Final Report Summary - EUROHOPE (EuroHOPE – European Health Care Outcomes, Performance and Efficiency)

Executive Summary:
For the first time in Europe, it was possible to compare what happens to all patients with specific conditions between countries and regions within a one-year follow-up after onset of the disease. The remarkable main impacts of the project are a) to guide how to make comparable figures from health care performance, b) correct currently unrobust and unreliable comparisons made by e.g. OECD where the data is rarely comparable yet results widely in use for decision making, and d) to motivate and give incentives for harmonizing the registries in Europe. The project also gives a ground for common efficiency and outcome comparisons between European countries, regions and hospitals and wakes-up decision-makers to understand performance analyses are essential for them to improve the processes.

The results show that there are differences in the performance of the different health care systems in all of the analysed subgroups. All the EuroHOPE countries had wide regional- and hospital-level differences indicating possibilities to intensify their treatments. Generally, health outcomes were good in Italy, Norway and Finland.
and Sweden in all of the analysed patient groups. The Netherlands had an average performance in these patient groups. Health outcomes in Finland were roughly on the same level as in Norway and Sweden, with the exception of acute myocardial infarction where Finland performed worse. The ranking of Scotland varied between conditions. Length of stay of the first hospital episode was shortest in Hungary for hip fracture and ischaemic stroke patients. However, for AMI and VLBW and VLGA infants, opposite results were found, with length of stay relatively long in Hungary in these patient groups. In the care of AMI, length of stay was shortest for Scotland, while Scotland showed long length of stay for ischaemic stroke and VLBW and VLGA. Unfortunately, in the project it was impossible to capture the extent to which ambulatory and home care was used as alternatives to admission in different countries.

Notably, no clear relationship between health care financing and performance was found while a prospective activity based hospital reimbursement seemed to increase the use of immediate percutaneous coronary intervention among the AMI patients. However, the reimbursement mechanism was not related to better outcome for AMI patients. There were both well- and poor-performing countries and regions both among social insurance and tax-based health care systems. The differences in performance between regions and hospitals were not explained by the analysed demand and supply factors either.

No apparent positive relationship between quality and use of resources except for the care of AMI patients in Finland and Hungary. No correlation was found either between the quality of care of AMI, ischaemic stroke and hip fracture indicating impossibilities to use quality of one specific treatment as a measure of an overall quality at the hospital level. Reliable benchmarking requires performance measures based on several health conditions.

The Nordic hospital comparison did not indicate clear productivity difference neither association between productivity and quality. Though, productivity in Swedish hospitals was about 20% lower than Nordic average.

The results importance and impact is powered by their applicability to all countries. The project sets prerequisites for adequately developed health care information systems and a unique personal identification. Know-how on efficient use of registries is funneled through the project to all European countries making it possible to even materialize the knowledge for exporting it as a product outside Europe. Meanwhile, learning from the best practices saves all European citizen payroll-tax to be used more efficiently.

Project Context and Objectives:
The objective of FP7-Health-2007 -3.2-2: The “health outcomes and cost-benefits” call was to investigate the relationship between quality of care and costs, efficiency, and accessibility, by identifying and assessing existing approaches; to develop/identify approaches/models to describe the balance between quality of care and costs, while taking into account patient satisfaction; and to analyse how the transition of changing treatment practice is being tackled across Europe. A multilevel approach was needed that also took into account the diversity of European systems. The findings should support national decision-makers who are reforming health systems and aim to support developments of the European community Health Indicators.

Following a period of cost control in health care, there is presently increasing interest in efficiency. The
recent health care system reforms have also been deployed to improve efficiency. Efficiency becomes ever more important as the share of the GDP spent on health care increases. Most health systems have developed measures and health statistics for use of resources and costs as well as indicators for outcomes and quality. Still, there is a gap when it comes to comparing outcome of care against costs, which exacerbates performance efficiency analyses. By using available databases as well as by collecting additional data on health-related quality of life measures (enables the use of Quality Adjusted Life years, i.e. QALYs, as an outcome measure) and patient satisfaction (including expectations) EuroHOPE project’s main objective was to evaluate the performance of European health care systems in terms of outcomes, quality, use of resources and cost.

There are various definitions of quality, outcome, accessibility, efficiency that differ between scientific disciplines and approaches. At EuroHOPE we understood by the term outcome those measures of health improvements (or decrease in deterioration) attributable to health care. Quality was defined as "the degree to which health services increase the likelihood of desired health outcomes and are consistent with current knowledge". Accessible care implied that for people who need care, access to it is timely and without major barriers. At the general level, performance measures evaluate the extent to which a health system meets its key objectives, such as quality, outcomes, accessibility or equity. The concept of efficiency relates these objectives to resources: for example by considering the optimal use of the available resources to yield the maximum attainment for the objectives. EuroHOPE’s focus in performing such research was that it should help to identify areas where possible improvements in efficiency can be made. It was also important that results from EuroHOPE could be easily translated into policy.

International comparison demonstrates that at present there are countries that obtain a comparable outcome (avoidable mortality) with fewer resources as well as countries that obtain a lower avoidable mortality with comparable resources. Thus, efficiency measured at the system level varies. The average increase in health expenditure in some countries does not appear to be applied optimally for a decrease in the average avoidable mortality, because the decrease in avoidable mortality is slightly lower than average. Similar patterns have been found between regions (hospital districts) within a county: in Finland, one-year mortality and cost of care among risk-adjusted AMI patients are not clearly related to each other. However, this kind of information is not strictly enough to guide health policy. In addition, there exists need to evaluate what are the reasons behind these variations. This can be considered in two ways. Firstly, we can investigate which measurable and policy driven factors explain the differences. The other approach is best-practice-benchmarking, in which the various aspects of organizations' processes will be evaluated in relation to the best practice, relative to their own field. The two approaches complement each other and require that the indicators are describing well the aim of health policy. This is also closely related to the current Tallinn Charter (WHO 2008), where the member states in the European Region committed them selves to promote transparency and be accountable for health system performance to achieve measurable results.

The measurement and international comparison of performance and efficiency can proceed at three different levels: system wide, by disease, and by subsector. There are pros and cons associated with each but when health outcomes are the main interest the disease-based approach will be the most suitable. In subsector analysis (such as in hospital care), meaningful international comparative work can be done only if case mix (patient heterogeneity) can be standardised using the same instrument.
There are various approaches to developing information systems for international efficiency comparisons. The first approach relies on developing a coherent conceptual framework for information collection, analysis and dissemination. Another approach is more opportunistic, seeking merely to assemble readily accessible data, often by-products of existing national data collection, such as hospital discharge registers, as well as work that has been done for other purposes. It is of course helpful to maximize the effectiveness of existing data resources. This 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 are an essential element of such approaches. In this case the precise definition, collection and scrutiny of the data are left to expert groups to determine.

During recent years, many efforts have focused on using health care in national accounts as part of a top-down approach. The current attempts at implementing SHA (System of Health Accounts), however, include only expenditure, leaving the methodological framework for price and output measurement still at the development stage. Eurostat (2001) and the Atkinson Review (2005) have suggested including the quality aspect into output measurements and some illustrative calculations have been done in the UK, but there are still many theoretical and practical problems associated with developing outcome-based measures. One can assume that, at least within the short run, only minor advances can be achieved in developing a measure that can be used to evaluate the impact of the whole health system on health outcomes. Thus at the current stage, the national accounting approach is not suitable for international efficiency comparisons and not even the quantitative data on outputs are comparable between countries.

EuroHOPE study was based on the bottom-up approach. We applied the disease-based model of the health care system i.e. a microeconomic approach based on individual patient records of whole population. It is based on modelling the natural progress of a disease, with specific interest in the role of health services as a determinant of the progress. The main idea of the approach is that it analyses time trends by using more detailed data pertaining to specific health conditions to illuminate the interconnected aspects (i.e. financing, organisational structures, medical technology choices) responsible for health system performance (i.e. health outcomes and expenditure). The main innovation of our approach was that we used individual level data available from registers that allowed us to measure the outcome (by following what happens to patients) and the use of resources (such as number of hospital days, use of specific procedures and drugs) in the selected risk-adjusted and well-defined patient groups. Thus, we were not only interested in a specific treatment or hospital stay (measured e.g. in term of DRGs) per se, but outcome- and cost-related to the whole cycle of care. Our approach is also stressed in the Tallinn Charter, where it is stated that health systems should integrate targeted disease-specific programmes into existing structures and services in order to achieve better and sustainable outcomes.

One of the origins of the disease-based approach is the development of a productivity index for treating specific health problems, as has been done in the USA. The index compares, on an annual level, the value of change in health status (due to healthcare) with the costs of producing these health effects. For heart attack patients the outcome was evaluated by comparing health gains measured by quality-adjusted life years (QALYs) of similar (risk-adjusted) patients in different years. The approach has also been applied to low-weight infants, depression, cataracts, and breast cancer in the USA as well as heart attacks in Canada. In contrast, the PERFormance, Effectiveness and Cost of Treatment episodes (PERFECT)
project shows that, e.g. in Finland, the regional and hospital variations in outcomes and costs of treating seven diseases are much higher than the overall annual variation. Thus, in EuroHOPE we applied the idea of comparing countries and even regions and hospitals within each country to reinforce the accuracy and applicability of the results.

Our approach was based mainly on use of national health registers, which meant that we included only those countries and regions into EuroHOPE that had extensive registers and linkage possibilities. However, the results and methods from the study can be applied to (and by) other countries after they have adequately developed their information systems, as many countries are in the process of introducing unique identification numbers to be used in health care registers. In addition, an electronic patient record system (including all health care activities) is under development in many countries and will give new, unique possibilities for the development and applicability of our approach. EuroHOPE project also utilised the experience of previous projects: OECD aging-related disease (ARD) project, Technological Change in Healthcare (TECH) project and PERFormance, Effectiveness and Cost of Treatment episodes (PERFECT) projects and develops and expands the disease-based approach to European countries. Due to its nature, measuring costs and outcomes and especially the investigation of the relationship between the two topics EuroHOPE requires micro (patient level) data. In addition, EuroHOPE will also supplements several on-going projects: OECD Health Care Quality Indicator Project (HCQI), EU Public Health Outcome Research and Indicators Collection (Euphoric), and The EuroBASKET as well as EuroDRG. For each presented project above, there exists a partner in EuroHOPE who has been involved or is still involving with the project. This facility gave much experience to the project and enabled co-ordination and co-operation with the existing projects.

EuroHOPE applied mainly the disease-oriented approach. The disease-oriented approach was supplemented with an analysis of one of the subsectors where it was possible to relate costs with quality indicators. The chosen approach yielded the following specific aims for the project:

- Develop methods to measure outcomes and costs of care of specific diseases that can be used for routine evaluation of care given in the whole cycle of care (not only for specific procedures or short episodes)
- Develop methods to measure quality, access, outcomes and cost of (care) that can be used for routine evaluation and monitoring of the performance. We will give a recommendation of a list of indicators to be routinely collected and published by the EU (as a part European community Health Indicators)
- Develop methods for international comparative health service research using register data
- Investigate the relationship between outcomes (or quality) with costs (or other measures of use of resources) between European countries, regions and providers applying a multilevel approach
- Explore reasons behind difference in outcomes and costs. In particular, the interest will be on policy driven factors (such as treatment practices, use of medicines and modern technology, waiting times, financing, organisation of delivery, and reforms).
- Give proposals concerning the data content of national level registers and outcome measurements in order to improve the continuous monitoring of performance on an intra- and international level.
- Implement European-wide benchmarking of outcomes, quality and costs, which will enable decision-makers as well as health professionals at different levels to learn from the best practices.

Project Results:
Since benchmarking is the underlining feature of EuroHOPE, comparability in the results is the focal point.
In the analyses, case-mix adjustments have tackled heterogeneity between units by using registers together with robust coding (ICD-10, ICD-9). Also, carefully and exclusively selecting patient groups with extensive data on risk-adjustment has in itself maximized the comparability. Finally, follow-up across a wide time bracket has improved the credibility of outcome measurement and has intensified the transparency of the effects of the whole treatment chain on the outcome. The analyses were made robust by making use of the latest econometric knowledge and solid statistics know-how. The usability of the results was improved through standardisation and by modelling and computing also the confidence intervals for the standardized indicators. For the risk-adjustment variables, information from drug prescriptions and diagnoses were used to infer co-morbidities. Naturally, the age and gender of the patients were also taken into account.

The hospital-level Nordic study aimed at expanding country and hospital comparisons to include all care given to patients (i.e. all diagnoses) in the hospitals, covering both the costs and the quality of care measured by selected quality variables. The type of patient classification system varies between the EuroHOPE countries, but the four major Nordic countries (Norway, Sweden, Finland and Denmark) all have nationwide patient registers applicable for use in the same hospital-wide case-mix system.

Data were collected on hospital costs and patient data in each diagnosis-related group (DRG) for a total of 160 acute hospitals in 2008–2009. Operating costs were collected using harmonized definitions, and nominal numbers deflated to a common basis to adjust for differences in input price levels. Patient-register-based measures of quality, such as readmissions, mortality (in hospital or outside) and patient safety indices, were developed and case-mix adjusted.

Are country-level differences related to health care systems?
When one finds differences in the data between countries, the first natural question is whether the observation is due to different features of the health care systems. Of the seven countries included in the study, five can be considered tax-based systems, while two countries rely on social health insurance (SHI). Two of the tax-based systems — that of Norway and Scotland — mainly rely on central taxation; those of Finland, Italy and Sweden on the other hand rely in various ways on regional and local taxes. The two countries with social insurance systems also differ, with the Netherlands relying on a system with multiple insurers, whereas the Hungarian system is a social health insurance system with a sole insurer.

The main differences between the two groups of systems are the organization of the provider side, where the tax-based system has a long tradition of integrated public providers, whereas the SHI countries have independent providers with a length-of-arm relationship. In the Netherlands the major part of the hospital sector is private non-profit and in Hungary the regional authorities are responsible for the hospital sector during the study period.

The Finnish system is the most decentralised. Responsibility for hospital care is given to 21 hospital districts, which are federations of municipalities. In the Finnish system, hospital care is most concentrated at regional level, since most hospital districts have one central hospital that is responsible for all acute care. In this respect the Finnish system differs from the decentralised Swedish system, where there exist many providers within each county (21) responsible for arranging hospital services. In Italy 19 regions and two autonomous provinces have responsibility over the organization and delivery of health services. However, some regions in Italy are greater in population size than the individual Nordic countries. In
Norway and Scotland the central government holds the regulatory power to provide investments and thus also to maintain the provider structure of the hospital sector. In Scotland, hospitals are part of integrated healthcare systems and Boards have some flexibility about how to utilise central funding. All seven countries applied prospective payment systems, some with elements of cost compensation. The following countries used activity-based funding systems: Hungary, Italy, the Netherlands, and Norway. The Netherlands used DRG-based funding, while the Norwegian and the Italian models combined activity-based funding based on the DRG system of global budgets. In Norway, the global budgets were risk-adjusted. Finland and Scotland used fixed payment systems (global budgets). In Sweden, the reimbursement system differed between the counties using global budgets and activity-based funding based on DRGs.

In addition to health system characteristics, the overall economic situation of the countries may also affect performance. In 2008, GDP per capita was clearly highest in Norway, followed by the Netherlands and Sweden. Finland, Scotland and Italy (the regions included in the study) represented average countries in this study in terms of GDP, whereas Hungary has the lowest GDP per head.

Mortality variation

The most important outcome measures are mortality at the 30-day, 90-day and one-year follow-up after the onset of disease. Mortality rates for AMI, ischaemic stroke, hip fracture and VLBW and VLGA infants varied to a similar extent, i.e. with a 10 to 15% difference between the best-performing and worst-performing country. Hungary had the highest mortality for AMI, hip fracture and VLBW and VLGA infants. For ischaemic stroke, in Scotland and Hungary the figures were about the same. Italy obtained better results, having the lowest mortality rates in all conditions, except for AMI (where mortality was lowest in Norway). Mortality was quite low in Sweden in all the conditions. The Dutch system was performing at about average outcome levels in AMI and ischaemic stroke. Finnish AMI care seemed to give poorer outcomes compared to the other Nordic countries with the exception of ischaemic stroke, for which mortality was about the same as in Sweden. Breast cancer mortality was lowest in the Nordic countries and Italy. VLBW and VLGA infants showed a somewhat different pattern, with high mortality for Hungary, followed by Finland, Italy, the Netherlands and low mortality for Scotland and Sweden. It should be mentioned that the Italian data were not representative and covered relatively wealthy Italian regions (city of Turin and the Lazio region).

Length of the first hospital episode and use of procedures

The episodes for AMI, ischaemic stroke, and hip fracture start with an acute phase in the hospital, usually occurring immediately after the event. The first hospitalisation terminates on the day of the first discharge either to home, death, or is censored after a specified time of continuous inpatient care, depending on the disease. In order to achieve better comparability, we defined a first “acute” care episode, which excluded rehabilitative and nursing services given during the continuous treatment given in hospitals. Age- and sex-standardised length of first acute hospital episode and their 95% confidence intervals of AMI, ischaemic stroke and hip fracture patients by country in 2008 (Norway 2009)

There was no clear pattern of length of stay between countries and conditions. In Finland and Hungary length of stay was quite short in two of the three conditions. In Scotland and Sweden, ischaemic stroke and hip fracture patients had a considerably long acute first hospital episode. In Italy the length of stay of hip fracture patients was extremely high compared to other countries.
Variation in the use of procedures was analysed for AMI, where percutaneous coronary intervention (PCI) within 2 days after infarction was committed most often in Sweden followed by the Netherlands and Hungary (the country with the highest mortality). The country ranking changed when both PCI and CABG are considered after 30-day follow-up. Now the highest figure was in Italy together with Norway and Sweden. In all countries except Scotland, over half of the patients received a cardiovascular procedure within 30 days.

In summary, the country differences in outcomes and treatment patterns cannot be easily explained by health system characteristics. In addition, country differences in the use of PCI within 2 days were not associated with differences in outcomes. On the other hand, the variation in outcome may have reflected differences in general health status between the countries. In the case of Hungary, these may be associated with socioeconomic conditions i.e. its relatively low GDP and high income inequity.

How much do regional level characteristics explain variation in health care performance?
In addition to exploring the variations in mortality at the country level, we examined the variations also on a regional level in each country. As shown below, there was great variation in all of the analysed conditions within every country. The existence of regional variations has been discussed for decades, but the definite reasons behind them remain unknown. Using our data, we studied whether selected regional-level characteristics were associated with the observed differences. The regional characteristics include factors such GDP per capita, unemployment, education, population density and age structure, concentration of hospital care (Herfindal-Hirschman index), as well as condition-specific measures of supply of services. The regional analysis was based on patients’ place of residence. Each country has defined the partition of its regions to be suitable for benchmarking. In Finland, Italy, Norway, Scotland and Sweden the regions describe local authorities who are responsible for health care, while in social health insurance countries the regions are based on regional governmental or sub-national authorities that are not responsible (the Netherlands) or are responsible only in part (Hungary) for health care. In the two last-mentioned countries, the average population size of the regions is much greater than in the Nordic countries and Scotland. From Italy only 6 regions were defined. In the analysis of very low birth weight (VLBW) and very low gestational (VLGA) infants’ data from smaller areas were pooled into larger geographic entities in Finland, Sweden, Norway and Italy.

Regional variations in outcome
In all cases for one-year mortality after acute myocardial infarction (AMI), ischaemic stroke and hip fracture patients, respectively, the regional differences were larger compared to between-country variation, although region by region comparisons (within countries) had overlapping confidence intervals in most areas. The degree of variation between regions was rather similar across countries. Regarding AMI patients, most of the Italian and Swedish and all the Norwegian regions performed better than average regions for all countries in one-year mortality, whereas some Finnish, most of the Scottish and all the Hungarian regions performed poorer than average. Among ischaemic stroke patients, four of the Italian regions, about half of the Swedish counties and some of the Finnish regions performed better than average, taking into account the confidence intervals. In hip fracture, well-performing regions were found—in addition to Italy and Sweden—from Norway, and in VLBW and VLGA also from Scotland. The analysis of regional-level factors related to the regional variation of mortality after AMI was focused on the following characteristics: hospital location (urban/rural), average age of patients admitted (self-reported), gender mix of patients admitted (male/female), and hospital size (number of beds).
the use of percutaneous coronary intervention (PCI) and its effect on outcome. We found that the reimbursement system had an impact on procedure intensity: the two-day PCI rate was about 17 per cent higher in countries and areas with an activity-based reimbursement system. GDP per capita was negatively associated with 30-day mortality. Also here the use of PCI had a negative but not statistically significant effect at the regional level.

However, at the individual level, higher PCI use was associated with lower mortality, while a smaller effect of PCI on mortality was found in a country with the lowest mortality. Hungary showed the highest mortality in combination with high PCI rate. The results of more detailed analyses of data from Finland and Norway suggested that the effects of socioeconomic factors on mortality through the use of PCI were small.

Regarding ischaemic stroke and hip fracture patients, the regional differences in length of stay and mortality were not related to any of the analysed regional level factors. Only GDP per capita was positively associated with lower mortality among ischaemic stroke patients.

Regarding VLBW and VLGA infants, socio-economic variables at regional level appeared to have an impact on mortality in Hungary but not in the whole sample. Also the concentration of care and the capabilities in neonatal care (NICU level), the level of the delivery hospital did not appear to have an impact on mortality and length of stay when data for four countries were combined. However, in Hungary and Finland these organizational variables had significant coefficients showing that being born or treated in a tertiary-level hospital was associated with lower mortality. On the other hand, length of stay also tended to be higher among infants born in these hospitals in Scotland, Italy and Hungary.

Summarizing, the results of a number of regression analyses showed that various demand and supply side variables could not explain much of the regional variation in mortality, length of stay (LOS) or utilisation of procedures. The combination of large differences in health outcomes and use of resources (LOS), and a lack of demand-side variables to explain the variation indicate room for improvement in health care performance. In addition, we may not have captured all of the important differences in the comprehensiveness of care provided in different regions. This also could be the result of variation in the adoption of effective technologies, in the quality of doctors and other health care providers, or in physician beliefs about treatment effectiveness. Another explanation is that differences in institutional factors do not explain performance as much as theory would suggest, which would be in accordance with the results of the OECD study (Joumard et al. 2010).

Comparing use of resources between countries
For policy-makers who aim at improving health care, it is crucial to derive an understanding of the reasons behind variations in health care costs, both within and across countries. Variations in health care cost due to differences in access and treatment intensity would require a different response than that for differences in productivity in the production of single services. An improved understanding of the background for variation in health care costs requires micro-data at the level of the individual patient.

To conduct across-country comparisons of treatment cost, four major challenges must be handled: firstly defining treatment episodes in a comparable way across countries, secondly the development of methods for calculating resource use; thirdly modelling the distribution of the estimated risk-adjusted cost function; and, finally, finding a method for the ranking of outcome and cost in order to determine differences between countries (regions).

Indicators of resource use
Cost figures are only rarely provided at the individual patient level (bottom-up approach). Hence, one often

...
has to rely on figures derived from a top-down approach, perhaps supplemented with information from hospitals that make use of bottom-up cost-per-patient (CPP) figures. Alternative methods for cost-calculations may result in variations in the cost figures and may potentially have a considerable impact on cost estimations and comparisons.

In EuroHOPE a register-based approach to identify items of resource use has been utilised. We use two specific approaches in EuroHOPE that are intended to supplement each other.

Approach I: All countries have in their discharge registers and pharmaceutical prescription databases registrations that indicate the main components of resource use (services). The registered components are mainly related to procedures and hospital length of stay. The relative cost of the different components of resource use is approximated by data from the cost-per-patient (CPP) database by the Swedish Association of Local Authorities and Regions (SALAR). Cost in Swedish Kronor (SEK) is then converted to Euros by means of the input–based Purchasing Power Parity index (developed by Eurostat) for hospital services.

Approach II prescribes that each country contributes with their best cost estimate based on their own system of cost calculations. In the majority of countries, cost estimates generated by variants of the DRG system are used and costs of medicines based on data from the prescription register are added. The different approaches have different characteristics with regard to the type of across-country variation that is considered. In approach I, only variation in the procedures and length of stay create variation in resource use across countries. Approach II also takes also the variation in cost of producing a particular service into account. The problem with Approach II is that the system of cost assignment is likely to vary across countries.

Empirical specification of the cost function

Given adequate measures of resource use, there still exist challenges to estimating health care costs while taking patient heterogeneity into account. In EuroHOPE we are mainly interested in mean costs accrued in hospitals and their differences between countries. To meet this purpose, we selected a model based on various goodness-of-fit measures. Based on the preferred model, we studied differences in costs between regions and countries.

Ranking of countries according to costs

Data describing the use of resources are more complete for AMI than for the other diseases. So far, ranking of mean treatment cost across countries is done only for AMI.

Considering treatment cost according to Approach I, during first hospital episode, we found that Hungary had the highest mean cost followed by Finland, Sweden and Norway. We also found that the ranking of countries depends on the cost indicator used. According to Approach II, during the first hospital episode Sweden had the highest mean treatment cost followed by Finland, Hungary and Norway. The ranking also depended on the length of the observation period (first hospital episode or 365 days after the index day). According to Approach I and one-year cost, Sweden had the highest mean cost followed by Norway, Finland and Hungary. A relatively higher one-year cost than the first hospital episode cost could stem from more hospital care during the follow-up of AMI patients in specialist care compared with other countries and a higher frequency of other treatments.

Summarising, the analyses provided several conclusions with important relevance for health policy. First, the hospital discharge registers did not contain sufficient information on treatment procedure to calculate cost estimators for all diseases. AMI and hip fracture had the best procedure information. Second, risk
adjusters were able to explain only a small proportion (5–10%) of the variation in the calculated cost across patients. Third, the ranking of countries depended on the cost indicator used. Fourth, the ranking of countries depended on the length of the time-period taken into account. And finally, the ranking of countries did not depend on the risk-adjusters included or the specification of the cost function. This means that the ranking of countries according to crude cost gives the same result as ranking of countries according to the estimated expected cost adjusted for variation in comorbid conditions.

A policy implication is that ranking of countries could be done by comparing mean cost as calculated in each individual country. Even though the Netherlands and Scotland are not included in the present study because of data sharing restrictions, they would still be able to calculate their crude mean cost according to regions.

Quality, use of resources and their interrelationship in hospitals

The analysis of hospital-level differences was focused on AMI, ischaemic stroke and hip fracture patients from five European countries (Finland, Hungary, Italy, Norway and Sweden). The comparison of quality and cost was based on hospital-level random effects models using individual patient-level data, which allowed us to take into account patient- and hospital-level heterogeneity. We also explored whether hospitals’ quality and cost variation could be explained by hospital- and health-system-level characteristics. After examining outcomes and costs for hospitals separately, we analysed the existence of a cost–quality trade-off by comparing hospital-level survival rates and costs.

Survival

For the empirical Bayes estimates of hospital random effects for quality, as obtained from the basic model, where age, comorbidities and transfers to a higher-level hospital were taken into account. Quality of care was measured by 30-day survival (i.e. a complement of 30-day mortality). Hospitals to the left of the graph have lower than average survival than hospitals to the right. Bayesian 95% confidence intervals were formed from the posterior distribution of each provider effect. The provider effects do not as such have exact practical interpretation. However, by calculating indirectly the standardised ratio and multiplying the ratio by the mean survival in the sample, we estimated that survival difference between the worst and best hospital was 30 percentage points (min 67.5 max 97.5) in the care of AMI. The corresponding figures for ischaemic stroke were 27.5 percentage points (min 69.2 max 96.7) and for hip fracture 16.4 percentage points (min 80.8 max 97.2).

In the treatment of AMI, the Hungarian and Finnish hospitals were performing poorly compared to hospitals in other countries. In Hungary, the hospital-level variation was higher than in the other countries. Most of the Hungarian hospitals were performing below the average level (random coefficient below 0) of all hospitals but the best performing hospitals in the country are at the same level as the best performing Finnish hospitals, which in turn are at about the same level as in poorly performing Swedish, Norwegian and Italian hospitals. The performance of most Finnish hospitals did not differ statistically significantly from the average level, while among some of the Italian, Norwegian and about half of the Swedish hospitals, performance was better than average when confidence intervals are taken into account.

The hospital- and regional-level variables explained only a small part of the country differences. Survival was positively related to the existence of a catheterisation laboratory in all countries except Italy. In Hungary and Norway, a lower concentration of AMI care was associated with better survival. In addition, GDP per capita was positively associated with survival in Hungary and Finland.
Country differences in survival were clear also in care after ischaemic stroke while the within-country hospital differences were considerable. The variation between the hospitals was again highest in Hungary but now lowest in Finland. The best performing Hungarian hospitals were at the same level as the best hospitals in Finland and Sweden. In Hungary survival was higher in university hospitals and in Italy in hospitals with a stroke unit.

Compared to the two conditions, hospital-level variations in hip fracture were lower and confidence intervals wider. The performance of all Hungarian hospitals was poorer than average whereas most Swedish hospitals were performing better than average. The hospital level differences were not related to hospital or regional variables.

Use of resources
Our cost measure describes the use of resources (Approach I above) during the first acute hospital episode. In all three conditions it was based on the number of inpatient days and for AMI patients also on the use of cardiovascular procedures (PCI, CABG) and for hip fracture patients on the type of surgery. In the care of AMI, cost variation within countries was much higher than between countries. The costs were highest in Italy and Hungary. Costs were higher in all countries for a hospital with a catheterisation laboratory. The concentration of AMI care within regions decreased the cost in all countries except in Italy, where its effect was the opposite. Norway was the only country in which population density reduced the costs.

Moreover, in the care of ischaemic stroke, the cost variation within countries was higher than between countries, though now Finnish hospitals were operating clearly at a lower level of resource utilisation. The university/teaching status of a hospital increased costs in Sweden. In Hungary and Finland, an increased concentration of stroke care had a strong negative effect on cost.

Country differences in costs were more systematic in the care of hip fracture than in the two other conditions. The costs were lowest in Finland and Norway. University/teaching status increased the cost in Sweden and higher volume decreased the costs in Italy. Concentration of care for hip fracture patients decreased costs in Finland.

Cost–quality trade off
An important policy question is whether the costs of a hospital are related to quality. If there is a choice between minimising cost and maximising quality, there is a cost–quality trade-off, i.e. better quality may be provided by increasing costs. On the other hand, the absence of the relationship would indicate a potential for improving performance by containing cost with no reduction in quality or improving quality without increasing costs.

In the care of AMI we found positive correlations between cost and quality in the analysis using both pooled and separate country data. The effect was strongest and most systematic in Hungary and Finland. In the care of ischaemic stroke and hip fracture we did not find clear evidence of a cost–quality trade-off. In summary, our results show significant differences between hospitals and countries in both survival and cost. Again the findings cannot be easily explained by the characteristics of the health care system. However, we found some evidence supporting an increasing horizontal integration in care for the three conditions. An increase in the concentration of the regional hospital system was associated with a decrease in costs. The effect was found in all countries except Italy. But the effect varied between countries and conditions. In Finland (a country with the highest average concentration) the effect was found for all three conditions, in Hungary in the care of AMI and ischaemic stroke, and in Sweden and
Norway it was found only in AMI care. However, in Norway an increase in the concentration was associated with a decrease in survival of AMI, indicating that cost savings achieved by increasing concentration could be related to a possible decrease in outcomes. Our results concerning the cost–quality trade-off corroborate those of recent studies that have suggested that the cost–quality association is inconsistent and is present for certain treatments or for some patient groups, though not in all countries. This implies potential exist for improving hospital performance by containing cost or improving quality without increasing costs.

Productivity and quality in the Nordic hospitals
In EuroHOPE project the disease-based analysis of performance was supplemented with a hospital-level analysis focused on four Nordic countries. Previous Nordic comparisons have indicated that Finnish hospitals have had significantly higher average productivity than hospitals in Sweden, Denmark and Norway, while also revealing substantial variation within each country. Controlling for within-country variations in activity-based reimbursement, length of stay (LOS), outpatient shares, university hospital status or capital region only contributes to a small portion of these differences. The aim of this analysis was to examine whether quality differences can form part of the explanation for productivity differences and attempts to uncover any cost-quality trade-off at the hospital level.

Quality of hospital care
We developed patient-register-based measures of quality such, as case-mix-adjusted readmissions, mortality (in hospital or outside) and patient safety indices. For the emergency readmissions the confidence intervals were very narrow, which means that there were significant differences between most pairs of hospitals. There was mostly a clear ranking of hospitals within countries, since each hospital performance measure was mainly outside the range of other hospitals’ confidence intervals. Denmark had the lowest rates, but there was some overlap with the Finnish and Norwegian hospitals. It was not possible to calculate this indicator for the Swedish hospitals.

For 30-day mortality the confidence intervals were wider, but most pairs of hospitals were still significantly different from the mean and from each other. Most Norwegian hospitals had significantly lower 30-day mortality than hospitals in the other countries.

Productivity
DEA productivity estimates of the hospitals sorted by country confirm previous results showing that Finnish hospitals were on average more productive than in the other Nordic countries, though Denmark was almost as productive. Even Norway had not much of a cost disadvantage in this analysis, a clear catching up from previous studies. Sweden, however, still lags behind, which was verified also using statistical criteria.

Quality–productivity trade-off
When productivity estimates were plotted against the two of the performance measures, one finds no strong correlations. There seemed to be a positive correlation (r=0.674) between productivity and emergency readmissions, implying a trade-off between high quality and high productivity. There was a slight tendency for low readmission rates to go together with high productivity in Finland, but the main impression is of a large dispersion. For 30-day mortality there was a clear negative correlation between
productivity and performance measures.
In sum, the results show that there were significant differences between countries on most measured quality indicators. There were also significant differences between hospitals within countries, but only the readmission and mortality measures showed enough differences to rank the majority of hospitals. While previous findings on the relative productivity of the hospitals in the Nordic countries were confirmed, there was no clear pattern that any country had higher or lower quality on all measures. This may be because the treatment patterns and practices vary a lot between countries, even for countries that are as similar as Denmark, Finland, Norway and Sweden.

The evidence for a trade-off or a positive association between quality and productivity varies between the different performance measures. There seemed to be a trade-off between productivity and better (lower) inpatient readmission rates, but high productivity was associated with lower mortality rates. This effect was most important in Finland. For mortality at least, there seemed to be a possibility of improving both quality and productivity.

Measuring health related quality of life and patient satisfaction
Measuring health related quality of life and patient satisfaction is fascinating but hard to conduct reliably. EuroHOPE attempt is strengthens the fact. In EuroHOPE, a feasibility to measure HRQoL and patient satisfaction of stroke and breast cancer patients in seven countries (six for the stroke patients) was studied. The main findings are discouraging to conduct international data collection in HRQoL and patient satisfaction. Firstly, to avoid country-level patient selection problems a common survey protocol was needed but it exposed to be almost impossible to construct such that would still collect all the required information. The reasons were multi-folded from the international differences in ethical legislation to varying treatment practises. Secondly, participating hospitals are hard to find. This is mainly due to many on-going surveys whose aims are often even overlapping causing too much pressure on personnel and not least to patients in severe medical states. Thirdly, surveys are always suffering dramatic selection problems at the patient level causing unreliable data. However, all those countries that were able to conduct the data collection are currently working on data-analyses and towards scientific publications.

General results and foregrounds
In general, EuroHOPE developed common protocols and data extraction protocols to enable comparable data collection and indicator production that is at the spot for benchmarking. EuroHOPE provides the audience also on variety of levels of information through its Internet-site. All the published papers as well as EuroHOPE Discussion Paper Series are available.

In addition to that the benchmarking and the results transparency has been enriched by making use of an Atlas-map application in which EuroHOPE patient groups treatment outcomes and other selected indicators can be compared on international and regional level. Since by the EuroHOPE protocols the information has been collected and processed into indicators by using the same manners in all countries, the information is comparable and reliable to compare contrary to many OECD indicators.

Potential Impact:
EuroHOPE advanced the state of the art by allowing fair performance comparisons between European countries. So far there is no common ground for efficiency and outcome comparisons between countries, regions and hospitals between different European countries. However, performance analysis is essential to provide decision-makers with the necessary means to improve processes that could finally lead to a more efficient allocation of health care resources in the EU Member States. The project contributed also to the ongoing discussion about a possible trade-off between costs and outcomes. Observations of worse outcomes of care in one country with lower costs or shorter length of stay should have an important implication for European health care. Although the study was based on only a few countries with good register data covering whole populations, its results are ready to be applied to (and by) other countries once they have adequately developed their health information systems and adopted a unique personal identification that will enable register linkages. In addition, the approach adopted here can later apply to other diseases. In particular, the development of primary health care registers enables the application of the approach to chronic disease as well as preventive actions.

So far most of international performance-comparisons of health systems are based on the system level, using aggregate-level data. Thus, the outcome measurement is very difficult and at least within the short run, only minor advances can be achieved in developing the measures. Due to this, EuroHOPE provided decision-makers of the EU member states with the necessary means to perform fair benchmarking. This should enable decision-makers at different levels to learn from each other and to improve the efficiency of health care processes. The applied approach of EuroHOPE thus produced benchmarking data, in addition to at the national level, also at the regional level and even further, at the producer level (hospitals). Benchmarking and analysis at the disease level was very fertile and enabled us to analyse how the transition of changing treatment practice could possibly tackled across European countries, and their impact on costs and outcomes. More generally, our project yielded two new dimensions in benchmarking. Firstly, indicators developed in the project help the decision-makers in a new way, since they can compare their own performance not only by using cost or process indicators, but also indicators of outcomes and information on the relationship between costs, process and outcomes. Secondly, local decision-makers are now able compare their performance not only within their own country but also with regions and hospitals in other countries. This kind of information has already affected action at the local level for low birth weight infants in Finland. For managerial purposes, the disease based approach will have the great potential and when it is linked with the bottom-up approach, it motivates data collection.

Development of the scientific evidence-base supports the Member States to better organise their health systems. The scientific evidence from EuroHOPE should make a major impact on the way Member States organise their health systems in order to achieve better health with the current or even less resources. Member States are increasingly viewing the Open Method of Co-ordination (OMC) as an appropriate way in which they can learn from each other in order to improve their health systems and services. The OMC in relation to health services will depend on methodologically sound systems for comparing outcomes and costs. The link of the project to these essential principles was underlined by one specific work package, which addresses the relationship between risk-adjusted outcomes and costs. The experience from Finland indicates that among the AMI and Stroke patients, about 20–30% of costs can be contained if all regions (hospital districts) had the same cost as that of the cheapest region in terms of risk-adjusted measures.

The enhancement of co-operation between researchers in Europe and other geographic regions promoted
integration and excellence of European research in the field through active involvement of these countries in the project’s developments and invitations to its workshops. EuroHOPE provided unique opportunities for high quality and policy-relevant research, where clinical experts directly collaborated with health economists and other experts in health research. This has also increased clinical experts’ understanding of the importance of performance evaluation, in addition to their more individual-patient-oriented research. The wide and deep co-operation implies that we are developing indicators that are acceptable and understandable also at all levels of the decision-making process in health care. This co-operation will also promote the understanding that efficiency and good performance should be the starting point of all decision-making in health care.

The results impact is powered by their applicability to all countries. The project sets prerequisites for adequately developed health care information systems and a unique personal identification. Know-how on efficient use of registries is funneled through the project to all European countries making it possible to even materialize the knowledge for exporting it as a product outside Europe. Meanwhile, learning from the best practices saves all European citizen payroll-tax to be used more efficiently.

Previous studies comparing regions or countries in the fields of medicine covered by EuroHOPE were often restricted to selected hospitals or diseases, or to ‘metadata’, or to only one of the aspects of outcomes or health care pathways. A noteworthy exception is the recent study by Chung et al. (2013) on AMI that used nationwide registries with detailed patient-level information on all hospital admissions. Unfortunately, such registries currently exist only in the UK and Sweden and only for AMI. In addition, linkage with other registers, such as those on medication use and preferably costs, is needed in order to comprehensively assess the cost-effectiveness of health care systems.

The EuroHOPE case studies are unique in having collected nationwide data at the level of the patient, for several diagnoses, and with well-defined criteria for selecting patients at first hospital admission and following them up until one year after the index admission. Linkage of records made it possible to clearly delineate episodes of care and assess vital status. This creates comprehensive information on regional and provider variations and health care performance.

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. Other countries must first develop their information systems, while laws that might hinder available data linkages may need to be addressed. For example, an electronic patient 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 data using standardised and internationally comparable definitions of activities and classifications describing the treatments (i.e. diagnosis, procedures) to be nationally available for research, thus enabling an evaluation of performance across countries, regions and producers.

Future studies that build upon this approach could focus on acquiring additional information that was lacking in the current study. In particular, disease-specific patient characteristics (especially on severity of the condition) and quality of care indicators would prove beneficial in giving better insight into the causes of variation in performance.
of regional variation and into the performance of regions. Additionally, it seems important to improve the registration of diagnostic and treatment procedures that determine treatment outcome and cost. Furthermore, it would be useful to have a better understanding of differences in coding practices across countries.

In addition, it seems that outcomes and relationships between outcomes and explanatory factors may vary across levels of analysis (national, regional, hospital, and individual). For a better understanding of regional variations, it is worth analysing such ‘inconsistencies’ across levels in more detail. Also, research could be extended to other diseases or regions to validate the findings.

Although administrative data may provide a large and possibly relatively cheap information source, substantial effort was required in the EuroHOPE project to create comparable datasets that cover the health care pathway of individual patients as well as health outcomes. Moreover, privacy issues prevented the sharing and pooling of national datasets into a single EuroHOPE database, limiting the possibilities of e.g. risk-adjustment or multilevel modelling. In addition, the performance at hospital level could not be studied in all countries, since it was not permissible to share outcomes at hospital level. Such experiences should be taken into account in future studies, especially given that possibilities for linking and sharing data appear to vary widely between countries (OECD, 2013). Nevertheless, as this type of research may provide the necessary step forward in the monitoring and evaluation of health care systems and policies, these data infrastructure issues require close attention.

EuroHOPE supplemented also other international projects (EURPHORIC, EuroBASKET, EuroDRG, OECD Quality indicators). For each of these, there exists a partner in EuroHOPE which has been involved or is still involved in the project.

EuroHOPE has used all possible levels to disseminate the information. By organizing specific sessions in yearly Health Economics conferences (2011-2014) the academic audience has reached information and EuroHOPE on its part received fresh ideas from academics. In the spring 2014, The Lancet made an observation on importance of the EuroHOPE which can be taken as an outstanding compliment and exposure that dissemination has reached its target. Also, national seminars in partner countries to decision makers and other stakeholders have taken places. In addition, one policy brief for EU-decision makers in Brussels, Belgium and final seminar at Karolinska Insitutet, Stockholm that was also stream-casted for remote audience were well noted in newspapers and other media. EuroHOPE has established its own discussion papers series that is available at the project’s Internet-site and there are 5 clinical articles submitted, 2 published articles at Health policy, methodological paper in SAGE journals and 10 articles Special Issue at Health Economics on its ways to make the impact ever greater in various fields of profession.

Finally, impact of EuroHOPE has already realized. Namely, there are various project that are willing to use the methodology developed in the EuroHOPE. Also, ERIC-project is providing base for possible continuation for the work that started in EuroHOPE and is possibly extending it to deeper analyses as well as wider patient selection by choosing new patient groups into it.

List of Websites: