Health Policy 112 (2013) 100–109

Contents lists available at ScienceDirect

Health Policy

journal homepage: www.elsevier.com/locate/healthpol

Health care performance comparison using a disease-based approach: The EuroHOPE project☆

Unto Häkkinen[a,*,1], Tor Iversen[b,1], Mikko Peltola[a,1], Timo T. Seppälä[a,1], Antti Malmivaara[a,1], Éva Belicza[c,1], Giovanni Fattore[d,1], Dino Numerato[d,1], Richard Heijink[e,1], Emma Medin[f,1], Clas Rehnberg[f,1]

[a] National Institute for Health and Welfare, Centre for Health and Social Economics (CHESS), Helsinki, Finland
[b] The University of Oslo, The Department of Health Management and Health Economics, Oslo, Norway
[c] Semmelweis University, Budapest, Hungary
[d] Bocconi University, Department of Policy Analysis and Public Management, CERGAS – Centre for Research in Healthcare Management, Milano, Italy
[e] National Institute for Public Health and the Environment, Centre for Prevention and Health Services Research, Bilthoven, The Netherlands
[f] Medical Management Centre, Karolinska Institutet, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 1 November 2012
Received in revised form 12 April 2013
Accepted 17 April 2013

Keywords:
Performance measurement
Health care
Disease-based approach
Episode
Outcome
Cost

ABSTRACT

This article describes the methodological challenges associated with disease-based international comparison of health system performance and how they have been addressed in the EuroHOPE (European Health Care Outcomes, Performance and Efficiency) project. The project uses linkable patient-level data available from national sources of Finland, Hungary, Italy, The Netherlands, Norway, Scotland and Sweden. The data allow measuring the outcome and the use of resources in uniformly-defined patient groups using standardized risk adjustment procedures in the participating countries. The project concentrates on five important disease groups: acute myocardial infarction (AMI), ischemic stroke, hip fracture, breast cancer and very low birth weight and preterm infants (VLBWI). The essentials of data gathering, the definition of the episode of care, the developed indicators concerning baseline statistics, treatment process, cost and outcomes are described. The preliminary results indicate that the disease-based approach is attractive for international performance analyses, because it produces various measures not only at country level but also at regional and hospital level across countries. The possibility of linking hospital discharge register to other databases and the availability of comprehensive register data will determine whether the approach can be expanded to other diseases and countries.

1. Introduction

International comparison of performance can proceed at a number of possible levels, including system wide, by disease, and by sub-sector (such as hospital or nursing homes) [1]. There are arguments for and against each, but when it comes to health outcomes the disease-based approach is the most suitable, since the health gains of the activities can be measured quite accurately at the disease level [2]. Compared with system wide comparison, this approach reduces heterogeneity of the population studied

☆ Open Access for this article is made possible by a collaboration between Health Policy and The European Observatory on Health Systems and Policies.

* Corresponding author at: National Institute for Health and Welfare, Centre for Health and Social Economics (CHESS), Mannerheimintie 166, 00270 Helsinki, Finland. Tel.: +358 29 524 7327; fax: +358 29 524 7537.
E-mail address: unto.hakkinen@thl.fi (U. Häkkinen).

[1] On behalf of the EuroHOPE Study Group.

0168-8510
http://dx.doi.org/10.1016/j.healthpol.2013.04.013

Table 1
EuroHOPE in a nutshell.

Project name	European Health Care Outcomes, Performance and Efficiency (EuroHOPE).
Main task	To contemplate and compare national as well as international differences in five economically important patient groups with respect to effectiveness and efficiency of their whole cycle of care.
Patient groups	Acute myocardial infarction, stroke, hip fracture, breast cancer, and very low birth weight and very preterm infants.
Coordinator	National Institute for Health and Welfare, Finland.
Partners	Karolinska Institutet, Sweden; Ragnar Frisch Centre for Economic Research, Norway; University of Oslo, Norway; University of Edinburgh, UK (Scotland); Universita Commerciale Luigi Bocconi, Italy; Semmelweis University, Hungary; National Institute of Public Health and Environment, Netherlands.
Duration	4 years, 2010–2013.

and allows more precise identification of the characteristics of the health care process that have an effect on outcomes.

All international comparisons require suitable information systems. These systems have until now been developed using two different approaches. The first approach relies on developing a coherent conceptual framework for information collection, analysis and dissemination. An example of this is national accounts, in which health care is dealt as a part of the whole economy. Another approach assembles readily accessible data, often the by-products of existing national data collection, such as hospital discharge registers, as well as work that has been done for other purposes. The bottom-up approach relies on individual experts, provider organizations, and countries engaging in quality and efficiency improvement initiatives. Micro-level comparative data on clinical actions, costs and outcomes represents an essential element of such an approach. In this case, the precise definition, collection and scrutiny of the data are left to expert groups to determine [3].

A desirable health care performance measure at the disease level is one that reliably and accurately reflects the process, costs and outcome of care [4]. Such a measure provides valuable information for improving treatment processes and for steering at national, regional as well as hospital levels. In addition, measures that enable reliable comparisons across providers might encourage them to develop their treatment processes to attain better positioning in benchmarking.

By making use of available databases through a microeconomic disease-based approach,[2] the EuroHOPE project (European Health Care Outcomes, Performance and Efficiency, Table 1) evaluates the performance of European health care systems in terms of outcomes, quality, use of resources and cost. EuroHOPE uses patient-level data available from linkable national or regional registers and other data sources that allow for measuring the outcome (by following what happens to patients) and the use of resources (such as cost, number of hospital days, treatment with specific procedures and drugs) in the selected well-defined and risk adjusted patient groups.[3] Thus, EuroHOPE is not only focusing on a specific treatment or hospital stay (measured e.g. in term of DRGs) [5] per se, but rather on outcomes and costs related to the complete clinical pathway.

Using register data it is not possible to consider patient or provider experiences. For example, no information on health-related quality of life nor patient satisfaction is routinely available in the registers and thus cannot be used for enriching comparisons between regions or hospitals. In order to fill the gap, feasibility studies making use of patient surveys on health-related quality of life and patient satisfaction in selected hospitals in each country have been initiated in EuroHOPE.[4]

The project concentrates on five disease groups: acute myocardial infarction (AMI), ischemic stroke, hip fracture, breast cancer and very low birth weight and preterm infants (VLBWI). The approach is based on analysing the progress of a disease, with specific interest in the role of health services and health care policy as a determinant of the progress. The main idea of the approach is that it analyses performance by using detailed data pertaining to specific health conditions to illuminate the interconnected aspects (i.e. financing, organizational structures, medical technology choices) that are responsible for health system performance (i.e. health outcomes and expenditure). Of the seven countries participating in EuroHOPE, five countries are considered as tax-based systems and two countries rely on social health insurance. Two of the tax-based systems, Norway and Scotland, mainly rely on central taxation, whereas the other three, Finland, Italy and Sweden, in various ways rely on regional and local taxes. The two countries with social insurance systems also differ as The Netherlands rely on a system with multiple insurers and regulated competition among them, whereas the Hungarian system is a social health insurance system with a sole insurer with a monopsony power.

This article describes the methodological challenges of international comparisons related to a disease-based approach and how they are addressed in the EuroHOPE project. The project is based on merged register data. Thus, this article shows how existing data sources can be

[2] The EuroHOPE project applies sub-sector approach to Nordic hospitals. This will be discussed elsewhere in this issue [29].

[3] This idea is not novel, however. For example, Eurostat [30] has defined the output in health care: "Health output is the quantity of care received by patients, adjusted to allow for the qualities of service provided, for each type of health care. The quantities should be weighted together using data on the costs or prices of the health care provided. The quantity of health care received by patients should be measured in terms of complete treatments."

[4] Feasibility will be studied in two different dimensions: in establishing common protocols for survey data collection in differing countries, and in collecting information in differing patient groups. For health related quality of life two different measures EQ-5D and 15D and for patient satisfaction EORTC-IN-PATSAT32 have been utilised together with collection of rich background information.

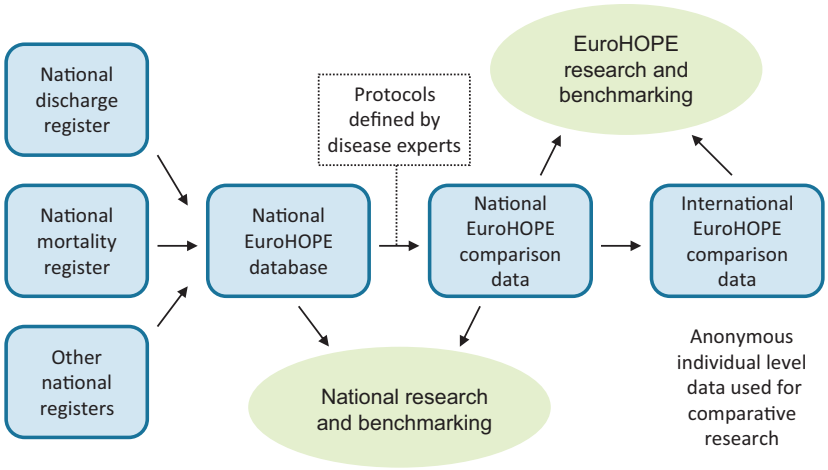

Fig. 1. EuroHOPE databases.

exploited in performance analysis using the bottom-up approach.

2. Methodology and data

A performance measure must be carefully constructed and be appropriate for comparisons in order to be useful. There are many important and well-documented methodological and practical questions that need to be considered when administrative data are used for performance measurement [6–9]. Performance indicators will vary due to type of hospital, regional or individual level variations or random variation. The focus of our interest is in variation at the hospital, regional level and country level. In EuroHOPE, cross-country heterogeneity (e.g. in comorbidity conditions) is reduced by using proper risk adjustment methods for individual-level data.

Implementing register-based performance evaluation requires – in addition to the availability of comprehensive data – methodological understanding and a multidisciplinary involvement. Health system knowledge is essential for deciding the scope and the specific questions to be addressed, and needs to be supplemented with understanding of the possibilities and limitations of register information [10]. Clinical knowledge is needed when appraising details of the indication and management of a disease, and economic, epidemiological, statistical and data mining expertise are required to ensure that the methodology is appropriate. Finally, all the aforementioned must be integrated during the entire process. In EuroHOPE, the performance indicators have been developed in collaboration with clinical experts for all the diseases in focus as well as with experts in health economics, epidemiology and statistics.

The five EuroHOPE disease groups were chosen because they have high prevalence, health impact or high economic burden in all developed countries (AMI, ischemic stroke, hip fracture, breast cancer [11]), or they are very resource intense (VLBWI). Different populations are in focus, such as the elderly (AMI, ischemic stroke, hip fracture), middle aged (breast cancer) and newborn (VLBWI) or the risk for the conditions differ by sex (AMI, hip fracture and breast cancer). The diseases differ in the involvement of different types of specialities, regarding either surgical treatment or outpatient pharmaceutical treatment, and also with respect to technological change. As a result, the impact of health system characteristics (financing, organization, technology, etc.) can be tested across a diverse set of areas, creating a stronger, more comprehensive evidence base. The choice is also based on a pre-enquiry and the assumption that these diseases offer the best possibilities to access internationally comparative data, since in all of the diseases the main responsibility for patient treatment falls to acute hospital care.

We developed an international comparative database that allows performance analysis, research and use indicators calculated at national, regional and hospital levels (Fig. 1). The disease-based approach requires patient-level data covering the whole population and the possibility to deterministically link records in different national registers. In the six included countries (Finland, Hungary, the Netherlands,[5] Norway, Sweden and Scotland) included in the EuroHOPE project it was possible to link national hospital-discharge registers with mortality registers and in five countries (excluding Scotland) also with registers of prescribed medicines (Table 2). In Italy, similar data were available for two geographical areas. All databases present population data reflecting patterns of care and outcomes of the entire population residing in the defined territories.

For each disease, the building of the database is based on several stages: (i) definition of the patient population, (ii) collection of the register material for the patient population at hand, (iii) definition of the start and end of the episode (by defining and using the index admission and deciding how referrals should be treated) for the patients and from the necessary variables concerning the care given, (iv) checking history and follow-up of the use of health care services in order to define state and time variables for the

[5] In the Netherlands the data is available of about 85% of the patients, since collection of national discharge data is voluntary.

Table 2
Registers used in EuroHOPE.

Country	Register name (year since included in EuroHOPE)	Sub-studies in which the register is used
Finland	Hospital discharge register	1–5
	Cause of death register	1–5
	Drug prescription register	1–5
	Cancer register	4
	Medical birth register	5
	Data file on small preterm infants	5
	Register of congenital malformations	5
Hungary	Hospital discharge register	1–5
	Death register	1–5
	Prescribed medicine register	1–5
	Cancer register	4
	Child birth register	5
Italy	Hospital discharge register	1–5
	Causes of death register	1–5
	Register of medicine prescribed and dispensed by the Italian NHS	1–5
	Outpatient services in specialist care register	1–5
	Cancer register (only city of Turin)	4
	Medical birth registry (only province of Rome)	5
The Netherlands	Hospital discharge register	1–3,5
	Causes of death register	1–3,5
	Prescribed drug register	1–3,5
	The Netherlands perinatal registry	5
Norway	Norwegian patient register	1, 3–5
	Norwegian prescription database	1, 3–5
	National cancer register	4
	Cause of death register	1, 3–5
	Medical birth registry of Norway	5
	The Norwegian hip fracture register	3
Scotland	Scottish morbidity record 00 – outpatient attendance	1–5
	Scottish morbidity record 01 – general acute inpatient and day case	1–5
	National records of Scotland death extract	1–5
	Scottish open cancer registration and tumor enumeration system	4
	National records of Scotland birth extract	5
	Scottish birth record	5
	Scottish morbidity record 02 – maternity inpatient and day case	5
Sweden	National patient register	1–5
	National cause of death register	1–5
	National prescribed drug register	1–5
	National cancer register	4
	National medical birth register	5

Sub-studies: 1 AMI, 2 stroke, 3 hip fracture, 4 breast cancer, 5 very low birth weight and very preterm infants.

patients, (v) construction of the comorbidity variables, (vi) calculation of the direct health care cost, and finally, (vii) combination of the information of the previous stages in order to generate the comparison database.

3. The development of performance measurement

3.1. Episode of care

The concept of an episode has been used to distinguish between discharge and single intervention. The idea of an episode approach is not new [12,13], but the implementation of this concept has been challenging in practice [14] and has not been done before in such a large scale.

An episode of care refers to the entire treatment pattern from the beginning of the disease (e.g. time of diagnosis) to the end of the treatment across organizational boundaries to face the health problem at hand in a specific time frame. Thus the protocol for an episode includes the definitions of starting and finishing dates (follow-up time) as well as inclusion and exclusion criteria, which are used when constructing a comparison data set for a specific disease group.

In EuroHOPE, the follow-up data covers at least one year for each patient. Main observable events in register data are process measures such as admissions, procedures, and discharges as well as outcomes measures such as survival. Secondary observable events are outpatient visits and prescribed medication purchases. In addition to the follow-up data, we have similar information on the history of service use of the patients. Using the available data it is possible to reconstruct treatment pathways that describe what has happened (before and) after e.g. an operation on a daily basis [10] (Fig. 2).

Episode of care provides a framework that can be operationalized in terms of linkable register data. In order to

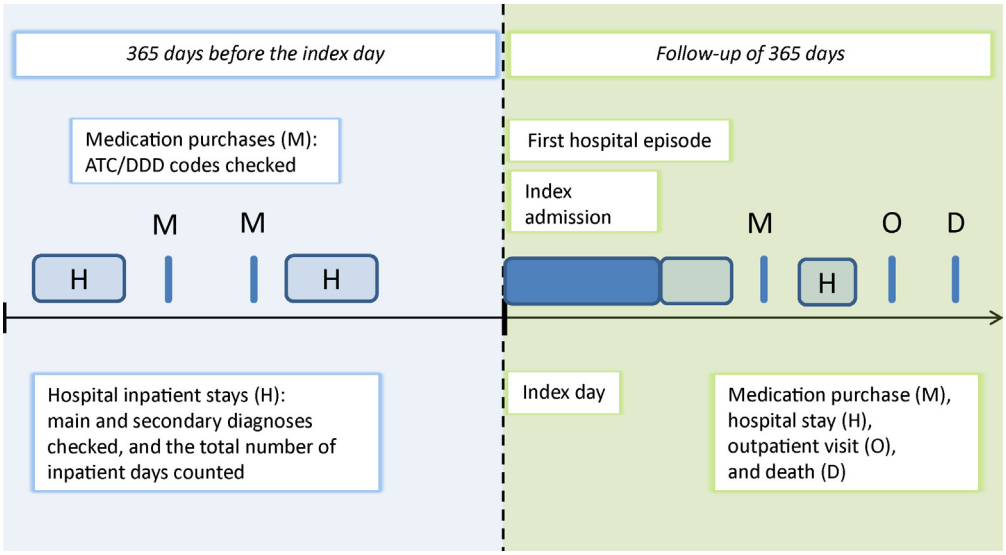

Fig. 2. Diagram of data utilization in EuroHOPE.

model the episode of care and calculate the measures of performance, the nature of the disease and the medical history and comorbidities of the patients must be taken into account, all key factors affecting treatment decisions, delivery of care, and outcomes.

The episodes for AMI, ischemic stroke, and hip fracture start with an acute phase in the hospital, usually occurring immediately after the event. In the case of VLBWI and breast cancer the episode starts with birth or a disease event, respectively, followed by hospital use. The hospital admission that is the first in the episode of care is in our terminology defined as the index admission, and the admission day the index day. The first hospital episode starts on the index day and terminates at the day of the first discharge to home, or by death, or after a specified time of continuous inpatient care, depending on the disease. The first hospital episode describes the acute-phase of treatment. The follow-up period ends one year after the index day or at death. The total episode of care includes all the health service used during the follow-up period (Fig. 2). For AMI, ischemic stroke and hip fracture patients, the index day (i.e. the beginning of episode) is defined by means of the main diagnosis of the hospital stay (AMI ICD-10-codes I21-I122, ischemic stroke I63, hip fracture S72.0-S72.2) using at least two exclusions. First, we exclude all patients who have had a hospital admission related the disease in question during the previous 365 days before the index day. Second, we exclude non-residents and patients with an incomplete personal identity number (PID).

3.2. Indicators

The benchmarking task of the project is carried out through basic reports (available on the internet) which include performance indicators at national and regional level. The indicators presented in the disease-specific basic reports can be roughly categorized into four different groups: baseline statistics, process indicators, cost indicators, and outcome indicators.

In each disease-specific analysis, the baseline statistics of the patient population in each country and region are given. These include, e.g. the number of patients affected by the disease, mean and median age and proportion of males and females.

3.2.1. Process

The process indicators describe health care service use during the episode of care. The indicators include measures such as length of stay in hospitals (in first hospital episode and during the follow-up period), procedures and other treatment practices, and use of prescribed medicines.

3.2.2. Costs

The cost indicators describe the cost of first hospital episode and the total episode of care. Ideally, there would be detailed cost calculations for each individual patient available and standardized method of cost calculations across countries. Although many of the countries are using Diagnostic Related Groups (DRGs) or a similar system to calculate resource use related to standardized patients, there is no common method of grouping patients [15] and DRG tariffs may not accurately reflect costs. Since Geue et al. [16] show that the costing method may have an important impact on how the effects of explanatory variables on cost are assessed, it is important to find measures that are considered to be valid and comparable. As a prerequisite each country should be capable of providing the necessary data input. In EuroHOPE, the cost and utilization measures are in general limited to hospital care (inpatient and outpatient) and pharmaceuticals dispensed outside hospital.

Two approaches that are designed to complement each other are set up. In the first approach, the essential components of resource use during the hospital treatment are defined and extracted from register data at the individual

Table 3
Description of regions used in EuroHOPE reporting.

Country	Description	Number of regions	Average population size
Finland	Hospital districts and hospital regions responsible for providing specialized health care. Smallest districts combined.	19	280,000
Hungary	19 counties and Budapest area providing self-governmental administrative duties (not health care).	20	500,000
Italy	City of Turin and province of Rome (divided into two smaller units: city of Rome and outside of Rome municipalities).	3	1,630,000
The Netherlands	Provinces responsible for matters of subnational or regional importance (not health care).	12	1,380,000
Norway	Hospital trusts responsible for providing specialist health care in their geographical areas.	20	250,000
Scotland	Health boards responsible for health care. Smallest boards combined.	14	370,000
Sweden	Counties responsible for providing health care.	21	450,000

patient level. Then a resource weight is attached to each component. Resource weights are based on patient level cost data from Sweden [17]. The first cost component is the procedure/treatment given during the hospital stay. For instance, for AMI, treatment with coronary artery bypass surgery (CABG) and percutaneous coronary intervention (PCI) are registered and resource weights are attached to them. Another component is the number of hospital days, with a cost weight attached that describes the basic cost of a hospital day. Inpatient stays for main diagnoses other than AMI are assessed by a cost per day corresponding to the mean cost per day for all admissions registered in the cost per patient database. All outpatient consultations are registered and weighted by the mean cost of an outpatient consultation. Eurostat purchasing power parities (PPP) for hospital services [18] are used to convert costs to the cost level in the EU15.

In the second approach, each country makes use of their best available data to calculate hospital costs at the individual patient level. Some countries may take advantage of cost data at the individual patient level, while others are using classification systems (e.g., DRGs) and related tariffs as proxies of costs. Results of approach one and approach two are supposed to supplement each other in the comparative analyses of costs across countries. In both approaches, the cost of pharmaceuticals dispensed outside hospitals is recorded at pharmacy selling price as registered in the national pharmaceutical databases. Then, the PPP for each country's Gross Domestic Product (GDP) is used to adjust for cost level in each country.

3.3. Outcomes

In the EuroHOPE project outcome indicators are considered those measures of health that are attributable to health care. The most important outcome measures are mortality at the 30-day, 90-day and one-year follow up. In addition, we use other outcome measures calculated from registers such as readmission and complications, since it is generally agreed that these have an impact on patient outcomes.

3.3.1. Regional comparison

We calculate indicators by countries and within countries by regions. The regional indicators are based on patients' place of residence. Each country has defined the partition of its regions to be suitable for benchmarking (Table 3). In Finland, Italy, Norway, Scotland and Sweden the regions describe local authorities responsible for health care, while in social health insurance countries (the Netherlands and Hungary) the regions are based on regional governmental or sub-national geographical division where public authorities have no (Netherlands) or limited responsibility of health care (Hungary). In the last mentioned countries, the average population size of the regions is much greater than in the Nordic countries and Scotland. However, in all countries the metropolitan areas are constructed so that they enable international comparison.

3.4. Risk adjustment

When comparing countries, regions, hospitals, and yearly patient cohorts, patient-associated factors must be accounted for. EuroHOPE has endeavored to ensure meaningful comparisons using three steps. Firstly, the disease groups have been defined so that they are as comparable and homogeneous as possible. Secondly, information on risk factors has been gathered from the patients' medical history. Thirdly, statistical models have been applied to adjust the indicators and calculated their 95% confidence intervals.

3.4.1. Risk adjustment variables

One of the most commonly-used, and clinically important, risk adjustment variable is comorbidity. Numerous measures of comorbidity are available when using administrative data (for a review see e.g. [19–21], the most common being the Elixhauser method [22] and the

Table 4
Comorbid diseases in EuroHOPE AMI study.

Comorbidity	ICD-10	ICD-9	ATC/DDD
Hypertension	I10*–I15*	40*	C03*, C07* (with neither coronary artery disease nor atrial fibrillation indicates hypertension), C08*, C09*
Coronary artery disease	I20*–I25*	410*–414*	N/A
Atrial fibrillation	I48*	4273*	N/A
Cardiac insufficiency	I50*	428*	N/A
Diabetes mellitus	E10*–E14*	250*	A10A*, A10B*
Atherosclerosis	I70*	440*	N/A
Cancer	C00*–C99**, D00*–D09*	140*–208*	L01* (except L01BA01)
COPD and asthma	J44*–J46*	4912*, 496*, 496*	R03*
Dementia	F00*–F03*, G30*	290*, 3310*	N06D*
Depression	F32*–F34*	2962*, 2963*	N06A*
Parkinson's disease	G20*	332*	N04B*
Mental disorders	F20*–F31*	295*–298* except 2962* and 2963*	N05A* (except N05AB01 and N05AB04), and no dementia
Renal insufficiency	N18*	585*	N/A
Alcoholism	F10*–F19*	291*, 304*, 305*	N/A
Stroke	I60*, I61*, I63*, I64*, G45*	430*–438*	N/A

* Indicates that all subgroups are included i.e. I63* = I63.0–I63.9.

Charlson Comorbidity Index (CMI) [23]). Besides comorbidity chartings based on information about secondary diagnoses in the hospital discharge data, the Anatomical Therapeutic Chemical (ATC) classification system has been used to identify pharmaceutical treatments that are common for a selection of relevant comorbid diseases [24–26]. Using the data available and experience of these measures, as well as statistical testing, our disease-specific expert groups separately tailored a set of conditions for each disease that were potentially used for risk adjustment. In the final risk adjustment only those conditions whose prevalence exceeded 1% in each of the participating countries' data for that particular disease were included.

Two different databases were used to identify for comorbidity information in patients' records: the hospital discharge register and the register of prescribed medicines. The two sources overlap in terms of the content of information, i.e. a person hospitalized with a specific comorbidity has a high probability of having purchased a prescribed medicine for that illness. On the other hand, the registers complement each other so that the information of the comorbidities of patients becomes reliable. This is relevant, in particular because of the expected undercoding of secondary diagnoses in discharge registers. The diagnoses (main and secondary diagnoses separately) of patients' inpatient hospital treatments in the previous year (i.e. 365 days before the index admission) are checked. The assumption made is that if a person has been hospitalized with a diagnosis of a comorbidity condition before the start of the episode, the patient has had the comorbidity at the time of the event under observation (except among low birth weight infants). Similarly, purchases for prescribed medicines are checked for the last 365 days before the start of the episode of care (Fig. 1). Hospital discharges and medicine purchases during the event or afterwards are not checked for comorbidities. In addition to checking the diagnoses of a patient, we have counted the number of hospital inpatient days in the year preceding the index admission. Age and sex of the patient are also considered in the risk adjustment.

The comorbidities available for AMI patients are presented in Table 4. For each indicator we perform risk adjustment with three different sets of confounders: (1) age and sex (M1), (2) age, sex, hospital days in the previous year, and comorbid diseases based on hospital discharge data (M2), and (3) age, sex, hospital days in the previous year, and comorbid diseases based on both medication and hospital discharge data (M3). The comorbidities used in each disease group vary; the table provides a broad perspective on the comorbidities selected in the study and in the marking criteria.

3.4.2. Risk adjustment modeling

There are many possible methods that could be used for risk adjustment, such as methods related both to observable confounders (standardization using different approaches, such as nonlinear regressions, propensity score, confidence intervals using shrinkage estimators and other Bayesian methods) and unobservable confounders (instrumental variable methods and two-stage methods) [8]. In practice, the selection of appropriate methods is based on balancing what can be done on a routine basis with adequate methodological aspects. As the number of indicators in EuroHOPE is high, the development of refined statistical models for each indicator separately is not feasible when it comes to reporting and benchmarking purposes.

Thus, we chose to carry out an indirect standardization for measures of incidence, while for all other indicators a modeling strategy is adopted: logistic regression for dichotomic responses (e.g. mortality), generalized linear modeling for continuous responses (e.g. for costs logit-link with gamma distribution), and negative binomial modeling for discrete responses (e.g. length of stay). The simplified approach chosen is justified on the practical grounds mentioned above. In addition, our methodology is more accessible to a wider audience than are the more complex alternatives, although we recognize that more advanced methods exist and that the reporting might benefit from those.

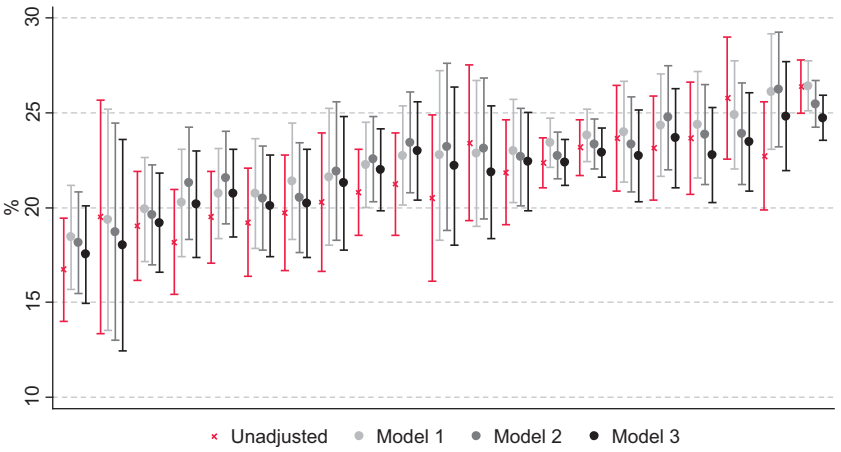

Fig. 3. One-year mortality of AMI patients in Sweden 2007. Regional unadjusted and risk-adjusted shares and their 95% confidence intervals.

In the estimation of the risk adjustment models a complication arises from the involvement of many different countries. Ideally, the individual-level data from all participating countries would be pooled before estimating the risk adjustment models However, not all countries have permission to share the individual level data which is why the risk adjustment is based on data from around half of the included countries.

In each disease, the parameter estimates for the confounding factors are first estimated for every process or outcome measure using the broadest possible data for the disease in question [e.g. for AMI the international EuroHOPE comparison data, available from the year 2007 from Finland, Hungary, Sweden, Norway (data of 2009) and two regions of Italy]. The estimations are made by weighting the data so that each country has the same weight. Then, the coefficients of each model are made available to all partners who then calculate individual-level-predicted values for the indicators. The predicted values are then summed up to country and regional level. The ratio of the observed value and the expected value of the dependent variable in the comparable unit can be multiplied with the average value of the indicator in the pooled data to constitute the risk adjusted indicator.

3.4.3. Standardization in data processing and calculations

From a practical point of view, the whole process of risk adjustment and reporting of the indicators has been automatized as much as possible. Each partner was individually responsible for producing its own national EuroHOPE comparison data, with the principles stated in the disease-specific study protocols.[6] After this, the partners executed a fixed file in a common statistical program (Stata) which automatically processed the data, found the coefficients for the models from the EuroHOPE server, and calculated the predicted values and the risk adjusted values at all levels. Finally, the descriptive statistics along with the country-, regional- and hospital-level indicators and their confidence intervals were automatically transferred to a reporting template.

In addition to all this, the common statistical programs also included material (for example, modeling with the country-specific data, graphing etc.) which produced additional material for the study. The automatization of the indicator production minimized the chances for human programming errors, but most importantly it saved time and resources of all partners. Also, this enabled the use of individual-level data to the fullest extent in all countries.

4. Illustrative examples

The results of analyses for AMI patients are used to illustrate the methods for risk adjustment. In Fig. 3, the three alternative risk adjustment approaches (M1, M2 and M3) for the one-year mortality are compared with crude mortality between the Swedish regions (counties). It seems that the different risk adjustment methods generate similar county rankings: The correlation coefficient between the risk adjusted measures varied from 0.99 to 0.96. The crude mortality was less clearly associated with the risk adjusted figures (correlations 0.90–0.80).

Fig. 4 describes regional variation in age- and sex-standardized one-year mortality of AMI patients in the seven countries. Mortality as well as its regional variation was much higher in Finland, Hungary and Scotland than in the three other countries.

5. Discussion

The disease-based approach is attractive for international comparative performance comparison, because it produces various measures not only at country level but also at regional and hospital level across countries. Its usefulness depends, nevertheless, on comprehensive register data being available and the possibility of linking the hospital discharge register to other databases. The EuroHOPE

[6] The study protocols for each disease are available at http://www.eurohope.info.

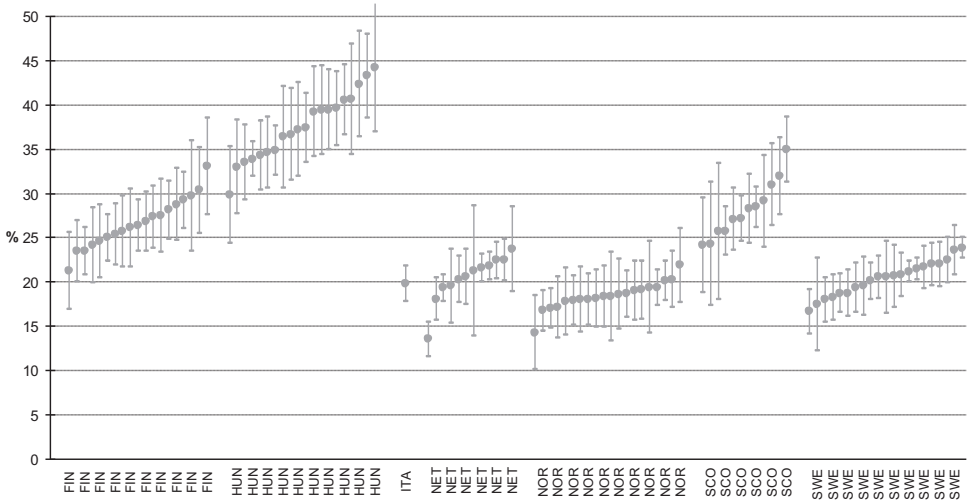

Fig. 4. One-year mortality of AMI patients per region by country, adjusted for age and sex, with confidence intervals.

project is one of the first international projects to discover the possibilities of the approach at this scale, aiming to link the data-sources using entirely standardized procedures across countries. The benefits are already apparent from the preliminary results, which indicate clearly that the approach gives new possibilities for performance comparison of certain patient groups. A variety of indicators could be used for benchmarking and data were provided at various relevant levels, that is, at the levels where most of the actions in policy making are taken. Comparison with other national and international quality databases – such as those of the OECD Health Care Quality Indicator Project (HCQI) and the Agency for Healthcare Resources and Quality (AHRQ) – reveals that the EuroHOPE project has grasped subtle and richer information on the selected diseases including both long-term outcomes and costs. In addition, the follow-up of patients extends through the whole clinical pathway.

There are still challenges to be met. For some of those participating countries, the formal and technical processes of having access to data are too slow and cumbersome. In these countries, public authorities need to be reminded by researchers that domestic and international comparisons require access to detailed patient-level data. As highlighted by the Hospital Data Project [27] there are differences in coding practices across countries and the quality of data is not always comparable. For example, the incidence of AMI (new AMI patients/population) was clearly higher in Sweden and Norway compared to Finland and Hungary. At this stage, it is not clear whether this reflects true differences in the use of acute hospital care of the patient group or coding differences of the main diagnosis. Another potential bias is related to the fact that patients are identified from the hospital discharge register. Some patients may die (or recover) before they are diagnosed in a hospital. If the proportion of these patients varies between countries, there is a potential bias in the outcome and incidence comparison. In addition, organizational structure as well as the availability of linkable data on long-term and rehabilitation care varies between countries. However, since we are utilizing individual-level data we can perform a sensitivity analysis to validate our main findings. For example, we can calculate the length of stay as well as cost both including and excluding rehabilitation and long term care. In addition, the merged international EuroHOPE comparison data enables us to use more sophisticated methods (such as propensity score matching) to increase the comparability of performance measurement [28].

The EuroHOPE project utilized mainly available register data in which outcome can be evaluated using various mortality measures or intermediate indicators such as complications, readmissions etc. These measures are crude and do not cover all dimensions of outcomes such as the effects of care on quality of life or functional status. For these measures patients are the only valid source of information. Gathering data directly from patients includes considerable expenses and feasibility challenges. Thus, the pilot studies for stroke and breast cancer patients evaluate the feasibility of establishing international comparisons on measuring health related quality of life and patient satisfaction in demanding patient environment.

The EuroHOPE project is based on data gathered from seven countries. The aim of the project is to develop methods for performance assessment that can be used for routine evaluation. Documentation with the publicly available study protocols, programming and reporting material make entry into the EuroHOPE group potentially easy. New countries must first develop and adjust their information systems while laws that might hinder available data linkages may need to be addressed. For example, an electronic record system (including all health care activities) is under development in many countries and will give new, path-breaking possibilities for the development of the disease-based approach. This requires that data using standardized and internationally comparable definitions of activities and classifications describing

the treatments (i.e. diagnosis, procedures) are nationally available for research and thus enable evaluation of the performance across countries, regions and producers.

Acknowledgements

This project was undertaken within the European Union 7th Framework Programme European Health Care Outcomes, Performance and Efficiency (EuroHOPE), Contract no. 241721

References

[1] Häkkinen U, Joumard I. Cross-country analysis of efficiency in OECD health care sector: options for research. OECD, Economics Department Working Papers, No. 554; 2007.
[2] Häkkinen U. The PERFECT-project: measuring performance of health care episodes. Annals of Medicine 2011;43(Suppl. 1):S1–3.
[3] Smith PC, Häkkinen U. Information strategies for decentralization. In: Saltman R, Bankauskaite V, Vrangbaeck K, editors. Decentralization in health care. Berkshire: Open University Press; 2007. p. 206–24.
[4] Street A, Häkkinen U. Health system productivity and efficiency. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, editors. Performance measurement for health system improvement. Experiences, challenges and prospects. Cambridge: Cambridge University Press; 2009. p. 222–48.
[5] Busse R. Do diagnosis-related groups explain variations in hospital cost and length of stay? Analyses from the EuroDRG project for 10 episodes of care across 10 European countries. Health Economics 2012;21(Suppl. 2):1–5.
[6] Wray NP, Ashton CM, Kuykendall DH, Hollingsworth JC. Using administrative databases to evaluate the quality of medical care: a conceptual framework. Social Science & Medicine 1995;40:1707–15.
[7] de Pouvourville G, Minvielle É. Measuring the quality of hospital care: the state of the art. What information should be made available to the public? In: Measuring up: improving health system performance in OECD countries. Paris: Organisation for Economic Co-operation and Development; 2002. p. 251–75.
[8] Iezzoni LI, editor. Risk adjustment for measuring health care outcomes. Chicago: Health Administration Press; 2003.
[9] Terris DD, Aron DC. Attribution and causality in health-care performance measurement. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, editors. Performance measurement for health system improvement: experiences, challenges, and prospects. Cambridge: Cambridge University Press; 2009. p. 311–38.
[10] Sund R. Utilisation of administrative registers using scientific knowledge discovery. Intelligent Data Analysis 2003;7:501–19.
[11] Heijink R, Koopmanschap MA, Polder JJ. International comparison of cost of illness. RIVM reports; 2006.
[12] Solon JA, Feeney JJ, Jones SH, Rigg RD, Sheps CG. Delineating episodes of medical care. American Journal of Public Health 1967;57:401–8.
[13] Hornbrook M, Hurtado A, Johnson R. Health care episodes: definition, measurement and use. Medical Care Review 1985;42:163–218.
[14] Rosen AK, Mayer-Oakes A. Episodes of care: theoretical frameworks versus current operational realities. The Joint Commission Journal on Quality Improvement 1999;25:111–28.
[15] Kobel C, Thuilliez J, Bellanger M, Preiffer K. DRG systems and similar patient classification systems in Europe. In: Busse R, Quentin W, Wiley M, editors. Diagnosis-related groups in Europe. Moving towards transparency, efficiency and quality in hospitals. Berkshire, UK: Open University Press; 2011. p. 37–58.
[16] Geue C, Lewsey J, Lorgelly P, Govan L, Hart C, Briggs A. Spoilt for choice: implications of using alternative methods of costing hospital episode statistics. Health Economics 2012;21:1201–16.
[17] Cost per patient (KKP – Kostnad per patient) project. See http://www.skl.se/kpp
[18] Eurostat, OECD. Methodological manual on purchasing power parities. Luxembourg: European Communities. OECD; 2006.
[19] de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. Journal of Clinical Epidemiology 2003;56:221–9.
[20] Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical Care 2005;43:1130.
[21] Holman C, Preen D, Baynham N, Inn J, Semmens J. A multipurpose comorbidity scoring system performed better than the Charlson index. Journal of Clinical Epidemiology 2005;58:1006–14.
[22] Elixhauser A, Steiner C, Harris R, Coffey R. Comorbidity measures for use with administrative data. Medical Care 1998;36:8–27.
[23] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Diseases 1987;40:373–83.
[24] Maio V, Yuen E, Rabinowitz C, Louis D, Jimbo M, Donatini A, et al. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. Journal of Health Services Research & Policy 2005;10:232–8.
[25] Lamers L, van Vliet R. The pharmacy-based cost group model: validating and adjusting the classification of medications for chronic conditions to the Dutch situation. Health Policy 2004;68:113–21.
[26] Naughton C, Bennett K, Feely J. Prevalence of chronic disease in the elderly based on a national pharmacy claims database. Age and Ageing 2006;35:633–6.
[27] Hospital Data Project. Phase 2. Final report; 2008.
[28] Schreyögg J, Stargardt T, Tiemann O. Costs and quality of hospitals in different health care systems: a multilevel approach with propensity score matching. Health Economics 2011;20:85–100.
[29] Medin E, Häkkinen U, Linna M, Anthun KS, Kittelsen S, Rehnberg C. International hospital productivity comparison: nordic experience. Health Policy 2013, http://dx.doi.org/10.1016/j.healthpol.2013.02.004.
[30] Eurostat. Handbook on price and volume measures in national account. Luxemburg: European Communities; 2001.