20 years of Nordic comparative health economics research

UNTO HÄKKINEN[1]
TERJE P. HAGEN[2]
SVERRE A. C. KITTELSEN[2,3]

[1] Finnish Institute of for Health and Welfare (THL), Helsinki, Finland
[2] Department of Health Management and Health Economics, University of Oslo, Oslo, Norway
[3] Frisch Centre, Oslo, Norway

Abstract: Nordic comparative health economics research stands out internationally both by its access to excellent patient data and its long-time commitment to rigorous analyses. In this article, we present the methodological foundations and the results from two types of performance analyses – comparative analyses of health care outcomes and costs at hospital level and similar analyses at the disease level. In the concluding part, we discuss strengths and weaknesses of the Nordic comparative analyses, and how we should develop Nordic comparative health economics research further.

JEL classification: I18, L30, P47

Keywords: Nordic comparison, health care performance, hospital performance, disease-based analysis

1 Introduction

Spending on health care keeps increasing and induces countries to improve system efficiency. Under these circumstances, countries might benefit from international comparisons of health system performance, as benchmarking provides opportunities to identify and learn from the best practices, increases the sample size and thus statistical precision, and can provide control groups for evaluation of reforms.

While health economists in the Nordic countries can look back on regular workshops in the Nordic Health Economics Study Group (NHESG) for more than 40 years (Christiansen, 2014), comparative Nordic health economics studies have been rare until the

**** Correspondence to:*** Sverre A.C. Kittelsen, Frisch Centre, Gaustadalléen 21, 0349 Oslo, Norway. Email: sverre.kittelsen@frisch.uio.no.
Published: Online December 2023. dx.doi.org/10.5617/njhe.10133

turn of the millennium.[1] At that time, there were a few Nordic comparative studies of procedures and outcomes from multi-centred trials in the medical literature, as well as some studies in economics on health care expenditure or behaviour with country level data (e.g. Gerdtham et al., 1994). The introduction of large administrative databases from the early 1990s, while a common western trend, was particularly advanced in the Nordic countries. These databases included patient registers, initially for somatic hospitals, and extended gradually to psychiatric specialised care and to primary care at later stages. This article attempts to summarize the challenges and results of comparative Nordic register-based studies in health economics[2].

During the last two decades, the Nordic comparative health economics research has focused on two topics: hospital comparisons and disease-based performance analyses. The implementation of a hospital benchmarking system started in Finland in 1996 (Linna and Häkkinen 2008). Building on separate efficiency analyses by Miika Linna in Finland (e.g., Linna, 1998; Linna and Häkkinen, 1998) and by Jon Magnussen in Norway (e.g., Magnussen, 1994, 1996, Mobley and Magnussen, 1998), Linna et al. (2006) presented a seminal comparison of hospital efficiency in Finland and Norway. The analysis used state-of-the art Data Envelopment Analysis (DEA) methods and pooled data based on patient registers and hospital accounts from both countries, taking advantage of similar organisational structures and common standards for recording utilization and outcomes.

Linna et al. (2006) formed the starting point for the Nordic Hospital Comparison Study Group (NHCSG), which analysed productivity differences across acute hospitals in the four largest Nordic countries; Denmark, Finland, Norway, and Sweden. The group has been through three stages. In the first stage, the emphasis was on production performance, i.e., productivity and technical efficiency, and using the analyses to inform policy evaluations. Similar definitions for cost and output measures have been used to calculate aggregate indices of efficiency in the production of patient care. In the second stage the focus moved to explaining the efficiency differences, and in the third stage the efficiency measures have been extended to also include quality and outcome measures. We present the methodological foundation of the productivity analyses in Section 2 of the article.

Recognising the value of Nordic comparisons, Tor Iversen encouraged further use of the unique possibilities for comparative research across the Nordic countries in editorials in the Nordic Journal of Health Economics (Häkkinen and Iversen, 2012) and in the Nordic Economic Policy Review (NEPR; Iversen and Kittelsen, 2012). Both editorials emphasised the possibility of extending this comparative approach to other topics, among them pharmaceuticals, disease studies and primary care. In the same NEPR issue, Rehnberg and Häkkinen (2012) summarised the studies so far and pointed out that further directions of research should include the impact of fund-holding, contractual relations and incentives between the levels of public governments, as well as including quality indicators in efficiency analysis.

An early comparative Nordic analysis at a disease level was the decomposition of cancer costs in the NHCHG report in Norwegian by Kalseth et al. (2011), with a condensed version published later in Health Policy as Kalseth et al. (2014). However, this study only considered costs and activity levels and not health outcomes. Nordic comparative analyses

[1] As far as we know one of the first attempts was a paper by Brommels M, Häkkinen U, Lindgren J and Roos P: "Jämförande studie över produktiviteten i Svensk och Finländsk hälso- och sjukvården" presented in NHSEG meeting Odense 1-2,9 1986.

[2] Google Scholar searches were conducted using terms such as Nordic/Scandinavian, Comparison, Health care, Hospitals, but the large number of hits were primarily from the medical literature. The choice of articles is therefore heavily biased towards the studies which the authors have worked on and therefore only in English or Nordic languages.

of health care outcomes and costs at the disease level started therefore with the EuroHOPE project, an EU-funded project that encompassed seven countries. The disease-based approach was adopted for evaluating the performance of European health care systems (including those of Finland, Hungary, Italy, the Netherlands, Norway, Scotland, and Sweden) using the experiences of the Finnish PERFECT project (Häkkinen, 2011). In Section 3, we present the methods and indicators that were developed for international register-based health care performance measurement and comparison. Recently, the disease-based approach has been extended to include primary health care and social services using data from the capital areas of Denmark, Finland, Norway, and Sweden, an extension that makes it is possible to describe and analyse the whole cycle of care (Häkkinen, et al., 2018a, 2020).

The next sections describe methods, data and main results for the hospital level analyses (Section 2) and the disease level analyses (Section 3). We provide a discussion in Section 4 while Section 5 concludes by suggesting that more quality measures need to be included and the communication between researchers and policy makers should be improved.

2 The Nordic hospital productivity analyses

2.1Measuring productivity and efficiency

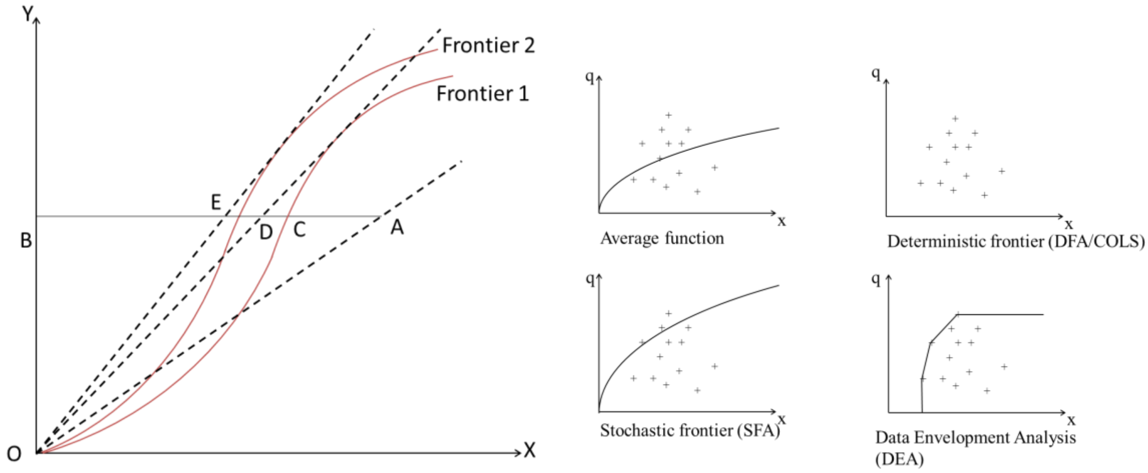
The comparative hospital level productivity and efficiency studies are methodologically based in production economics, with emphasis on frontier methods for estimating production functions and cost functions that started with Farrell (1957) and have been refined in various directions later (see e.g., Fried et al., 2008). Most analyses have been conducted using variants of Data Envelopment Analysis (DEA) that has its basis in Farrell (1957), with linear programming formulations in Charnes et al. (1978) and Banker et al. (1984).

Initially, the NHCSG used DEA, and was an early adopter of the bootstrapping tools for evaluating significance and confidence intervals for DEA efficiency estimates developed by Simar and Wilson (1998, 2000). Building on another strain of the frontier methods based in econometrics, the NHCSG has also used Stochastic Frontier Analysis (SFA) to estimate cost functions, following a tradition from Aigner and Chu (1968). Often, the SFA analyses have been used as robustness tests for the main DEA results. The main characteristics of each NHCSG study are tabulated in Appendix 1.

2.1.1 Theoretical concepts

These methodological traditions estimate the frontier as the boundary of a Production Possibility Set (PPS), sometimes termed a technology, which is defined as the set of input and output levels where it is feasible to produce the output levels using those input levels. In Figure 1 panel a), Frontier 1 is the boundary of the PPS (in period or country 1) and the feasible combinations of inputs (X) and outputs (Y) are to the right and below this boundary. A hospital observation could be represented by the point A in the interior of the PPS. When there is only one output the frontier is also called the production function, but the methodology is easily extended to multiple inputs and outputs. Thus, it is necessary to consistently define input and output vectors that cover all the activity of a hospital without overlapping categories.

Figure 1: Frontier methods. Panel a) The components of hospital total technical productivity in input–output space. For observation A in country 1, Total technical productivity (TTP) = BE/BA, Technical efficiency (TE) = BC/BA, Technical productivity (TP) = BD/BA, Scale efficiency (SE) = BD/ BC and Country productivity (CP) = BE/BD. (Kittelsen et. al., 2015b). Panel b) Four main methods for frontier estimation.



Productivity is conventionally defined as the ratio of output(s) to input(s). In Figure 1 panel a), the dashed lines from the origin represent all points with equal productivity and one can see that the maximum productivity consistent with Frontier 1 is the slope of the dashed line through OD. A normalisation of productivity is therefore the ratio of the productivities at A and at D, which geometrically can be showed to equal BD/BA. Since one does not use prices for this calculation, relying instead on properties of the technology, this measure is called *Technical productivity*. In this tradition, *Technical efficiency* is commonly defined as the ratio of actual productivity to maximal feasible productivity given the scale of production, or equivalently the ratio of necessary to actual inputs for a given level of output (Y). Since the curved production function has variable returns to scale (VRS), the point D is not actually feasible; to produce the quantity OB, one needs to use at least BC of the input. Thus, *Input Technical efficiency* is the ratio BC/BA. The reason why technical efficiency here is higher than technical productivity is that the hospital A is operating below the optimal scale, and the ratio BD/BC is therefore called *Scale efficiency*. All these measures are generally normalised so that a fully efficient or productive hospital will have a score of 1.0 (or equivalently 100%), while any score less than 1.0 implies that the hospital is inefficient.

2.1.2 Empirical estimation

The methods differ in the way they estimate the PPS or its frontier as illustrated in panel b) of Figure 1. Largely abandoned in this literature, the average regression function estimate assumes that all unexplained variation is stochastic noise. Shifting the estimated regression function to the best performing unit, the Deterministic Frontier Analysis (DFA) at the other extreme assumes that all unexplained variation is inefficiency, while retaining restrictive assumptions on the functional form. The SFA estimate emanates from a regression approach but decomposes the error term into a symmetric stochastic error and a one-sided (positive) inefficiency term (Aigner et al., 1977). Proceeding from the basic axioms of free disposal, feasibility and convexity, the DEA estimate is the minimum extrapolation set, which will then be piecewise linear and fitting as close as possible to the observations (Banker et al., 1984). Thus, the methods each have drawbacks, principally the assumption of the absence of measurement error in DEA and the strong assumptions on the functional form of the production function (or cost function) and the error terms in SFA (see Fried et al., 2008, for a good overview of the methods).

2.2 Major data considerations

Using the Finland-Norway comparison of Linna et al. (2006) as a starting point, the first task of the NHCSG was to build a database of costs and outputs for Nordic hospitals using a common set of standards and definitions. In spite of similar health systems, which all used variants of the Nordic Diagnosis Related Groups (DRGs, defined below), this entailed a number of difficulties. Among these were the weighting of the DRGs in output aggregates, the lack of capital costs, the deflator to use to make costs comparable across time and countries, and the divergent treatment of costs and outputs related to research and education.

2.2.1 Measuring output

The Nordic comparisons have been greatly facilitated by using a common Nordic grouper for the Diagnosis Related Groups (DRGs), maintained by the Nordic Casemix Centre (http://www.nordcase.org/). The DRG system was invented in a US setting to facilitate the payments for hospital services, aiming at grouping patients that are homogenous in resource usage, and later used in many countries in Europe (Busse et al., 2011). The Nordic NorDRG system originally grouped all inpatients into about 450 groups based on ICD diagnosis and procedure codes (Linna and Virtanen, 2011). From 2008/9 the outpatients have also been grouped in DRGs.

There have been (for the most part small) national divergences in the groupings, but initially it was possible to use the same DRGs across countries. However, from 2003 Denmark considerably revised its DRG system, making direct comparison impossible. For this reason, Linna et al. (2010) only used data for 2002, and Kittelsen et al. (2008) only used data for 2002 for Denmark, while using more years for the other countries. In the subsequent studies, the Danish DRGs were mapped to the closest corresponding NorDRGs, while for the EuroHOPE data the admissions were available at the patient level allowing the regrouping of all Nordic patients using the common Nordic NorDRG grouper.

A related problem was the weighting of the DRGs in the aggregated outputs, since each country had their own weights with different calculation bases. As in other indexes, the aggregates are not very sensitive to the exact weighting scheme used, and for the most part the studies have used weights from the Helsinki hospital district, which had arguably the best cost per patient data (Kautiainen et al., 2011).

While diagnosis and procedures are everywhere coded according to the international standards, Sweden does not code the acute/elective dichotomy, nor any measure of waiting times prior to hospitalisation. Therefore, the analyses in several papers were conducted with and without Swedish data, but results are mostly consistent across specifications.

2.2.2 Measuring costs

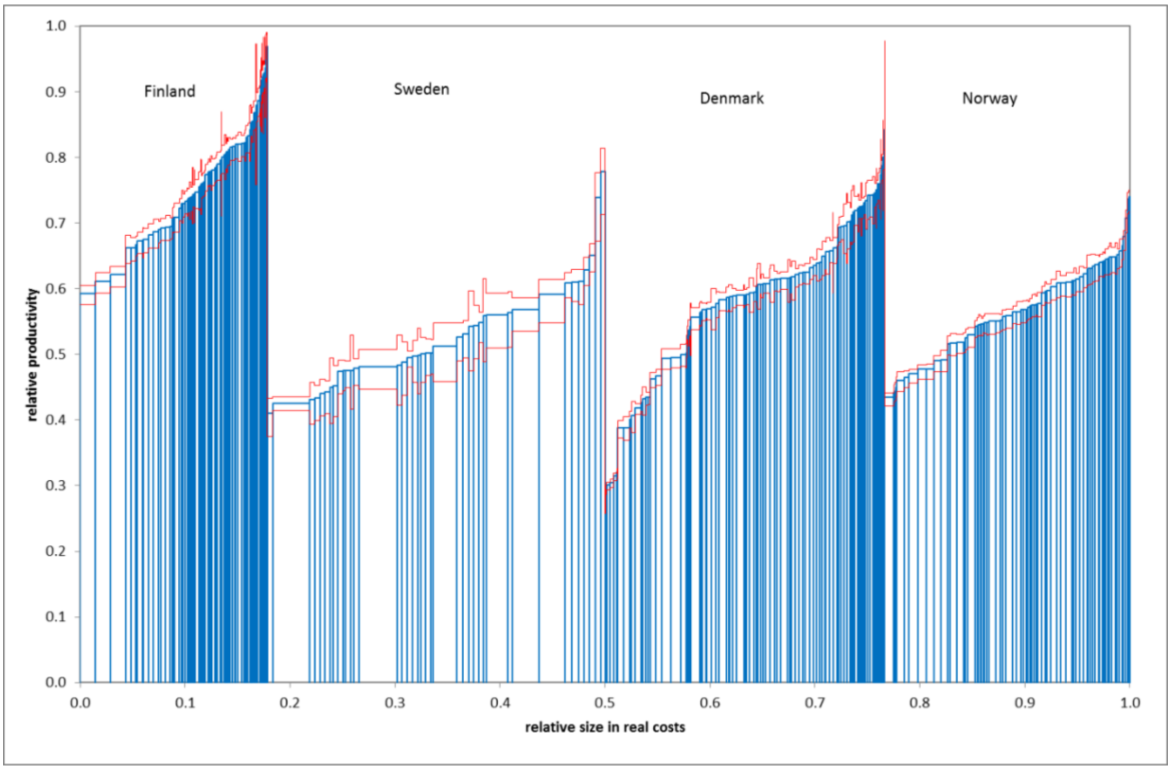
Sweden is also different in their treatment of cost data, since these are not routinely collected at the hospital level. Instead, the Swedish participants in NHCSG have collected cost data at the regional owner (Landsting) level, which has meant that the observation units are generally different from the hospitals in the other countries. The Norwegian hospital reform of 2002 contributed to this confusion, since the cost data unit, and therefore the unit of observation, changed to health enterprise level. The health enterprises comprise several physical locations and the change reduced the number of annual Norwegian observations across the different studies from 43 to 21.

Capital costs are hard to come by in any country, and closer analysis in Kittelsen et al. (2009) and Kalseth et al. (2011) concluded that those that existed were obviously defined differently and therefore not comparable between countries. The unfortunate approach used in all studies has therefore been to exclude capital costs and use net operating costs as an input. Similarly, cost attributed to research and education may have large definitional differences between countries and are therefore excluded in most of the analyses.

2.3 Main productivity and efficiency findings

The Finland-Norway comparison of Linna et al. (2006) found that the Finnish hospitals on average were considerably more efficient than the Norwegian ones, with a margin of 17–25 percentage points depending on model specification. With the inclusion of data for Sweden and Denmark, the relative efficiency of Finland was confirmed under assumptions of both Variable Returns to Scale (VRS) technical efficiency and Constant Returns to Scale (CRS) technical productivity (Linna et al., 2010). In Denmark, the average efficiency was closest to the Finnish average, with a difference of only 0–9 percentage points. Sweden appeared to have the least efficient hospitals with a difference of 13–20 percentage points compared to Finnish hospitals.

Figure 2: Technical Productivity as reported in Kittelsen et al. (2015b). The height of each bar is the bootstrapped productivity estimate for each observation with 95% confidence interval, and width is proportional to the observation size measured by real costs.



2.3.1 Explaining country differences

The country differences in average efficiency prompted an examination of the components of the differences between countries, with a special focus on what made the Finnish hospital sector perform better. In a Norwegian language reports from NHCSG using data for 2005 to 2007, Kittelsen et al. (2009) and Kalseth et al. (2011) systematically looked at differences in structure, financing, governance, and professional division of labour between the countries. The main results were published in Kittelsen et al. (2015b) which showed that the largest productivity differences were not in technical efficiency or scale efficiency within each country, but rather frontier differences stemming from characteristics that vary systematically between countries. Such characteristics may include the financing structure, ownership structure, regulation framework, quality differences, standards, education, professional interest groups, and work culture, but the importance of each of these sources could not be established statistically as these were generally fixed factors within each of only four countries. The decomposition of relative productivity is illustrated in panel a) of Figure 1 with two country-specific frontiers, and the relative productivity results for each Nordic hospital is shown in Figure 2.

Interestingly, while Sweden was estimated to have increasing returns to scale, the other countries showed decreasing returns to scale, even though the observational units in fact were already larger for Sweden. A possible interpretation of this paradox is that while the optimal size of a hospital (or provider) is quite small, the optimal size of an administrative region (or purchaser), such as the Swedish Landsting, is quite large.

There were and are some obvious differences between the countries, i.e. in wage and income levels as well as population density and topographic and climatic conditions. Utilisation rates for hospital services per capita were lowest in Finland and highest in Norway, particularly for the oldest patients. Unless older patients are very resource demanding, lower utilisation rates should mean more severe and therefore costly patients, lowering productivity measures. However, Finland also had the highest outpatient shares, which seems to increase the productivity estimates. Other contributing factors could be that the role of different professions was less controversial in Finland, and that Finland had municipal ownership of hospitals while Sweden and Denmark had regional ownership.

Norway changed ownership structure in 2002, from county to central government ownership, but delegated to five (later four) regional health enterprises. The NHCSG was able to use the Nordic data to evaluate the effect of this reform in a difference-in-difference analysis of hospital productivity (Kittelsen et al., 2007a, 2007b, 2008) of the Nordic hospitals in 1999–2004, in essence using the other countries as a control group for evaluating the Norwegian reform. With hospital fixed effects and annual time dummies in both a DEA second stage regression and an SFA simultaneous model, they found robust productivity gains from the reform in the order of 4 per cent.

Among variables that had been omitted in these studies due to lack of data were research and education outputs. This especially affects the performance of university hospitals. In a separate bootstrapped DEA study of university hospitals which contribute to education and research in addition to patient treatment, Medin et al. (2011) collected and used data on the education of interns and residents as well as citation data. All variables were significant in explaining resource usage. As before, Finland was found to be strongly (16 percentage points), and Norway (11 points) and Denmark (10 points) moderately more productive than Sweden in the patient care model, but in the models with teaching and research outputs included, there were no statistically significant country differences.

Rehnberg and Häkkinen (2012) concluded that higher efficiency levels among Finnish hospitals do not seem to be explained by differences in use of market mechanism and reimbursement systems between the countries. Instead, this might be linked to more detailed differences in the structure of financing, regulatory framework, and organisational arrangements. The authors proposed that the superior performance of Finnish hospitals was related to the methods and arrangements for the allocation of resources between different health services and the trade-off against other public duties. In Finland municipalities were (until the beginning of 2023) purchaser of hospital services and providers of various other public services (such as primary care, social services as well as school services). This was in contrast to the other Nordic countries where resource allocations to the various public services were made in separate organisations.

The NHCSG also undertook a cost of illness analysis in a comparison of cancer treatment costs in the Nordic countries at aggregate whole country level (Kalseth et al., 2014). This was not based on frontier methods, nor on the disease-based approach detailed in the next section, but demonstrated the added value of decomposing documented costs in interpreting national differences. Differences in the per capita costs of cancer treatment between the Nordic countries were found to be as much as 30%, and these are driven mainly by differences in case-mix adjusted prevalence rates, service utilisation rates and productivity.

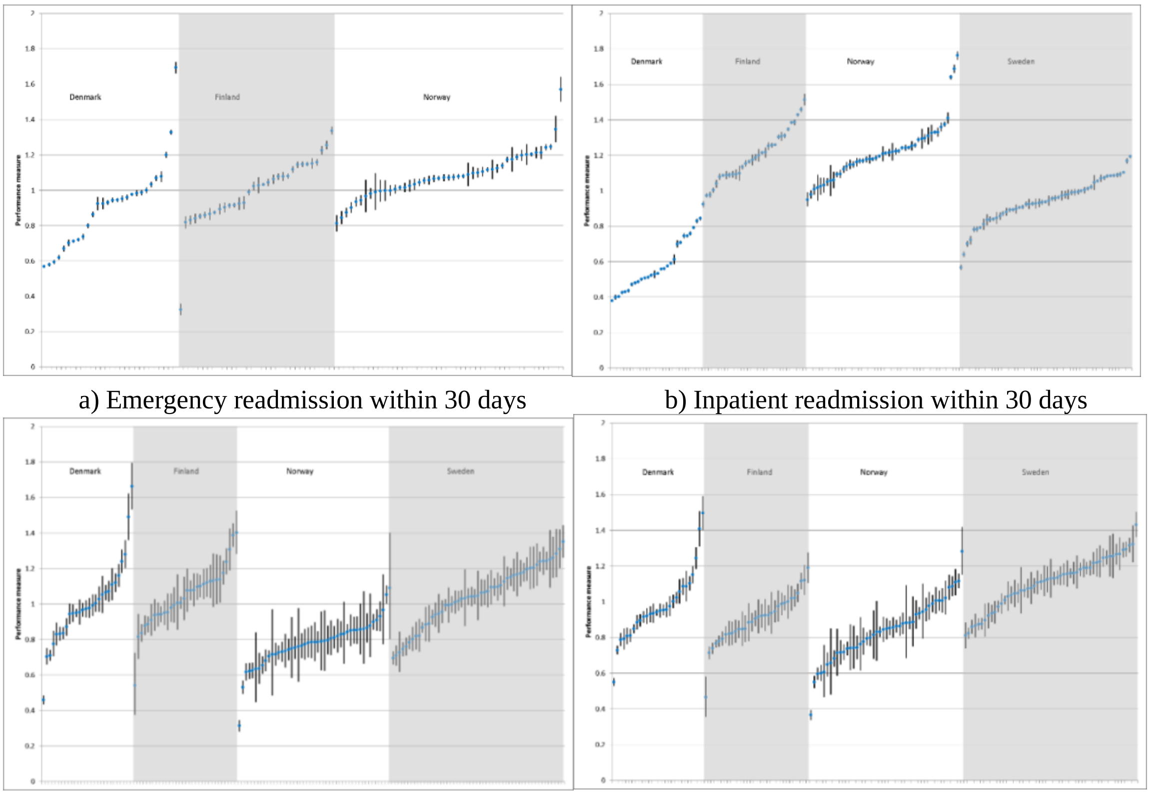
2.3.2 Quality and scale

With the advent of the EuroHOPE project, the NHCSG was able to extend the search for explanations in productivity differences to measures of the quality of outcomes. Motivational work was published in Medin et al. (2013), the major data and methodological considerations are found in Anthun et al. (2012), while results appeared in Kittelsen et al. (2015a). For this purpose, 60 million patient records were anonymised and pooled across countries to calculate quality measures. Patient register data with attached out-of-hospital mortality data made it possible to calculate mortality rates, readmission rates and patient safety indices (PSIs) for each patient group in each hospital. Using indirect standardisation based on logistic regressions on age, gender, and transfer status (to/from home, nursing home or other institution) within in each DRG, and aggregating these across DRGs for each hospital, case mix performance measures were constructed. Length of stay (LOS) and municipal characteristics were tested but not found to influence performance rates.

Case-mix adjusted quality performance measures were significantly different between hospitals and countries, with Denmark having fewer readmissions, and Norway having lower mortality. The PSIs were not indicative for quality at the hospital levels due to each of them affecting only a very small share of the patients. Figure 3 shows the distribution of the main measures across hospitals grouped in countries. Productivity was again highest in Finland, but only Sweden had clearly lower productivity. There was a slight tendency for a trade-off with productivity positively associated with inpatient readmissions, but not with emergency readmissions. For mortality, there was a significant negative association with productivity, indicating that it would be possible to reduce mortality and increase productivity at the same time, but this was mainly driven by the Danish observations. The quality performance measures did not explain the country differences in productivity.

Empirical analysis of hospitals in production economics often find little or no evidence of scale economies and quite small optimal sizes. Medical literature on the other hand provides evidence of better results for hospitals with a large volume of similar procedures. Based on the EuroHOPE data on hospitals and patients, Kittelsen et al. (2018) found that the inclusion of quality variables in the production models did not change the scale elasticity in an SFA estimation of a Cobb-Douglas cost function. This may be because medical volume effects are confined to few patient groups or possibly even offset by effects on other groups, where quality is reduced by volume. There was a weak positive effect of mortality on costs, but no effect of readmissions. Surprisingly, the scale elasticity was significantly larger than 1.0, a result that contradicts previous studies which have mostly found decreasing returns but is in line with other more recent articles that find increasing returns to scale (see e.g., Hagen , 2021).

Figure 3: Selected case-mix adjusted performance measures for hospitals sorted by country, with 99% confidence intervals. Lower numbers indicate better quality. (Kittelsen et al., 2015a).



a) Emergency readmission within 30 days

b) Inpatient readmission within 30 days

c) Death within 30 days of last admittance

d) Death within 365 days of last admittance

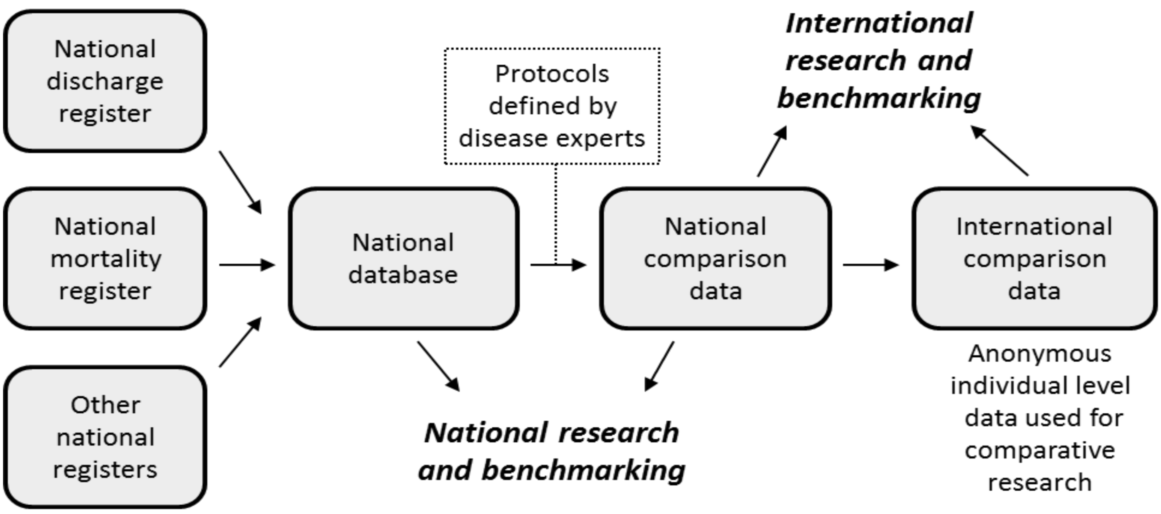
3 Disease based analysis

3.1 Background

All international comparisons require suitable information systems. These systems have until now been developed using two different approaches. The first approach relies on developing a coherent conceptual framework for information collection, analyses, and dissemination. An example of this is National Accounts, in which health care is dealt with as a part of the whole economy. Another approach is more pragmatic and based on secondary analyses of existing data. The approach assembles readily accessible data, often the by-products of existing national data collection, such as hospital discharge registers and financial reimbursement systems. This bottom-up approach relies on the work of individual experts, provider organizations, and governmental bodies engaging in quality and efficiency improvement initiatives. Micro-level comparative data on clinical actions, costs and outcomes represents an essential element of such an approach. Accurate definition, collection and scrutiny of the data are in these cases left to expert groups to determine (Smith and Häkkinen 2007, Häkkinen et al., 2013).

Important outputs from these bottom-up processes are performance measures at the disease levels. A desirable health care performance measure at the disease level is one that reliably and accurately reflects the processes, costs, and outcomes of care (Street and Häkkinen, 2009). Such measures provide valuable information for improving treatment processes and for administration at the national, regional as well as provider levels. In addition, measures that enable reliable comparisons across providers might encourage them to learn from each other and to develop their treatment processes to attain a better position in benchmarking.

Figure 4: Schematic presentation of disease specific data



3.2 Data and methodology

During the last 10 years, comparative databases have been developed and updated by using common disease specific protocols. The disease-based approach requires patient-level data covering the whole population and the possibility to deterministically link records from different national registers. The protocols define how comparison data have been constructed, based on hospital discharge registers, mortality registers, and other available administrative health care registers (medication use, specialist visits, etc.). The international comparative databases were formed using these national databases (Figure 4).[3]

For each disease, the construction of the database is based on several general stages: i) definition of the patient populations, ii) collection of the register data for the specific patient population, iii) definition of the start and end of an episode (by defining and using the *index admission* and deciding how referrals should be treated), iv) examination and coding of the patients' medical history and follow-up use of health care services in order to define specific state and time variables for the patients, v) construction of the comorbidity variables, vi) calculation of the direct health care costs, and finally, vii) the combination of the information of all stages in order to generate the comparison databases.

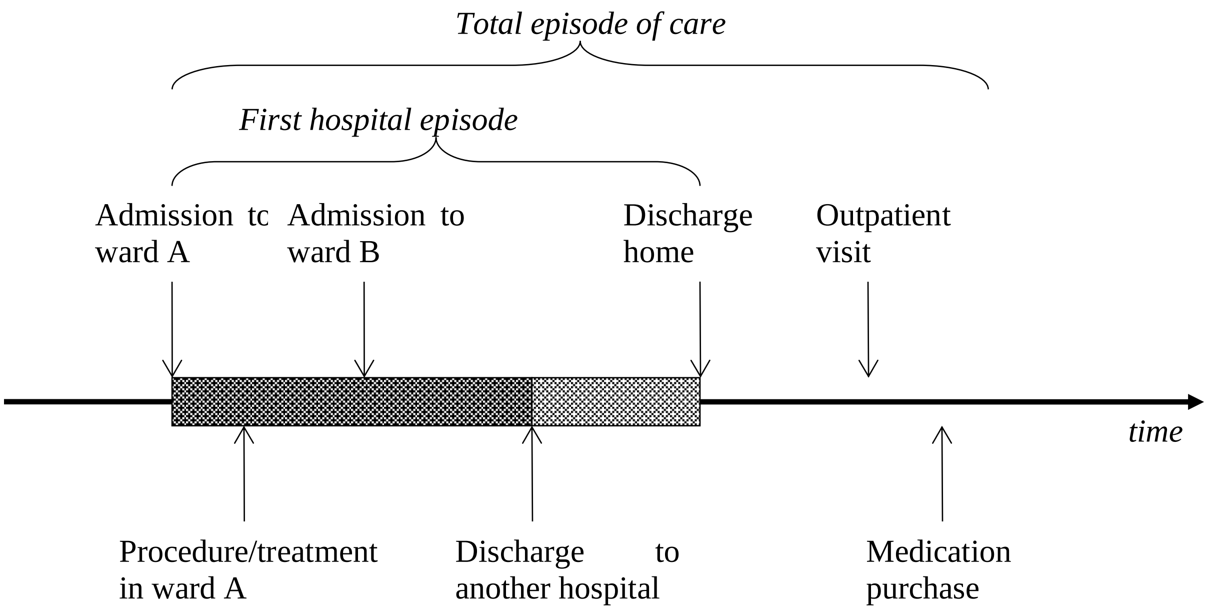
The protocols for each of the diseases defined indicators at the national level and on the regional and hospital levels within countries. The indicators included basic information

[3] The existing protocols can be downloaded from http://www.eurohope.info/. The ongoing NORCHER project has started to update protocols for ACS, stroke, and hip fracture.

on the patients (the number of patients, demographic characteristics, co-morbidity), on the content of care (the use of services and procedures, costs, treatment practices, process indicators) and on outcomes (mortality, recurrence, rehospitalisation, complications). The protocols have been developed in collaboration with clinical experts for the diseases in focus, as well as with experts on health economics, epidemiology, and statistics. The implementation of the benchmarking role of the database is carried out through basic reports (available at http://www.eurohope.info/), which include performance indicators at national and regional levels (EuroHOPE 2016). Figures 6–10 below give examples of outcome indicators that can be found in the Atlas map reports in EuroHOPE (2016).

In the following, we will describe more in detail some issues that have been important when constructing the databases and the studies as summarised in Appendix 2.

Figure 5: A schematic presentation of the follow-up of patients throughout the treatment pathway demonstrating the definitions of the first hospital episode and the total episode of care.



3.2.1 Episode of care

An episode of care refers to the entire treatment pathway from the beginning of the disease (e.g., time of diagnosis) and to the end of the treatment, crossing organizational boundaries, to deal with the health problem at hand in a specific period, see Figure 5. Thus, the protocol for an episode includes the definitions of start and end dates (follow-up time), as well as inclusion and exclusion criteria.

In EuroHOPE, the follow-up data covered at least one year for each patient. The main observable events in the register data were the time and type of admissions, the type of procedures performed, and the times of discharges as well as death dates. The secondary observable events were outpatient visits and prescribed medication purchases. In addition to the follow-up data, we had similar information on the medical history of the patients, usually for the last year before the index admission. Using the available data, it was possible to reconstruct treatment pathways that described events and localizations of the patients before and after the index stay on a daily basis. To track the patients' movements along the care pathway during the total episode, we constructed a 'state variable' that describes the services the patient received for each day of the 365 days before and 365 days after the index day (Sund and Häkkinen, 2016; Moger and Hagen, 2017).

Within the two-year period (one year before and one year after the index day), we distinguished between the *first hospital episode* and the *first institutional episode*. The first hospital episode covers all care given to a patient during the first hospital stay, which can include transfers between hospitals. Consecutive hospital discharges are included in the same hospital episode if the preceding hospital stay's discharge date is the same as the following discharge's admission date or the admission date is the next date after the preceding discharge date. The *first institutional stay* covers all care given to patients as an inpatient in a hospital or other institution, like a rehabilitation institution and/or a nursing home. The first institutional hospital episode ends when the patient is discharged to home (and is at home for at least one day), or the patient dies. The *total episode of care* also includes follow-up treatment and care after *the first institutional episode*, i.e., care and treatment in the patient's home.

3.2.2 Cost indicators

Initially, three complementary approaches to the construction of cost indicators were used (Iversen et al. 2015). Firstly, resource use was expressed in terms of weighed procedures and hospital days based on Swedish cost-per-patient (CPP) data. Secondly, resource use was expressed in terms of national DRG systems with their respective weights. Finally, resource use was expressed in terms of the common Nordic DRG system with Finnish cost weights. Each approach has its strengths and weaknesses. In the most recent studies (Häkkinen et al., 2018a, 2018b, 2020; Iversen and Häkkinen, 2018), the first approach was applied using the standard Finnish cost of specific cost items. In-hospital care costs of individual Finnish patients with credible cost data were regressed on intermediate products, and costs were then predicted for all patients using the estimated coefficients and the magnitude of the intermediate products.

3.2.3 Risk adjustment

When comparing countries, regions, hospitals and cohort years of patients, the characteristics of the patients must be accounted for. To ensure meaningful comparisons, three steps were taken. Firstly, we defined the disease groups so that they were as comparable and homogeneous as possible. Secondly, we gathered information on risk factors from the patients' medical history. Thirdly, we applied statistical models to risk-adjust the performance indicators and calculated their 95% confidence intervals. We defined the specific disease groups by the ICD 10 system. As an example, patient with acute coronary syndrome included the ICD 10-groups I20.0, I21 and I22, all registered at hospital discharge. We assumed that ICD 10 coding is homogeneous across countries, as there are European guidelines which are implemented in all Nordic countries for the diseases we have been working on.[4]

Age and sex are commonly used variables in risk-adjustment. In addition, depending on available variables and the possibility to link them with hospital discharge data, numerous variables were used. Comorbidities were assessed using patients' medical records of the previous year from two data sources: i) based on the primary or secondary diagnoses recorded during hospital admissions within 365 days prior to the index admission, and ii) based on the purchase of medications that can be linked to specific diagnoses (Moger and

[4] See for example guidelines from European Society of Cardiology: https://www.escardio.org/Guidelines/Clinical-Practice-Guidelines/Acute-Coronary-Syndromes-ACS-in-patients-presenting-without-persistent-ST-segm

Peltola, 2014). In addition, in the Nordic capital area comparison (see below), the number of days a patient had been in different service settings before the index day were used in the risk adjustment.

In international and regional comparisons we used an indirect standardisation of measures of incidence, while for all other indicators we adopted a modelling strategy: a logistic regression for dichotomic responses (e.g., mortality), a generalized linear modelling (logit-link, with gamma distribution) for continuous variables (e.g., costs), and a negative binomial modelling for discrete variables (e.g., length of stay).

In the estimation of the risk adjustment models a complication arises from the involvement of many different countries. Ideally, the individual-level data from all participating countries should be pooled before estimating the risk adjustment models. However, not all countries give permission to share the individual level data. In the initial comparisons performed, the confounding factors were first estimated for every process or outcome measure using the broadest possible data for the disease in question (e.g., for AMI the international comparison data, available from the year 2007 from Finland, Hungary, Sweden, Norway (data of 2009) and two regions of Italy). The estimations were made by weighting the data so that each country had the same weight. Subsequently, the coefficients of each model were made available to all partners who then calculated individual-level-predicted values for the indicators. The predicted values were finally summed up to country and regional level. The ratio of the observed value and the predicted value of the dependent variable in the comparable unit could be multiplied with the average value of the indicator in the pooled data to constitute the risk-adjusted indicator.

In most of the later comparisons we did not have ethical approval from all the partners to pool the individual-level data from all countries. In this case the parameter estimates for the confounding factors were first estimated for each performance indicator using the data from Finland (from the Helsinki area in the Nordic capital area comparison). The coefficients of each model were made available to all partners, who then calculated individual-level predicted values for the indicators.

3.3 Extension to primary and social care

An important extension of the disease-based approach was to include data from primary health care, including care services for elderly and disabled. The pilot, financed jointly by EU (The BRIDGE project) and the Norwegian Research Council (Comparative effectiveness analyses of coordinated care initiatives in three Nordic countries), made it possible to link data from specialist and primary care for the four Nordic capital areas, Helsinki, Oslo and Stockholm for the years 2009 to 2014 and Copenhagen for the year 2014 (Häkkinen, et al., 2018a, 2020). For Finland and Norway, we also included socioeconomic variables provided by the national statistical offices. New variables describing the timing of discharge to home and institutionalization, as well as variables describing the use and cost of primary and social hospital services, were included. Risk adjustment was performed with four different sets of confounders.

Since an important outcome measure is the time from discharge to home, in some analyses it was reasonable to consider only patients who were not institutionalised before the start of their hospital episode. The ‘state’ variable mentioned above was based on the idea that a patient can be at only one specific place each day (but the patient can receive an outpatient visit at the same day as s/he stays at home). The variable conveys information about 1) the patient’s fundamental state (dead, alive at home, or alive at an institution), 2) the type of care (hospital, rehabilitation, nursing home, home nursing, or others), 3) the main diagnosis and intensity of the treatment (i.e., acute care or non-acute), and 4) the type of outpatient visit.

Figure 6: National age- and sex-standardised 30-day mortality and confidence intervals of Ami patients, 2009-2014 (%). Reference population: Finnish AMI patients. In Italy Friuli-Venezia-Giulia region.

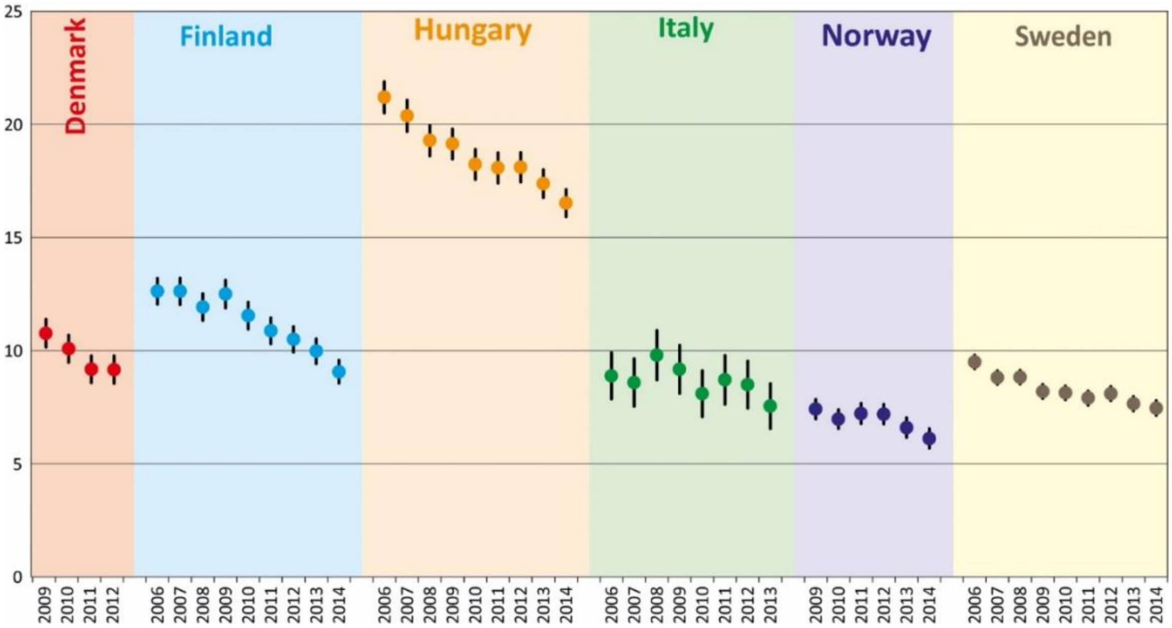
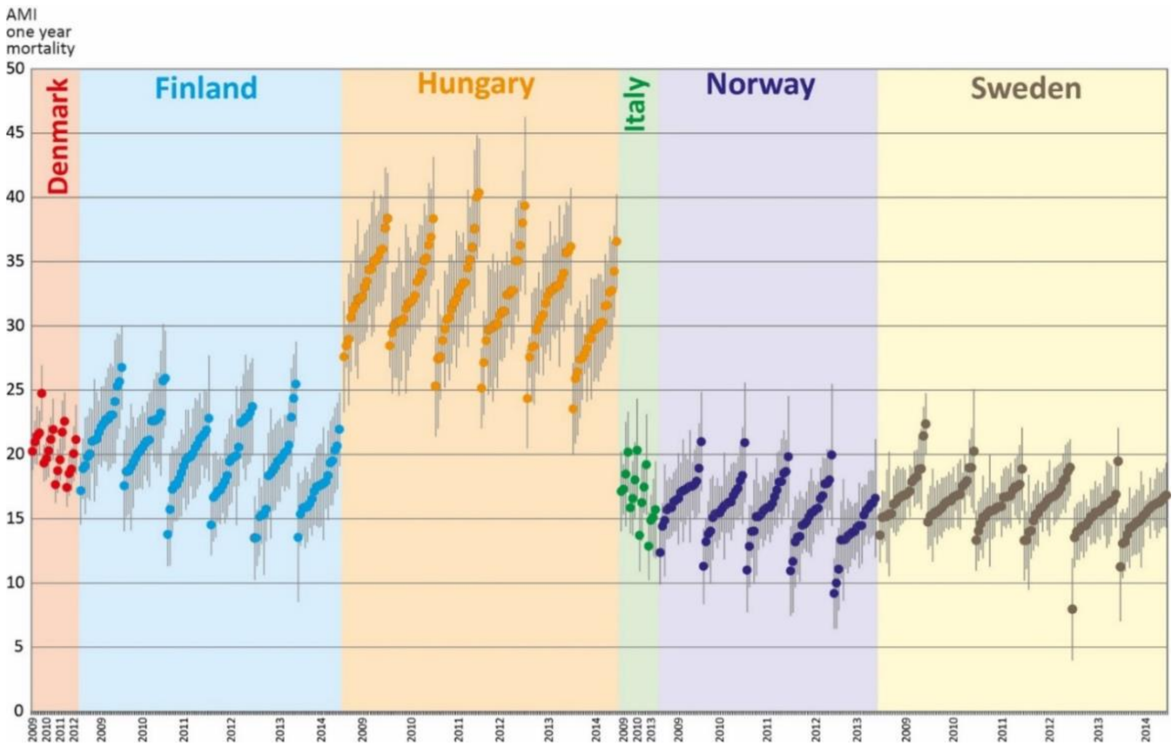


Figure 7: Regional age- and sex-standardised one-year mortality and confidence intervals of AMI patients, 2009-2014 (%). Reference population: Finnish AMI patients. In Italy, four provinces of the Friuli-Venezia-Giulia region



3.4 Main findings of the diseased-based approach[5]

1.1.1 Acute Myocardial Infarct (AMI)/Acute Coronary Syndrome (ACS)[6]

The mortality of AMI patients has decreased in all Nordic countries during the years 2009–2014, but country differences have persisted both for 30-days and one-year mortality (EuroHOPE 2016, Figures 6–7). Variation within countries has been higher than between countries. Risk adjusted 30-day and one-year mortality was higher in Finland than in the other Nordic countries. However, in the Nordic capital area comparison, most of mortality indicators were somewhat higher in Stockholm compared to the other capital areas (Häkkinen et al., 2018a). During the years 2008–2009, the survival of AMI patients at the hospital level was lower in Finland than in Sweden and Norway (Häkkinen et al., 2015) with the average survival rate in Finnish hospitals at the level of the worst performing hospitals in Sweden and Norway.

In a specific Finland-Norway comparison, we were *not* able to explain the mortality differences of ACS patients between countries using variables describing differences in the hospital system and organization of care, including variables describing level of centralization. Norway and parts of Finland had centralized, while the rest of Finland had more decentralized percutaneous coronary intervention (PCI) facilities. After adjusting for all variables (including organizational factors) the marginal country differences were still highly significant, with 3.4–4.0 percentage points lower mortality in Norway than in Finland (Moger et al., 2018).

Using cost of care data for AMI patients from 2009, Iversen et al. (2015) found that the overall pattern of the estimated country-level effects depended both on the cost indicator used and on the length of the observation period (the first hospital episode or one year after the index date). A study on use of resources at hospital level (Häkkinen et al., 2015) found that Norwegian hospitals used 9% more resources and Finnish hospitals 7% less resources than hospitals in Sweden during the first acute episode. University hospitals had higher costs compared to other hospitals in Finland and Sweden whereas regional concentration of AMI treatments care decreased hospital cost in Norway and Sweden. In Finland, survival was positively associated with regional GDP per capita.

Recent comparison of costs between Norway and Finland found that the adjusted costs of the first hospital episode were higher in Finland than in Norway, while one-year hospital costs were higher in Norway, which was accompanied by a comparatively lower mortality rate in Norway. In the capital area comparison, patients in the Helsinki area had higher total costs than those in Oslo during the one-year follow-up (Iversen and Häkkinen, 2018).

Hagen et al. (2015b) examined whether differences in one-year mortality between socio-economic groups in Finland and Norway were related to the use of PCI. The analysis indicated that socio-economic differences in the use of PCI intervention were only marginally associated with differences in mortality, while there were stronger effects between socio-economic variables and mortality. The study concluded that to reduce socio-economic inequalities in AMI mortality, policies should focus on the living conditions and lifestyles of patients with low income and low education.

[5] Here we focus on studies that are relevant for Nordic comparison and do not discuss all studies (Hagen et al. 2015a;Medin et al. 2015;Peltola et al. 2015) included in the Appendix 2.

[6] ACS is broader definition of acute coronary diseases including unstable angina.

Figure 8: National age- and sex-standardised 30-day mortality and confidence intervals of ischemic stroke patients, 2009-2014. Reference population: Finnish stroke patients. In Italy Friuli-Venezia-Giulia region

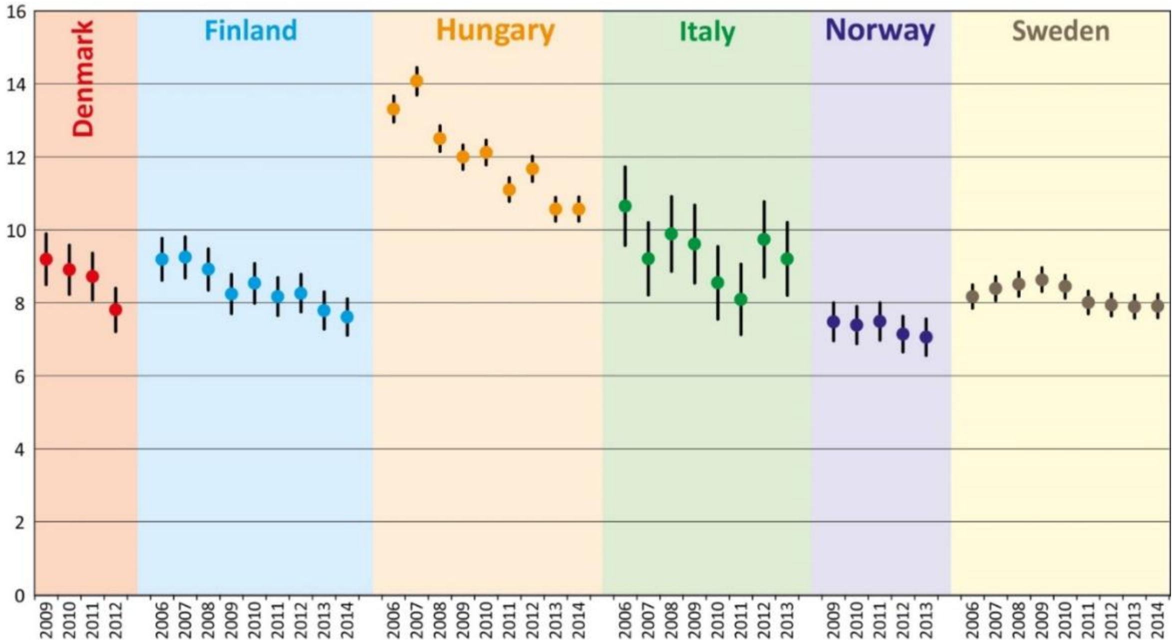
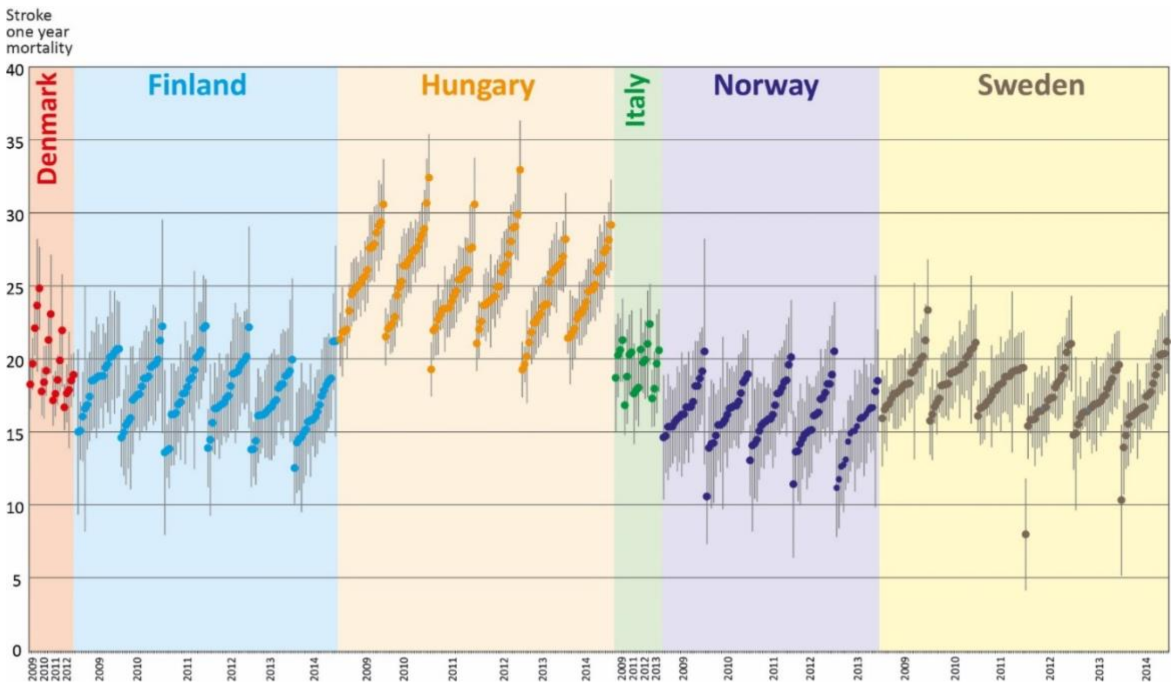


Figure 9. Regional age- and sex-standardised one-year mortality and confidence intervals of ischemic stroke, 2009-2014. Reference population: Finnish stroke patients. In Italy, the four provinces of the Friuli-Venezia-Giulia region

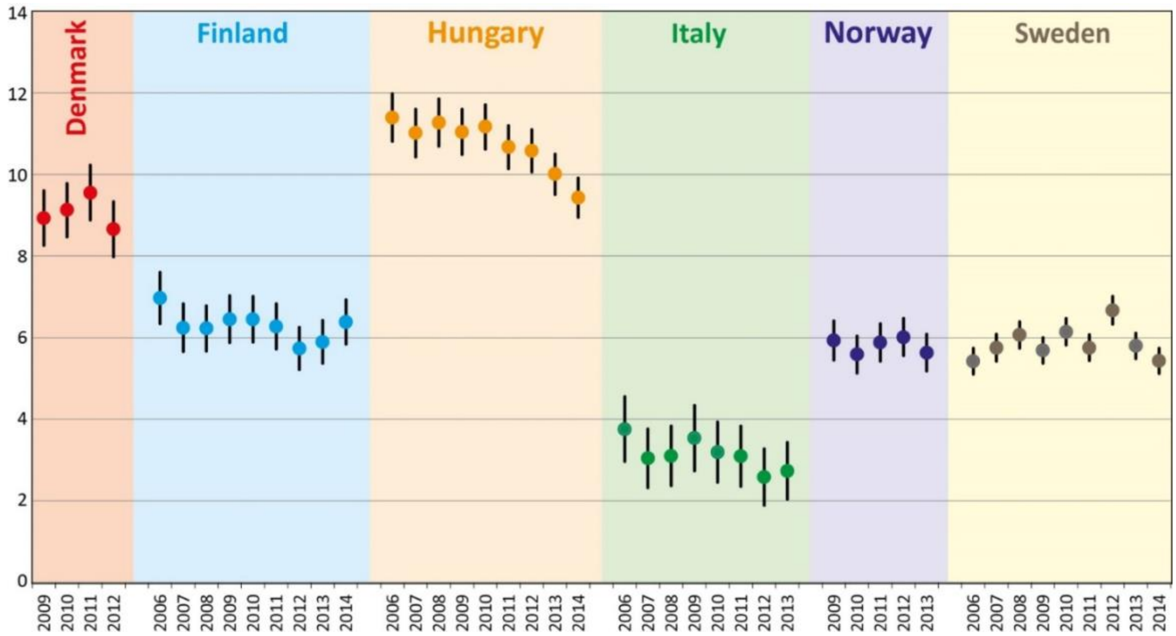


3.4.2. Ischemic Stroke

Age and sex standardised mortality of ischemic stroke patients during the years 2009–2014 was somewhat lower in Norway compared to the other Nordic countries (EuroHOPE, 2016, Figures 8–9). Except for Sweden, mortality had been decreasing in all countries. Country differences were maintained. As for AMI/ACS, regional variation within countries was higher than between countries and the regional differences remained high in all countries. In the Nordic capital area comparison, most of 30-day, 90-day and one-year mortality indicators were significantly higher in Stockholm compared to the Helsinki area in the analysis covering years 2009–2014 (Häkkinen et al., 2020). Compared to Oslo, the mortality figures were also somewhat higher in Stockholm, but the differences were not significant.

In the hospital-level analysis for the years 2007–2008, survival was about the same in Finland and Sweden, but the length of stay of the first acute hospital episode was shorter in Finland compared to Sweden (Häkkinen et al., 2015). In Finland, survival was positively associated with regional GDP per capita, while length of stay was negatively associated with regional concentration of stroke care. In Sweden, the length of stay was positively correlated with university status of a hospital and regional GDP per capita, and negatively correlated with regional population density.

Figure 10. National age- and sex-standardised 30-day mortality and confidence intervals of hip fracture patients, 2006-2014. Reference population: Finnish hip fracture patients. In Italy Friuli-Venezia-Giulia region



3.4.3 Hip fracture

During the years 2009 to 2012, 30-day mortality was somewhat higher in Denmark compared to the other Nordic countries as seen in Figure 10 (EuroHOPE, 2016). During the study period, 30-day mortality had not changed in any of the Nordic countries. In addition, in the Nordic capital area comparison (Häkkinen et al., 2018a), 30-day mortality was higher in Copenhagen than in the other Nordic capital areas.

The hospital level analysis from years 2007 to 2008 did not show any differences in survival between Finland, Sweden, and Norway, while use of resources during the first acute hospital episode was significantly higher in Sweden compared the two other Nordic countries (Häkkinen et al., 2015). The use of resources was positively associated with university hospital status in Sweden, with regional GDP per capita in Norway, and negatively associated with regional concentration of care of hip fracture in Finland.

3.4.4.Extension to primary care

The Nordic capital area comparison found considerable differences in age- and sex-standardised incidence numbers, as well as in the age structure of ACS, stroke, and hip fracture (Häkkinen et al., 2018a, 2020). For example, the age and sex-standardised number of patients per capita was highest in Oslo for all disease groups. In addition, the share of patients who were permanently institutionalised before the onset of their disease was much higher in Oslo than in the Helsinki area. This may reflect differences in the role of acute care in treating patterns, as well as a more aggressive and resource-intensive treatment in Norway.

In Oslo, the ambulatory services included mainly services from GPs, while in Copenhagen and Stockholm ambulatory services included both GPs and hospital outpatient services, and in the Helsinki area merely hospital outpatient services. In addition, the use of home help services both before and after the onset of the diseases was considerable higher in Copenhagen and Oslo compared to the Helsinki area. This may indicate more developed primary and home help services in Norway and Denmark compared to Finland and to some extent Stockholm, as assumed based on descriptive information on the primary care systems (Häkkinen and Jonsson, 2009). However, these differences were not associated with lower mortality in Copenhagen and Oslo compared to the Helsinki area.

The ranking of various performance measures between the Nordic capital areas was sensitive to the risk-adjustment method used. Risk- adjustment based on age, sex and even comorbidities in the medical history of patients may not be enough for a reliable international performance comparison of diseases affecting older persons. The studies indicated that functional ability (such as measures of activities of daily living) before the onset of disease or valid measures of severity should be considered in the risk adjustment formulas.

4 Discussion

The first 20 years of Nordic comparative health economic research have produced significant results, of which many have been presented in this article. The main results from the hospital efficiency analyses indicate that the Finnish hospitals perform better than Swedish hospitals, with Danish and Norwegian hospitals in-between. The most recent study indicates economics of scale in the Nordic hospital sector. The disease-based approach indicate a significantly higher AMI/ACS-mortality in Finland and significantly higher mortality for hip fracture patients in Denmark than in the other countries. There are significant regional variations in mortality within the four countries for the three diseases that have been presented in this article and as the main rule, these differences are higher than the differences between countries. Many explanations for the differences have been tested – only a few have been confirmed.

The Nordic comparative analyses have several strengths compared to similar research done among other countries. The Nordic health registers used as data sources for the analyses stand out by their comprehensiveness and quality. This enables us to analyse performance at the regional and hospital levels. So far, most international performance comparisons have been made at the national level (OECD, 2021), without possibilities of considering patient pathways.

Table 1: Availability of national registers in Nordic countries in 2022.

Country	Denmark	Finland	Norway	Sweden
Information on health care utilisation and outcomes				
Somatic hospital care (in and out-patient)	Yes, from 2019 difficult to distinguish in- and outpatients	Yes	Yes	Yes
Psychiatric care	Yes, from 2019 difficult to distinguish in- and outpatients	Yes	Yes	Yes
Primary care	Yes	Available on services provided by health centres and private providers but not from occupational care	Yes	Available from most of regions but no national register
Prescription medicines	Yes	Yes	Yes	Yes
Formal long-term care (institution)	Yes	Yes	Yes	Available from most of municipalities but no national register
Other formal long-term care (home care)	Yes	Yes	Yes, but inconsistencies in the data structure	Available from most of municipalities but no national register
Rehabilitation (inpatient)	Included in hospital discharge register but not possible to identify	Included in hospital discharge register but difficult to identify	Included in hospital discharge register but difficult to identify	Yes – specialised rehab.
Rehabilitation (outpatient)	Yes	Yes	Yes	Yes – specialised rehab.
Other social services (e.g. sheltered housing, social assistance)	Only few data	Social assistance from whole country, outpatient social services available from some regions	Yes	Available from most of municipalities but no national register
Quality register for hip fracture patients	(Yes)	No	Yes	Yes
Quality register for stroke patients	Yes	Only from some hospitals	Yes	Yes
Quality register for ACS patients	Yes	National register in process of developing.	Yes	Yes

Country	Denmark	Finland	Norway	Sweden
Socio-economic and demographic data				
Mortality register	Yes	Yes	Yes	Yes
Demographic information	Yes	Yes	Yes	Yes
Employment and unemployment	Yes	Yes	Yes	Yes
Income and pensions	Yes	Yes	Yes	Yes
Cost data at different levels of aggregation				
Acute hospital /ward or speciality /patient level	Available only at hospital level. At ward/speciality and patient level cost can be estimated based on DRGs using national cost weights	Available from hospitals and ward (speciality level), and at patient level from several hospital	Available at level of health enterprises and at patient level (Cost per patients, KKP) from most acute hospitals	Available at regional level but not hospital level, and at patient level (Cost per patients, KKP) from most acute hospitals
Primary inpatient and outpatient care at local /ward or provider/patient level	Only information on national tariffs and charges which can be used to estimate provider, regional and patient level provider reimbursements	Only at municipal level, for some private services information on prices	Available at municipal level, some resource use information available at institution/provider level	Available at regional and at patient level (Cost per patients, KKP) from some regions
Social inpatient and outpatient care at local /ward or provider/patient level	Only at municipal level	Only at municipal level	Only at municipal level	Only at municipal level

In addition, the development and regular updating of the NorDRG system (Linna and Virtanen, 2011) provides internationally unique possibilities to compare productivity and cost of hospital care between the Nordic countries. As far as we know, other international hospital efficiency and productivity studies have been based on very crude measurement of output (e.g., Varabyova and Schreyögg, 2013; Mateus et al., 2015).

The national hospital, mortality and medication registers are also similar in their structure (variables and coding) across the Nordic countries, which facilitates the definition of common protocols and the interpretation of results. In all Nordic countries, it is also possible to link the health registers to registers of socioeconomic variables. This facilitates risk adjustments and enable sub-group analyses, for example structured by education or income classes.

However, there are some concerns for the comparability and comprehensiveness of registers. In Finland for example, there is no register data on occupational health care, which provides ambulatory health care services for most of employed population (Holster et al., 2022). In addition, it is not currently possible to distinguish between home nursing and practical assistance in Finnish registers. The Norwegian register for municipal health and care services (KPR) in not fully consistent when it comes to home services. In Sweden, national registers do not cover primary care, long-term care and social services (Table 1).

The most important data challenges include difficulties in pooling individual level data between countries due to different legal and procedural prerequisites. In the analyses done during the first 20 years of comparative research, there were only three cases of use of pooled individual patient data (international comparison data) from several countries: a) studies using the years 2008–2009 from Finland, Hungary, Italy and Sweden and year 2009 from Norway on the three diseases (Hagen et al., 2015b; Häkkinen et al., 2015; Iversen et al., 2015), b) the hospital level comparison on the 2008–2009 data from Denmark, Finland, Norway and Sweden (Kittelsen et al., 2015), and c) the more recent Finnish-Norwegian studies on ACS/AMI and hip fracture patients (Moger et al. 2018; Häkkinen et al.2018b; Iversen and Häkkinen 2018). Current technology (e.g. TSD–Services for sensitive data at the University of Oslo and FIONA–remote access system at Statistics of Finland) allows a pooled analysis to be done in a secure data portal to protect data privacy.

A few practical problems should also be mentioned. Importantly, to get access to data is a time-consuming process. In Sweden, access to data requires consent from all counties. In Norway, the process of acquiring data may take up to 2 years due to inefficient administrative processes.

Lack of cost data at patient level has been a weakness so far (Table 1). However, all countries are now in a process where they are establishing cost per patient data. Whether these data will be made available for research purposes remains to be seen. In addition, the existence of detailed data on capital costs need to be examined to improve the comparability between countries.

For the further development of the Nordic comparative health economic research, we believe that the following element should have priority. Firstly, the pilots where data from primary and social care are linked to hospital data, i.e. the capital area analyses, should be extended to the country level and the efficiencies of alternative pathways of patient treatments carefully evaluated. In Finland and Norway, country specific analyses already exist. Secondly, more quality measures need to be included in the analyses of performance and efficiency at all levels. This can be done by utilizing the quality registers that are developed in Sweden, Denmark, and Norway and that now are piloting in Finland. The cancer registers can also be used in this respect. Thirdly, productivity analyses should be extended to primary care, mental health care and long-term care of the elderly.

Of particular interest are the analyses of reforms and interventions made in one country at specific points in time, for example the Finnish health and social care reform of 2023. Due to the design of these reforms (implementation all over the country at the same time) it will be difficult to claim causality in analyses of the reform and outcome. However, by using other Nordic countries as control group causality can be discussed more meaningfully. So far this approach has applied to evaluate impacts of the 2002 Norwegian hospital reform (Kittelsen et al., 2008) and the 2012 Norwegian coordination reform (Häkkinen et al.,2018b).

5 Conclusion

There is great scope for utilising international comparisons of health service performance at many levels, at the level of wards and institutions, regions, health subsectors and sector-wide health systems. Such studies enable researchers to a) increase the number of observations and thus get estimates that are more precise, b) provide control groups for evaluating reforms and changes in one or more countries, and c) identify and learn from the best practices.

This is particularly important for the quite small Nordic countries, which have few institutions of each type, and reforms that are implemented across whole countries at the same time. The quality of administrative registers and similarity of health systems and coding standards make such comparative studies potentially very powerful in the Nordic countries. The feasibility of future studies depends crucially on the continued work on standards, as well as on the possibilities of pooling data from all countries and/or the development of statistical methods for analysis of partially pooled and standardised data. This can provide the basis for the extension to more and better measures of quality that must be included in future Nordic comparisons.

While this research needs long-term funding for the standardisation and data pooling exercises, the research community must also improve the feedback to policy makers, health service provider institutions, and health personnel at all levels, in order to strengthen the relevance of the research and contribute to better health services in all countries.

Acknowledgments

The authors would like to thank Clas Rehnberg and Marie Kruse providing information to Table 1. Thanks also to participants at the Iversen Workshop in Oslo 23rd September 2022 in honour of Tor Iversen's 70th anniversary and the editors of the NJHE for valuable comments.

References

Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, pp. 21–37.

Aigner, D.J. and Chu, S. (1968). On estimating the industry production function. *American Economic Review*, 58, pp. 826–39.

Anthun, K.S., Goude, F., Häkkinen, U., Kittelsen, S.A.C., Kruse, M., Medin, E., Rehnberg, C. and Rättö, H. (2012). *EuroHOPE Hospital level analysis: Material, Methods and Indicators*. EuroHOPE Discussion Papers. http://www.eurohope.info/doc/EHDP10_MMI.pdf

Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies. *Management Science*, 30, pp. 1078–92.

Busse, R. A., Geissler, W., Quentin and M. Wiley (Eds.). (2011). *Diagnosis-related groups in Europe: moving towards transparency, efficiency and quality in hospitals*. Maidenhead: Open University Press.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), pp. 429–44.

Christiansen, T. (2014). The Nordic Health Economists' Study group (NHESG) 35 years. *Nordic Journal of Health Economics*, 2 (1).

EuroHOPE (2016). Atlas map report available at http://www.eurohope.info/

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120, pp. 253–81.

Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (2008). *The measurement of productive efficiency and productivity growth.* Oxford: Oxford University Press.

Gerdtham, U.-G. and Jönsson, B. (1994). Health care expenditure in the Nordic countries. *Health Policy*, 26 (3), pp. 207-220.

Hagen, T. P. (2021). Er det stordriftsfordeler i sykehus? Praktisk økonomi & finans, 37(3), pp. 233–241. doi: https://doi.org/10.18261/issn.1504-2871-2021-03-06.

Hagen, T.P., Häkkinen, U., Belicza, É., Fattore, G. and Goude F. (2015a). Acute myocardial infarction, use of percutaneous coronary intervention, and mortality: A comparative effectiveness analysis covering seven European countries. *Health Economics*, 22 (Suppl. 2), pp. 88–101.

Hagen, T.P., Häkkinen, U., Iversen, T., Klitkou, S.T. and Moger T. (2015b). Socio-economic inequality in procedures and mortality among AMI patients: quantifying the effects along different paths. *Health Economics*, 22 (Suppl.2), pp. 102–115.

Holster, T., Nguyen, L. and Häkkinen, U. (2022). The role of occupational healthcare in ambulatory healthcare in Finland. *Nordic Journal of Health Economics*. Online July 2022. https://journals.uio.no/NJHE/article/view/8561/8121

Häkkinen, U. (2011). The PERFECT project: measuring performance of health care episodes. *Annals of Medicine*, 43 (Suppl 1), pp. S1–3.

Häkkinen, U. and Iversen, T. (2012). Editorial. *Nordic Journal of Health Economics*, 1 (2), pp. i-iii.

Häkkinen, U. and Jonsson P.M. (2009). Harnessing diversity of provision, in Magnussen, J., Vrangbæk, K. and Saltman, R.B. (Eds): *Nordic health care systems: Recent reforms and current policy challenges*. Berkshire: Open University Press, pp. 198-213.

Häkkinen, U., Iversen, T., Peltola, M., Seppälä, T., Malmivaara, A., Belicza, É., Heijink, R., Fattore, G., Numerato, D., Medin, E. and Rehnberg, C. (2013). Health care performance comparison using a disease-based approach: the EuroHOPE project. *Health Policy*, 112(1–2), pp. 100–109.

Häkkinen, U., Rosenqvist, G., Iversen, T., Rehnberg, C. and Seppälä, T. (2015). Outcome, use of resources and their relationship in the treatment of AMI, stroke and hip fracture at European hospitals. *Health Economics*, 24 (Suppl. 2), pp. 116–139.

Häkkinen, U., Engel-Andreasen, C., Goude, F., Hagen, T.P., Kruse, M., Moger, T., Rehnberg, C. and Peltola, M. (2018a*). Performance comparison of patient pathways in Nordic capital areas – A pilot study*. Discussion paper 22/2018. Helsinki: THL.

Häkkinen, U., Hagen, T.P. and Moger. T.A. (2018b). Performance comparison of hip fracture pathways in two capital cities: Associations with level and change of integration. *Nordic Journal of Health Economics*, 6 (2), pp. 80–98.

Häkkinen, U., Goude, F., Hagen, T.P., Kruse, M., Moger, T.A, Peltola, M. and Rehnberg C. (2020). Performance comparison of patient pathways in Nordic capital areas – a pilot study for ischaemic stroke patients. *Scandinavian Journal of Public Health*, 48, pp. 275–288.

Iversen, T. and Häkkinen, U. (2018). Comparative treatment costs for patients with acute myocardial infarction between Finland and Norway. *Nordic Journal of Health Economics*, 6(2), pp. 58–79.

Iversen, T. and Kittelsen, S.A.C. (2012). Challenges in health care financing and provision. *Nordic Economic Policy Review*, 2, pp. 7–22.

Iversen, T., Aas, E., Rosenqvist, G. and Häkkinen U. (2015). Comparative analysis of treatment costs in EuroHOPE. *Health Economics*, 22 (Suppl. 2), pp. 5–22.

Kalseth, B., Anthun, K.S., Hope, Ø., Kittelsen, S.A.C. and Persson, B. (2011). *Spesialisthelsetjenesten i Norden. Sykehusstruktur, styringsstruktur og lokal arbeidsorganisering som mulig forklaring på kostnadsforskjeller mellom landene*, Rapport A19615. Trondheim: SINTEF Health Services Research.

Kalseth, J., Halvorsen, T., Kalseth, B., Anthun, K.S., Peltola, M., Kautiainen, K., Häkkinen, U., Medin. E., Lundgren, J., Rehnberg, C., et al. (2014). Cross-country comparisons of health-care costs: The case of cancer treatment in the Nordic countries. *Health Policy*, 115 (2–3), pp. 172–9.

Kautiainen, K., Häkkinen, U. and Lauharanta, J. (2011). Finland: DRGs in a decentralized health care system. In R. Busse, A Geissler, W. Quentin and M. Wiley (Eds.), *Diagnosis-related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. Maidenhead: Open University Press, pp. 321–338.

Kittelsen, S.A.C., Magnussen, J. and Anthun, K.S. (2007a). *Har sykehusreformen ført til økt produktivitet.* Resultatevaluering av sykehusreformen. Oslo: Norges forskningsråd.

Kittelsen, S.A.C., Magnussen, J. and Anthun, K.S. (2007b). Sykehusproduktivitet etter statlig overtakelse: En nordisk komparativ analyse. *Working paper* 2007/1, Oslo: Health Economics Research Programme at the University of Oslo (HERO).

Kittelsen, S.A.C., Magnussen, J., Anthun, K.S., Häkkinen, U., Linna, M., Medin, E., Olsen, K. and Rehnberg, C. (2008). Hospital productivity and the Norwegian ownership reform – A Nordic comparative study, *STAKES discussion paper* 8/2008, Helsinki: STAKES.

Kittelsen, S.A.C., Anthun, K.S., Kalseth, B., Halsteinli, V. and Magnussen, J. (2009). *En komparativ analyse av spesialisthelsetjenesten i Finland, Sverige, Danmark og Norge: Aktivitet, ressursbruk og produktivitet 2005–2007,* Rapport A12200, Trondheim: SINTEF Health Services Research.

Kittelsen, S.A.C., Anthun, K.S., Goude, F., Huitfeldt, I.M.S., Häkkinen, U., Kruse, M., Medin, E., Rehnberg, C. and Rättö, H. (2015a). Costs and quality at the hospital level in the Nordic countries. *Health Economics*, 24(Suppl. 2), pp. 140–63.

Kittelsen, S.A.C., Winsnes, B.A., Anthun, K.S., Goude, F., Hope, Ø., Häkkinen, U., Kalseth, B., Kilsmark, J., Medin, E., Rehnberg, C., et al. (2015b). Decomposing the productivity differences between hospitals in the Nordic countries. *Journal of Productivity Analysis*, 43(3), pp. 281–93.

Kittelsen, S.A.C., Anthun, K.S., Häkkinen, U., Kruse, M. and Rehnberg, C. (2018). Scale and quality in Nordic hospitals. *Nordic Journal of Health Economics*, 6(2), pp. 29–44.

Linna, M. (1998). Measuring hospital cost efficiency with panel data models. *Health Economics*, 7(5), pp. 415–27.

Linna, M., and Häkkinen U. (1998). A comparative application of econometric frontier and DEA methods for assessing efficiency of Finnish hospitals, in Zweifel, P. (Eds.), *Health, the medical Profession and Regulation*. Boston/Dordrecht: Kluver Academic Publishers, pp. 169–187.

Linna, M, and Häkkinen U. (2008). A comparative application of Benchmarking Finnish Hospitals, in Blank, J. and Valdmanis, V. (Eds), *Evaluating hospital policy and performance: contributions from hospital policy and productivity research*. Advances in health economics and health services research, 18. Oxford: Elsevier, pp. 179–190.

Linna, M., Häkkinen, U. and Magnussen, J. (2006). Comparing hospital cost efficiency between Norway and Finland. *Health Policy*, 77(3), pp. 268–78.

Linna, M. and Virtanen, M. (2011). NordDRG: the benefits of coordination. In Busse, R., Geissler, A., Quentin, W. and Wiley, M. (Eds.), *Diagnosis-related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. Maidenhead: Open University Press, pp. 293–300.

Linna, M., Häkkinen, U., Peltola, M., Magnussen, J., Anthun, K.S., Kittelsen, S.A.C., Roed, A., Olsen, K., Medin, E. and Rehnberg, C. (2010). Measuring cost efficiency in the Nordic Hospitals – a cross-sectional comparison of public hospitals in 2002. *Health Care Management Science*, 13(4), pp. 346–57.

Magnussen, J. (1994*). Hospital Efficiency in Norway – A Non-Parametric Analysis*. Dissertations in economics. 6, University of Bergen.

Magnussen, J. (1996). Efficiency measurement and the operationalization of hospital production. *Health Services Research*, 31, pp. 21–37.

Mateus, I., Joaquim S., and Nunes C. (2015). Measuring hospital efficiency — comparing four European countries. *European Journal of Public Health*, 25 (Supplement 1), pp. 52–58.

Medin, E., Anthun, K.S., Häkkinen, U., Kittelsen, S.A.C., Linna, M., Magnussen, J., Olsen, K. and Rehnberg, C. (2011). Cost efficiency of university hospitals in the Nordic countries: a cross-country analysis. *European Journal of Health Economics*, 12(6), pp. 509–19.

Medin, E., Häkkinen, U., Linna, M., Anthun, K.S., Kittelsen, S.A.C. and Rehnberg, C. (2013). International hospital productivity comparison: Experiences from the Nordic countries. *Health Policy*, 112(1), pp. 80–87.

Medin, E., Goude, F., Melberg, H.O., Tedioisi, F., Belicza, É., and Peltola, M. (2015). European regional differences in all causes mortality and stay for patients with hip fracture. *Health Economics*, 24(Suppl. 2), pp. 53–64.

Mobley, L. and Magnussen, J. (1998). An International Comparison of Hospital Efficiency. Does Institutional Environment Matter? *Applied Economics*, 30, pp. 1089–100.

Moger, T. A. and Hagen, T. P. (2017). Analyser av tjenestebruk, reinnleggelser og mortalitet blant infarkt, slag-, hoftebruddpasienter i Oslo, 2009–2014. *Working paper* 2017/1, Oslo: Health Economics Research Programme at the University of Oslo (HERO).

Moger, T.A. and Peltola. M. (2014). Risk adjustment of health care performance measures in a multinational register-based study: A pragmatic approach to a complicated topic. *SAGE Open Medicine*, January–December 2014 vol 2.

Moger, T.A., Häkkinen, U. and Hagen, T.P. (2018). Higher mortality among ACS patients in Finland than in Norway: Do differences in acute services and scale effects in hospital treatment explain the variation? *Nordic Journal of Health Economics,* 6(2), pp. 80–98.

OECD (2021). *Health at a Glance. OECD Indicators*. Paris: OECD Publishing. https://doi.org/10.1787/ae3016b9-en.

Peltola, M., Seppälä, T., Malmivaara, A., Belicza, É., Goude, F., Fletcher, E., and Heijinik R. (2015). Individual and regional -level factors contributing to variations in length of stay after cerebral infarction in six European countries. *Health Economics*, 24(Suppl. 2), pp. 28–52.

Rehnberg, C. and Häkkinen, U. (2012). Productivity differences in Nordic hospitals: Can we learn from Finland? *Nordic Economic Policy Review*, 2, pp. 279–315.

Simar, L. and Wilson, P.W. (1998). Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. *Management Science*, 44(1), pp. 49–61.

Simar, L. and Wilson, P.W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productive Analysis*, 13(1), pp. 49–78.

Smith, P.C. and Häkkinen U. (2007). Information strategies for decentralization, in Saltman, R.B., Bankauskaite, V. and Vrangbæk, K. (Eds), *Decentralization in Health care*. Berkshire, Open University Press, pp. 206–224.

Street, A and, Häkkinen U. (2009). Health system productivity and efficiency, in Smith, P.C., Mossialos, E., Papanicolas, I. and Leatherman, S. (Eds.), *Performance measurement for health system Improvement. Experiences, challenges and prospects*. Cambridge, Cambridge University Press, pp. 222–248.

Sund, R and Häkkinen U. (2016): Using registry data to compare health care efficiency, in Cylus, J., Papanicolas, I. and Smith, P.C. (Eds), *Health System Efficiency. How to make measurement matter for policy and management.* European Observatory on Health Systems and Policies, Policy Series 46, pp. 53–74.

Varabyova, Y. and Schreyögg, J. (2013): International comparisons of the technical efficiency of the hospital sector: Panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy* 112, pp. 70– 79.