Use of statistical data concerning hospital morbidity in epidemiological studies on rare rheumatic diseases

Krzysztof Kanecki¹, Aneta Nitsch-Osuch¹, Paweł Goryński², Piotr Zbigniew Tyszko³

¹ Department of Social Medicine and Public Health, Medical University, Warsaw, Poland
² National Institute of Public Health – National Institute of Hygiene, Warsaw, Poland
³ Witold Chodźko Institute of Rural Health, Lublin, Poland


Abstract

Introduction. Data validity constitutes a fundamental problem in epidemiology. An accurate estimate of the occurrence of a rare disease may require a large population study. In the case of rare diseases, all available data resources are used to estimate the occurrence rate, including statistical data concerning hospital morbidity. For life-threatening systemic vasculitides an analysis of its epidemiology is of particular importance.

Objective. The main aim of the study was to present the possibilities for the use of data from hospital databases in an analysis of the epidemiology of rare diseases, using the example of particular systemic vasculitides.

Results. The study shows the results of analyses performed on a population-based administrative data set concerning systemic vasculitides taken from the national hospital morbidity database prepared by the Polish National Institute of Public Health in 2004–2013. The study also discusses the main advantages and limitations of the adapted methodology. Furthermore, it proves that data from hospital databases can be used for analyzing disease incidence and prevalence, gender and territorial distribution, analysis of hospital mortality (fatality) rates, and trends in long-term observations.

Conclusion. An analysis of the epidemiology of rare rheumatic diseases can contribute to a better understanding of their specificity, which can be useful in everyday clinical practice.

Key words

hospitalization, rare disease, national databases, rheumatic diseases, systemic vasculitides

INTRODUCTION

In the EU, a disease is considered to be rare when the number of people affected is lower than 5 per 10,000 [1]. In the United States, a rare disease (RD) is defined as a condition that affects fewer than 200,000 people. This definition was created by Congress in the Orphan Drug Act of 1983 [2]. Because of the rarity and complexity of RDs, scientific knowledge regarding this type of diseases in Poland and other European countries can be limited or fragmented.

Due to the rarity and diversity of RDs, research may be difficult, as commercial interest in such research among private sponsors and institutions is scarce. In addition, lack of competition between potential private investors can result in the lack of impetus for innovative research. As a result, only public research funds can bridge the critical gap in research on RDs [3]. Understanding factors that can affect the development of rare diseases, such as systemic vasculitides, can be very helpful for understanding other, more common, systemic connective tissue diseases. Research on RDs has proved to be highly useful for the better understanding of the mechanism of common conditions, such as obesity and diabetes, as they often represent a model of dysfunction of a single biological pathway [4].

Rheumatic diseases, also called musculoskeletal diseases, are characterized by pain and a consequent reduction in the range of motion and function in one or more areas of the musculoskeletal system. In some diseases, signs of inflammation can occur, including swelling, redness and warmth in the affected areas. Rheumatic diseases can also affect internal organs [5]. In one study from Poland it was suggested that, on average, old patients hospitalized in the rheumatic diseases ward suffer from a different and more diverse spectrum of diseases comparison to their younger counterparts [6].

Diseases of the musculoskeletal system and connective tissue are coded M00-M99 according to International Statistical Classification of Diseases and Related Health Problems (10th Revision) [7]. Systemic vasculitides can play a significant role among rare diseases. These disease can develop quickly and involve many organs. In this group of diseases, an important element of diagnosis and treatment is the recognition of genetic factors as well as environmental and individual factors that increase the likelihood of the occurrence of the disease. Understanding of these factors may require an evaluation of a large population, based on data collected systematically in national hospital morbidity databases.

A retrospective analysis of hospitalizations with Kawasaki syndrome among children <18 years of age in the United States was based on the Kids’ Inpatient Database (1997, 2000, 2003, and 2006) and the Nationwide Inpatient Sample (1998–2007) [8]. In Sweden, a national cohort was created by compiling nationwide registers. Patients with giant cell arteritis were identified in the Swedish Hospital Inpatient and Outpatient Registers [9]. In a study from Korea, aimed at the analysis of information on the epidemiology of Takayasu’s arteritis, such as the incidence, prevalence, survival and the
cause of death, data were taken from a national, population-based database maintained by the Rare Intractable Disease (RIRD) registration programme and the Health Insurance Review and Assessment Service. Data from Statistics Korea were compared with the RIRD database to confirm the survival rate and the cause of death [10].

In the Polish health care system, hospitalization is widely accepted for diseases that require many or advanced diagnostic or therapeutic procedures. Systemic vasculitis usually requires hospitalization, and therefore an analysis of hospitalization cases can provide a good estimate of the disease incidence. Hospital discharge data were used for estimation of the incidence of vasculitides, such as Kawasaki disease in the United States [11], and granulomatosis with polyangiitis in Finland [12]. There is a growing interest in the development of tools and methods for the surveillance of chronic rheumatic diseases with the use of existing resources, such as hospital morbidity databases [13]. For instance, the prevalence of polymyositis and dermatomyositis was determined on the basis of population-based administrative data obtained from the Quebec (Canada) physician billing and hospitalization databases [14]. Data obtained from the 2010 national database of the hospitalized patients of the Thai Health Coding Centre were used to establish the admission rate, disease determination, hospital mortality rate, length of hospital stay and hospital charges among patients diagnosed with systemic connective tissue disorders [15]. In another study, data on patients hospitalized with rheumatic diseases in California was abstracted from the state hospitalization database [16].

Occurrence of rare diseases. Surveillance of population health and well-being is one of 10 essential public health operations (EPHO) [17]. The purpose of the EPHO is to provide information and data for the assessment of health needs and health impact, as well as for healthcare planning. In the case of RDs:

Patient registries and databases constitute key instruments to develop clinical research in the field of rare diseases, to improve patient care and healthcare planning. They are the only way to pool data in order to achieve a sufficient sample size for epidemiological and/or clinical research. They are vital to assess the feasibility of clinical trials, to facilitate the planning of appropriate clinical trials and to support the enrollment of patients [18].

Although there is enough justification for creating registers of rare diseases, monitoring of their occurrence varies across Europe. In 36 countries included in the report, there are 496 national, 69 regional and 61 European registers. The largest number of registers is in France – 105, whereas in some countries there is only one register. In Poland there are 6 national registers of rare diseases:
- Polish Cystic Fibrosis Patient Registry;
- Polish Registry of Inherited Tubulopathies (POLTube);
- Polish Registry of Patients with Neuromuscular Diseases;
- Polish Registry of Primary Immunodeficiencies;
- Polish severe chronic neutropenia registry, Polish Registry of Congenital Malformations (PRCM) [18].

The total number of diseases reported in the registers of 36 countries is 702 [18], which is about 10% of identified rare diseases [19].

To a large extent, the epidemiology of rare diseases is based on the estimation of disease prevalence and incidence in populations. In the case of the RDs, data on the number of disease cases are published. If the disease is of a family nature, the number of families affected by this disease is published. Collection of data necessary for the epidemiological characterization of RDs is of hybrid character: all available sources of information are used. In the reports prepared jointly by Orphanet [20] and Orphadata [21] the following sources of information are used: registers of rare international diseases of individual countries, and information collected by national and international institutes and agencies. The strategy of studying the epidemiology of RDs involves using all useful information sources and techniques, including statistical studies on hospital morbidity, in order to estimate the value of epidemiological measures.

Hospital morbidity study in Poland. One of the sources of information on rare diseases can be a regular public statistics survey. Since 1979, a hospital morbidity study has been carried out in Poland by the Polish National Institute of Public Health in Warsaw. This is part of research within the framework of the public statistics system [22]. From the perspective of research into the epidemiology of RDs, the current study is important for the following reasons:
- it includes all RD cases;
- it includes diagnoses in accordance with the International Statistical Classification of Diseases and Related Health Problems (ICD-10);
- information on each disease case, information format and the process of including this information in the register kept by the National Institute of Public Health are normalized.

Data on all hospitalized cases have been collected since 2000. A special website for sharing data has been designed, offering the possibility to create result reports for external users [23].

In recent years, the so-called General Hospital Disease has been included into the minimum data set, thus enabling better comparison of data from different countries of the European Community. Similar data are collected in all countries of the Organization for Economic Co-operation and Development (OECD) and used in the analyzes of this organization [24].

OBJECTIVE

The main purpose of the study was to present methodological premises for the use of statistical data on hospital morbidity in the study on the epidemiology of selected rare systemic vasculitides, as examples of rare rheumatic diseases.

Methodological premises and the results of the analyzes of population-based statistical databases prepared by the Polish National Institute of Public Health in 2004–2013 are presented. A substantive analysis is made of the possibilities and conditions for using data on hospital morbidity in the epidemiological studies on rare rheumatic diseases.
POSSIBILITIES FOR USING STATISTICAL DATA ON HOSPITAL MORBIDITY

Data from the hospital morbidity database useful for epidemiological purposes. The full scope of information on hospitalization cases available in the hospital morbidity register is much broader than that presented below. However, the nature of the study conducted by the authors required the use of specific data. Data on rare rheumatic diseases that are useful for epidemiological purposes include:
- date of admission to hospital and discharge from the hospital;
- date of birth, age and gender of the hospitalized patient;
- code of the commune competent for the place of residence;
- hospital department code;
- diagnosis of the underlying disease according to the current International Statistical Classification of Diseases and Related Health Problems (ICD-10);
- commodities according to the current International Statistical Classification of Diseases and Related Health Problems ICD-10 (max. 25 commodities);
- primary and secondary cause of death.

Use of data from the statistical hospital morbidity study in an analysis of rare rheumatic diseases.
1. Date of admission to hospital and discharge from the hospital can assist in determining an approximate onset of the disease and establish the incidence, prevalence and morbidity rate, seasonality of the disease, and the dynamics of the disease both in a long-term period and within one year.
2. Date of birth, age and gender of the hospitalized patient. Data can be helpful in studies on the dynamics of hospitalization with regard to age and gender, as well as in studies on the relationship between the frequency of hospitalizations and gender and age. This type of data is also used in group analysis, where the differentiating factor is either gender or age.
3. Code of the commune competent for the place of residence, date of birth, gender, date of admission to hospital. The hospital morbidity database does not include patients’ names. It is impossible to identify patients directly using the data from the database. Thus, it is difficult to compile the hospitalizations of an individual patient. The authors’ original solution assumes, with a probability close to 100%, that different hospitalizations can be assigned to one person based on the date of birth, code of the commune and gender. In addition, in the case of patients with rare diseases, the risk of matching two different patients to one hospitalization is very small and can be ignored in the epidemiological analyzes. Two analytic samples were considered in the studies: all hospitalizations and first-time hospitalizations. In addition to data from the hospital morbidity study, demographic data for the general Polish population were obtained from the Polish Central Statistical Office. The incidence and prevalence rates were calculated using the number of patients and corresponding census data.
4. The hospital department code can be useful in estimating the dominant symptoms in hospitalized patients. It can be assumed that the department where a patient was hospitalized was related to dominant symptoms of the disease or diseases.
5. Diagnosis of the underlying disease and commodities according to the current International Statistical Classification of Diseases and Related Health Problems (ICD-10). All disease identifying data are helpful when a patient is hospitalized for a different reason, but a rare disease was indicated among co-morbid diseases.
6. For the primary and secondary cause of death, data are needed to determine the cause of death and determine the average survival rates in patients with a rare disease.
7. Data on admission mode, length of stay, hospital procedures and number of deaths per hospitalizations are helpful in an analysis of health care system efficiency and rare disease treatment.

Strengths and limitations of the study according to the use of data from the statistical hospital morbidity study.
The authors believe that the study has major strengths. Since it was performed in a big group of the Polish population, it produced large, unselected cohorts, useful for an analysis of the epidemiology of rare rheumatic diseases.

Suspected systemic vasculitis requires diagnostic procedures or an advanced treatment under hospital conditions. With regard to systemic vasculitides, a long-term (at least five-year) analysis of hospital morbidity databases can be regarded as an analysis of all disease cases. However, it is impossible to completely rule out cases where the registration of a patient with a rare disease was not included in the study.
Methods of incidence analysis used in the presented study allowed for an anonymous and accurate (though not 100% accurate) analysis of the first cases of hospitalizations in patients with rare systemic vasculitides.

The study also has some limitations, one of which is that it was based on a retrospective review and relied on discharge records from inpatient hospitalizations. This source of information does not include patients who were treated only as outpatients, or whose condition was stable and did not require hospitalization in the period of time defined as the observation period of the study. The first diagnosis of systemic vasculitis in the hospital morbidity database is not necessarily the date of the first diagnosis, which may lead to miscalculation of the incident cases. Nevertheless, a long observation period can minimize the overestimation in this study.

It was assumed that the information on diseases used in this study was reliable and based on current standards for diagnosing systemic vasculitides. However, the accuracy of diagnoses was not analyzed with the use of available medical documentation. A detailed analysis of data correctness can require significant resources, and potential benefits may not have a significant impact on the estimation of the epidemiology of these rare diseases.

Another potential problem could be failure to identify all comorbidities in cases where the reason for hospitalization was not the rare disease in question. A longer observation period can reduce the likelihood of the rare disease being overlooked when reporting the cases to the national registers.

RESULTS

Table 1 presents the results of conducted and published studies on the epidemiology of rare rheumatic diseases, including rare systemic vasculitides. The studies concerned
diseases such as Behçet’s disease [25], Churg-Strauss Syndrome [26], Kawasaki disease [27], Polyarteritis nodosa [28], and Granulomatosis with polyangiitis [29]. The results obtained in these studies were positively verified against data from other sources of world scientific literature.

The results show that the method presented in the study can be useful in the research into rare rheumatic diseases, in particular rare systemic vasculitides. Hospitals in Poland are legally obliged to collect data and submit them to the hospital morbidity register. Polish databases are a part of the European database of hospital morbidity and have to fulfill the requirements of reliability defined by the European institutions dealing with the collection of data on hospitalization in the EU countries, including Poland.

It should be emphasized that in the absence of other data resources that could be used to assess the occurrence of rare rheumatic diseases, including systemic vasculitides, statistical data on hospital morbidity, in practice are the only available and reliable source of information for epidemiological purposes.

CONCLUSIONS

1. In the research into the occurrence of rare rheumatic diseases, including systemic vasculitides, it is advisable to use databases from statistical studies on hospital morbidity. If specific methodological assumptions are made, an analysis of these databases allows for an identification of disease cases and a descriptive epidemiological analysis.

2. The greatest advantage of using hospital morbidity database in the studies on rare diseases is that this database covers almost all hospitalization cases over a long period of time that are registered in a uniform manner. This results from the obligation to submit data to national registers and from the requirements of European institutions dealing with medical statistics.

3. The results concerning the incidence and prevalence of rare systemic vasculitides, which are obtained on the basis of an analysis of statistical data on hospital morbidity in Poland, are reliable and comparable with data from resources published in other European countries.

4. Hospital morbidity databases can be useful for the development of clinical trials on rare diseases, improvement of patient care and organization of the health care system.

REFERENCES


2. FAQs About Rare Diseases. Genetic and Rare Diseases Information Center (GARD) – an NCATS Program. 2018.05.17. https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases (access: 2018.05.17).


5. 10 things you should know about rheumatic diseases. https://www.eular.org/myUploadData/files/10%20things%20on%20RD.pdf (access: 2018.05.15).


15. Foocharoen C, Thavornpitak Y, Mahakankanukruath A, Suwannaroj S, Nanagar A. Admission rate and characteristics of hospitalized systemic
Wykorzystanie statystycznych danych o chorobowości szpitalnej w badaniach epidemiologicznych rzadkich chorób reumatycznych

**Streszczenie**


Cel pracy. Głównym celem pracy jest przedstawienie możliwości wykorzystania danych dotyczących statystycznego badania chorobowości szpitalnej do celów analizy epidemiologii chorób rzadkich, na przykładzie wybranych układowych zapalenń naczyń.

Stan wiedzy. W pracy przedstawiono doświadczenia z przeprowadzonych analiz danych z krajowego statystycznego badania chorobowości szpitalnej, prowadzonego w latach 2004–2013 przez Narodowy Instytut Zdrowia w odniesieniu do rzadkich chorób reumatycznych na przykładzie układowych zapalenień naczyń oraz omówiono główne zalety i ograniczenia przyjętej metodologii. Na podstawie dotychczas przeprowadzonych badań wykazano, że dane pochodzące z rejestrów szpitali krajowych mogą być wykorzystywane do analizy chorobowości, zapadalności, częstości występowania choroby, rozkładu płci i rozmieszczenia terytorialnego oraz analizy współczynników umieralności (śmiertelności) szpitalnej i tendencji w obserwacjach długoterminowych.

Wniosek. Analiza epidemiologii tych chorób może przyczynić się do lepszego zrozumienia ich specyfiki, co może być przydatne w codziennej praktyce klinicznej.

**Słowa kluczowe**

hospitalizacja, choroba rzadka, krajowe rejestrzy, choroby reumatyczne, układowe zapalenia naczyń

---