.g persons with two or more telephones);

4) Incorrect auxiliary information: persons with wrong information.

Coverage errors can lead to bias and underestimation of the variance. Overcoverage, multiple listings, and incorrect auxiliary information can be avoided checking the information about each unit of population. Undercoverage is more difficult to detect and specialised frame quality reviews are necessary to discover them.

The main methods to evaluate completeness are qualitative and quantitative.

B3.1 Assessment of completeness: qualitative methods

Qualitative methods estimate the degree of completeness by comparison of own study data to other data sources or over time.

- Historic data analysis: they include the comparison of data with those observed in other populations that have been expected to manifest similar rates. Differences from regional standards may reflect specific local variations in prevalence of risk factors or in the use of different intensity in implementing the screening for some high risk conditions; anyway systematic discrepancies (across several sites) provide evidence of possible under-registration (or over registration due to inclusion of duplicate records);
- Historical verification: proportion of cases of disease with no other information except death certificate; it is often cited as an indicator of incompleteness of registries; these cases represent the residuum of cases for which no other information, other than death certificate, could be obtained;

- Ratio between mortality and incidence rate: it is an example of independent case ascertainment method; when the ratio between mortality and incidence values are greater than expected, they lead to a suspicion of incompleteness (incident cases missed by the registry);

- Number of sources/notifications per case: using as many sources as possible reduces the possibility of unreported cases, increasing the completeness of the registry data. Efficient record linkage is essential.

B 3.2 Assessment of completeness: quantitative methods

For registries or HIS/HES, quantitative methods assess the extent to which all eligible cases have been registered.

- Independent case ascertainment: to recheck the sources of information in order to detect any case missed during the registration; to use one or more

independent sources of cases and compare the databases (cases recruited in international clinical follow-up study, patients enrolled into a multicentre clinical trial, databases of General Practitioners, patients enrolled in cohort studies, patients enrolled in multihospital case-control studies, patients enrolled in community screening).

- Capture recapture method: this method was originally developed as a method for estimating the size of a closed animal population. The procedure can be described as follows: at one time as many animals as possible in a defined area are captured, tagged and released (capture stage); at a later time this procedure is repeated (recapture stage). The number of animals in each sample and the number of animals common to both samples (recaptured) are used to estimate the overall number in the total population (assuming that capture and recapture are independent); capture and recapture methods have been advocated to be used in epidemiology to estimate the completeness of ascertainment of disease registries [21]. Typical applications include estimating the number of people needing particular services (i.e., services for children with learning disabilities, services for medically frail elderly living in the community), or with particular conditions (i.e., illegal drug addicts, people infected with HIV, etc.). To take a concrete example, in order to implement a register of children with Type 1 diabetes, children are identified from hospital admission records, from GPs (family doctors) databases, and from the records of the local Diabetes Association. None of these sources have a complete list, but by putting them together it is possible to obtain and to estimate how many children are identified in total, how many children with Type 1 diabetes are living in the vital community, and the completeness of ascertainment of each data source.

Examples of measures of completeness for registries and HIS/HES:

- Population based registries: cancer registry performs the percentage of cases without microscopic diagnosis (completeness of information); cardiovascular registry performs number of fatal cases with insufficient data (completeness of information), number of patients admitted in hospital out of the surveillance area (completeness of cases);
- Health interview/Health examination survey: assessment of the proportion of persons examined by the eligible population (participation rate); incompleteness of collected information or exams (percentage of missing data for each variable).

B4 Consistency

This dimension describes the plausibility of data within the agreement with values in other data, and same database.
Consistency of data can be checked within a variable (internal consistency) or between different variables or at two or more point in time (historical consistency). Most of quality checks performed for single variables concern for format and allowed values, but can be also more specific and referred to dates and classifications. Consistency can be affected by processing errors that can occur between data collection and the beginning of statistical analysis. Processing errors can involve each single step: coding, data entry, data editing, imputation, etc.
To evaluate the impact of the error on the final statistics, data should be recoded/re-entered to the computer or re-imputed and errors corrected. A correction can be performed by assigning multiple imputed values to wrong or missing data.
Sometimes it is difficult to distinguish between processing and measurement errors since they can overlap.

Example of internal consistency (single variable):

- Dates: if this variable contains day/month/year all together, it is necessary to check each of them. If the month is January, March, May, etc., the range values for the days must be (1-31). If the month is April, June, September, the range values for the days must be (1-30). For example, if 31/09/2016 is found, there is a mistake, because this date doesn't exist;
- Age at diagnosis: it is usually calculated as year of diagnosis minus year of birth; this procedure rounds age by 6 months. The age range values must be positive and between 0 and 100.

Example of consistency between variables:

- Population–based registries
    - consistency between dates (birth, diagnosis, death, autopsy): date of birth should be <= date of diagnosis; date of birth should be <= date of death; date of death should be <= date of autopsy;
    - consistency between age and sex: some diseases/conditions occur almost exclusively in specific age groups (children, adult, elderly); others only in men or in women (i.e., prostate hypertrophy, menopause).
- Health interview/Health examination Survey

- consistency concerning a condition: the consistency can be checked by assessing related information in cross tables (e.g., a never smoker cannot have information on current number of cigarettes/day);
- consistency between ages: the computed age at diagnosis should be identical to the registered age.

C. TIMELINESS AND PUNCTUALITY

Timeliness relates to the rapidity of data collection, to processing and reporting reliable and complete data, and to length of time between collection of data and dissemination of results. Speedy access to data and indicators is a priority and of clear benefit to health providers and researchers. Early provision also enhances the reputation of the registry and the data source. However there is a trade-off between timely data and the extent to which they are complete and accurate. Punctuality is referred to the lag time between the scheduled date, established in a calendar (reference date) and the actual delivery dates. Usually data collections have predefined time intervals (year) that should consider a balance between timeliness and completeness. Timeliness and punctuality influence the frequency of released statistics that depend on the time needed to plan and perform the survey or to collect complete data for a registry, on the time to perform the quality control of data, statistical analyses and interpretation of results.

In population based Registries, timeliness is referred to the length of time between the occurrence of the event and the available statistics; e.g., registries are constantly updating their databases as they receive reports, but some notifications, especially those from death certifications, arrive long after the case was diagnosed. Timeliness in the publishing of results can be related to completeness: registries may have a tendency to delay the dissemination of their results in order to achieve better completeness, so that there can be competition between the two goals. Availability of quantitative methods that allow to estimate completeness at a given stage of the registration process could help registries to decide in a more rational way when data can be considered ready for publication.

In *surveys*, timeliness is referred to the length of time between the collection of data and dissemination of the results (e.g. prevalence and means). Usually, in surveys, timeliness and punctuality may be affected to delay in the collection of data and quality control of data. A delay of some years has usually less importance in the etiological study implemented to support policy-makers in the planning of preventive actions, than in surveys aiming at evaluating efficacy of preventive or health care performance.

D. ACCESSIBILITY AND CLARITY

Accessibility is referred to the facility of access to statistics by users as well as the suitability of the form or medium through which the information can be accessed. Therefore accessibility is related to different aspects of the dissemination such as the distribution channel, marketing conditions (i.e., copyright, etc.), ordering and delivery procedures, pricing policy, availability of micro or macro data, formats (i.e., paper, files, CD-ROM, Internet, etc.), etc.

Clarity is referred to the presentation of statistics in an understandable and clear manner. Clarity presuppose that statistics are accompanied by textual information and explanations, graphs, figures and other illustrations, offered as assistance to users by the data provider. Usually documents tend to be written using specific language and communication way understandable only to experts. A rigorous scientific communication is essential to share results among scientists and to provide the evidence for guidelines, national and international reports, and publications, and to support policy makers in the planning of prevention and care programs. Efforts should be also pursued to make results user-friendly also to other stakeholders, i.e. health operators, journalists, patient associations and interested persons. If data cannot be accessible or the associated metadata are not understandable, the most accurate and coherent data have a little value too.

Accessibility and clarity are the most neglected quality dimensions, even though the value of a registry or a survey is limited by the extent that the information collected is fully analysed and then clearly disseminated to relevant audiences.

E. COMPARABILITY

Comparability is the extent to which differences between statistics from several geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics [17]. Comparability of data is a crucial aspect to allow for reliable conclusion and benchmarking between countries/regions and over time. Factors responsible of loss of comparability are related to: 1) use of different definitions, or 2) use of different procedures or measuring tools.

Comparability can be ensured with proper approaches for standardization and harmonization of collected data. These have to be clearly defined and described in detail such as in the report from HA4: Standardisation methods for the collection of health information.

For disease specific data the basic requirement is the standardization of the case definition. For example, in the WHO registry for AMI (1976), in the WHO-MONICA (MONItoring trends and determinants in CArdiovascular diseases [19]) Project (1980-2000), and in the EUROCISS (European Cardiovascular Indicators Surveillance Set) [22,6]) recommendations (during 2000s), three different definitions of events were adopted:

- WHO registry for AMI since 1976; only anamnestic history was collected and the electrocardiogram (ECG) was not used for the identification of the event.
- The WHO MONICA Project where disease specific data collection was based on hot and cold pursuit for the identification of events: enzymes, evolution of ECG, symptoms, and autopsy were used as diagnostic criteria for the validation of the event [19].
- The EUROCISS Project recommended to follow a step-wise procedure based on: standardised data collection, appropriate record linkage between hospital records and mortality or other sources of information (e.g., GPs, drug dispensing register), a selection of a random sample of suspected events and their validation applying the MONICA and the new diagnostic criteria of European Society of Cardiology/American College of Cardiology (ESC/ACC), assessing the positive predictive values (PPVs) for the International Classification of Diseases (ICD) codes reported in death certificates and in hospital discharge records, and using PPV to estimate the number of current events [22,6].

Precise knowledge of current and historical registration procedures, methods and definitions are also of great importance in the analysis of the geographical and temporal variation. The geographical comparability is referred to the comparison of similar surveys, analysing the same phenomena, but involving population of different geographical areas or conducted by different organisations, eventually including different times too.
Regarding comparability over time, data collected in a specific reference period cannot be fully comparable with data of following periods if changes occurred and consequently a break in the time series is introduced. Changes due to modification of references, concepts or measurement process should be documented and their impact should be assessed. An example is reported for Acute Myocardial Infarction (AMI): incidence of the disease has changed and evolved over time, because severity of non-fatal disease is changed for the availability of more sensitive tests; more treatments for the acute phase are available today and detection of mild events (Acute Coronary Syndromes) is more feasible in comparison to the past (troponine test).

To assure comparability, particular attention should be given to:

- definition of area under surveillance and target population;
- used case definitions (e.g., ECHIM indicators, international guide-lines), rules for coding events (e.g., sources, algorithms, version of international classification of diseases), definition of multiple events in the same individual (e.g., for coronary or cerebrovascular recurrent events the threshold of the 28^{th} day is internationally used), and the date when the disease becomes an event -'onset of the event'- (first date of admission at hospital, first prescription, diagnosis by the GP, until death date);

- asymptomatic event detection: incidental detection of event (e.g., asymptomatic myocardial infarction or cancer can be detected during a Health Examination Survey or a screening program);
- autopsy performed with or without consent: in some countries autopsies are performed frequently for medical, scientific or educational reasons without consent; in these countries or regions the number of events may be higher.

In order to improve comparability of data, extensive description of methods used to produce data, data sources, and methods used to compute indicators should be published and updated periodically. Description of methods should be published in web-sites but also in the research institute reports and national and international journals, and project protocols in order to avoid loss of methodologies adopted over time (the pages of the web may change over time).
Statistics from different domains can be compared taking into account that different concepts could be used (definition of characteristics, reference period, etc). All the differences should be reported and their effects evaluated.
Measures and remarks on comparability of data should go together with the list of health indicators; e.g., remarks on comparability between countries, over time and with national data, are provided for indicators in the ECHI shortlist, where applicable. Comparability of data is a crucial aspect to allow reliable conclusions and to perform benchmarking between countries/regions and periods.

F. COHERENCE

Coherence of registries and surveys results and statistics is their adequacy to be reliably combined in different ways and for various uses and purposes. The coherence of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time [23].
The use of standard concepts, classifications and target populations promotes coherence, as does the use of a common methodology across registries and surveys. Coherence does not necessarily imply full numerical consistency. Coherence reflects the degree to which the data and information from a single study are brought together with other data and information, and how they are logically connected and completed. Fully coherent data are consistent internally, over time and across products and programmes. Where applicable, the concepts and target populations used or presented are logically distinguished from similar concepts and target populations or from commonly used notions and terminology.

Coherence can be assessed for different areas:

a) coherence between temporary and final statistics: establish if the difference between temporary and final statistics has effectively a meaning;

b) coherence of annual and short term statistics;

c) coherence of the statistics in the same domain: when a group of statistics, possibly of a different type, measures the same phenomenon with different approaches;

d) coherence among survey statistics compared at national level.

Coherent statistics validate data each other and have the potential to be validly combined and used jointly.

VII. Quality report and quality indicators

In order to assess data quality, first of all a clear picture of data quality is needed. Definitions and dimensions discussed in the previous section VI are preconditions. Secondly a report on data quality is indispensable, which reflects data characteristics by quality components and presents data features according to data quality requirements.
The quality report summarises the most important information on quality. The measurable aspects of the quality can be characterised by indicators and the information in the report helps to understand the limitation of a given product. Self-assessment, audits and peer reviews are based on information from quality indicators and reports, process variables, and user surveys. Furthermore, they sometimes might use specifically designed checklists in order to present the information needed in a more structured and accessible way. According to the Eurostat Standard quality report for Labour Force Survey, a detailed quality report should include:

- Administrative information
  - The name, the reference period and the periodicity of registries or surveys
- General Description
  - The design and methods used for the survey
  - A description of the methods used during the survey process (classification, sampling design, data collection process, etc.)
- Relevance
  - A description and the classification of the users
  - A description of the variety of the users' needs

- Main results regarding the satisfaction of users
- *Indicators:*
  - user satisfaction index
  - rate of available statistics

- Accuracy
  - Sampling errors
    - Order or magnitude (or at least sign) of the bias of the main variables
    - Methodologies applied for variance estimation
    - *Indicators:*
      - estimated coefficients of variation (CV) for the statistics
      - imputation rate and ratio
  - Coverage errors
    - Type and size of coverage errors
    - Information about the frame: reference period, updating actions, quality review actions
    - *Indicators:*
      - over-coverage and misclassification rates
      - geographical under-coverage ratio
  - Measurement errors
    - The measurement errors identified and their extent
    - Indications about the causes of measurement errors
  - Processing errors
    - A summary of the processing the data are subjected between collection and production statistics
    - Processing errors identified and their extent
    - *Indicator:*
      - average size of revisions
  - Non-response errors
    - Non-response; unit and item non-response rates for the main variables, both unweighted and weighted
    - Imputation methods used (if any)
    - Indications about the causes of non-response

- Timeliness and Punctuality
  - *Indicators:*

        - the average timeliness of data (time lag between the end of reference period and the date of first results; time lag between the end of reference period and the date of the final results)
        - The data frequency and average data freshness
        - Punctuality of time schedule of effective publication
    - The reasons for late delivery

- Accessibility and Clarity
    - A summary description of the conditions of access to data: media, support, marketing conditions, existing service-level agreement, etc.
    - A summary description of the information accompanying the statistics (documentation, explanation, etc.)
    - *Indicators:*
        - Number of publications disseminated and/ or sold
        - Number of accesses to databases
        - Rate of completeness of metadata information for released statistics

- Comparability
    - The reference period of the survey where the break occurred
    - The differences in concepts and methods of measurement before and after the break
    - *Indicators:*
        - Length of comparable time-series
        - Number of comparable time-series
        - Rate of differences in concepts and measurement from European norms
        - Asymmetries for statistics mirror flows

- Coherence
    - Coherence of statistics in same domain: summaries of the mirror statistics
    - Coherence with National Accounts: a summary of the comparison
    - *Indicators:*
        - Rate of statistics that satisfies the requirements for the main secondary use

- Cost and Burden
    - Cost supported by National Statistical Institute (NSI).
    - Response burden.

Quality indicators are specific and measurable elements of statistical practice that can be used to characterise the quality of statistics. The indicators as simplified and generally quantified measures – calculated according to clear rules – intend to characterise quality features of data.

Quality indicators make the description of a product by quality components more informative and increase transparency. The statistician or the user can assess the quality of different surveys or the same data in different periods by using the quality indicators.
Some quality indicators should be produced for each output in line with the frequency of production or publication (for example, standard errors should be calculated related to each new estimate). However, some quality indicators should be produced once for longer periods, and should only be rewritten when major changes occur (e.g. lag time between the end of the reference period and the date of first results). The frequency of the indicators calculation depends on the purpose of quality indicators (e.g., monitoring the quality over time) or on the survey or publication frequency.
Measuring quality should be an inherent part of any statistical production process and should not be a separate activity carried out after the statistics are produced or when someone needs it. This is not only for cost and time efficiency, but for the purposes of quality improvement.
The way for the dissemination of quality reports will have to be further integrated in the dissemination policy of registries or surveys. Registries and surveys managers and experts should work together in developing the quality reports and indicators. In the short-term, when a detailed quality report is available for internal use, it might be reasonable to extract information that are possibly useful for external users and to prepare standard explanation notes how to exploit this information. Quality measurement burden should remain at an acceptable level both in terms of expenses and time use. A good decision is to start with some components and indicators in some main dimensions. Indicators may be misleading or focus only on a part of the phenomena. For example concerning accuracy, generally much more attention is given to sampling error than non-sampling error; however, the latter may have a dominant role, even though it is difficult to be measured. Later on, the use of this simplified quality report can be extended to all dimensions. Finally the level of detail and content of the report has to be improved to meet the special needs of the users of the given products. Preparation of quality reports needs time and efforts; they are efficient only if they are used in a proper way: their level of detail, structure and form must be suitable for the targeted users (not too long for managers, easy to understand for public users, importance of a component should be weighted up by considering which users will directly or indirectly require information on it, etc.). When quality indicators are used to inform users on the quality of statistics, it is recommended to include qualitative statements helping to interpret quality information and to summarise the main effects on the usability of the statistics.
The use of the same quality report template with a limited set of standard indicators for different products will support transparency. Quality indicators and quality reports include the most important information on the quality; this is the

reason why they are used very often as the basic documents for self-assessments or audits.

VIII. Quality methods developed in EU projects: implications and limitations

Some of the EU funded projects on health information are based on ad hoc survey (primary use of data) or on integrated health information systems (secondary use of routine data) or on linkage of different sources of information and ad hoc survey (secondary and primary use of data), therefore a short description of the projects and literature review is presented together with possible implications, limits and recommendations according to quality dimensions.

Other projects did not collect data, but simply recommended procedures and methods to collect high quality and comparable data and methods to improve and maintain quality over time.

Table 2 describes the information collected by the questionnaire sent to the principal investigators within BRIDGE Health regarding data collection tool, quality methods described in the manual of operations and references.

Table 2. Quality methods in the EU Projects. Sources: BRIDGE Health questionnaire and references

EU Project	**Data collection tool**	**Primary or secondary use of data**	**Quality methods adopted (manual operations, websites literature)**	**References**
EHLEIS	HIS, Registries	Secondary	*Completeness* of information and events; *Data consistency/coherence;* *Internal validity;* *Personnel training.*	
	Different data sources	Secondary	*Completeness* of information and events; *Data consistency/coherence;* *Internal validity;* *External validity.*	
EHES	Population based health examination survey	Primary	*Completeness* of information and events; *Personnel training.*	Tolonen H [24] Tolonen H [25]
Injury Surveillance-JA on Monitoring Injuries in Europe	Routine care data and some disease registries	Primary and Secondary	*Completeness* of information and events; *External validity;* *Personnel training.*	Nwaru BI [26]
Injury surveillance platform/EuroS	Routine/ad ministrative data sources	Secondary	*Completeness* of information and events; *Internal validity;*	EuroSafe [27]

afe			*External validity.*	
EUROCISS	Register data on AMI and Stroke	Primary	*Representativeness;* *Standardised procedures and methods;* *Completeness* of information and events; *Internal validity;* *External validity;* *Personnel training.*	Madsen [22] Giampaoli [28]
	Health examination survey	Primary	*Representativeness;* *Completeness* of information and events; *Internal validity;* *External validity;* *Personnel training.*	Primatesta [29]
COPHES/DEMOCOPHES OBELIX/ENRIECO FLEHS	HIS; Examinations surveys (hair and morning urine)	Primary	*Completeness* of information; *Internal validity;* *Personnel training.*	Becker et al (2014) Casteleyen et al (2015) Esteban et al (2015) Exley et al (2015) Fiddicke et al (2015) Schindler et al (2014)
EURO-PERISTAT	Different data sources	Secondary	*Completeness* of information and events; *Internal validity;* *External validity;* *Training of personnel*	Gissler M [30] Euro-Peristat [31]
ECHO-European Collaboration for Health Care Optimization	Routine/administrative data	Secondary	*Completeness* of information and events; *Internal validity;* *External validity.*	ECHO [32] ECHO [33] ECHO [34]
EUROHOPE	Routine/administrative data	Secondary	*Completeness* of information and events; *Internal validity;* *External validity.*	Häkkinen U [35] EuroHOPE [36]
ECHI-1, ECHI-2, ECHIM, JA-ECHIM	Different data sources	Secondary	*Completeness* of information and events; *Internal validity;* *External validity;* *Personnel training.*	ECHI [37]
EuroREACH	Different data sources	Secondary	*Completeness* of information; *Internal validity.*	EuroREACH [38]
EUBIROD Network	Different data sources	Secondary	*Completeness* of information and events; *Internal validity;* *External validity.*	Carinci F [39] Cunningham SG [40]

A. EURO-PERISTAT

EURO-PERISTAT Project monitors health and care of mothers and babies during pregnancy, delivery and post-partum period [30]. Thirty-one countries currently participate, including all current EU member states, Iceland, Norway, and Switzerland. Bulgaria and Croatia joined the network in 2015.

Indicators are grouped in four categories: neonatal health; maternal health; population characteristics or risk factors; health care services. Each category contains core indicators, recommended indicators, and further indicators in development. The indicator definitions and data are available from the Euro-Peristat website (www.europeristat.com)

Table 3. EURO-PERISTAT, list of indicators, updated 2012

Category	**Core**	**Recommended**	**Further development**
Neonatal health	C1-Fetal mortality rate by gestational age, birth weight, plurality C2-Neonatalmortality rate by gestational age, birthweight, plurality C3-Infant mortality rate by gestational age, birthweight, plurality C4-Birthweight distribution by vital status, gestational age, plurality C5-Distribution of gestational age by vital status, plurality	R1-Prevalence of selected congenital anomalies R2-Distribution of APGAR score at 5 minutes R3-Fetal and neonatal deaths due to congenital anomalies R4-Prevalence of cerebral palsy	F1-Severe neonatal morbidity among high risk infants F2-Prevalence of neonatal encephalopathy F3-Causes of fetal and neonatal death other than CA
Maternal health	C6-Maternal mortality ratio by mater age	R5-Maternal mortality ratio by cause of death R6-Prevalence of severe maternal morbidity R7-Prevalence of tears to the perineum	
Population characteristics or risk factors	C7-Multiple birth rate by number of fetuses C8- Distribution of	R8-Percentage of women who smoke during pregnancy R9- Distribution of	

	maternal age C9- Distribution of parity	mothers' education R10-Distribution of households' occupational classification R11-Distribution of mothers' country of origin R12-Distribution of mothers' body mass index (BMI)	
Health care services	C10-Mode of delivery by parity, plurality, presentation (of fetus), previous caesarean section	R13-Percentage of all pregnancies following sub fertility treatment R14-Distribution of timing of 1st natal visit R15-Distribution of births by mode of onset of labour R16-Distribution of place of birth by volume of deliveries R17-Percentage of very preterm infants delivered in units without a NICU R18-Episiotomy rate R19-Birthswithout obstetric intervention R20-Percentage of infants breast-fed at birth	F4-Neonatal screening Policies F5- Content of antenatal care

Implications

This Projects demonstrates the feasibility and value of using indicators to monitor perinatal health at a European level as data on these indicators have been collected for three publications (for the year 2000, 2004 and 2010). However, the results of this project also illustrate that continuing international collaboration is needed to improve the consistency of definitions and to prioritise the development of methods for collecting data for many perinatal health indicators. While all indicators can be provided by at least a few countries, no country can provided the full set of indicators. The core indicators are more widely available.

Quality is ensured by having common pre-established definitions, collecting data as numbers of births (as opposed to percents already calculated), so that numbers can be cross-checked across indicators (numbers of live births, stillbirths, etc) and so that the coordination team can be sure that the percentages are calculated in the same manner. The number of missing data are also collected for each indicator and reported in the tables.

When countries cannot provide the Euro-Peristat data using the agreed definition, they are requested to provide data using their national definition and describe their definition.

For each data source, are requested information on the population, coverage and whether evaluations of its quality (coverage, completeness, external validity) have been undertaken.

Finally, quality is maintained by using the data to produce scientific articles. While analysing the specific indicators and comparing them with others as well as the scientific literature, outliers are identified and discussed with the scientific representatives and other experts. This allows the group to get a better sense of the strengths and limitations of our data. More than 50 articles have been published in peer reviewed journals using data from the Euro-Peristat project.

Limits

The quality method problems in PERISTAT are related to case definition and coverage of data collection, which influence completeness of data collection.

Case definition: WHO criteria for stillbirth is foetus with a birth weight of 500 gr, or -if this is missing, a gestational age of 22 weeks. This legal limit for registration is not always respected by different countries that used more than 22 weeks (24 weeks Hungary, Portugal, UK, 25 weeks + 5 days Italy, Luxembourg, Spain, 28 weeks Greece, Sweden).

A standard definition of "stillbirth" is essential when international comparisons are made. For livebirth most countries have no limits for weights and gestational age, while others have defined values for inclusion.

However, Euro-Peristat collects the information required to create comparable indicators by using thresholds that can be applied in all countries [41,42].

Another discrepancy depends on the different criteria used by civil registration and health registration system for the inclusion of non – residents. Civil registration is limited to permanent residents only, while health system comprises all the events. This difference can influence the data where a large number of non-residents are present (refuges, immigrates, visitors, asylum seekers).

Recommendations

Given the large proportion of deaths before 28 weeks (over one-third of all deaths), it is essential to improve information systems in Europe by developing common guidelines for recording births and deaths at 22 weeks.

Continued collection of the full set of Euro-Peristat indicators provides an impetus for countries to improve their national systems so that they can produce key indicators which are available in other countries and learn from comparison of their perinatal health systems and health with their neighbours.

B. EUROREACH PROJECT

EuroREACH is a project with the objective to improve access to and use of healthcare data and to enhance cross-country comparisons of health system performance [43]. It has produced the "Handbook to access health care data for cross-country comparisons of efficiency and quality" and then a digitised form of the Handbook, the Health Data Navigator – HDN (www.healthdatanavigator.eu) [44] to facilitate the dissemination of the outputs to the research community. The EuroREACH criteria appraisal used by HDN were: governance, access to database, coverage, linkage, data quality, strengths and weaknesses. Referring to data quality the following points are considered: entry errors (multiple common entries, redundancies that lead to incorrect/incomplete data), breaks (changes of standards that lead to incompatibility); consistency of terminology (differences in how data are described when collected).

The EuroREACH Case Study on diabetes care has pointed out a number of problems that can have an effect on comparability of data from different health systems.

a) The source of data can be delivered by different systems and this influences the comparison. A system with a centralized laboratory is more likely to provide laboratory data that allows for quality monitoring than a system in which laboratories contract with payers.
b) There are differences in coding systems between countries. Therefore a complete list of all codes that can identify a specific diagnosis should be available.
c) The type of hospital reimbursement can influence the level of detail with which diagnoses and procedures are coded. Prospective payment formula of hospital reimbursement generates data with more precise diagnosis and procedure coding than that using a per-day payment rate.

The EuroREACH case study for diabetes is an useful example to explore practical problems faced when administrative data are used to compare the performance of chronic disease care in an international context.

Due to problem of comparability and availability, only three countries participated (Finland, Estonia, Maccabi - Israel) in the study comparison of diabetes care at international level.

The study has revealed differences in the health care system/health care definition of diabetes depending on severity or disease stage. Among the participating countries of the EuroREACH diabetes study, Estonia did not have any disease Registry (except cancer) and provided data from the administrative health service Database of Estonia Health Insurance. No data linkage was necessary for the identification of the cohorts. Instead, Finland has a system of medical registers and administrative database. Linkage was performed by a unique national personal

identification number. It is necessary to underline that Registers for primary care have been created recently and quality of diagnosis coding is not always satisfactory. Consequently, the definition of diabetic population requires the use of medication data from the Social Insurance Institution. For this specific study, a linkage-based research data base of Finnish diabetic patients was already available, hosted by the National Institute for Health and Welfare and its partners. Israel has a patient system covering visits to hospitals, physicians, prescription drug purchase, laboratory testing (with results), imaging studies. For the international data comparison, all databases were linked using the national identifier.

Problems in data comparability can be tackled by using the instrument HDN. They can be grouped in 1) identification of data generating process (data sources: administrative system or health record); 2) case definitions (use of primary classifications: data specifications defined nationally by the authorities are not similar); 3) Comparison of cost and resource used; 4) Stratification/risk adjustment [43,44].

C. EHES - European Health Examination Survey

European Health Examination Survey is an initiative to set up a system of standardized, representative health examination surveys of the adult general population of the European countries. The core measurements are height, weight and waist circumference to measure body composition, blood pressure, total and HDL-cholesterol and fasting glucose. EHES includes also a self-reported questionnaire with important items needed to support objective measurements as well as the health status and health perception. EHES is a survey with primary use of data collection which provide comparable indicators on risk factors, chronic disease prevalence, physical performance, cognitive function, etc.

Implications

With the pilot study conducted in 14 countries, the Project has demonstrated the feasibility of a survey conducted in samples of the general adult population with standardized procedures and methods in different countries.

Data quality is assured by the training and testing of the personnel involved in the fieldwork, applying international standardized procedures and methods and with internal and external quality control. With these, study of time trends and geographical variation can be assured. Completeness is assessed by participation rate, which depends on age, sex, season of survey, at home or at centre examination, time and day of the week, etc. Audit and site visits are used to check quality and improve standardization. Risk factors, risk conditions and prevalence use primary data collection.

Limits

Some countries have a long tradition in health examination survey, therefore they prefer using methods or devices from their previous survey to measure risk factors and occasionally these differ from provided standardized recommendations. Changes in procedures and methods can affect the study of time trends of risk factors. A problem revealed by the project is the comparability of the national HESs due to differences in age groups and survey methods. Some of the measurements are sensitive to the used measurement protocols and devices. An example is the blood pressure measurement, which is influenced by participant's activity before measurement, posture during the measurement and the used device for the cut-off level (see example *Table 4* at page 41).

Recommendations

To minimize the differences between the different national HES and optimize the comparability, a proper standardization of the measurement protocols, training of the personnel, and periodical internal and external audits are required.

D. EUROCISS - European Cardiovascular Indicators Surveillance Set

The EUROCISS project, aims to prioritise the aspects of cardiovascular diseases of major interest at European level, to provide a list of recommended indicators, sources of information, case definition and quality methods for monitoring Acute Myocardial Infarction/Acute Coronary Syndrome (AMI/ACS), stroke and of cardiovascular diseases surveys.

The main objective was to prepare the manuals of operations which provide simple and comparable tools to support and stimulate implementation of surveillance systems in those countries which lack them by using administrative databases such as mortality and hospital discharge records and validating a random samples of current event. Starting from a minimum data set and following a step-wise procedure, EUROCISS provides a standardized model for an efficient implementation of a validated surveillance system at reasonable cost. To set public health priorities and determine appropriate actions, a standardized definition of event is crucial and indicators should be comprehensive, valid (sensitive and specific), standardised, and meet quality criteria. The definition of the event must take into account either the ICD codes reported in hospital discharge diagnoses (main or secondary) and in causes of death (underlying or secondary) and the duration of event (28 days). This definition is of particular importance since myocardial infarction may occur more than once, then it is necessary to consider both first and recurrent events. Inaccuracy increases with age. A unique person identification number (PIN) for each subject is a strong tool in linkage procedures between hospital discharge diagnoses and death certificate data; alternatively, multiple variables (e.g. name, date of birth, sex, residence)

can be used for record linkage (deterministic and probabilistic).The high cost of registers limits their implementation at national level; therefore they should be established in representative areas of a country (regions, macro-areas, etc.).

Implications

Population-based registers are the best data source for cardiovascular surveillance as they include morbidity and in- and out-of-hospital mortality. They provide estimates of key indicators such as attack rate and case fatality. Attack rates consider both first and recurrent events; case-fatality considers in-hospital mortality and sudden death, i.e. those serious cases not able to reach in medical services. Incidence can be assessed if information on first event is available. If survival rates are available, also prevalence can be assessed.

Limits

Limits of this surveillance system are: lack of registration of non-fatal events occurred outside the surveillance area or outside hospitals (nursing homes, clinics) and of the registration of asymptomatic events.

Recommendations

A simplified method, based on a record linkage of hospital discharge diagnoses and death certificates, with validation of a sample of events according to standardized diagnostic criteria, as already adopted in many countries, is the key recommendation of EUROCISS Project and might be applied in those countries which do not have registers. This method uses sources of information and data bases currently available in public health services and aims to identify the current numbers of fatal and non-fatal major coronary events.

E. EUBIROD - European Best Information through Regional Outcomes in Diabetes

The EUropean Best Information through Regional Outcomes in Diabetes-EUBIROD Project is a European diabetes registry, based on integrated health information systems on type 1 and type 2 diabetes of existing national/regional frameworks which use the BIRO technology; this system automatically generates local statistical reports and safely aggregate data to produce international reports of diabetes indicators.

The Biro project delivered 79 indicators on demographic, clinical, health system characteristics, risk adjusted indicators. 19 countries provide data from different studies based on local data source. For each indicator, *consistency* with EUBIROD definition and *completeness*, are assessed and combined to provide the *Overall Quality Score-OQS*; moreover, for each indicator, some parameters are calculated

(e.g. *Percentage Recorded-PR* as percentage of data sources with data item recorded) and used as axes in a graphical plot of Feasibility vs Validity for each data item [40].

Data dictionary (metadata repository) is a central repository of information about data (meaning, relationship to other data, origin, usage, format). The data dictionary is realised through the BIRO common data set and the EUBIROD survey. The BIRO Common dataset is defined, assessed, and periodically revised by clinical experts, epidemiologists, statisticians, and IT experts. The EUBIROD survey, conducted across EUBIROD diabetes registers, contributed to assess the consistency of standard definitions with local practices.

Implications:

- the BIRO is a system which helps to centralise, and aggregate databases to a central server, as an essential element for a secondary use of health data;
- the BIRO System is an open source suite of integrated software tools distributed as a complete Linux operating system running on any platform (Windows, Linux);
- BIRO Academy organizes annual residential courses (practical/theory) for use of diabetes data and application of BIRO system. Training materials (software, video, demo applications) are available at the academy web site.
- Respect of privacy: a novel method of Privacy Impact Assessment which ensured complete privacy protection without hampering the information content for public health, is adopted. In EUBIROD the clinician which collected data can correct and control data elements at the basis of diabetes indicators.

Limitations:

- in the Common Dataset event definitions, procedures and methods for data collection, are not standardised. Each country adopts own proper clinical judgement and sources of information;
- considering that EUBIROD is focused on diabetes, unawareness of diabetes should be included among indicators in the EUBIROD dataset (ad hoc HES-Health examination Survey are not considered);
- the parameter assessment criteria involve subjective judgement of the local clinical reviewers and the qualification of reviewers may not be consistent across all centres.

Recommendations: data dictionaries and data standards can be used to improve quality, consistency and comparability of national information if case definition is standardized (not depending on local practices and clinical reviewer). This procedure may be recommended to assess quality of parameters from different databases.

F. JAMIE - Joint Action on Injury Monitoring in Europe

The JAMIE project run from 2011 to 2014 and aimed at having by 2015 a common hospital-based surveillance system for injury prevention in all Member States (MSs). Such a system should report on external causes of injuries due to accidents and violence as part of the Community Statistics on Public Health.

The project was carried out by a consortium of centres of excellence in injury surveillance based in the EU region (Austria, the Netherlands, Hungary, UK, Germany). The EuroSafe organisation provides leadership to the project.

Implications

Injury data collection efforts should include all acute physical injuries attending Emergency Departments for diagnosis, investigation or treatment, which fall into the nature of injury categories listed in the dataset.

The IDB-JAMIE data source has been judged as credible and sustainable enough to be included into the ECHI. With respect to injuries there are a few indicators related to home and leisure injuries - reported by survey or from registries - and indicators related to road traffic injuries, work related injuries, and suicide attempts. The home and leisure injury indicators are defined as injuries that have occurred in and around home, in leisure time and at school resulting in an injury that required treatment in a hospital. These data are expected to be provided from national hospital discharge information systems as well as national emergency departments-based injury data in line with the IDB-JAMIE methodology.

- All countries should implement the core IDB-Full Data Set (FDS) in a representative sample of emergency departments (the IDB FDS is based on a systematic injury surveillance system that collects accident and injury data from selected emergency departments of MS hospitals, existing data sources, such as routine causes of death statistics, hospital discharge registers and data sources specific to injury areas, including road accidents and accidents at work). Where possible this should be based on injuries from all external causes. In some circumstances where this is not possible it may be limited to home and leisure related injuries only.
- Where FDS has not been previously implemented and resources are scarce each country should implement the FDS in at least one hospital.
- In addition all countries should widely implement the Minimum Data Set (MDS) (intent, location and setting, activity, and mechanism) unless the FDS in operation provides a sufficiently large and representative sample at a country level. In which case there is no need for an additional MDS to be collected.

Limitations

- The heavy work load in accident and emergency departments and the limited time for/interest in administrative work, puts severe pressure on the quality of reporting and completeness of information.

- The exclusion of many specific and non-specific codes provides a potential for some biases in recording (e.g. road traffic injury is included within the major mechanism category because of the importance of monitoring and supporting road traffic injuries in almost all settings. Of course, road traffic injuries occur due to a variety of mechanisms including cutting/piercing, burns, sheering stresses but the vast majority are due to blunt force from contact with hard objects).

IX. How to improve data quality

The first step to plan and organize data collection in surveys, registries, and administrative databases, is to prepare the manual of operations. It includes all specific information on methods and procedures, as definition of target population, sampling, measurements, questionnaire, communication to participants, data analysis, data storing. If the aim of data collection is also to build one or more indicators, it is important to describe how to process and compute data for building indicators. All steps included in the manual of operations should be checked and quality of performance and of collected data should be tested. Country-specific conditions and health monitoring needs should be regarded in this context.

Studies based on administrative data, where two or more datasets are linked, require standardization of the event definition, harmonization of data for a reliable and comparable definition, description of the sources of information and procedure to aggregate and elaborate data for generating indicators.

The following steps are crucial to improve data quality:

A. TRAINING AND TESTING

The staff, which performs measurements or collects data should be qualified through training and testing. After a complete information of the study (objectives, procedures and methods for measurements) usually given by an expert, the staff attend practical sessions. In this activity, firstly staff observe standardized procedures and methods for measurement and/or questionnaire administration, for results codification and computer input performed by an expert; then the staff replicate the same procedures under the supervision of the expert. The staff are trained when they are able to perform procedures as indicated in the manual of operations. The testing process is the agreement of staff performance results within the established values. Training and testing can improve the reliability of data.
E.g. *Method to train and test the staff for measuring blood pressure before the health examination survey and to assess the quality of data collected.* Training includes explanation of the reason why standardized blood pressure measurements

are needed, conditions requested to the examined person for the measurement (abstain from doing heavy physical activity, eating, smoking and avoid exposure to cold temperature for one hour before measurement), number of measurements needed (two/three consecutive measurements), position of the personnel and position of examined person (sitting position), key steps in the measurement procedure (selection of the cuff, high level to inflate, deflate the cuff at a rate of 2mmHg per second), how results should be recorded (as a mean of two/three measurements); how the results should be explained to the participant.
Practical training includes adequate number of measurements under the supervision of an expert and feedback session to discuss errors during the measurements. The double stethoscope is used to check the readings of blood pressure measurement; difference of more than two mmHg between trainer and trainee is not allowed. At regular intervals, the following quality checks are needed: distribution of last digits for systolic and diastolic measurements to test the accuracy of measurements, and proportion of identical measurements in the same participant in order to test that three measurements of blood pressure are really done. Site visit and audit are important to maintain quality control.

B. DATA COLLECTION

Procedures, methods and tools used during the study have an impact on the final results of the study. They should be standardized, comply with the purpose of the study, match the population under evaluation, and be completed. They should provide good quality data, do not overload the participant (in the case of ad hoc survey), comply with ethical and data-protection requirements, and have limited costs. Rules for and comments on the implementation of the data collection should be fixed in a written form and made available to the data collection personnel [45]. An ad hoc survey should be based on measurements and/or questionnaires, while population based registries or other routinely database studies, should be based on integration or linkage of several data sources (e.g. hospital discharges, death certificates, drug prescriptions, etc.).

Examples of quality control of blood pressure measurements during the European health examination survey are shown in table 4 [46]:
"In nine out of 12 surveys, more than one cuff size was available. Arm circumference was measured in ten surveys. When comparing measured arm circumferences to the size of the cuff used for the measurement, the miss-cuffing (use of too small or too large cuff) was observed only in 1-5% of the subjects, except in one survey where only one cuff was available and 20% of subjects were miss-cuffed. In three surveys which did not measure arm circumference, the occurrence of miss-cuffing would have been more likely to happen,
especially if the used cuffs did not have markings to indicate correctness of the cuffs for the specific arm circumference. In all these three surveys, cuffs with indicators to

assess the correctness of the cuff size were used. The proportion of identical sequential measurements was lower between the first and the second measurement than between the second and the third measurement for both systolic and diastolic blood pressure in three surveys using simple mercury sphygmomanometers. In each survey, the proportion of identical readings was higher for diastolic than for systolic blood pressure. Overall, the proportion of identical sequential measurements was high only in two surveys (28% or over) using simple mercury sphygmomanometer."

Table 4: EHES: Blood pressure measurements in 12 surveys. Recording of cuff size used and measured arm circumference, proportion of miss-cuffed subject, and proportion of identical readings between 1st-2nd and 2nd-3rd measurements.

Pilot survey	Recorded cuff size used	Arm circumference measured and recorded (M = measured, not recorded, B = measured and recorded, N = Not measured)	Mean arm circumference (min,max)	Proportion of miss-cuffed subject (based on optimal arm circumference reported on the cuff)	Proportion of identical readings between 1st and 2nd measurement		Proportion of identical readings between 2nd and 3rd measurement	
					SBP	DBP	SBP	DBP
A	No	B	30.8 cm (24.5,43.5)	#	8%	29%	16%	28%
B	Yes	B	31.6 cm (23.0,48.0)	1%	10%	29%	31%	41%
C	Yes	B	28.3 cm (20.0, 40.0)	0%	10%	13%	19%	19%
D	Yes	N	§	#	4%	10%	7%	12%
E	Yes	B	29.9 cm (22.0,41.5)	0%	5%	9%	6%	11%
F	Yes	B	28.9 cm (21.0,41.5)	2%	3%	9%	9%	11%
G	Yes	M	§	#	7%	12%	8%	13%
H	Yes	B	30.1 cm (21.0,48.0)	20%	3%	8%	5%	9%
I	Yes	B	30.9 cm (22.0,46.0)	0%	8%	10%	8%	10%
J	Yes	N	§	#	6%	14%	9%	15%
K	Yes	B	30.4 cm (21.5,43.3)	5%	4%	10%	7%	10%
L	Yes	B	30.4 cm (24.0,40.0)	1%	5%	9%	5%	11%

§ Arm circumference not measured, # Not possible to calculate.

Source: Tolonen H, Koponen P, Naska A, Männistö S, Broda G, , Kuulasmaa K7; EHES Pilot Project. Challenges in standardization of blood pressure measurement at the population level. BMC Med Res Methodol. 2015 Apr 10;15:33 [46].

C. DATA INPUT AND DATA DELIVERY

Collected data by questionnaire, measurements or routinely sources sometimes need to be codified (process that assigns a value to a response) and input into a computer system. Usually data entry of questionnaires and routinely collected sources are computer assisted and are automatically recorded and codified in an electronic way. This method reduces cost and enhances accuracy of the results in comparison to manually codifying and entering data.

Coding errors could be minimized by training of personnel and involving more than one person to enter the same data. Double entry of subsamples can be considered.

To identify errors due to wrong coding or data entry, preliminary quality checks should be performed assessing:

- *for quantitative variables:* format, plausibility of values, anomalous data, range, distribution, number of missing values, and consistency with other related variables;
- *for qualitative variables*: format, allowed codes, distribution, missing values, and consistency with other related variables.

Identification and correction of the errors due to wrong coding or data entry could be done by verifying original data stored on paper or electronic support, or deciding to replace wrong values as missing.

Arbitrary decisions and interventions on databases should be as rare as possible and completeness of data should be reported in the results of the statistical analysis.

In studies where data from different sources are included, data may need to be transferred to a central entity. In this case, a clear definition of the procedures of data delivery is desirable. E.g. in the Data Delivery Guidelines for the European Health Interview Survey, the way of transmission of the data is described in detail, including validation rules to be applied to the delivered data. [47]

The validation rules define how to perform checks for allowed codes and values, skip checks, and consistency checks.

D. DEFINITION OF NEW VARIABLES OR INDICATORS

The definition of new variables or indicators should be clear and precise. Changes to the variable values or the building of new variables are to be documented in writing in each individual case [45]. The process of selecting a new variable's definition depends on the nature of the collected data and on the nature of the new variable. The latter could be a simple transformation of a continuous variable using a different unit of measurement, or a new codification of a quality variable aggregating possible answers, or an aggregation of quantitative and qualitative variables. Great attention should be given to the possible values of the new variables, and particular care is needed in the use of missing values, especially when more variables are aggregated.

An example of aggregation of quantitative and qualitative variables is the definition of hypertension (yes/no) in which three variables collected during the survey may be considered: systolic and diastolic blood pressure (quantitative variable) and use of antihypertensive treatment (qualitative variables). Presence of hypertension may be defined as systolic blood pressure>=140 mmHg or diastolic

blood pressure>=90mmHg or under specific treatment. Hypertension could be defined in several ways, for example the following three:

- systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg or under specific treatment, excluding those subjects with one or more missing variables;
- systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg or under specific treatment, excluding those subjects with all three missing variables;
- systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg not considering specific treatment.

The definition of hypertension and use of missing values is crucial for the estimation of the prevalence of hypertensives.

When data of two or more surveys/databases are pooled, the knowledge of procedures and methods to measure blood pressure, questions used to record treatments and definition of hypertension used (including use of missing values), is crucial. When two or more surveys/databases are aggregated, it is better to include all variables collected during the data collection (in the mentioned example, blood pressure measurements and use of antihypertensive medications) than to report only the new variable "hypertension", in order to be sure to use a common definition for all involved databases.

In a population based register, an example of aggregation of variables collected within the study could be the process of event validation. It is based on a precise and complex flow chart starting from a record linkage of hospital discharges and death certificates that allow to build the variable "current event" (yes/no), and then to start a standardized process of validation of this information to build the new variable "validated event" (yes/no). This example underlines the importance of collecting and checking completeness and coherence of data collected within the registry and of specifying a precise process to build new variables (current event and validated event).

X. Conclusions and recommendations

Quality should concern data, data sources and indicators. This report contains an overview of the quality methods related to these three topics.

The assessment of data quality is crucial for the reliability and comparability of data among countries and across regions, for monitoring time trends, for building health information systems at national and European level.

Without good data, quality of indicators, quality of studies and therefore decisions on planning and evaluation of preventive programs, health care delivery, resource allocation, and research are severely impaired.
A detailed description of data, data sources, their size and characteristics, process to compute indicators, and all related measures of quality is fundamental for aggregation, harmonization and comparison of indicators.

XI. References

1. https://stats.oecd.org/glossary/detail.asp?ID=2620
2. Bergdahl M, Ehling M, Elvers E, Földesi E, Körner T, Kron A, Lohauß P, Mag K, Morais V, Nimmergut A, Sæbø HV, Timm U, Zilhão MJ. Handbook on Data QualityAssessment Methods and Tools. EUROSTAT. 2007. M Ehling and T Körner (eds) pg 1-139
3. ESS handbook for quality reports. EUROSTAT Manuals and guidelines. European Statistical System. 2014 pg 1-164
4. Signore M, Brancato G, Carbini R, D'Orazio M, Simeoni G. Linee guida per la qualità dei processi statistici. 2012. ISTAT.pg 1-52
5. AHRQ Quality Indicators—Guide to Prevention Quality Indicators: Hospital admission for Ambulatory Care Sensitive Conditions. Rockville, MD: Agency for Healthcare Research and Quality. Revision 3. (January 9, 2004). AHRQ Pub. No. 02-R0203
6. EUROCISS- European Cardiovascular Indicators Surveillance Set, FINAL REPORT; 14 February 2008, http://www.cuore.iss.it/eurociss/en/projectII_report2007.asp
7. EHES-http://www.ehes.info/
8. Non Communicable Diseases Risk Factor Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19· million participants. Lancet 2016; 387: 1377–96
9. Non Communicable Diseases Risk Factor Collaboration 2. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. Lancet 2016; 387: 1513–30
10. Eurostat/A4/Quality/03/General/Definition Working Group "Assessment of quality in statistics" Sixth meeting Luxembourg, 2-3 October 2003 ITEM 4.2 Methodological documents - definition of quality in statistics
11. http://ec.europa.eu/eurostat/ramon/nomenclatures/index
12. OECD :glossary of statistics terms: htpps://stats.oecd.org
13. http://www.nationaalkompas.nl/object_binary/o14707_ECHIM-final-report_Part-II_PDF.pdf
14. Tolonen H, Koponen P, Mindell JS et al. Under-estimation of obesity, hypertension and high cholesterol by sef-reported data: Comparison of self-reported information and objective meaures from health examination surveys. Eur J Public Health 2014;24:940-947
15. www.cuore.iss.it
16. http://www.epicentro.iss.it/passi/dati/sovrappeso.asp
17. http://www.istat.it/it/files/2016/07/Fattori-di-rischio_salute_def.pdf?title=Fattori+di+rischio+per+la+salute++-+26%2Flug%2F2016+-+Testo+integrale.pdf

18. Donfrancesco C, Lo Noce C, Brignoli O, Riccardi G, Ciccarelli P, Dima F, Palmieri L, Giampaoli S, Oprg OP. Italian network for obesity and cardiovascular disease surveillance: a pilot project. BMC Family Practice 2008, 9(1): 53
19. Tunstall-Pedoe H, Kuulasmaa K, Tolonen H, Davidson M, Mendis S with 64 other contributors for The WHO MONICA Project. MONICA Monograph and Multimedia Sourcebook. Geneva: World Health Organization; 2003
20. C. Batini, M. Scannapieca. Data Quality: Concepts, Methodologies and Techniques, Springer, Heidelberg 2006
21. Porta M. (2008) A dictionary of Epidemiology, 5th ed., Oxford University Press, New York
22. Madsen M et al: Population based registry of acute myocardial infarction : manual of operations. Cardiovascular surveillance: The EUROCISS Project. Europ J Cardiov Prev Rehabilit 2007 14: 3-22
23. http://ec.europa.eu/eurostat/statisticsexplained/index.php/Glossary
24. Tolonen H (ed.) (2013a) EHES Manual, Part A. Planning and preparation of the survey. National Institute for Health and Welfare. Directions 2013_001. URN:ISBN:978-952-245-842-1, URL: http://url.fi/URN:ISBN:978-952-245-842-1
25. Tolonen H (ed.) (2013b) EHES Manual, Part B. Fieldwork procedures. National Institute for Health and Welfare. Directions 2013_002. URN:ISBN:978-952-245-843-8, URL: http://urn.fi/URN:ISBN:978-952-245-843-8
26. Nwaru BI, Mukherjee M, Gupta RP, Farr A, Heaven M, Stoddart A, Bandyopadhyay A, Fitzsimmons D, Shields M, Phillips C, Chamberlain G, Fischbacher C, Dibben C, Aftab C, Simpson CR, Lyons R, Strachan D, Davies G, McKinstry B, Sheikh A. Challenges of harmonising data from UK national health surveys: a case study of attempts to estimate the UK prevalence of asthma. J R Soc Med 2015. OnlineFirst, published on October 2, 2015 as doi:10.1177/0141076815600909
27. EuroSafe (2014). IDB-JAMIE Full data set (IDB-FDS) Data Dictionary. Available at: http://www.eurosafe.eu.com/uploads/inline-files/IDB%20JAMIE%20FDS%20Data%20Dictionary%20MAR14.pdf
28. Giampaoli S, Hammar N, Adany R et al. Population-based register of stroke: manual of operations. Eur J Card Prev and Rehabil 2007; 14 (Suppl 3): S23-S41
29. Primatesta P, Allender S, Ciccarelli P et al. (2007) Cardiovascular surveys: manual of operations. Eur J Card Prev and Rehabil 2007; 14 (Suppl 3): S43-S61
30. Gissler M, Mohangoo AD, Blondel B, Chalmers J, Macfarlane A, Gaizauskiene A, Gatt M, Lack N, Sakkeus L, Zeitlin J; Euro-Peristat Group. Perinatal health monitoring in Europe: results from the EURO-PERISTAT project. Inform Health Soc Care. 2010 Mar;35(2):64-79. doi: 10.3109/17538157.2010.492923
31. Euro-Peristat (2012) Euro-Peristat list of indicator, updated 2012. Available at: http://www.europeristat.com/images/doc/updated%20indicator%20list.pdf
32. www.qualityindicators.ahrq.gov
33. www.atlasvpm.org

34. http://www.echo-health.eu/handbook/infrastructure.html
35. Häkkinen U, Iversen T, Peltola M, Seppälä T, Malmivaara A, Belicza É, Heijink R, Fattore G, Numerato D, Medin E, Rehnberg C. Health care performance comparison using a disease-based approach: the EuroHOPE project. Health Policy 2013; 112(1-2): 100-9
36. http://www.eurohope.info/
37. http://ec.europa.eu/health/indicators/echi/index_en.htm
38. http://www.healthdatanavigator.eu
39. Carinci F, Di Iorio CT, Massi Benedetti M, Standardized information exchange in diabetes: integrated registries for governance, research and clinical practice, in Bruttomesso D, Grassi G (eds), Technological Advancements in the Treatment of Type 1 Diabetes, (Frontiers in Diabetes, vol 24), p.236-249, Basel, Karger, 2015
40. Cunningham SG, Carinci F, Brillante M, Leese GP, McAlpine RR, Azzopardi J, Beck P, Bratina N, Boucquet V, Doggen K, Jarosz-Chobot PK, Jecht M, Lindblad U, Moulton T, Metelko Ž, Nagy A, Olympios G, Pruna S, Skeie S, Storms F, Di Iorio CT and Massi Benedetti M, Defining a European diabetes data dictionary for clinical audit and healthcare delivery: core standards of the EUBIROD project, Methods of Information in Medicine, 2015 Dec 15;55
41. Mohangoo AD, Buitendijk SE, Szamotulska K, Chalmers J, Irgens LM, Bolumar F, Nijhuis JG, Zeitlin J. Gestational age patterns of fetal and neonatal mortality in europe: results from the Euro-Peristat project. PloS one. 2011;6(11):e24727
42. Delnord M, Hindori-Mohangoo AD, Smith LK, Szamotulska K, Richards JL, Deb-Rinker P, Rouleau J, Velebil P, Zile I, Sakkeus L, Gissler M, Morisaki N, Dolan SM, Kramer MR, Kramer MS, Zeitlin J, Euro-Peristat Scientific C. Variations in very preterm births rates in 30 high-income countries: are valid international comparisons possible using routine data? BJOG. 2016
43. http://www.euroreach.net/compendium
44. The Health Data Navigator toolkit www.healthdatanavigator.eu
45. http://dgepi.de/fileadmin/pdf/GEP_LL_english_f.pdf
46. Tolonen H, Koponen P, Naska A, Männistö S, Broda G, , Kuulasmaa K7; EHES Pilot Project. Challenges in standardization of blood pressure measurement at the population level. BMC Med Res Methodol. 2015 Apr 10;15:33
47. https://circabc.europa.eu/webdav/CircaBC/ESTAT/healthtf/Library/technicalsgroupshis/2013_%2026%20February/Presentations/TG%20HIS%202013%20-%204_5.pdf.