POPULATON-BASED REGISTER OF STROKE: MANUAL OF OPERATIONS

Simona Giampaoli[1], Niklas Hammar[2], Roza Adany[3]; Juliette Bloch[4], Christine De Peretti[4] on behalf of the EUROCISS Research Group

1 Istituto Superiore di Sanità, Rome, Italy
2 Karolinska Institutet, Stockholm, Sweden
3 University of Debrecen, Debrecen, Hungary
4 Institut de Veille Sanitaire, Saint-Maurice, France

This manual was reviewed by Professor Birgitta Stegmayr
Deptartment of Public Health and Clinical Medicine
Umeå University
Umeå
Sweden

EUROCISS Project Research Group:

The EUROCISS Project (European Cardiovascular Indicators Surveillance Set http://www.cuore.iss.it/eurociss/eurociss.asp) was set up in 2000 by a partnership of 18 European Union countries under the coordination of the Istituto Superiore di Sanità, Rome, Italy.

S. Giampaoli (Project coordinator), P. Ciccarelli, L. Palmieri, S. Panico, D. Vanuzzo, ITALY

K. Steinbach AUSTRIA

M. Kornitzer/G. de Backer BELGIUM

J. Holub CZECH REPUBLIC

M. Madsen DENMARK

V. Salomaa FINLAND

J. Bloch/C. De Peretti FRANCE

A. Doring GERMANY

D. Trichopoulos/A. Trichopoulou GREECE

R. Adany HUNGARY

V. Gudnason ICELAND

M. Verschuren THE NETHERLANDS

S. Graff-Iversen NORWAY

A. Pająk POLAND

E. Rocha PORTUGAL

S. Sans SPAIN

N. Hammar SWEDEN

P. Primatesta UNITED KINGDOM

S. Allender EUROPEAN HEART NETWORK

CONTENTS

ABBREVIATIONS

1. INTRODUCTION AND RATIONALE
1.1. Burden of disease
1.2. Disease register
1.3. Historical background
1.4. Existing registers in Europe - an overview

2. OBJECTIVES

3. STRATEGY FOR SURVEILLANCE
3.1. Surveillance methods and types of registers
3.2. Target population
3.3. Data sources

4. METHODS
4.1. Definition of events
4.2. Indicators
4.3. Data collection methods

5. QUALITY CONTROL
5.1. Completeness of cases and completeness of information
5.2. Internal validity
5.3. External validity (representativeness)
5.4. Methods to evaluate diagnostic quality

6. ETHICAL ISSUES

7. ECONOMIC CONSIDERATIONS

8. IMPLEMENTATION – step-wise procedure

TABLE 1. National population-based stroke registers

TABLE 2. Regional population-based stroke registers

TABLE 3. Examples of healthcare services-based stroke registers in countries participating in the EUROCISS project

TABLE 4. Methods for surveillance of cerebrovascular disease in the population

TABLE 5. Sources of information

REFERENCES

ABBREVIATIONS

CSF = Cerebrospinal Fluid
CT-Scan = Computed Tomography Scan
CVD = Cardiovascular Disease
DRG = Diagnosis Related Group
EROS = European Register Of Stroke
ECHIM = European Community Health Indicators Monitoring
EU = European Union
EUROCISS = European Cardiovascular Indicators Surveillance Set
EUROSTAT = Statistical Office of the European Communities
GP = General Practitioner
HDR = Hospital Discharge Records
HF = Heart Failure
HES = Health Examination Survey
HIS = Health Interview Survey
ICD = International Classification of Diseases
IHD = Ischaemic Heart Disease
MONICA = MONItoring trends and determinants of CArdiovascular diseases
MRI = Magnetic Resonance Imaging
OECD = Organisation for Economic Cooperation and Development
PIN = Personal Identification Number
TIA = Transient Ischaemic Attack
WHO = World Health Organization

1. INTRODUCTION AND RATIONALE

1.1 Burden of disease

The most frequent forms of cardiovascular disease (CVD) are those of an atherosclerotic origin, mainly Ischaemic Heart Disease (IHD), stroke and Heart Failure (HF).

More than 1.9 million people die every year from CVD in the European Union (EU)[2]. Nearly half (42%) of all deaths (46% of deaths in women and 39% deaths in men) are from CVD [1].

CVD clinically manifests itself in middle life and older age after many years of exposure to unhealthy lifestyles (smoking habit, unhealthy diet, physical inactivity) and risk factors (total and low-density lipoprotein cholesterol, blood pressure, diabetes). CVD accounts for over 225,000 premature deaths before the age of 65 in the EU: 7% of all men and 3% of all women die from CVD before the age of 65 [1].

Even though clinical onset is mainly acute, stroke often evolves gradually, causes substantial loss of quality of life, disability, and life long dependence on health services and medications. The societal costs are substantial and they are not only those directly related to healthcare and social services, but also include those linked to a) illness benefits and retirement; b) impact on families and caregivers; and c) loss of years of productive life [1].

Stroke is the second leading cause of death in the European Union accounting for 490,000 deaths each year. Over one in eight women (13%) and one in ten men (9%) die from this disease and many more suffer from non-fatal events [1].

In most Western European countries death from stroke has declined by 30-50% since 1975, but in the countries of Eastern Europe stroke mortality has remained stable or slightly increased over the same period of time [2-5]. Despite the decline in mortality in Western Europe, the annual number of cases of stroke is expected to increase

[2] 25 member States: Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, United Kingdom.

within the next few decades, mainly due to a 30% growth in the elderly population, which will lead to an increase in the health burden of stroke and consequent increase in economic costs [6].

In the last decade, innovations in diagnostic technologies in the cardiovascular field have facilitated diagnosis at earlier phases in the course of the natural history of disease or in presence of less severe tissue damage. The use of diagnostic technologies, such as Computed Tomography Scan (CT-Scan) and Magnetic Resonance Imaging (MRI), has greatly improved the accuracy of diagnoses of hospitalised cerebrovascular events allowing delineation of the location and type of lesion.

The World Health Organization – MONItoring trends and determinants of CArdiovascular diseases (WHO-MONICA) project [7] has demonstrated a large variation between countries in case fatality rates (the proportion of fatalities occurring within 28 days after onset of acute stroke), ranging from 15% in Northern countries to 50% in some Eastern European states. The implications of these findings are that the quality of acute stroke care varies between countries and that an improvement in initial diagnosis, treatment and rehabilitation programmes may reduce case fatality rates [6].

Lifetime costs of first-ever stroke are estimated at between 31,440 euro in the Netherlands and 63,000 euro in Sweden, of which hospital costs account for 45% in the first year after a stroke [8,9]. It is estimated that hospital costs attributed to stroke will increase by 1.5% per year [9].

Across Europe with its ageing population there is a pressing need to cope with costs increase and make stroke prevention and treatment a priority to reduce the growing health burden and lessen its socio-economic impact [10].

According to the Organisation for Economic Cooperation and Development (OECD), it does not appear inevitable that longer life leads to higher costs. This is one of the reasons why the health system should be largely oriented to work on preventive actions. Epidemiological studies have shown that stroke is preventable to a large extent. Different preventive strategies can be implemented to a) reduce the

occurrence and impact of stroke (through, for instance, the identification of individuals at high risk of stroke such as hypertensives, diabetics and smokers); b) intensify treatment in people who have already experienced a stroke or Transient Ischaemic Attack (TIA); or c) improve rehabilitation.

At the European level, WHO, OECD and EUROSTAT (Statistical Office of the European Communities) collect simple indicators (mortality, hospital discharge rates) and process them into tables available on web-site (www.euro.who.int/hfadb; www.oecd.org; www.europa.eu.int/comm/eurostat). These data are rarely comparable due to the different methodology and the peculiar health system of each country.

1.2 Disease register

The objectives of a stroke population-based register is to (a) evaluate the frequency, distribution and prognosis of the disease providing indicators such as attack rate, incidence rate, prevalence and case fatality; (b) compare trends in different countries; (c) evaluate trends and changing pattern, outcomes and treatment effectiveness; and (d) monitor disease prevention programmes.

Focusing on the general population, a stroke register may provide a comprehensive picture of stroke in the community, highlight problem areas and suggest where there are population groups at high risk and where treatment facilities are most in need of improvement. It may provide information needed to plan healthcare services and to develop and test which methods are most useful as a basis for preventive action.

The register includes all cases 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 would also include rapidly fatal cases unable to reach the medical service.

It is important that collection of information on suspected events and application of diagnostic criteria follow a standardised methodology in order to enable data comparison in different areas of the same country or between different countries.

To summarise, a population-based register is intended for health professionals and policy makers and provides the means to understand the characteristics, the burden and the consequences of the disease in the population through:

- the monitoring of the occurrence of the disease (i.e to assess population differences and trends in attack and incidence rates and in mortality over time);
- the understanding of the differences and changes in the natural disease dynamics between genders, age groups, social classes, ethnic groups, etc.;
- the identification of vulnerable groups;
- the monitoring of in- and out-of-hospital case fatality;
- the assessment of relations between disease incidence, case-fatality and mortality;
- the monitoring of the consequences of disease in the community in terms of drug prescriptions and rehabilitation;
- the monitoring of the utilisation of new diagnostic tools and treatments and their impact.

This is crucial in order to:

- develop health strategies and policies;
- plan health services and health expenditures;
- improve appropriate allocation of resources;
- evaluate the effectiveness of interventions.

In order to provide this, a register must be validated. Validation provides the means to:

- take into account bias from diagnostic practices and changes in coding systems;
- trace the impact of new diagnostic tools and re-definition of events;
- ensure data comparability within the register (i.e. different sub-populations, different time points, etc);
- ensure data comparability with other registers within and between countries.

1.3 Historical background

The *WHO Stroke Register* was the first attempt to collect data on stroke in the community in a uniform manner from countries with different social, cultural, and

environmental background. It lasted from May 1971 to September 1974 and was a joint undertaking of WHO and 15 collaborating centres in 10 countries from Asia, Africa and Europe. About 2 million people were under surveillance and data was obtained from 6,395 new cases of stroke (3,270 men and 3,125 women).

Fourteen of the centres covered the general population in defined geographical areas and one centre covered an occupational group consisting mainly of men below the age of 55 years. No limitations of age and gender were set in the study areas, except for two centres in Sweden and Japan.

A stroke was defined as rapidly developing clinical signs of focal (at times global) disturbance of cerebral function, lasting more than 24 hours or leading to death with no apparent cause other than that of vascular origin [11].

The *WHO MONICA* Project [12,13] was started in the first half of the 1980s and lasted until the first half of 1990s. Stroke registers were established in 17 centres in 10 countries.

Study populations were residents in geographically defined areas and included men and women ages 35 to 64 years, with an optional inclusion of the 65 to 74 years decade.

All stroke events in defined populations were ascertained and validated according to a common protocol and uniform criteria. Almost 25,000 stroke events in more than 15 million person-years were analysed.

Stroke was defined “as rapidly developed signs of focal (or global) disturbance of cerebral function, lasting more than 24 hours (unless interrupted by surgery or death), with no apparent nonvascular cause”. This category included patients presenting with clinical signs and symptoms suggestive of subarachnoid haemorrhage, intracerebral haemorrhage, or cerebral ischaemic infarction. This definition excluded patient with TIA or stroke events in cases of blood disease or brain tumors. Secondary stroke caused by trauma was also excluded.

Up to 6-fold differences were observed in stroke mortality. Mortality declined in 8 of 14 populations in men and in 10 of 14 populations in women. An increase in

mortality was observed in Eastern Europe. In the populations with a declining trend, about 2/3 of the change could be attributed to a decline in case fatality. In populations with increasing mortality, the rise was explained by an increase in case fatality.

1.4 Existing registers in Europe – an overview

The data collection for the international MONICA study ended in 1994/95. Some countries continued to collect data every year, while others only periodically (every 5 years).

Presently, the existing registers in Europe adopt different data collection procedures: some registers are based on the procedures used in the MONICA study, others on administrative databases with or without record linkage, some are national and some are regional. Different age groups are covered, the degree of validation of the diagnostic information varies and in most registers is much less intensive than in the MONICA study. The registers are used for different purposes and have different strengths and limitations [14].

Tables 1, 2 and 3 give a brief overview of the existing stroke registers in Europe. As shown in Table 1, Denmark, Finland and Sweden have national stroke registers, which are based on record linkage between hospital registers and cause of death registers.

Table 2 shows regional population-based stroke registers: most of them are based on a disease specific data collection comparable to the MONICA registers, while others are based on other data collection methods.

Table 3 shows examples of registers based on data from healthcare institutions such as General Practitioner (GP) and hospitals. These registers are not population-based since they do not include out-of-hospital cases or cases not seen by GP and thus they do not consider sudden death occurring out-of-hospital. These registers are not intended to assess disease occurrence but rather to evaluate outcome and survival of stroke patients.

It is worthwhile to mention the European Register Of Stroke (EROS), a 4 year prospective study across Europe aiming at estimating the impact of stroke

understanding the factors underlying variation in the quality of care and outcome after stroke, and answering unresolved issues with regard to the influence of socio-demographic, case-mix and stroke healthcare, quality factors on the variations in health or stroke patients around Europe. The cities of London, Helsinki, Glasgow, Edinburgh, St Petersburg, Kaunas, Warsaw, Dijon, Menora, Florence, Stockholm participate in EROS [15].

2. OBJECTIVES

The purpose of the EUROCISS Project is to provide a general guide and updated methods for the surveillance of stroke to those EU countries which lack appropriate surveillance systems and therefore wish to implement a population-based register in order to produce comparable and reliable indicators.

Taking into account developments in new diagnostic criteria, treatment and information technologies in recent years, this manual provides a standardised and simple model for the implementation of a population-based register. It recommends to start from a minimum data set and follow a step-wise procedure based on standardised data collection, appropriate record linkage and validation methods.

This manual is intended for investigators, health professionals, policy makers and data collection staff interested in the surveillance of stroke.

Although in many countries data extracted from some sources of information (mortality and hospital discharge records [HDR]) are now available thanks to the continuing process of computerisation, they are rarely reliable and comparable. These data can produce reliable indicators only if properly processed and validated by independent epidemiological sources.

This manual represents a valid tool to build the core indicators (attack rate, incidence, case fatality) recommended by the EUROCISS Project Research Group for inclusion in the short list of health indicators set up by the European Community Health Indicators Monitoring (ECHIM) Project. This Project was launched in 2005 with the aim of implementing health monitoring in EU [16].

3. STRATEGY FOR SURVEILLANCE

3.1 Surveillance tools and types of registers

Surveillance is the ongoing, systematic collection, analysis, interpretation and dissemination of health information to health professionals and policy makers. Surveillance, defined as a continuous, and not episodic or intermittent activity, differs from monitoring [17,18].

Disease surveillance in a population can be done using many different data sources (Table 4). Most countries have national databases on causes of death and on discharge diagnoses for hospitalised patients.

Mortality statistics have for many years been the main tool for comparing health and disease patterns among countries and today still remain the only source of information for some countries. They have also been used to monitor trends in cerebrovascular disease and compare mortality among countries. Since the 1950s, the cause of death has been registered according to the International Classification of Diseases (ICD) to make data comparable. Different classification of disease within versions and different methods of ascertainment have led to problems in comparison between different revisions of ICD and/or similar versions among countries.

In recent years, routine statistics also include discharge diagnoses from hospitalisation and, for some countries, visits to outpatient clinics coded according to the same international classifications as the mortality data. Stroke can be extracted for relevant populations and age groups and these routine statistics are still very important tools for monitoring the disease.

Many countries have also Health Interview Surveys/Health Examination Surveys (HIS/HES). These surveys are primarily used for monitoring disease prevalence (included cerebrovascular disease), prevalence of risk factors (health behaviour, social network, environmental risk factors) and of disease consequences (disability, reduced physical function, unemployment). They are described in detail in the Manual of Operations of CVD Surveys.

Few countries have an established disease-specific stroke register which ensures a more precise and valid monitoring of this disease.

A population-based register is usually formed through linkage of various sources of information (mortality data, hospital discharge and GP's records) and covers a defined population (entire municipalities, regions or whole country) and a specific age group (35 to 74 or 35 to 64 years or all ages).

A population-based register should be used for the surveillance of stroke morbidity and mortality since it considers both fatal and non-fatal events occurring in- and out-of-hospital; therefore, it provides estimates of key indicators such as attack rate and case fatality. Incidence can be assessed if information on first event is available. If survival rates are available, prevalence can be assessed as well.

Case finding and validation procedures depend on data collection methods, healthcare and financing system, and diagnostic criteria applied in the definition of events. The accuracy of rates produced is related to the completeness and quality control of the data collected for the numerator (death and hospital discharge registers) and the denominator (census or population register). Completeness also depends on tracing subjects treated outside hospital (nursing home, clinic, etc.) and outside the area of surveillance. The routine recording of diagnoses may be a problem for registration of stroke: a large proportion of "new stroke diagnoses" are merely sequelae of an old stroke. This problem increases with ageing.

The definition of the event must take into account both the ICD codes reported in the hospital discharge diagnoses (main or secondary) or in the causes of death (underlying or secondary) and the duration of the event. Stroke may occur more than once and therefore it is necessary to consider both first and recurrent events. In this context, deaths occurring within 28 days are usually considered to reflect the same event [17] (See the definition of recurrent events in paragraph 4.1).

A Personal Identification Number (PIN) is a strong tool in linkage procedures between hospital discharge diagnoses, GP's records and death certificates; alternatively, multiple variables (e.g. name, date and place of birth, gender, residence) may be used for record linkage.

Specific Stroke Register

The strength of this register lies in the possibility of validating each single event according to standardised diagnostic criteria and collecting disease-specific clinical and paraclinical data [19]. The weakness lies in the fact that data collection is expensive and this kind of register can usually be maintained only for a limited period of time in a defined population of reasonable size. Another limitation is that a local or regional register may not be representative of the whole country.

Identification of events can be obtained 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. Information bias is minimised by the *hot pursuit* approach as information is collected immediately after the event. The process is comparatively demanding in terms of resources.

Cold pursuit implies the use of routine and delayed procedures by means of hospital discharge and death records. The process is easier and less expensive than *hot pursuit*; the number of cases studied is typically smaller because discharge diagnoses are more precise and specific than those on admission, but there is a possibility of missing important information. Both methods are used to identify suspected events, which are subsequently validated using specific diagnostic criteria.

The specific stroke register is important since it collects fatal and non-fatal events; actually, official mortality statistics provide only a limited and sometimes biased picture of stroke in the population. A large proportion of stroke victims are left with permanent disability; economic and human consequences of stroke extend far beyond what emerges from routine mortality data. The specific stroke register, which allows to assess incidence and prevalence, reflects better than mortality the impact of stroke in the community. Monitoring non-fatal stroke is associated with a number of problems, the most important being the completeness of case finding, especially in areas where many stroke cases are not treated in hospital. An extensive review of stroke incidence registers showed that few of them provide reliable data [20]. Indeed, it has been claimed that most of the differences in stroke mortality and incidence rates reported to exist between populations are attributable to methodological bias.

A specific stroke register provides standardised and reliable epidemiological data for public health initiatives aimed at preventing the disease. It has been used in the WHO MONICA Project, where uniform criteria for recording cardiovascular disease have been applied to 14 populations in 9 countries [14].

Register based on routine administrative data

Identification of events is based on linkage of mortality data and HDR. The register based on routine administrative data has existed for many years in the Northern countries, where all individuals are identified by a PIN which allows record linkage between different information sources. This register is economical, covers the whole country, all age groups and collects large numbers of events. The main objective of administrative databases is to produce relevant statistics to plan health services and healthcare expenditure and to give internationally comparable data on mortality, causes of death and hospital admissions. The register based on routine administrative data is not primarily planned for research purposes but is increasingly used in epidemiological research. Its strength lies in the fact that it covers the whole country and the completeness is close to 100%. The weakness lies in the fact that data are not standardised to the same degree as in the disease-specific data collection and that clinical and paraclinical data available are limited. If used in research, this register needs to be carefully validated. Stroke registers based on administrative data, such as hospital discharges and deaths, have been employed in Denmark and Finland in order to obtain national rates of stroke incidence, mortality and case fatality [21,22]. A similar approach is being investigated for use in Sweden.

Studies on feasibility of combining data from routine hospital discharge and cause of death registers have been performed in Finland: over 90% of hospitalised acute stroke events (first and recurrent) included in the Finland MONICA Stroke register were found in the HDR with one of the stroke diagnoses. The missing events were mainly explained by errors in the PIN (leading to unsuccessful record linkage) and different practice of defining an event as hospitalised when death occurred in the emergency room (leading to exclusion from the HDR) [21].

In the past, hospitalisations for rehabilitation purposes were often coded using an ICD code for acute stroke; with the introduction of ICD-9 version, a separate diagnosis for acute events and sequelae was made possible. The definition of stroke death also differs between the specific stroke register and the mortality register: in the specific stroke register the death is very strictly defined as a death occurring within 28 days from the onset of event; on the contrary, deaths occurring after 28 days from the onset of symptoms are often coded as stroke in the mortality register [21].

In studies assessing trends in stroke subtypes the change in the use of neuroimaging examinations and autopsy frequency should be reported.

General Practitioner register

The great majority of health problems are managed in primary care and do not go further into other levels of the healthcare system. This is true especially for those less serious problems which do not require hospitalisation. The fact that primary healthcare is generally the first and most frequently utilised health service makes general practice a rich source of information. This further emphasises the need for monitoring health in primary care settings to have a full picture of health status of populations. This is particularly necessary for stroke, which occurs especially among elderly and in some countries patients with stroke are treated at home even during the acute phase: this makes the GP's register a valid source of information for monitoring stroke. Monitoring health in primary care should however not be seen in isolation from other sources of information about health.

Essentially, there are two models for collecting morbidity data in primary care. One is based on episodes of care, recording data on all doctor-patient interactions, gathering information on consultation rates and patterns of clinical management; the other focuses on specific disorders, using a limited number of standardised case definitions and attempting to assess the burden of disease attributable to those disorders in the population in question. The first model is exemplified by the English General Practice Research Database Programme [23,24], and the use of International Classification of Primary Care, ICPC codes [25], while the second one is illustrated

by the Morbidity Sentinel Stations Programme that is now operational in several European countries [26-28].

3.2 Target population

A population-based stroke register may cover a whole country; where this is not feasible, the population under surveillance would typically be residents of a defined region in the country. The target population should preferably cover a well defined geographical and administrative area or region for which population data and vital statistics are routinely collected and easily available each year. Both urban and rural areas should be monitored: differences often exist with regard to exposure to risk factors, treatment of predisposing disease and access to facilities.

It is important that all cases among those with residence in the area are recorded even if the case occurs outside the area (*completeness*). In the same way, all cases treated at hospitals within the area but with residence outside the area must be excluded. If this is not possible, it is important to give an estimate of the magnitude of the loss of cases and establish whether it could be changing and interfering with the validity of the observed trends in the rates over a period of years.

It is also important to consider to what extent an area is representative for the whole country (*representativeness*): it could be representative according to the CVD mortality rate, the distribution of risk factors (socioeconomic status and health behaviour) and the distribution of health services (specialised hospital, GP). In some countries it might be better to start with high risk area.

The population to be monitored should be selected in order to produce estimates of disease rates that are sufficiently robust from a statistical point of view, so that trends can be established and data comparability ensured.

In general, it is necessary to select more than one area representative for socio-economic or ethnic differences in order to have a comprehensive picture for the whole country, and a coordinating body between the areas is recommended to ensure comparability.

The target population should be selected taking the following parameters into account:

age: the age range covered by the MONICA Project was 35 to 64 years. As reported in the final report, the EUROCISS Research Group suggests the wider age range 35 to 74 years, or even up to 84 years of age when possible, considering that in patients above 65 years of age more than half of the stroke events occur. The age groups recommended from EUROCISS Project to present morbidity and mortality are decennia, in particular the age ranges 35 to 44, 45 to 54, 55 to 64, 65 to 74 and 75 to 84. If administrative routine data are used, all ages are automatically included, but for patients ages 85 and above the validity of the diagnostic information tends to be less reliable. Age-standardised rates (35 to 74 and 35 to 84) are recommended using the European Standard Population as reference.

Gender: stroke is an important cause of death and disability in men and women, and the population should include both genders. There are no major gender differences in stroke presentation or management; mortality and quality of life at 6 months are similar in women and men.

Population size: to be eligible to participate in a stroke population-based register, a minimum of 300 stroke events per year in the population ages 45 to 74 years is necessary. The size of the population under surveillance is determined by the number of fatal and non-fatal events and the event rate in the age group concerned. The minimum of 300 events (fatal and non-fatal) has been established in order to detect a decrease in mortality trend by 2% in event rate per year. This means that the population to be under surveillance could range between approximately 1,200,000 (all ages) in low incidence country like Italy and approximately 400,000 (all ages) in a high incidence country like Finland, basing the calculation on female attack rate usually lower than male attack rate. If more areas are enrolled, it would be desirable that the same number of 300 total events is considered for each single area.

Patient eligibility: an individual is considered eligible for inclusion in a stroke population-based register only if he/she is resident in the area under surveillance, meets the selected age and had a stroke event within the defined time period.

3.3 Data sources

To monitor stroke in the general population, the following sources of information should be available at a minimum: mortality records with death certificates; and, hospitalised discharge records with clinical information.

A special stroke register would typically include several sources of information.

Some events occur suddenly and are not able to reach the hospital and some non-fatal cases may not be referred to hospital for treatment, particularly those occurring to very old individuals. Therefore, additional sources are usually needed to achieve complete information on all fatal and non-fatal events: clinical pathology laboratory (autopsy register), nursing home, clinic, emergency or ambulance service, GP, radiology unit (Table 5).

Death Certificate

The death certificate provides complete data on fatal events and are collected in a systematic and continuous way in all EU countries. Mortality statistics are easily accessible in all countries but are usually available in a detailed and complete form after 2-4 years.

The format of the death certificate varies from country to country, but generally it includes personal identification data, date and place of death (i.e. municipality, nursing home, hospital or other) and causes of death (underlying, immediate and contributory). Causes of death are coded according to ICD. Problems of temporal and geographic comparisons derive from the different versions of the ICD adopted over time (7^{th}, 8^{th}, 9^{th}, 10^{th} revision) and from different coding practices in each country. Furthermore, diagnostic criteria for coding death certificates are not defined at the international level and the ICD nosologic and nosographic versions are updated every 10 years by the WHO.

Some countries code the underlying cause of death only.

The reliability of mortality data depends on the completeness and accuracy of the vital registration system as well as the registration and coding of causes of death. When the proportion of deaths coded as "unknown cause of death" is higher than 5%, cause-specific mortality data should be used with caution. The accuracy of the recorded causes of death depends on the autopsy rate. This rate varies largely between countries and over time. In some countries, the autopsy rate has declined in recent years, which is a problem for the use of mortality statistics in disease surveillance.

Hospital Discharge Records

HDR give the number of hospitalisations for stroke, which are absolutely necessary to monitor CVD. Moreover, clinical information and medical care reported in hospital documents are important for validation of events.

Hospital discharge data are available in most EU countries, but in some countries only as aggregated tables without detailed information on age and gender distribution and without haemorrhagic and ischaemic stroke as separate diagnostic categories.

HDR include personal identification data, admission date, type of hospitalisation (urgent, ordinary or transfer to other structure) and discharge diagnoses. Hospital discharge diagnoses are coded by ICD codes (currently ICD-9 or ICD-10). For some countries only a limited number of diagnoses is coded.

Problems in the assessment of a specific stroke event may arise when an acute event is followed by a period of rehabilitation or a transfer to other wards and the event could be counted more than once (sequelae). HDR do not include emergency room and private hospitals or nursing homes are only included in some countries.

Discharge diagnoses are not validated on a routine basis and validation studies are necessary in all countries to check the diagnostic quality. The validity of a hospital discharge diagnosis may vary on the basis of patient characteristics, geographical region and type of hospital or clinic.

Hospital admission policies vary over time and place; the registration of the most severe cases dying shortly after the arrival to the hospital differs between hospitals, depending on the administrative procedures connected to hospital admissions. HDR may also include patients not resident in the area under surveillance.

The adoption of new diagnostic techniques, such as MRI and CT-Scan, may cause major changes in event rates estimated from HDR. Therefore these techniques should be taken into account when interpreting trends.

A further problem may derive from the use of Diagnosis Related Group (DRG). In some countries, financing healthcare services is based on the DRG tariff system, which is built on equal-resources criteria and aggregates events in major diagnostic categories.

DRG may be useful in hospitals for acute events but are not reliable for chronic diseases requiring a long hospital stay and rehabilitation, such as stroke.

Countries using the DRG system are Denmark, Finland, France, Germany, Italy, Norway, Portugal, Spain and Sweden. In order to assess the occurrence of stroke,

HDR from all hospital departments should be used but if this is not possible at least the following departments must be taken into consideration:

- intensive care (an intensive care unit, including any type of acute medical unit);
- medical (a general medical ward, including a geriatric unit);
- neurological/neurosurgical (a general neurological ward);
- rehabilitation (a specialised rehabilitation unit, except a rehabilitation stroke unit);
- stroke (acute and rehabilitation stroke units);
- other (other units, e.g. radiology).

Autopsy register

Not all countries perform autopsy on suspected or sudden deaths on a routine basis. Autopsy is performed on violent deaths or on deaths occurring in hospital when clinical diagnosis is undetermined. The first one is performed by a forensic medicine specialist, the second one by a pathologist of the hospital where death occurred. Data from this register refer therefore to a low percentage of deaths but provide a more valid diagnosis to complement the information reported on the death certificate.

Nursing home and clinic

The nursing home and clinic mainly provide data on cases among older patients who sometimes get care from these institutions without being admitted to hospital. Therefore, information on events occurring in the nursing home can be critical, especially if the register covers elderly patients. In some countries rehabilitation after an acute event is provided by the rehabilitation clinic which may give information on patients who have received the acute care outside the region.

Emergency and ambulance services

Data provided by emergency and ambulance services are useful to integrate information for register implementation since patients dying suddenly or experiencing fatal stroke are not always able to reach the hospital. These services are able to provide data otherwise not obtainable, such as CT-Scan or MRI during the

acute phase of the event or blood pressure measurement, blood glucose, peripheral oxygen saturation, body temperature and fluid balance, level of consciousness (fully conscious; somnolent; semicomatose; comatose) and muscular deficit at the time of event occurrence in paucisymptomatic patients referring to emergency services. The need of very urgent medical treatment often makes information partial but the integration of these data with those from other sources of information contributes to the implementation of the register.

General Practitioner Register

In some countries a GP register can be useful when dealing with events not necessarily requiring hospitalisation. This is particularly important for the elderly population.

Radiology unit

The role of the radiology unit (CT-scan or MRI) is a support in the identification of non-hospitalised events, in the diagnosis of stroke type (haemorrhagic or ischaemic) and in treatment.

4. METHODS

4.1 Definition of events – Subtypes

There are three major stroke subgroups as follows: ischaemic stroke; intracerebral haemorrhage; subarachnoid haemorrhage

Type	Caused by	Diagnosis based on
Ischaemic stroke (ICD-9 434; ICD-10 I63)	Sudden occlusion of arteries supplying the brain, due to a thrombus formed: - directly at the site of occlusion (thrombotic ischaemic stroke), or - in another part of the circulatory system, which follows the blood stream until it obstructs arteries in the brain (embolic ischaemic stroke)	Neuro imaging recordings Note: it may not be possible to decide clinically or radiologically whether it is a thrombotic or embolic ischaemic stroke.
Unspecified stroke (ICD-9 436; ICD-10 I64)		
Intracerebral haemorrhage (ICD-9 431, 432; ICD-10 I61, I62)	Bleeding from one of the brain's arteries into the brain tissue	- Neuro imaging recordings
Subarachnoid haemorrhage (ICD-9 430; ICD-10 I60)	Arterial bleeding in the space between the two meninges, pia mater and arachnoidea. **Note:** Typical symptoms are sudden onset of very severe headache and usually impaired consciousness	- Neuro imaging, or - Lumbar puncture

Modified from WHO STEPS Stroke Manual V2.1

It should be noted that each type differs with respect to survival and long-term disability.

General major symptoms

Symptoms should be of a presumed vascular origin and should include one or more of the following definite focal or global disturbances of the cerebral function:

- unilateral or bilateral motor impairment (including lack of coordination);

- unilateral or bilateral sensory impairment;
- aphasia/dysphasia (non-fluent speech);
- hemianopia (half-sided impairment of visual fields);
- forced gaze (conjugate deviation);
- apraxia of acute onset;
- ataxia of acute onset;
- perception deficit of acute onset.

Other symptoms

Other symptoms that may be present but are not adequate for stroke diagnosis (often resulting from other diseases or abnormalities such as dehydration, cardiac failure, infections, dementia, and malnutrition) are as follows:

- dizziness, vertigo;
- localised headache;
- blurred vision of both eyes;
- diplopia;
- dysarthria (slurred speech);
- impaired cognitive function (including confusion);
- impaired consciousness;
- seizures;
- dysphagia.

Subarachnoid haemorrhage

For subarachnoid haemorrhage at least one of the following must be present in addition to the general major symptoms:

- recent subarachnoid hemorrhage, aneurysm or arteriovenous malformation (necropsy/autopsy);
- blood in the Fissura Sylvii or between the frontal lobes or in the basal cistern or in cerebral ventricles (CT or MRI);

- blood stained cerebrospinal fluid (CSF) (>2000 red blood cells per mm^3), aneurysm or an arteriovenous malformation (angiography);
- blood stained CSF (>2000 red blood cells per mm^3), also xanthochromic and intra-cerebral haemorrhage (necropsy or CT-Scan).

Stroke-like symptoms

A broad range of other diseases may cause similar symptoms, for example, HIV/AIDS, tuberculosis, syphilis, intracerebral cancer. These diseases are known to be able to cause focal neurologic disturbances and thereby mimic a stroke. Attention to the development of symptoms is an important factor to consider in order to avoid other diseases being misinterpreted as vascular disease and leading to ineffective preventive strategies.

Onset and survival

Stroke events are classified as *first ever* or *recurrent, with non-fatal* and *fatal* outcome:

- *First ever stroke event*: refers to people who have never had a stroke before.
- *Recurrent stroke event*: for a new episode of symptoms to be counted as a recurrent event, general stroke criteria must be met and either:
  - onset is day 1 (one);
  - a new stroke occurring after 28 days is a new event.

If a patient experiences further acute symptoms suggestive of stroke within 28 days of the onset of a first episode and in the same carotid or vertebral artery territory, this second episode is not counted as a new stroke event.

Equally, if a patient experiences further acute symptoms suggestive of stroke after 28 days of the onset of a first episode, this second episode is counted as a new stroke event.

- *Non-fatal stroke event*: refers to patient surviving at least 28 days after the onset of the stroke symptoms.

- *Fatal stroke event*: refers to stroke causing death within 28 days of symptoms onset.

It should be noted that each event is registered separately.

4.2 Indicators

Attack rate

Attack rate is the total number of new cases (separated into subtypes and summed) and recurrences per 100,000 target population over 1 year. It is calculated using either the main cause of hospitalisation or, in cases of out-of-hospital deaths, the underlying or contributory causes of death. It should be noted that in the case of stroke the hospital discharge can sometimes be quite distant from the onset of stroke event. Therefore, a hospital discharge register alone is not always an accurate source of information. Ideally, an in-patient inventory should be checked at the end of each year to identify patients who are hospitalised for stroke but not yet discharged [20].

Incidence rate

Incidence is the number of new cases per 100,000 target population over 1 year [20].

Case-fatality

Case fatality is the proportion of events that are fatal by the 28^{th} day.

The EUROCISS Project recommends for cerebrovascular events 7 day and 28 day case fatality. All in- and out-of-hospital fatal and non-fatal events are to be considered as denominator.

4.3 Data collection methods

The different types of registers described in section 3.1 use different data collection methods. Registers with disease-specific data collection can be divided into registers

based on routine administrative data using record linkage, disease specific registers using hot and cold pursuit and GP registers.

Stroke registers based on routine administrative data

In recent years, the development of computerised record linkage has made it possible to overcome obstacles in linking administrative database.

Record linkage methods can be summarised into three broad categories: *manual*, *deterministic* and *probabilistic*.

Manual matching is the oldest, most time-consuming and most costly method. In general it is not a feasible option when large databases are involved.

Deterministic linkage matches records from two data sets (or two records from different locations in a single data set) using a unique variable (e.g. PIN or hospital chart number) or by full agreement of a set of common variables (e.g. name, gender, birth date).

Probabilistic linkage [29] is used to identify and link records from one data set to corresponding records in another data set (or two records from different locations in a single data set) on the basis of a calculated statistical probability for a set of relevant variables (e.g. name, gender, date of birth). This type of record linkage links records with a specified high probability of match. The method requires detailed prior knowledge about various measures of the relative importance of specific identifier values in both files that are to be linked.

The main limitations of record linkage are the difficulty in:

- obtaining administrative files for research purposes: mortality data files are usually available at the National Institute of Statistics, while hospital discharge data are available at the Ministry of Health. These kinds of data are anonymous and therefore do not allow record linkage. Nominal files are available at regional level or at the sanitary units;

- combining data: missing events are mainly explained by errors in PIN or in name; they may lead to unsuccessful record linkage;
- defining and obtaining minimal data set (for mortality: PIN; family and first name; date and place of birth; gender; residence; date and place of death; underlying and secondary causes of death. For hospital discharge diagnosis the same variables should be considered together with admission date and hospital discharge diagnoses);
- obtaining necessary funds for processing large administrative files.

Nonetheless, record linkage studies provide evidence of the statistics that could become available with greater integration of administrative databases.

The national stroke registers in the Northern countries use record linkage between Hospital Discharge Registers and Causes of Death Registers as the basis for the register. The linkage as such is easy because of the PIN attached to every citizen in the country. However, the linkage has to be followed by many specific definitions of how to handle primary and secondary diagnoses, underlying and contributory causes of death, transfer between hospitals with difference in the diagnoses between the admitting hospital and the hospital where the patient is transferred, how to define date of attack, first time events and recurrences. Practical ways how to approach these problems have been suggested from work carried out in Finland [21,22].

It is usually difficult to detect the incident cases (first events): hospitalisation records within the previous 5-7 years are reviewed to check for disease; if no hospital admission for stroke is found, then the stroke case identified is considered a first event. Further problems may arise when estimating trends: for example the changes in the use of neuroimaging examinations and autopsy frequency can lead to an overestimation of the number of events or make the interpretation of stroke subtypes difficult.

Specific Stroke Registers

This kind of register uses hot and/or cold pursuit method for data collection.

Hot pursuit [30]

This method of detecting events involves identifying patients acutely in hospital by interviewing them directly. The problem with this method is that the data collection technique is very difficult to standardise (e.g. descriptions of symptoms may vary with the observer). Periods of staff shortages or holidays may lead to loss of cases that cannot be recovered and a large team is needed to search the wards for cases. However, some information may be more complete than that obtainable from case notes.

Notification of events should be instituted on a routine basis checking admission registers on the wards.

While the extreme forms of hot pursuit involve getting the information from the patient acutely, an alternative is to use the hot pursuit method to identify the patients of interest and to mark their notes or list them for review later. An efficient reliable routine is needed for picking up the case notes at an identifiable point in their processing.

A benefit of the hot pursuit method is that information on the diagnosis is collected soon after admission. This has its limitations, however, as initial diagnosis can sometimes be superseded by subsequent tests and other more detailed investigation.

Residents hospitalised outside the area will always have to be registered by cold pursuit, weeks or months later.

Cold pursuit [30]

Use of discharge diagnoses rather than hospital admissions is a more simple system of identifying events for the study. Its advantage is that it can be done months or years after the event but it is limited because the information in the case notes may not be complete and the notes themselves may not be accessible.

Once the event has been identified and validation is required, medical notes should be obtained in order to extract the necessary information. When a register is launched for the first time, a plan for future evaluation of trends is recommended. This can be achieved by continuous surveillance as part of a broader health information system or

annual register repeated at 5 to 10 year intervals. The minimum recommended period of observation is one complete calendar year because of possible seasonal variation.

Combined approach

A mix of hot and cold pursuit ensures the most complete identification of stroke events.

Some of the patients must have been identified as soon as possible after symptoms onset with the possibility of direct examination, while the remaining events are based on routine data.

It is difficult to check up on a hot pursuit system several months later, but discharge lists can be used as a backup method to ensure that the hot pursuit method had detected all the diagnosed cases. Residents hospitalised outside the area, and other late-detected cases mean that a proportion of events will always have to be registered by cold pursuit, weeks or months later.

5. QUALITY CONTROL

Quality control of registers is extremely important for a valid monitoring and comparison between regions and countries. The quality of the register depends on:

- completeness of coverage (sequence of events) and completeness of information;
- internal validity;
- external validity (representativeness).

The surveillance of stroke is complicated by the fact that a number of cases is not admitted to hospital, particularly in older age. The identification of cases in older populations outside hospital is essential for a precise determination of occurrence. These events are a combination of milder or more severe strokes than those admitted to hospital and, consequently, their inclusion influences incidence as well as case fatality.

5.1 Completeness of coverage and completeness of information

Completeness of coverage means that all stroke cases in the target population are included, i.e. events occurring independently inside or outside the region. The register has also to cover events whenever they occur during day/night or winter/summer as well as events occurring outside hospital (e.g. sudden death among patients who never reach the hospital).

Completeness of information means that all relevant information has been registered (e.g. place of treatment, date of admission, date of discharge, PIN, gender, hospital discharge diagnostic codes, intervention/procedure codes, department/ward, date of birth).

The most important source of systematic bias in estimating incidence is related to the coverage of event registration. The registration system must attempt to identify all possible cases of the disease that have come to the attention of the existing medical and medico-legal sources. The completeness of event identification (acute-care hospital, primary healthcare, nursing home) and the completeness and availability of

information, obtainable for each event recording and diagnosis, depend on the existing standard of medical care: if the medical care system misses or misdiagnoses cases, a register cannot remedy the omission.

When the event is defined (codes and duration), it may be possible to identify duplicate coding and to take out information for quality control purposes. Duplicate codes may include events transferred from one ward to another, e.g. for rehabilitation. In some cases, the duration of the admission is very short (< 2 days) either because of transferral or because of misclassification of the diagnosis. These cases may also be picked up for validation.

Cases not admitted to general hospitals are a problem for registration when the system is based only on hospital records. Another source of potential loss of identification is private practice: private physicians and hospitals may be less cooperative than those in the public system; in private hospitals the staff may be more sensitive to criticism and anxious to show how they register medical documents. GP case records are usually inadequate for full registration because patients are frequently looked after at home.

The identification of fatal events is in some way less difficult than that of non-fatal events. Whereas survivors may be lost in the totality of inhabitants of the surveillance area, death is unequivocal. However, the registration of causes of death may not be correct and needs to be validated. It is to be expected that some stroke deaths occur outside hospital. If the proportion of fatal events coded as hospitalised is very high it may indicate incomplete registration of out-of-hospital stroke deaths. High case fatality may indicate loss of non-fatal cases.

The identification of potential events may be based on many different data sources. This may involve a considerable amount of record linkage, which is facilitated if PIN is adopted.

Another problem relates to medical records, whose quality may be variable: younger patients may have had no other illness episodes and the records may be restricted to the relevant stroke event. In older patient, the identification of the event is more complicated due to the existence of comorbidities.

5.2 Internal validity

The most important question regarding validity concerns the diagnostic information. The diagnostic criteria for the event definition are valid if they measure the stroke they claim to measure. Validation preferably evaluates the sensitivity, specificity and predictive value of the registered diagnosis compared to a golden standard [19].

Validation studies of routine statistics have been carried out over the years with heterogeneous results due to differences in methodology or reflecting true differences in the validity of the routinely collected data between countries. Some studies have been carried out comparing community registers with national statistics and data from the MONICA project. These findings stress the importance of validating routine mortality and hospital statistics against the national register to determine whether and how they can be used to reflect true incidence and mortality [31]. Particular attention in this type of validation should be given to secondary discharge diagnosis or causes of death, especially for diagnostic codes, in order to detect potentially hidden cardiovascular diagnosis.

Consistency of coding with the diagnosis and consistency of coding/comparability of the information for different areas of the country and over time represent other problems for validation.

If it is not possible to validate all the diagnoses included in the disease register or in the mortality routine statistics, the objective for validation should be to evaluate a sample of events. The sample should be distributed along a full year in order to ensure that potential seasonal or other time related variations of diagnostic patterns are traced. The sample could include a feasible fraction of the 365 annual days (working and weekend days). For example in *n* days per month, all consecutive hospital admissions and deaths of eligible ICD codes may be validated.

5.3 External validity (representativeness)

It is not essential that the whole country is covered by a surveillance system but it is essential that the registration system of events is complete with regard to events occurring in the target population. It is important to know how representative the

register is for the whole country according to the CVD mortality rate, the distribution of age and gender and of health determinants (socioeconomic status and health behaviour) and the distribution of health service (specialised hospital, GP).

For the population chosen there must be good demographic data subject to at least annual revision; inaccuracy may become apparent years after the period being studied because of the results of a decennial national census.

A careful description of the population characteristics may help to describe how representative the target population is for the whole country.

5.4 Methods to evaluate the diagnostic quality

Using the diagnostic criteria it is possible to evaluate if the diagnostic tools used to establish application of valid methods are different if hot or cold pursuit is performed. Validation of the diagnostic information recorded in the register can include examination of all events or of random samples. The relevant register data must be checked periodically by sampling, as it is usually not feasible to check all the data [31]. Validation has to be carried out by an epidemiological team not involved in the patient's treatment. For local registers with a limited number of cases it may be possible to validate each single event, while national registers for practical reasons can only validate data based on random samples of suspected cases recorded during a selected period or during some days each month. A selection method consists of choosing some days each month and evaluate all events which have occurred in those days, extracted either from hospital discharge or mortality records, applying diagnostic criteria. In this way, seasonal variation can be traced.

The most important phase is the evaluation of the diagnostic information although other information in the register also needs to be included in the validation.

In order to produce valid indicators, a conditio sine qua non is to allow access to relevant medical records and routine raw data of health statistics.

In some cases it is possible to validate a register by linking the register to an independent data source, e.g. a high quality register for a small area within the region.

Criteria for validation of acute cerebrovascular events

This manual of operations does not aim to improve existing stroke definitions or formulating new ones but only to suggest a definition that already exist and to ensure comparability. According to the WHO criteria, stroke is defined as 'rapidly developing clinical signs of focal (or global) disturbance of cerebral function lasting more than 24 hours (except in cases of sudden death or if the development of symptoms is interrupted by a surgical intervention), with no apparent cause other than a vascular origin' [19,32]. Global clinical signs are accepted only in cases of subarachnoid haemorrhage or in patients with deep coma. Brain lesions detected by CT-scan but not accompanied by acute focal signs are not accepted as stroke, nor are extradural and subdural haemorrhages. Stroke cases with concomitant brain tumour, trauma or severe blood disorders are also excluded [19]. Therefore, key features of the clinical definition are as follows:

- sudden onset;
- neurological deficit;
- lasting 24 hours or longer;
- of presumed vascular origin.

The table below provides an example of some of the diagnoses that should be considered for stroke registration.

Stroke specific	Focal and global signs that could be caused by stroke
• (Acute) stroke *or* (acute) cerebrovascular episode • Cerebral *or* cerebellar embolus, thrombosis *or* infarction • Occlusion, thrombosis *or* embolus of carotid, (pre) cerebral *or* vertebral artery • Lacunar hemorrhage *or* stroke • Subarachnoid, (primary) intracerebral, cerebellar *or* pontine hemorrhage *or* stroke • Ruptured berry aneurysm	• (Acute) hemiplegia *or* (acute) hemiparesis • Faint, fit, funny turn, (acute) confusional state • Loss of consciousness • (Acute) dysphasia, dysarthria, dyspraxia • Homonymous hemianopia • Amaurosis fugax • Acute monocular blindness

A stroke case is recorded as fatal if death occurs within the first 28 days.

6. ETHICAL ISSUES

The Helsinki Declaration requires that biomedical research with human subjects must conform to generally accepted scientific principles.

The *"Recommendation n. R (97)5 of the committee of ministers to EU member states on the protection of medical data"* [33] gives guidelines to how medical data can be registered, stored and used in a way that ensure the rights and the fundamental freedoms of the individual and in particular the right to privacy. *(Adopted by the Committee of Ministers on 13 February 1997 at the 584th meeting of the Ministers' Deputies).*

In the following the most important recommendations are presented.

"*Medical data should be collected and processed only by health-care professionals, or by individuals or bodies working on behalf of health-care professionals. Individuals or bodies working on behalf of health-care professionals who collect and process medical data should be subject to the same rules of confidentiality incumbent on health-care professionals, or to comparable rules of confidentiality*."

Therefore it is essential that a neurological or stroke physician (or study nurse) with proven experience in the field of cerebrovascular is involved in the coordination of the stroke register.

"*Medical data shall be collected and processed fairly and lawfully and only for specified purposes."*

"Medical data may be collected and processed:

a. if provided for by law for:

i. public health reasons; or

ii. *subject to Principle 4.8*[*]*, the prevention of a real danger or the suppression of a specific criminal offence; or*

iii. *another important public interest; or*

b. if permitted by law:

i. *for preventive medical purposes or for diagnostic or for therapeutic purposes with regard to the data subject or a relative in the genetic line; or*

ii. *to safeguard the vital interests of the data subject or of a third person; or*

iii. *for the fulfilment of specific contractual obligations; or*

iv. *to establish, exercise or defend a legal claim; or*

c. if the data subject or his/her legal representative or an authority or any person or body provided for by law has given his/her consent for one or more purposes, and in so far as domestic law does not provide otherwise."

Whenever possible, medical data used for scientific research purposes should be anonymous. Professional and scientific organisations as well as public authorities should promote the development of techniques and procedures securing anonymity.

However, if such anonymisation would make a scientific research project impossible, and the project is to be carried out for legitimate purposes, it could be carried out with personal data on condition that:

a. the data subject has given his/her informed consent for one or more research purposes; or

b. when the data subject is a legally incapacitated person incapable of free decision, and domestic law does not permit the data subject to act on his/her own behalf, his/her legal representative or an authority, or any person or body provided for by law, has given his/her consent in the framework of a research project related to the medical condition or illness of the data subject; or

[*] Processing of genetic data for the purpose of a judicial procedure or a criminal investigation should be the subject of a specific law offering appropriate safeguards.

c. disclosure of data for the purpose of a defined scientific research project concerning an important public interest has been authorised by the body or bodies designated by domestic law, but only if:

i. the data subject has not expressly opposed disclosure; and

ii. despite reasonable efforts, it would be impracticable to contact the data subject to seek his consent; and

iii. the interests of the research project justify the authorisation; or

d. the scientific research is provided for by law and constitutes a necessary measure for public health reasons."

Record linkage between mortality and HDR is possible in countries which have adopted a PIN on a national level. Other nominal data (such as name, gender, date and place of birth) are usually available at a regional level.

Record linkage is important to match admissions and discharges or admissions and deaths, thus avoiding double counting which may occur when, for example, the same patient transferred to another ward (e.g. from neurology to neurosurgery and then to rehabilitation) is registered in the HDR more than once.

Moreover, the identification of patient is essential for the event validation when it is necessary to collect and examine the history and clinical documentation and to assess case fatality at different intervals (6 months, 1 year). Before starting any study, it is recommended to seek approval from the local ethics committee.

7. ECONOMIC CONSIDERATION

Stroke is a costly disease because of the large number of premature deaths, ongoing disability in survivors, impact on families or caregivers and on health services (treatment and rehabilitation).

Stroke is estimated to cost the EU economy over €34 billion a year: around one-fifth of the overall cost of CVD. Of the total cost of stroke in the EU, around 62% is due to direct healthcare costs, 18% to productivity losses and 20% to the informal care of people with stroke [1]. Cost considerations are essential before implementing a population-based register.

Without a valid surveillance system, it is not possible to plan and evaluate health services for populations, implement interventions for primary prevention, and identify "vulnerable subgroups" in terms of burden of disease such as the elderly, the young, the poor, the unemployed. Surveillance and evaluation mean a systematic way of learning from experience and using it to improve current activities and promote better planning by careful selection of alternatives for future actions and allocation of resources. The economic benefit of a good surveillance system clearly exceeds the cost of the registers.

A population-based register may be costly and to produce meaningful data it needs to be in operation for at least one year but preferably for some years. However, the importance of a valid and efficient stroke register justifies the high implementation costs and the consequent need to find adequate financing.

The register based on record linkage between administrative databases is the most cost-effective, but this register depends on the data quality of the Hospital Discharge Register and the Cause of Death Register and also on the possibility of a valid record linkage. In addition, methods need further evaluation and implementation. Notably, if the hospital discharge and mortality registers are available for record linkage, the costs for the linkage and dissemination of results are low. The main costs for using this methodology for assessment of stroke incidence in a defined population concerns the need to perform regular validations of the diagnostic information. It is

recommended to include a basic epidemiologic team in the cost. Sometimes access to data produces separate costs.

The register based on a disease specific data collection is more expensive especially if hot pursuit is used. Beside the cost mentioned above, this type of register also needs funding for the detailed prospective data collection and for validation of diagnostic information. The data collection includes: identification of patients, reading medical records, making inquiries to additional data sources, filing and validation of the data. This means that a team of epidemiologists, nurses, medical doctors and informatics dedicated to this work full time is needed. To give an example, resources needed to run the MONICA Project in Northern Sweden for the stroke registration included: 1 nurse working full time (full time i.e. 40 hours/week); 1 medical secretary working 25% of full time; and 1 internist working 5% of full time [34]. It should be recognised that this type of register usually collects information that permits analyses of research questions beyond the monitoring of stroke incidence, mortality and case fatality. This may concern the role of risk factors for disease occurrence or the role of treatment for survival in stroke patients. In the Northern countries registers based on disease specific data collection have for several years complemented national administrative registers in providing a comprehensive picture of the burden of stroke in the population.

8. IMPLEMENTATION – A STEPWISE PROCEDURE

This section describes the procedures needed to implement a stroke register taking into account the recommendations reported in this manual of operations.

STEP 1. Define target population and routine data

- Select a geographical administrative area with a population big enough to provide stable estimates. This means that a stable population in a representative area of the country with 300 fatal and non-fatal stroke events in the age range 45 to 74 should be chosen.
- Characterise population from a demographic point of view through a detailed description of the characteristics of the population under surveillance, in particular: demographic characteristics: (age and gender distribution); socio-cultural characteristics (educational level, occupation, social group, unemployment rate, migration, immigrants with or without citizenship); characteristics of the healthcare system (specialised hospital, GP, rehabilitation clinic); macro and micro areas (urban and rural). Disease frequency is often different in macro areas of the country; a description of difference in mortality and risk factors allows to select those areas to be included in the surveillance system. Within the population-based surveillance study, the phenomenon of immigration plays an important role, therefore immigrants coming from European and extra-European countries resident in the study area must be enrolled. Geographical or administrative borders of the surveillance areas must be clearly defined.
- Analyse existing Hospital Discharge and Mortality data: events in non-residents occurring in the study area or admitted to hospital in the study area do not qualify. Events of residents occurring out of the area do qualify. Efforts must be made to find them or to estimate the potential loss and whether or not it could be changing

and interfering with the validity of the observed trends in rates over a period of years.

- Identify problems with these data: coverage, ICD version, ICD codes, procedures, DRG, unit of analysis (number of events or discharges and/or number of patients), PIN, coherence with previous studies, etc. Data files are often available in detailed forms at the regional level.

When a register is launched for the first time, a plan for future follow-up to measure trends is recommended. This can be achieved by a continuous surveillance as part of a broader health information system or by annual register repeated at 5 to 10 years intervals.

STEP 2. Perform a pilot study and validate routine data

Before starting a stroke register or a large scale use of linked administrative data, a pilot study on available hospital discharge and mortality data in a small area is recommended in order to study the feasibility and to estimate internal validity.

Validation studies on available data include:

- estimation of coverage: comparison of different routine data sets (electronic or manual), number of patients treated in- and out-of-area, hospital/mortality ratios, age and gender ratios, principal vs. secondary and/or procedure diagnoses;
- validation of discharge diagnoses according to a standard method (including revision and abstraction of medical records) in a random sample or in all cases (including check of other related diagnoses);
- validation of mortality causes according to a standard method in a random sample or in all cases;
- analysis of demography and representativeness of the area in comparison with the region or country;
- selection of the age range of interest (35 to 74 or 35 to 84).

STEP 3. Carry out record linkage using administrative data

In the Northern countries where every citizen has a PIN included in national registers of hospital discharges and deaths, record linkage for the identification of stroke events is efficient and reliable. For countries which have not adopted the PIN it may be much more difficult to perform this step. Files have to be organised with the same format and variables (family name, name, date of birth, residence and place of birth).

It is recommended to:

- explore the feasibility of record linkage within hospital records - probabilistic or deterministic approach or using PIN (within the same hospital, among hospitals of the area, among hospitals at regional or national level). When hospital records are collected at regional or national level, it is possible to collect events that occur out-of-hospital;
- explore the feasibility of record linkage between hospital records and mortality register (probabilistic or deterministic approach or using PIN);
- explore the feasibility of linkage with other sources of information (e.g. GP, drug reimbursement register). Not all GPs are organised in networks, with computerised documentation of patient history; when they are, the definition of events rarely use the same diagnostic criteria.

STEP 4. Set up a stroke register

After performing STEP 2 and 3, it is possible to set up a population-based stroke register following A (record linkage between administrative registers) or B (specific stroke register).

A. Register based on routine administrative data based on record linkage:

- when the linkage procedure between hospital discharge and mortality records is feasible, it is important to define the event, the duration, how to handle transfer between hospitals with difference in the diagnoses between the admitting hospital

and the hospital where the patient is transferred, how to define first time events, recurrent events, fatal ad non-fatal events etc. (See paragraph 4.1). A linkage system and a control for duplicate records should be set up;

- validation of diagnostic information is recommended in a random sample of sufficient size of the identified events, with the estimation of sensitivity and specificity and positive predictive value of the defined events;
- target population data by age and gender are needed to estimate incidence, recurrence, attack rate, case fatality and mortality rates;
- periodic validations should be performed.

B. Specific Stroke Register:

- set up a pilot population-based register with proven standardised protocol for stroke and evaluate the pilot study results (coverage, completeness of information and diagnostic validity);
- based on the results of the pilot study, set up, if feasible, a full scale register and decide whether to use hot or cold pursuit;
- then, if feasible, design the full-scale register (target population, data collection methods and validation procedures).

To set up a full scale register:

- select one or more populations representative for the region or the country;
- for each selected population set up a population-based register with approved standardised protocol for stroke;
- write a detailed protocol for the data collection including validation procedures for each single case;
- evaluate the coverage, representativeness and completeness of information;
- use the results from the register to validate the administrative data.

STEP 5 Disseminate results

- Set up a strategy for analysis of data and for dissemination of results;

- indicators of attack rate, incidence, case fatality and other indicators defined in EUROCISS phase I should be published yearly, e.g. on a web-site, according to gender, age and other relevant characteristics;
- use data for research. This is very important to ensure a high quality of the register over time. And a high quality register can be the basis for good research.

TABLE 1. NATIONAL POPULATION-BASED STROKE REGISTERS

Country	Starting year	Last year available	Ongoing experience	Age range	Target population (x 1,000)		Access data
					Men	Women	
Denmark	1978	2001	yes	35 to 85+	2,677	2,734	NIPH
Finland	1991	2003	yes	35 to 85+	2,600	2,600	NIPH
Sweden	1994	2006	yes	all	4,589	4,523	NBHW

NIPH, National Institute of Public Health
NBHW, National Board of Health and Welfare

Source: European J of Public Health 2003; 13 (Suppl 3): 55-60 (updated 2006)

TABLE 2. REGIONAL POPULATION-BASED STROKE REGISTERS

Country	Area coverage	Starting Year	Last year available	Ongoing experience	Age range	Target population (x 1,000)		Access data
						Men	Women	
Finland	FINSTROKE	1993	1997		35 to 85+	93	103	NIPH
France	Dijon	1985	2004	yes	6 months→	69	81	CHU Dijon
Germany	Erlangen	1994		yes	18+	49	51	University of Erlangen
Greece	Arcadia	1993	1995	no	20+	42	39	Alexandra Hospital, University of Athens
Italy	8 areas (North, Centre and South Italy)	1998	1999	yes (every 5yrs)	35 to 74	4,500		Istituto Superiore di Sanità
Norway	3 counties	1972	2002	yes	all	1,000		Health Region West
Sweden	Northern Sweden	1985	ongoing	yes	25 to 74	160	162	Umeå University Hospital

NIPH, National Institute of Public Health
CHU, Centre Hospitalier Universitaire

Source: European J of Public Health 2003; 13 (Suppl 3): 55-60 (updated 2006)

TABLE 3. EXAMPLES OF HEALTHCARE SERVICES-BASED STROKE REGISTERS IN COUNTRIES PARTICIPATING IN THE EUROCISS PROJECT

Country	Area Coverage	1st Year	Age range	Access data
Greece (Athens)	Regional	1992	18+	Alexandra Hospital, University of Athens
Greece (Arcadia)	Regional	1993	20+	Alexandra Hospital, University of Athens
Hungary (HDR)	National	1996	all ages	The Centre for Health Information, National Health Insurance Fund, Department of Financial Informatics
Hungary (GP)	Regional	1998	all ages	School of Public Health, University of Debrecen
Poland	Selected hospitals	2001	all ages	Institute of Psychiatry and Neurology Warsaw
Sweden (Riks-Stroke)	all hospitals (85)	1995	all ages	Department of Internal Medicine, Norrland Umeå University Hospital

TABLE 4. METHODS FOR SURVEILLANCE OF CEREBROVASCULAR DISEASE IN THE POPULATION

Type of registers/health surveys	Data sources	Data collection	Indicators
Specific stroke registers	Mortality HDR GP Records Other sources	Collection of data including fatal and non fatal cases in and outside hospital by hot/cold pursuit	Attack rate / Incidence rate / Prevalence/ Case fatality rate Treatment Years of life lived with disability (YLDS) Estimate of long-term care needs
Registers based on routine administrative data	Mortality registers Hospital registers Drug-dispensing registers	Hospital discharge and mortality data unlinked with or without validation	Mortality Hospitalisation Length of stay Prescribed medication
		Extraction of hospital discharge and mortality data with record linkage and with or without validation of a sample	Attack rate / Case fatality rate
GP based-registers	GP reports to national centres	GP databases	Incidence rate / Prevalence
Surveys	Health interview and/or health examination	Questionnaire and medical examination of random population samples	Prevalence Risk factors

TABLE 5. SOURCES OF INFORMATION

Data sources	Routine administrative register	Specific stroke register
Death certificate	X	X
HDR	X	X
Autopsy register		X
Nursing home and clinic		X
Emergency and ambulance		X
GP register		X
Radiology		(X)

REFERENCES

1. Petersen S, Peto V, Rayner M, Leal J, Luengo-Fernandez R and Gray A. European cardiovascular disease statistics 2005. BHF: London.
2. Asplund K. Stroke in Europe: widening gap between East and West. Cerebrovasc Dis 1996; 6: 3-6.
3. Khaw K. Epidemiology of stroke. Journal of Neurol, Neurosurg and Psychiatry 1996; 61: 333-338.
4. Stegmayr B, Harmsen P, Rajakangas AM et al. Stroke around the Baltic Sea: incidence, case fatality and population risk factors in Denmark, Finland, Sweden and Lithuania. Cerebrovasc Dis 1996; 6:80–88.
5. Stegmayr B, Asplund K. Exploring the declining case fatality in acute stroke. Population-based observations in the northern Sweden MONICA Project. J Intern Med 1996; 240 (Suppl 3):143-9.
6. European Federation of Neurological Societies Task Force. Neurological acute stroke care: the role of European neurology. Eur J Neurol 1997; 4: 435-441.
7. Thorvaldsen PM, Asplund K, Kuulasmaa K, Rajakangas A-M, and Schroll M. Stroke incidence, case fatality, and mortality in the WHO MONICA project. Stroke 1995; 26: 361-367.
8. Asplund K, Marke L, Terent A, Gustafsson S and Wester PO. Costs and gains in stroke prevention: European perspective. Cerebrovasc Dis 1993; 3 (Suppl l): 34-42.
9. Bergman L, Van der Meulen JHP, Limburg M and Habbema JDF. Costs of medical care after first-ever stroke in the Netherlands. Stroke 1995; 26: 1830-1836.
10. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? Lancet 2005; 366: 1578-82.
11. The EUROCISS Working Group. Coronary and Cerebrovascular Population-based Registers in Europe: are morbidity indicators comparable? Results from the EUROCISS Project. European J of Public Health 2003; 13 (Suppl 3): 55-60.

12. Hatano S on behalf of the participants in the WHO Collaborative Study on the Control of Stroke in the Community. Experience from a multicentre stroke register; a preliminary report. Bull World Health Organ 1976; 54: 541-553.
13. Thorvaldsen P, Kuulasmaa K, Rajakangas AM, Rastenyte D, Sarti C, Wilhelmsen L. Stroke trends in the WHO MONICA project. Stroke 1997; 28 (Suppl 3):500-6.
14. Sarti C, Stegmayr B, Tolonen H, Mahonen M, Tuomilehto J, Asplund K; for the WHO MONICA Project. Are Changes in Mortality from Stroke Caused by Changes in Stroke Event Rates or Case Fatality? Results From the WHO MONICA Project. Stroke 2003; 34: 1833-41.
15. The European Register Of Stroke (EROS) www. http://www.ktl.fi/eros/index.html (accessed 21 March 2007).
16. European Community Health Indicators Monitoring (ECHIM) Project. http://www.echim.org/ (accessed 16 May 2007)
17. WHO STEPS Stroke Manual: the WHO STEPwise approach to stroke surveillance. http://www.who.int/chp/steps/Manual.pdf (accessed 21 March 2007).
18. Federici A. Le parole della nuova Sanità. Il Pensiero Scientifico Editore 2002.
19. Asplund K, Tuomilehto J, Stegmayr B, Wester PO, Tunstall-Pedoe H. Diagnostic criteria and quality control of the registration of stroke events in the MONICA project. Acta Med Scand Supplementum 1998; 728: 26-39.
20. The EUROCISS Project Research Group. EUROCISS Project final report 2003 http://www.cuore.iss.it/eurociss/en/rapporto03/rapporto2003.asp (accessed 17 May 2007).
21. Mähönen M, Salomaa V, Keskimäki I, Moltchanov V, Torppa J, Molarius A, Tuomilehto J, Sarti C for the FINMONICA Stroke Register Study Group. The feasibility of combining data from routine hospital discharge and causes-of-death registers for epidemiological studies on stroke. Eur J Epidemiol 2000; 16: 815-7.

22. Pajunen P, Pääkkönen R, Hämäläinen H, Keskimäki I, Laatikainen T, Niemi M, Rintanen H, Salomaa V. Trends in fatal and non-fatal strokes among persons aged 35-85+ years during 1991-2002 in Finland. Stroke 2005; 36: 244-248.
23. Fleming DM. The measurement of morbidity in general practice. J. Epidemiol. Community Health1991; 45: 180-183.
24. Walley T, Mantgani A. The UK General Practice Research Database. The Lancet1997; 350: 1097-99.
25. Brage S, Bentsen BG, Bjerkedal T, Nygard JF, Tellnes G. ICPC as a standard classification in Norway. Fam. Pract 1996; 13: 391-396.
26. Chauvin P, Valleron A. Participation of French general practitioners in public health surveillance: a multidisciplinary approach. J. Epidemiol. Community Health 1998; 52 (suppl 1): 2s-8s.
27. Milne RM, Taylor MW, Taylor RJ. Audit in populations in general practice: the creation of a national resource for the study of morbidity in Scottish general practice. J. Epidemiol. Community Health 1998; 52 (Suppl 1): 20s-24s.
28. Szeles Gy, Voko Z, Jenei T, Kardos L, Pocsai Z, Bajtay A, Papp E, Pasti G, Kosa Z, Molnar I, Lun K, Adany R. A preliminary evaluation of a health monitoring programme in Hungary. Eur. J. Pub. Health 2005; 15: 26-32.
29. Fellegi IP, Sunter AB. A Theory for Record Linkage. J Am Stat Ass 1969; 64: 1183-1210.
30. Tunstall-Pedoe H. Diagnosis, measurements and Surveillance of Coronary Events. Int J Epidemiol 1989; 18 (Suppl 3): S169-173.
31. Leppala JM, Virtamo J, Heinonen OP. Validation of stroke diagnosis in the National Hospital Discharge Register and the Register of Causes of Death in Finland. Eur J Epidemiol 1999;15(Suppl 2):155-60.
32. World Health Organization: Cerebrovascular Diseases. Prevention, Treatment, and Rehabilitation. WHO Tech Report Ser 1971; 469.
33. Recommendation no. R (97) 5 of the Committee of Ministers to Member States on the protection of medical data (Adopted by the Committee of Ministers on 13 February 1997 at the 584th meeting.