

Summary of the findings of the EuroHOPE project

The EuroHOPE
study group

euroHOPE

The EuroHOPE study group

Summary of the findings of the EuroHOPE project



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 241721.

1.4.2014

© EuroHOPE

Copy-Set Oy
Helsinki, Finland 2014

THE EUROHOPE STUDY GROUP

Eline Aas	Manon Lette
Kjartan S. Anthun	Antti Malmivaara
Helen Banks	Alison McCallum
Éva Belicza	Emma Medin
Grete Botten	Hans Olav Melberg
Harry Campbell	Peter Mihalizca
Anne Douglas	Tron Anders Moger
Peter Engelfriet	Dino Numerato
Giovanni Fattore	Mikko Peltola
Eilidh Fletcher	Clas Rehnberg
Fanny Goude	Gunnar Rosenqvist
Terje Hagen	Hanna Rättö
Richard Heijink	Timo T. Seppälä
Ingrid M.S. Huitfeldt	Colin Simpson
Unto Häkkinen*	Teresa Spadea
Tor Iversen	Sofia Sveréus
Merja Juntunen	Erika Takács
Sverre A.C. Kittelsen	Fabrizio Tediosi
Søren T. Klitkou	Amber van der Heijden
Roosa Kohvakka	Sarah Wild
Marie Kruse	Linda Williams

Partners

National Institute for Health and Welfare, Finland (Coordinator)
Semmelweis University, Hungary
Universita Commerciale Luigi Bocconi, Italy
National Institute of Public Health and the Environment, The Netherlands
Ragnar Frisch Centre for Economic Research, Norway
University of Oslo, Norway
University of Edinburgh, Scotland, UK
Karolinska Institutet, Sweden

*Correspondence: unto.hakkinen@thl.fi

www.eurohope.info

CONTENTS

Summary	7
EuroHOPE: Evaluating Outcomes, Performance, and Efficiency of European Health Care Systems.	9
Are country-level differences related to health care systems?11
Mortality variation	11
Length of the first hospital episode and use of procedures	15
How much do regional level characteristics explain variation in health care performance?17
Regional variations in outcome	17
Comparing use of resources between countries20
Indicators of resource use	20
Empirical specification of the cost function.	20
Ranking of countries according to costs	21
Quality, use of resources and their interrelationship in hospitals22
Survival.	22
Use of resources	24
Cost–quality trade off	25
Productivity and quality in the Nordic hospitals.26
Quality of hospital care.	26
Productivity	26
Quality–productivity trade-off	28
The way forward30
References31

SUMMARY

- Ever tightening budgets require more efficiency from the public health care sector, making learning from best practices invaluable. International comparison of performance can proceed at a number of possible levels, e.g. system-wide, by disease, and by subsector (such as hospital or nursing homes). 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 disease level. We evaluated the performance of seven European health care systems using the disease-based approach. In addition, we compared the productivity of Nordic hospitals against quality of care using the subsector approach.
- For the first time in Europe, we were able 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. We analysed patients suffering from acute myocardial infarction, ischaemic stroke, hip fracture, and breast cancer, as well as very low birth weight (VLBW) and very low gestational age (VLGA) newborn infants in Finland, Hungary, Italy, the Netherlands, Norway, Scotland, and Sweden.
- We found that there were differences in the performance of the different health care systems in all of the analysed subgroups. In addition, in all the countries there were wide regional- and hospital-level differences.
- Generally, health outcomes were good in Italy, Norway 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.
- The study did not find a clear relationship between health care financing and performance. There were both well- and poor-performing countries and regions both among social insurance and tax-based health care systems.
- A prospective activity based hospital reimbursement seemed to increase the use of immediate percutaneous coronary intervention among the acute myocardial infarction patients, but the reimbursement mechanism was not related to better outcome for patients.
- The differences in performance between regions and hospitals were not explained by the analysed demand and supply factors such as GDP per capita, unemployment, education, population density and age structure, concentration of hospital care as well as condition-specific measures of supply of services.
- Length of stay of the first hospital episode was shortest in Hungary for hip fracture and ischaemic stroke patients. However, for acute myocardial infarction 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 acute myocardial infarction, length of stay was shortest for Scotland, while Scotland showed long length of stay for ischaemic stroke and VLBW and VLGA. Unfortunately, we were not able to capture the extent to which ambulatory and home care was used as alternatives to admission in different countries.
- The results indicate that each country had the potential to improve efficiency, because of considerable variation both at the regional- and hospital-level. There was no apparent positive relationship between quality and use of resources except for the care of acute myocardial infarction patients in Finland and Hungary.
- At the hospital-level we did not find any correlation between the quality of care of acute myocardial infarction, ischaemic stroke and hip fracture. Therefore, the information on quality of treatment for one specific health problem (disease) cannot be used as the only source to compare the overall quality of care at the hospital level. Reliable benchmarking requires performance measures based on several health conditions.
- The Nordic hospital comparison did not indicate clear productivity difference between the Finnish, Danish and Norwegian hospitals. However, the Swedish hospitals' productivity was about 20% lower than in the other Nordic countries on average. There was no clear association between the productivity and quality of care among the Nordic hospitals, and thus productivity differences between the countries were not associated with differences in quality.

EuroHOPE: Evaluating Outcomes, Performance, and Efficiency of European Health Care Systems

There are three main reasons for measuring the outcomes, performance, and efficiency of European health care systems. First, there is the urge to improve efficiency when budgets are getting more binding. International comparisons of health system performance may provide lessons and best practices. Second, the efficiency of health care should be measured, since it has been placed high on international and European agendas. Third, while the need for international performance comparisons is self-evident, there is a lack of good quality research in the field. This is not due to the researchers, but due to incomparability of data. Patient- and sector-level data from various sources are not comparable as such, which reduces their feasibility in benchmarking and hence makes learning from best practices very difficult and may lead sometimes even to wrong interpretation.

From the set of available levels of analysis for efficiency comparisons, EuroHOPE (European Health Care Outcomes, Performance, and Efficiency) has applied both the disease and the sub-sector levels (Häkkinen and Jourmard 2007) in a four-year project funded by the European Commission. Disease-level analysis concentrates on different diseases, how those evolve and what impact an intervention has at different stages of the evolution on outcomes, while the sub-sector level analysis places heavy emphasis on how a health care subsector is organized (structure), which can influence how participants in the subsector behave (conduct), and ultimately how the subsector performs in the aggregate (and also typically stage by stage, or industry by industry).

EuroHOPE applied these levels of analyses to study five disease groups: acute myocardial infarction (AMI), ischaemic stroke, hip fracture, breast cancer and very low birth weight (VLBW) and very low gestational age (VLGA) infants. The data was obtained from a variety of national registers in Finland, Hungary, Italy, the Netherlands, Norway,

Scotland, and Sweden. Possibilities to link data from various sources at individual level have been utilised to enable rich risk-adjustment for controlling heterogeneity, to enable follow-up to seek for outcomes indicating effectiveness, and follow-down to ensure the patient cohorts consist of solely first-ever cases. The sub-sector level analysis has been used to compare the productivity of Nordic hospitals against their quality of care.

The main aim was to develop methods to measure the outcomes and costs of care of specific diseases in order to evaluate the care given in the whole treatment pathway. In addition, methods were developed also for aggregate hospital level analysis of quality and cost using patient level information on all patients. These methods can be used for routine performance evaluation and monitoring. A detailed list of the aims consisted of a) developing methods for international comparative health service research using register data, b) reflecting on the relationship between outcomes and use of resources (e.g. costs) and comparing them between European countries, regions and providers, c) exploring and revealing the reasons behind differences in outcomes and costs, d) comparing the quality and cost of acute hospital care in the Nordic countries, e) giving proposals concerning the data content of national level registers and outcome measurements in order to improve the continuous monitoring of performance on both national and international levels, f) establishing requirements and standards for European-wide benchmarking on outcomes, quality and costs, and g) facilitating decision-makers as well as health professionals at different levels to learn from best practices. In the wider perspective, the project also yields recommendations for lists of indicators to be routinely collected and published by the EU (as a part of European Community Health Indicators).

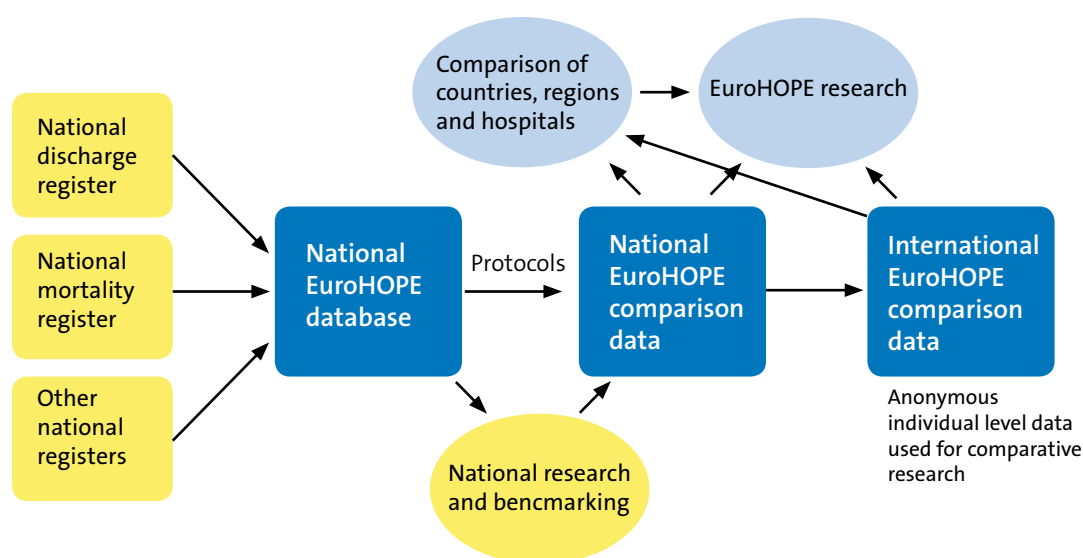


FIGURE 1. EuroHOPE data

To reach the targets, EuroHOPE has made use of a population-based cost-effectiveness approach (Häkkinen et al. 2013). A microeconomic disease-based strategy has been used that modelled the natural progress of a disease, with specific interest in the role of health services as a determinant in the progress. Collecting data uniformly from various sources requires data collection protocols. Clinical groups consisting of members from each participating country and disease-practice specialty funnelled their expertise so as to realise protocols for defining inclusion/exclusion criteria, the episode of care (when it starts, follow-up etc.), comorbidities (used in risk adjustment), and the specification of outcome measures.

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.

The purpose of this publication is to summarise the key points and main findings from the EuroHOPE project. All the results presented in this publication are based on the data and work of the EuroHOPE project. The structure of the remaining is the following. In the next chapter, the relation between country level differences and health care system is considered. That is followed by regional level analyses. The resource intensity is compared in the section thereafter, which is then followed by analyses where quality is reflected against resources. Finally, the hospital-level comparison for Nordic hospitals using sub-sector level analysis is studied, and observations that are useful for the future are made.

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 countries 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. Figures 2–6 describe these indicators for AMI, ischaemic stroke, hip fracture patients, breast cancer and very low birth weight infants, respectively.

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.

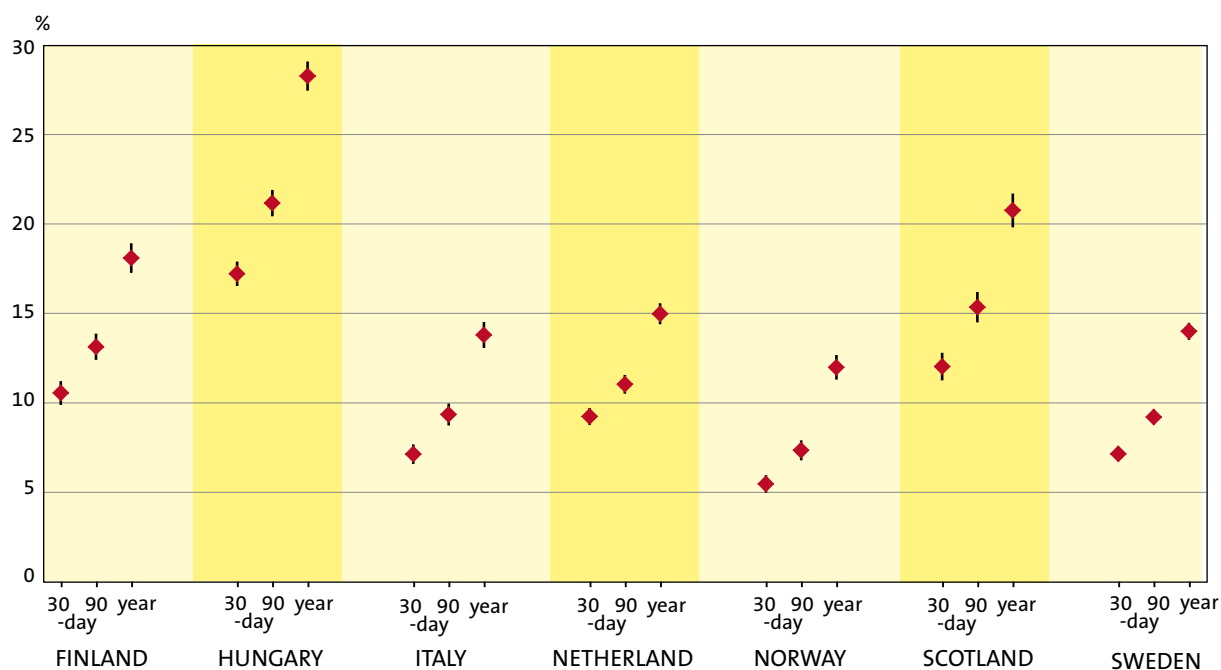


FIGURE 2. Age- and sex-standardised 30-day, 90-day and one-year mortality and their 95% confidence intervals of AMI patients by country in 2008 (2009 Norway)

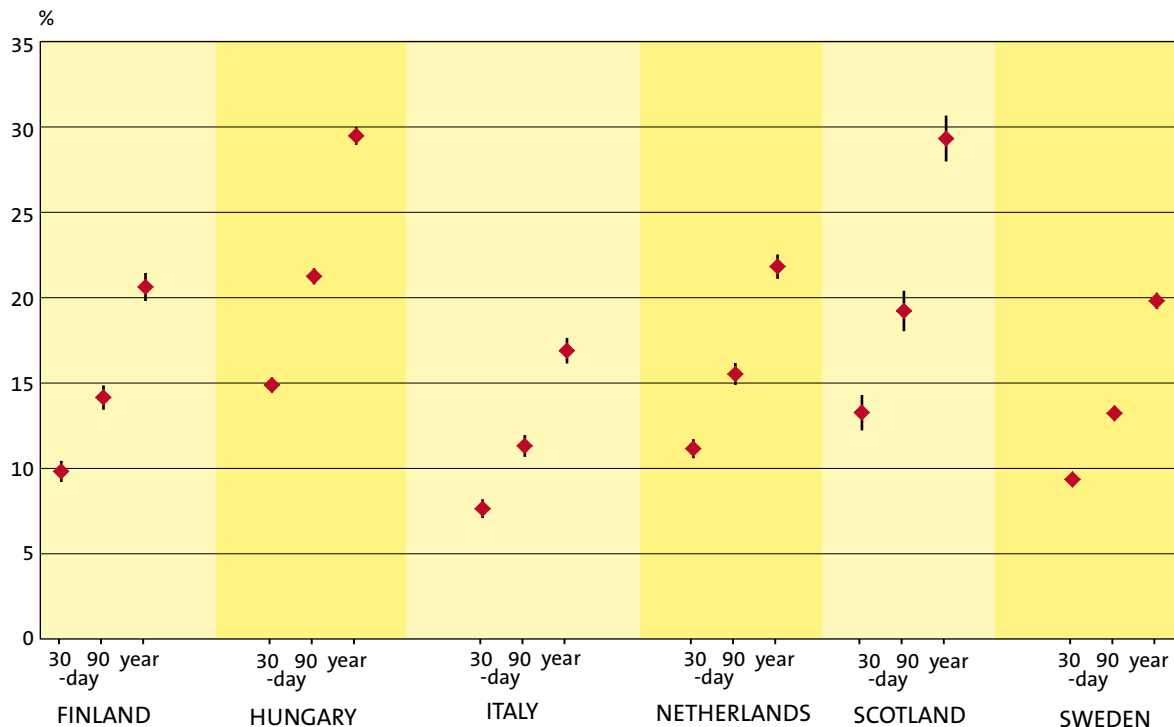


FIGURE 3. Age- and sex-standardised 30-day, 90-day and one-year mortality and their 95% confidence intervals of ischaemic stroke patients by country in 2008

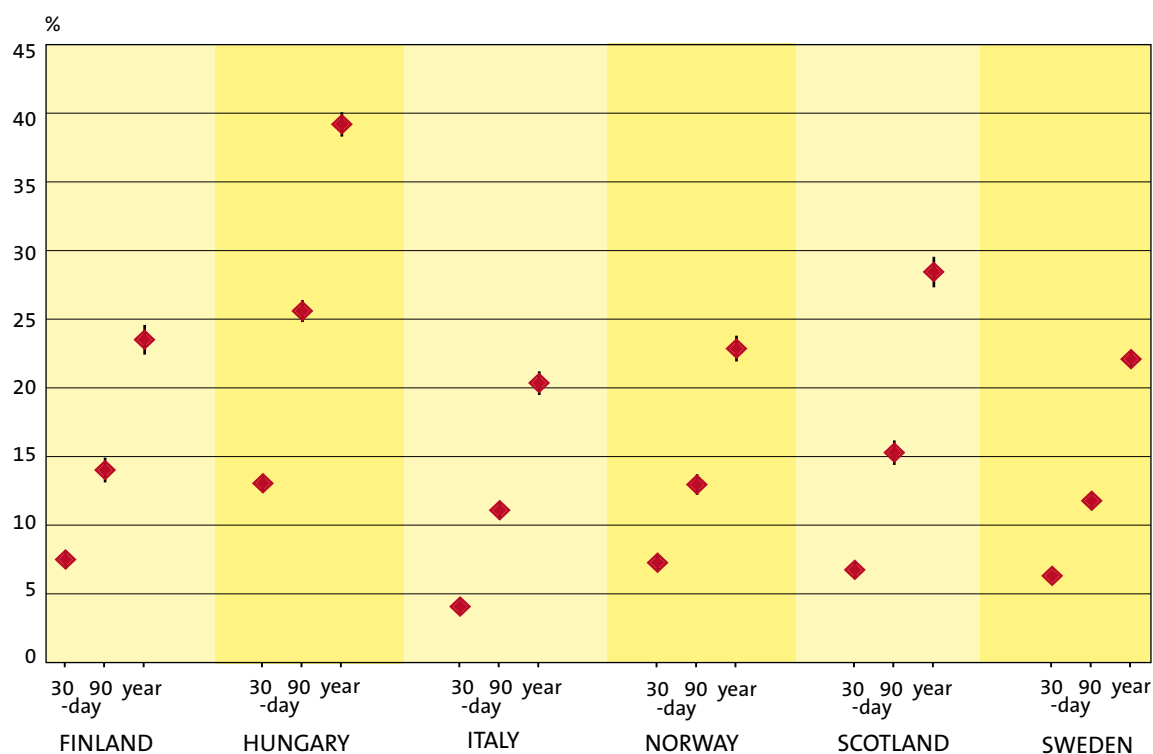


FIGURE 4. Age and sex standardized 30-day, 90-day and one-year mortality and their 95% confidence intervals of hip fracture patients by country in 2008 (2009 Norway)

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. Breast cancer mortality was lowest in the Nordic countries and Italy (Figure 5). 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 (Figure 6)¹. It should be mentioned that the Italian data were not representative and covered relatively wealthy Italian regions (city of Turin and the Lazio region).

¹ The ability to link birth registers to hospital discharge registers and mortality registers was problematic due to the time lag and inconsistencies in assigning personal identification numbers (ID) to newborn infants in virtually all countries. In particular, in Hungary personal IDs were not available and linkage was stochastic; linkage rates for Scotland, Sweden and Norway were estimated to be approximately 85%, 57%, and 59%, respectively, and that infants with poorer outcomes were more likely to be missing from the follow-up analysis. Whereas mortality rates were linked with medical birth registries in Sweden and Norway to try to correct for this bias, in Scotland this was not possible. Therefore, the mortality rates and length of stay measurements in these three countries should be interpreted with caution, especially Scotland.

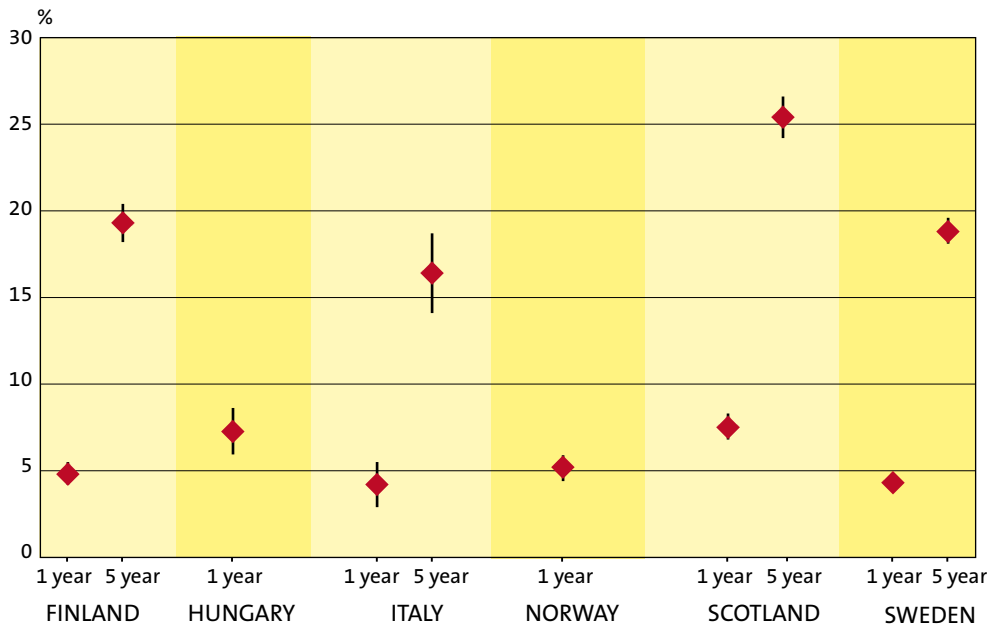


FIGURE 5. Age standardised one-year and five year mortality and their 95% confidence intervals of breast cancer patients by country in 2006 (Hungary 2009, Norway 2005)

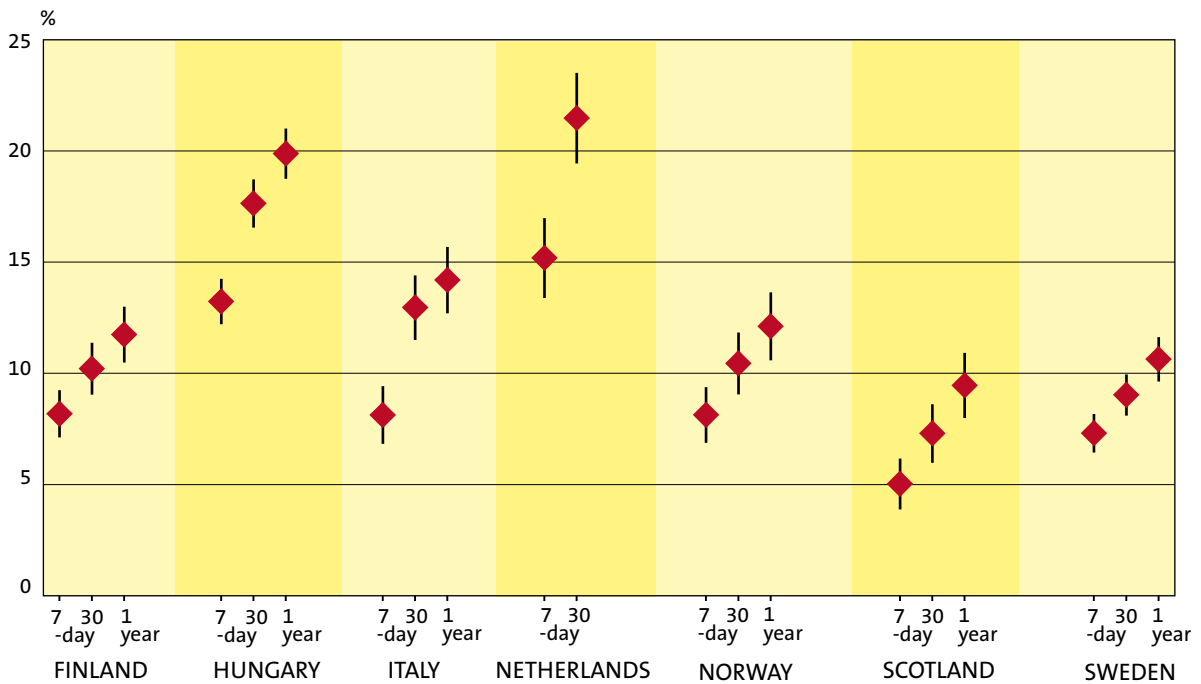


FIGURE 6. Risk-adjusted 7-day, 30-day and one-year mortality and their 95% confidence intervals of very low birth weight and very low gestational age infants by country in 2006–2008 (Netherlands 2005–2007, Norway 2008-2009)

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. Figure 7 describes the length of first hospital episode in the three conditions by country.

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. Also in Italy the length of stay of hip fracture patients was high.

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 (Figure 8). 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.

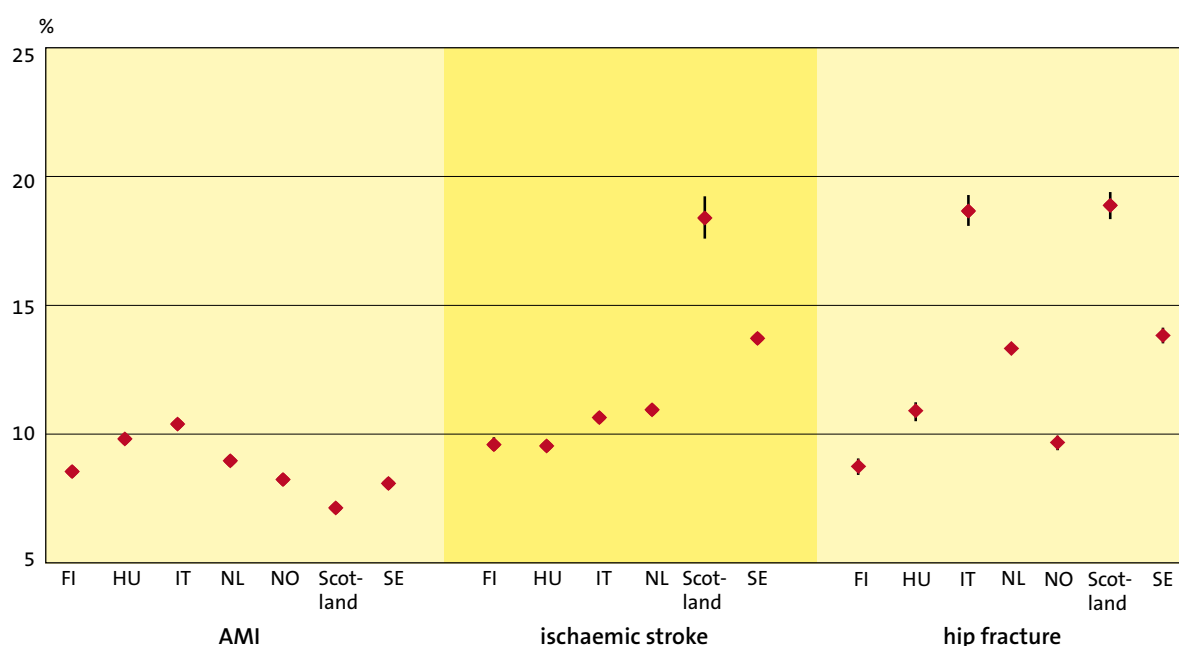


FIGURE 7. Age- and sex-standardised length of first acute hospital episode and their 95% confidence intervals of AMI, ischaemic stroke and hip fracture patients by country in 2008 (Norway 2009)

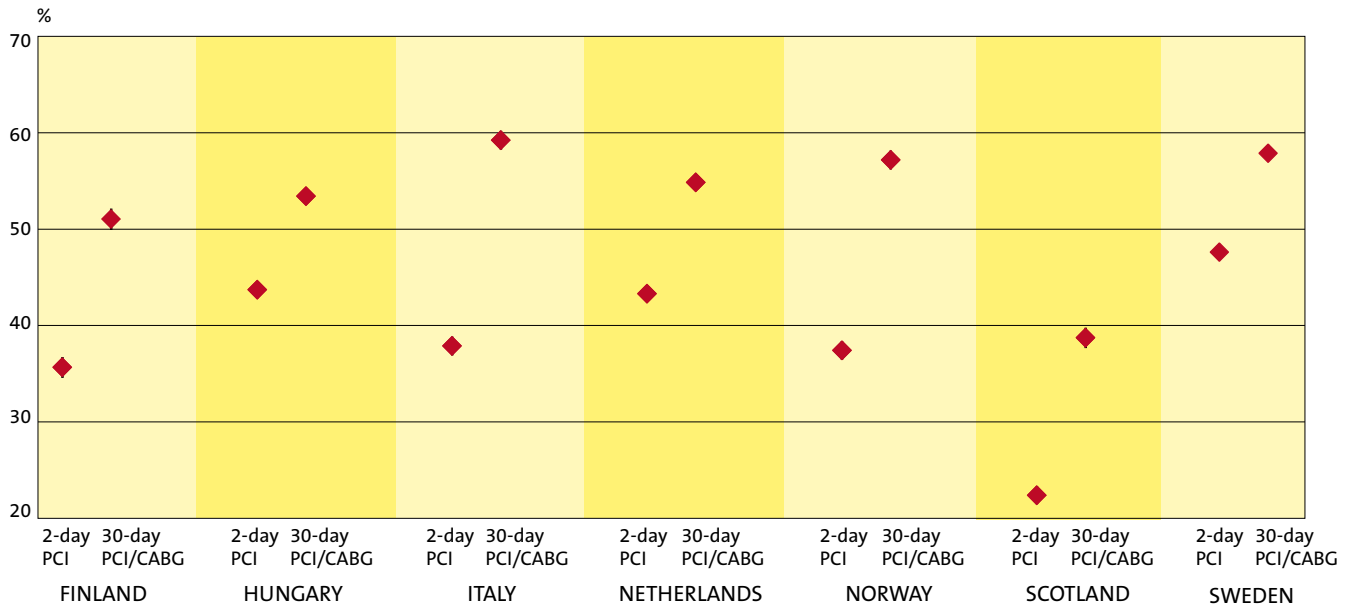


FIGURE 8. Age and sex standardized 2-day PCI rate and 30-day PCI/CABG rate and their 95% confidence intervals of AMI patients by country in 2008 (Norway 2009)

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 as 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

Figures 9, 10 and 11 describe the regional differences in one-year mortality after acute myocardial infarction (AMI), ischaemic stroke and hip fracture patients, respectively. In all cases, 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 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.

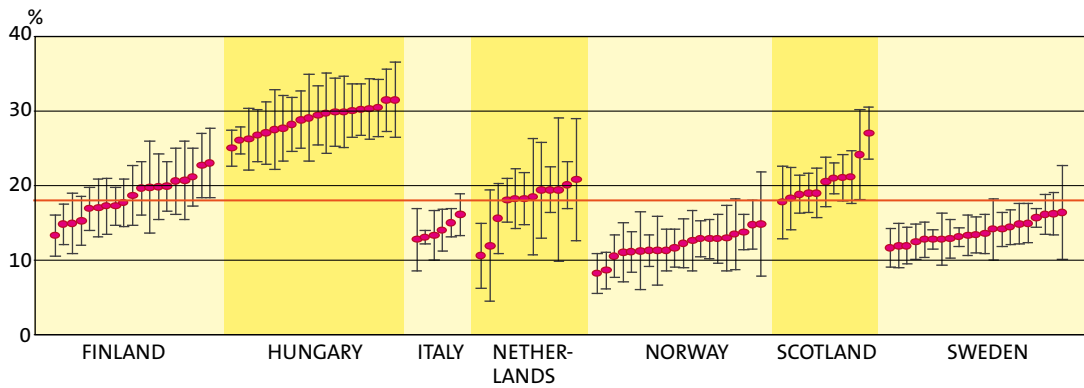


FIGURE 9. Age- and sex-adjusted one-year mortality by regions, AMI in 2008 (2009 in Norway).

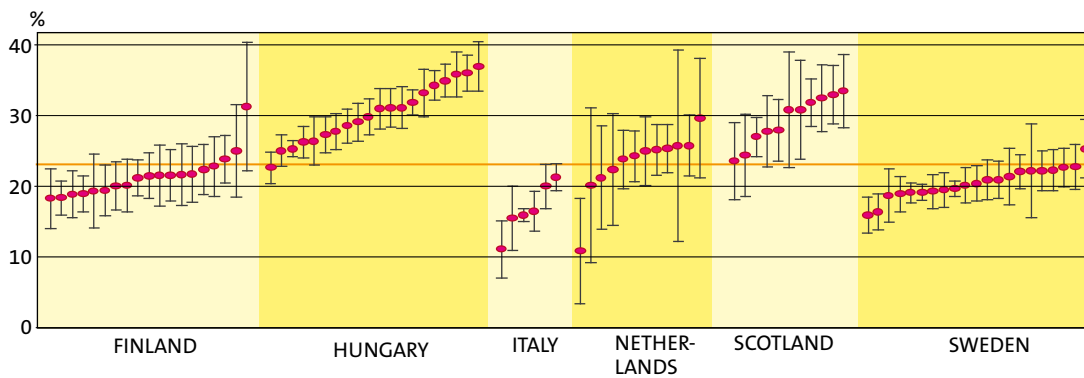


FIGURE 10. Age- and sex-adjusted one-year mortality by regions, ischaemic stroke in 2008.

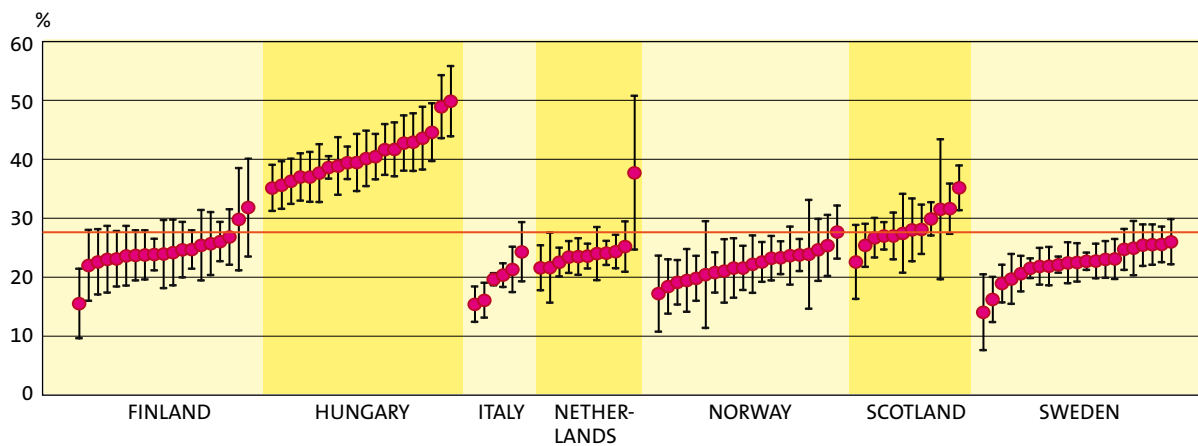


FIGURE 11. Age- and sex-adjusted one-year mortality by regions, hip fracture in 2008 (Norway 2009).

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

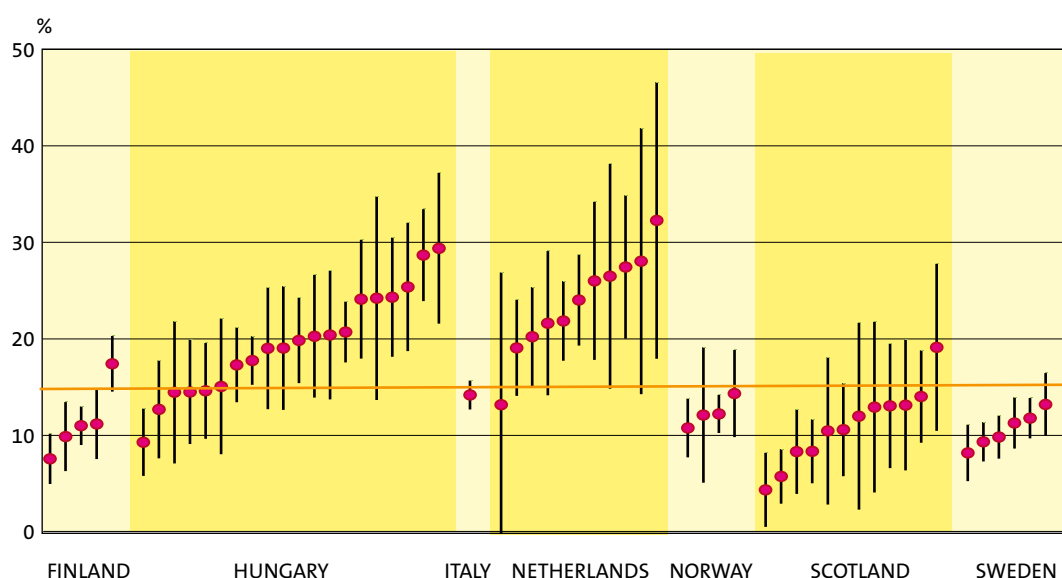


FIGURE 12. Risk-adjusted one-year mortality by regions, VLBW and VLGA infants in 2006–2008 (Netherlands 2005–2007, Norway 2008–2009)

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.

TABLE 1. Differences in the predicted cost of treating acute myocardial infarction across countries.

	First hospital episode			One-year cost		
	FIN	HUN	NOR	FIN	HUN	NOR
Approach I						
FIN						
HUN	HUN>FIN			HUN>FIN		
NOR	FIN>NOR	HUN>NOR		NOR>FIN	HUN>NOR	
SWE	FIN>SWE	HUN>SWE	SWE>NOR	SWE>FIN	HUN>SWE	SWE>NOR
Approach II						
FIN						
HUN	FIN>HUN			FIN>HUN		
NOR	FIN>NOR	HUN>NOR		NOR>FIN	NOR>HUN	
SWE	SWE>FIN	SWE>HUN	SWE>NOR	SWE>FIN	SWE>HUN	SWE>NOR

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. Results are shown in Table 1.

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

Figures 13–15 show 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 (Figure 16). 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 (Figure 14) 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 (Figure 15). 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.

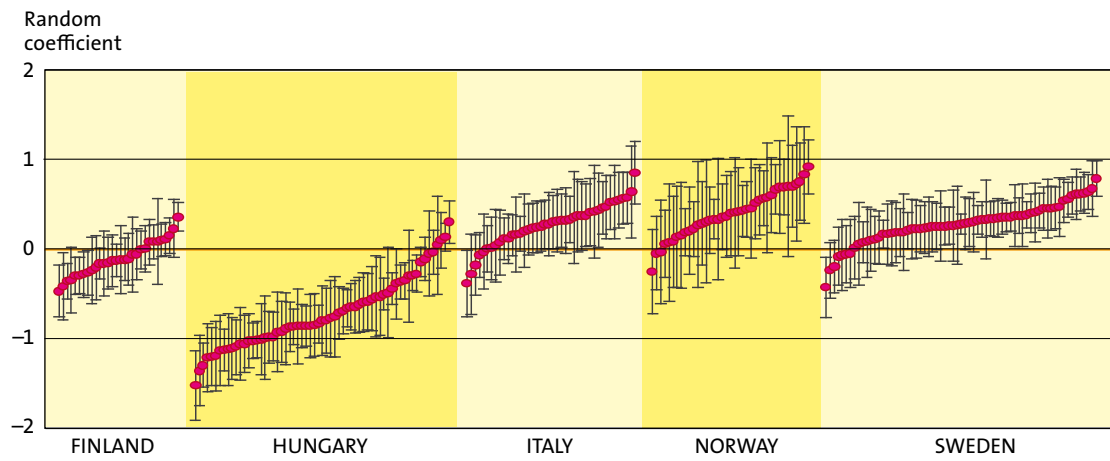


FIGURE 13. Hospitals' quality in care of AMI.

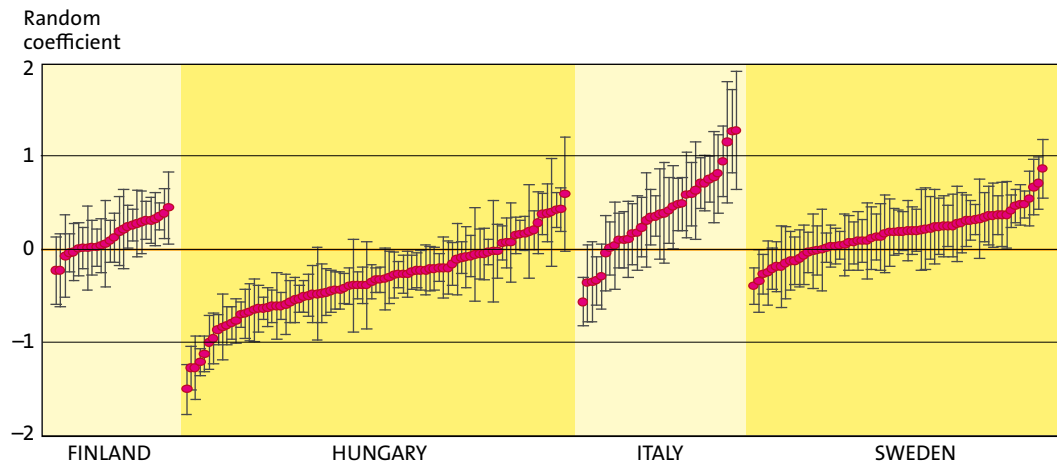


FIGURE 14. Hospitals' quality in care of ischaemic stroke.

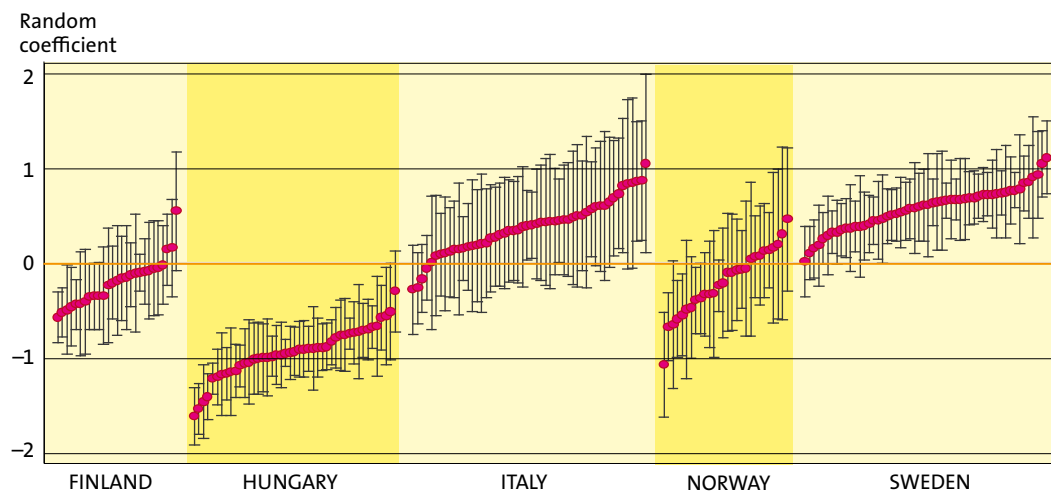


FIGURE 15. Hospitals' quality in care of hip fractures.

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. Figures 16–18 show hospitals' cost performance indicators and their 95% confidence intervals. The indicators describe how many percentage points the hospital's cost differs from the average cost for all hospitals.

In the care of AMI, cost variation within countries was much higher than between countries (Figure 16). 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 (Figure 17). The university/teaching status of a

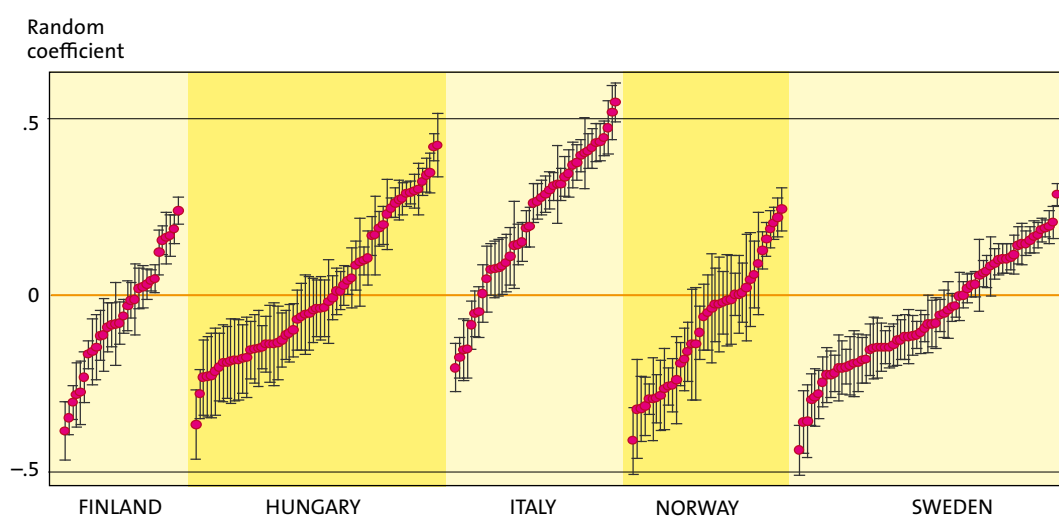


FIGURE 16. Hospitals' cost performance in the care of AMI patients

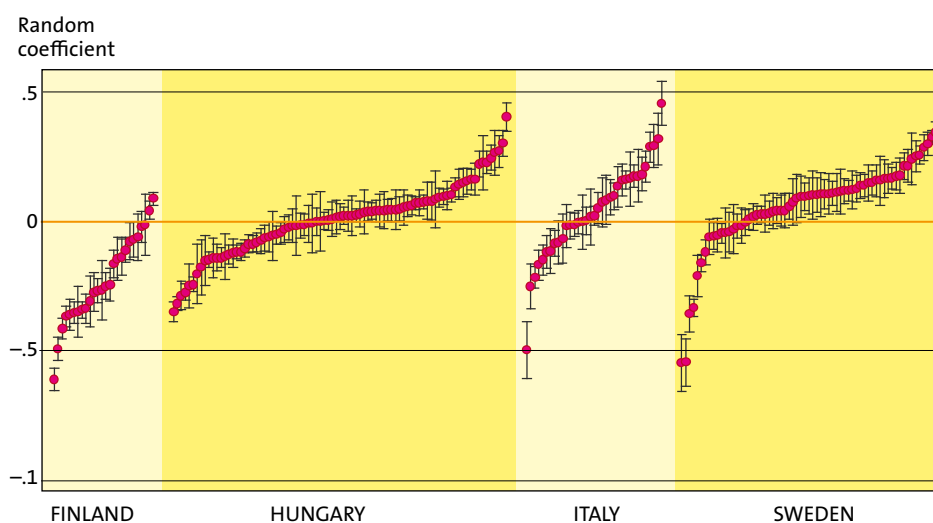


FIGURE 17. Hospitals' cost performance in the care of ischaemic stroke patients.

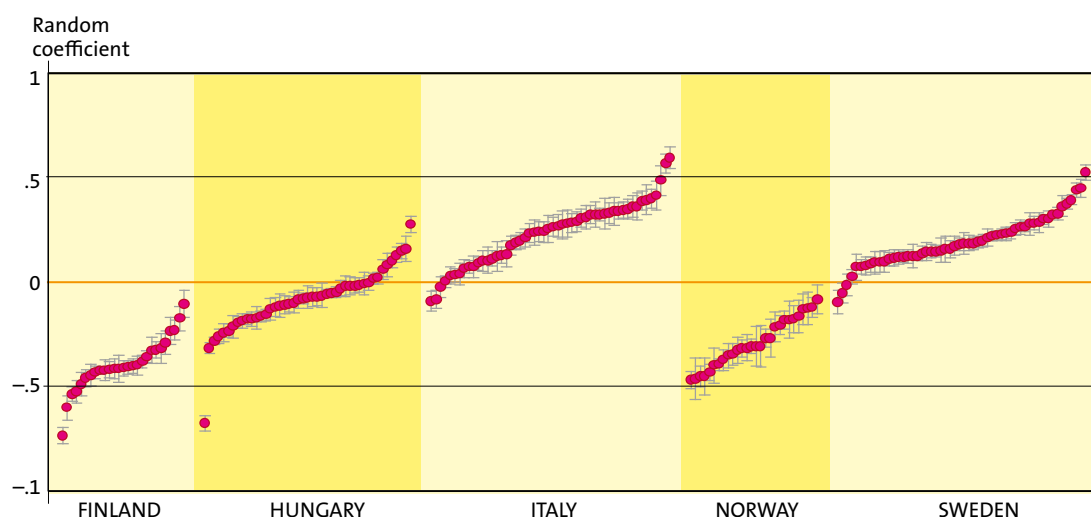


FIGURE 18. Hospitals' cost performance in the care of hip fracture patients

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 (Figure 18). 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. Figures 19 and 20 plot two of the performance measures and their 99% confidence intervals for the individual hospitals sorted by countries. 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 as shown in Figure 20 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

Figure 21 shows DEA productivity estimates of the hospitals sorted by country, with the width of the bars proportionate to hospital size. Bootstrapped 95% confidence intervals are also shown. The figure confirms 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.

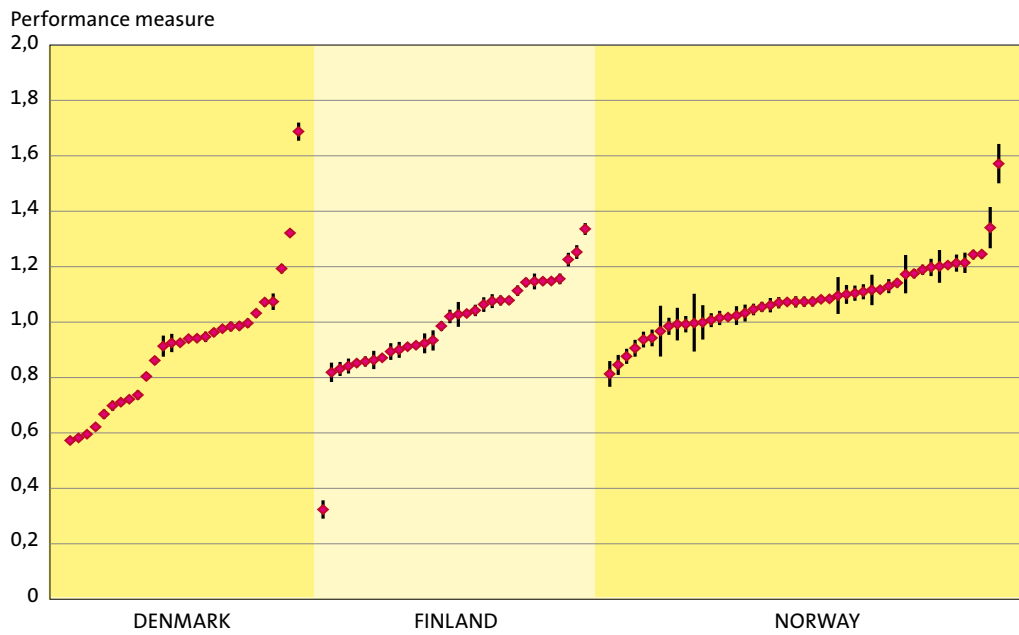


FIGURE 19. Hospital case-mix-adjusted performance measures for emergency readmissions within 30 days. Hospitals sorted by country, with 99% confidence intervals. Lower numbers indicate better quality. The mean of 1.0 corresponds to a rate of 5.62%.

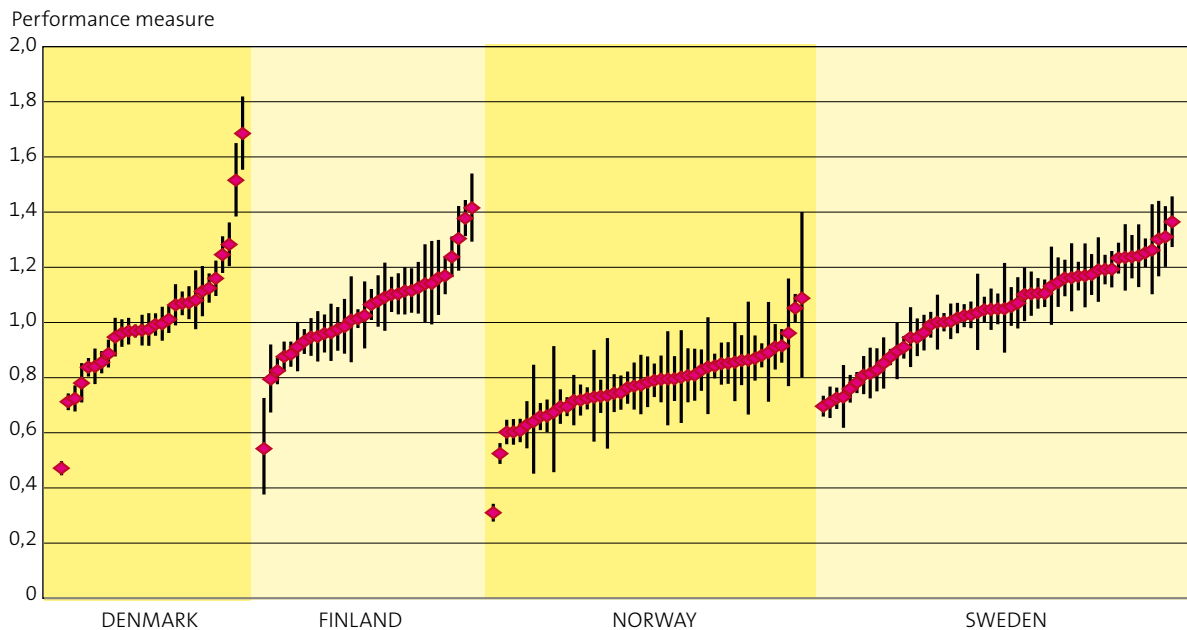


FIGURE 20. Hospital case-mix-adjusted performance measures for mortality within 30 days of last hospital admission. Hospitals sorted by country, with 99% confidence intervals. Lower numbers indicate better quality. The mean of 1.0 corresponds to a rate of 0.43%.

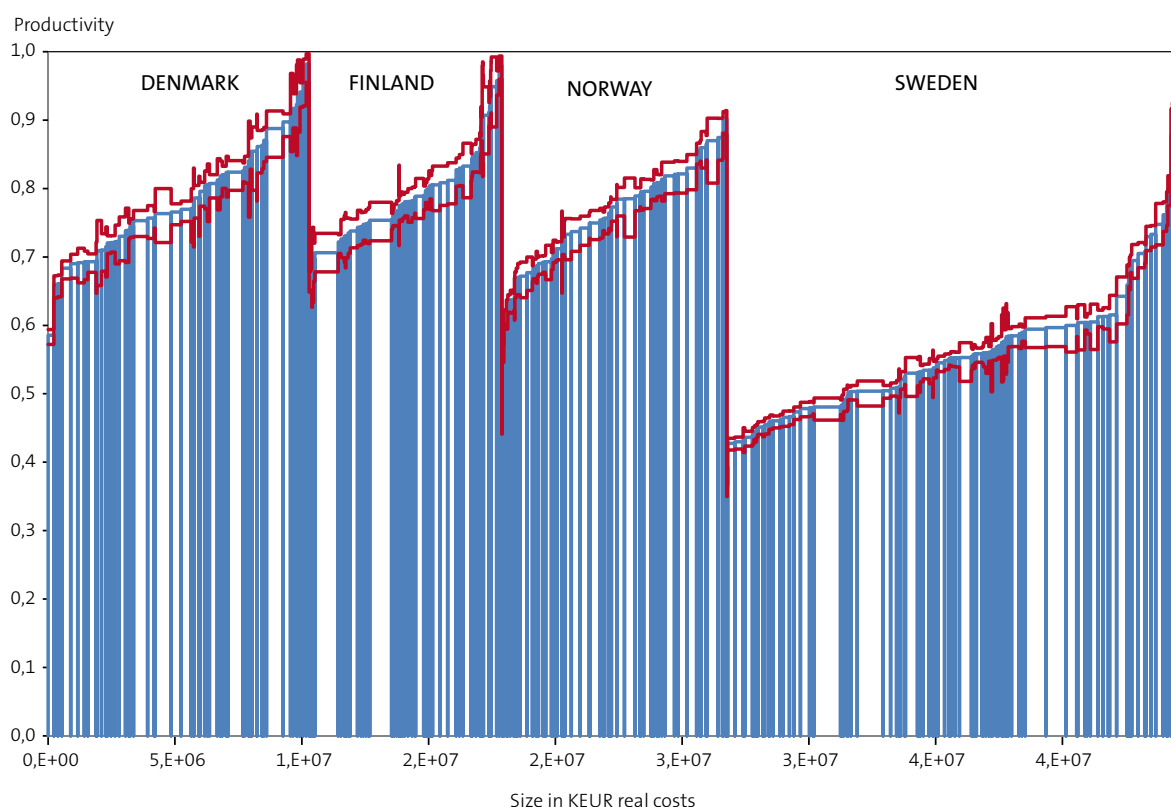


FIGURE 21. Salter diagram of bootstrapped DEA hospital productivity estimates sorted by country with 95% confidence intervals. The width of each column is proportional to hospital size, as measured by real costs.

Quality–productivity trade-off

When productivity estimates were plotted against the two of the performance measures in Figures 22 and 23, one finds no strong correlations. In both panels the optimal frontier would be at the lower right with highest productivity and lowest performance measure. In Figure 22 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.

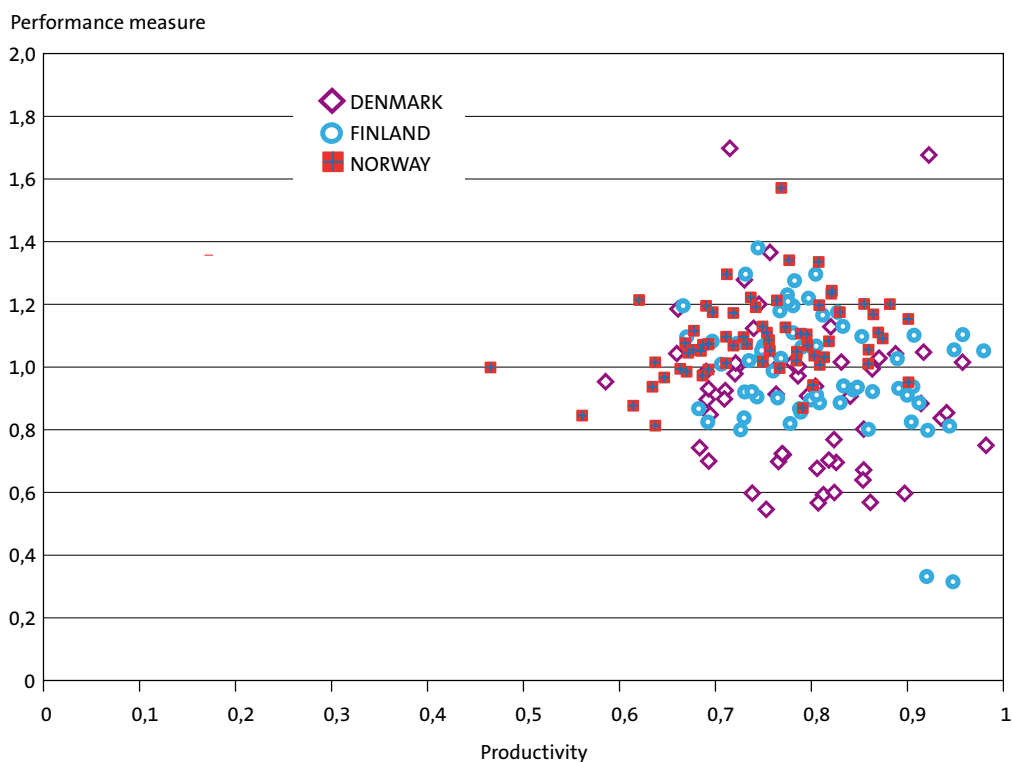


FIGURE 22. Hospital productivity estimates (horizontal axis) plotted against hospital performance measures for emergency readmissions within 30 days (vertical axis). Better joint performance is a low performance measure and high productivity (lower right).

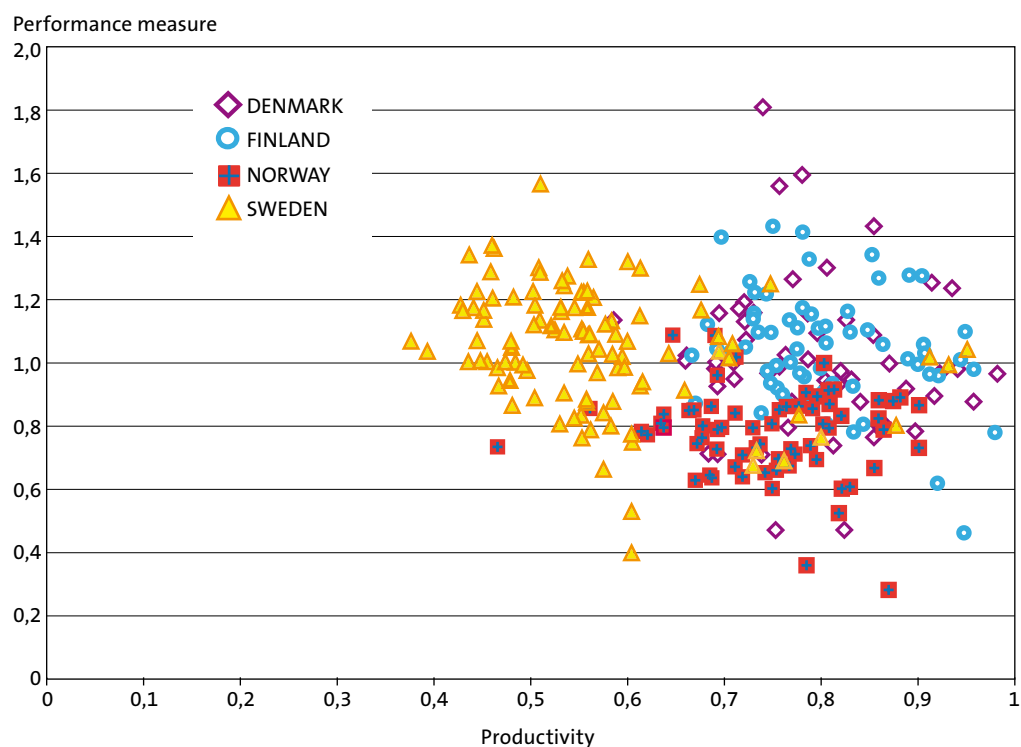


FIGURE 23. Hospital productivity estimates (horizontal axis) plotted against hospital performance measures for mortality within 30 days of last hospital admission (vertical axis). Better joint performance is a low performance measure and high productivity (lower right).

The way forward

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 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.

Finally, 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.

REFERENCES

- Chung, S. C., Gedeberg, R., Nicholas, O., James, S., Jeppsson, A., Wolfe, C., Heuschmann, P., Wallentin, L., Deanfield, J., Timmis, A., Jernberg, T. & Hemingway, H. 2014. Acute myocardial infarction: a comparison of short-term survival in national outcome registries in Sweden and the UK. *Lancet*. Published online January 23, 2014.
- Häkkinen U, Iversen T, Peltola M, Seppälä TT, Malmivaara A, Belicza É, Heijink R, Fattore G, Numerato D, Medin E, Rehnberg C. 2013. Health care performance comparison using a disease-based approach: the EuroHOPE project. *Health Policy* 112 (1-2):100-9.
- Häkkinen U, Joumard I. 2007. Cross-country analysis of efficiency in OECD health care sector: options for research. OECD Economics Department working papers, no. 554. Paris: Organisation for Economic Co-operation and Development.
- Joumard, I., André, C. & Nicq, C. 2010. Health care Systems: Efficiency and Institutions. Economics department working papers No.769. Paris: Organisation for Economic Co-operation and Development.
- OECD. 2013. Strengthening health information infrastructure for health care quality governance: Good Practices, New Opportunities and Data Privacy Protection Challenges. Paris: Organisation for Economic Co-operation and Development.