HEALTH SERVICES RESEARCH RELATED TO BENCHMARKING AND PERFORMANCE INDICATORS

A discussion paper

KLAZINGA Niek, FISCHER Claudia, PLOCHG Thomas

Corresponding author: Klazinga Niek n.s.klazinga@amc.uva.nl

Department of Social Medicine, Academic Medical Center University of Amsterdam

March 16, 2010
Disclaimer: The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement n°223248. Sole responsibility lies with the authors and the European Commission is not responsible for any use that may be made of the information contained therein.
Introduction

Benchmarking and performance indicators have become key themes in health care policy in European countries. Accountability and transparency are central notions in today’s ideas about health system governance and stewardship (EU, OECD). Whether looking towards health care from an economic perspective, a public health perspective, a medical perspective or a legal perspective; in all cases the actual performance of health care services and systems needs to be made explicit and used as the basis for optimization strategies. This assumes that the performance of health care services (meso level) and health care systems (macro level) should be measured by valid, reliable and relevant performance indicators (performance measurement) and that this information is interpreted correctly to assess the relative position of health services or health care systems towards each other (benchmarking) thus resulting in appropriate decision making to improve and optimize the outcome of health services and health systems. Measuring and managing performance is thus not only essential for policies that aim to improve the quality of health care (defined in this paper as effectiveness, safety and patient-centeredness) but is also a pre-condition for policies focusing on increasing efficiency and value creation.

Health Services Research in the countries in the European Union is related to the enrolling agenda of benchmarking and performance indicators in many different ways. The emphasis seems to be on performance measurement; trying to validate measures that tell us something about the performance of specific services or health care systems. These lines of research are closely linked to available health care statistics. Other lines of HSR focus on the actual embedding of performance measurement in policy making and health care management. These lines of research are closely linked to the wider agenda of implementation research in health care and the application of theories and methods developed in other industrial and public sectors to the health care sector.

The aim of this paper is not to give a complete overview of all ongoing Health Services Research activities in Europe related to benchmarking and performance indicators. It rather tries to identify the main themes and focuses on the opportunities to improve the HSR evidence base behind the policy developments in the field of benchmarking and performance indicators.
Two biases in this paper should be noted beforehand:

- when operationalizing performance the emphasis will be on quality (effectiveness, safety and patient centeredness) and, although through a lesser extent, on costs. These two domains, quality and cost, were chosen because these are the terms used in the definition on Health Services Research that is used in the EU project for which this paper is written and they cover the area of expertise of the lead author. As a consequence the paper does not address performance related to equity or access to health care.

- Emphasis will be put on the possibilities to improve the information infrastructure for measuring performance. Mortality statistics, specific registries, administrative databases, electronic health records and systematic population and patient surveys are essential data sources for indicators. National governments and international organizations such as the EU can play an important role in assuring that these data sources are actually fit for the performance measurement they want. The optimization of national information infrastructures is therefore seen as an important theme that can be tackled in the context of an EU conference on the future of HSR research related to health policy in Europe. As a consequence this paper does not address specific methodological issues related to study design and methods for analyses in quantitative and qualitative HSR that seem more appropriate for a scientific HSR conference.

The author hopes that readers of this paper will provide feedback on the overall analyses, will complement it with areas and arguments that are missed and add examples of HSR from their own country. This paper will form the input for the session on benchmarking and performance indicators during the HSR conference April 8-9 in The Hague and should finally, through discussion and feedback, result in a series of recommendations on the future of HSR related to benchmarking and performance indicators in Europe.

**Which type of research questions are asked in HSR related to benchmarking and performance indicators?**

In the Dutch handbook on Health Services Research, based on the work of AD de Groot, five types of questions are identified in Health Services Research; descriptive questions, explorative questions, instrument building related questions, hypotheses testing questions and theoretical/interpretative questions. When we look at HSR related to benchmarking and performance indicators it seems fair to state that at present a lot of the work in Europe can
be grouped under the heading of “research to build instruments cq measures”. Performance indicators are in essence measures and instruments for measurement. Thus research on performance indicators often addresses issues as validity (does this indicator actually measures what it is supposed to measure), reliability (quality of data-sources and thoroughness of data-collection methods) and relevance (usefulness for managers and policy makers). Especially when dealing with outcome indicators on effectiveness, research tends to focus on the need for case-mix –adjustments and ways how to present the indicator information in an easy to interpret way to policy makers (i.e. funnel plots, scoring systems). Descriptive research also takes place when performance in a specific area still needs to be operationalized. Deciding on concepts, definitions and categories forms an important part of this work. Different HSR methods are used to develop common descriptions of quality of care for example focus groups, nominal group techniques and concept mapping. An important lesson here is that for performance indicators to be useful the subjects being measured and/or the subjects being the potential users of the indicators should be actively involved in the process of indicator development. In all countries descriptive research efforts seem to be going on to develop commonly agreed on sets of performance indicators.

As part of the validation of indicator sets, research is also exploring the relations between measures on structure, process and outcome in health care. The type of HSR questions addressed here could be labeled as explorative.

HSR really focusing on hypotheses testing in the areas of performance indicators and benchmarking is rarer. It usually related to the evaluation of the effectiveness of specific strategies in which the use of indicators is embedded and thus related more to the overall area of implementation research in health care.

Theoretical interpretative work is also going on focusing mainly on the various concepts behind operationalizing (sets of) performance indicators and reflection on the normative and ethical aspects of trying to govern health care services and systems through performance measures.

In short, most HSR work related on performance indicators and benchmarking seems to address instrumental, descriptive and explorative research questions. Hypotheses testing research questions are less common.

**Which kind of research methods are used in HSR related to benchmarking and performance indicators?**

As a consequences of the type of research questions that should be addresses the appropriate methodologies focus on systematic group processes to identify relevant themes,
reaching agreement on definitions, development of data collection instruments (surveys, strategies to derive specific data-sets from existing administrative data bases, registries or medical records), statistical analyses to validate indicators and establish their discriminative power, determine the necessity of case-mix adjusters, exploring relations between various sets of indicators (structure, process, outcome) and evaluation of the effectiveness of strategies to use indicators for performance improvement (experimental with or without control groups). The competences to do this type of Health Services Research are quite distinct from the research competences needed for bio-medical research. The methodological research qualities needed for this type of research asks for a mix of (clinical) epidemiological and social-sciences skills. Furthermore a high sensitivity for application of findings in practice is needed as performance indicators and benchmarking should primarily be owned by the users, manager and policy makers, themselves. HSR can help them to develop the tools but health Services Researchers in this area should be aware of the various interests of parties in the outcome of their work. Thus the need to integrate the opinions and wishes of the subjects of measurement and the subjects who are going to use the measurement, in the research activities, this sensitivity for the use of research results in practice is a generic competence that all HSRers should have.

**Performance indicators and benchmarking related to mortality data**

Historically death statistics have been the basis for assessment of the performance of health care systems. Life expectancy and perinatal death are in many international reports still used to compare the relative performance of health care systems (WHO, Worldbank) and several adjustments have been made to refine the measures (for example the use of Disability Adjusted Life Years).

With respect to performance indicators and benchmarking the following two research lines are interesting to note.

- After initial research in the eighties (Holland et al) recently avoidable mortality has become again a topic of interest. Following research done by Nolte and Mc Kee, at present an EU funded project (AMIEHS) is looking again at the avoidable mortality lists and refining the measurement methodology and the international comparability. Several European countries seen to be using this method for comparison of performance within their country and the OECD has recently started to look into the opportunities of reporting on avoidable mortality rates as part of their international comparative work.
A second line of mortality statistics based research is the growing popularity of the Hospital Standardized Mortality rates. Initially developed by Jarman, this methodology to compare the performance of hospitals based on adjusted hospital mortality rates has been taken up in a series of countries. Methodological debates are focusing on the (lack of) international comparability of these data given the differences in hospital systems and recording. However, policy interest is substantial.

Possibilities to use mortality statistics for performance indicators and benchmarking seem to be hampered by still existing differences in coding practices around causes of death but also in the varying possibilities in countries to link for example mortality statistics from hospital administrative data bases (in hospital mortality) with the overall death statistics. Possibilities for using Unique Patient Identifiers and linkage of data-bases are the key for making further progress in this area.

**Performance indicators and benchmarking related to cancer care**

Apart from becoming the major cause of death in EU countries, cancer is also the area were relative complete statistics are available for the performance of cancer care for various types of cancer such as breast cancer, cervical cancer and colon cancer. Quite often these “outcomes” are related to discussions to have national screening programs when valid screening methods are available. Recent work of the CONCORD group has reported on cancer world-wide (Coleman) but also the EU (Eurocare) and the OECD are active in improving the measures for the international comparability on the performance of cancer care. For this work to progress, the quality of the (national) cancer registries is of key importance. As long as countries don’t have cancer registries that cover their whole population, valid performance data are difficult to produce. Furthermore for cancer data to be useful for benchmarking, apart from mortality data the cancer staging data are essential. And, as with mortality data, linkages between cancer registries and administrative-data bases such as on hospitals are essential to do meaningful research on the relation between the quality of cancer care and the use of services and resources. Although the coding practice in this field, in comparison with other disease areas has already been internationally standardized, further standardization is warranted to increase the potential for benchmarking.
Performance indicators and benchmarking on care delivered in hospitals

A lot of indicator development work is undertaken in the field of hospital indicators. Indicators are developed and tested (i.e. Spain, Italy, Portugal, France, Germany, UK, Netherlands, Belgium, Denmark). Popular categories for indicators are 30-day case fatality rates (for example in AMI and Stroke), re-admission rates, complication rates in surgery, hospital infection rates, bedsores, volume of specific treatments, waiting times, systematic measurement of patient experiences, systematic measurement of experiences of hospital staff. Most of the HSR research in this area focuses on the development and testing of indicators. Sources for indicators are mainly administrative data-bases and medical records. Generic problems identified in these projects and international comparative work of the WHO (PATH project) and the OECD (HCQI project) with respect to administrative data bases seem to be:

- Quality of coding practices for administrative data-bases (ICD9-10)
- Lack of (internationally) standardized procedure codes
- Lack of coding of secondary diagnoses
- Lack of coding whether a certain condition was present at admission
- Lack of opportunities to link the administrative data bases of individual hospital with other data bases; for example by using a Unique Patient Identifier.

Apart from the present limitations with using administrative-data bases for doing HSR on performance indicators and benchmarking, deriving the appropriate data from medical records also poses problems. Although the techniques of doing audits based on medical and nursing records have been improving, methodological flaws are still reported. The approach taken in the US by McGlynn and all in their study on the quality of care for adults holds important lessons for Europe. Furthermore, progress made with the implementation of Electronic Health Records is in most countries not developed enough to use these as a prime source for data to calculate performance indicators. Most of the problems around optimizing EHR’s for population based statistics are not technical but political. Privacy legislation and insufficient focus on standardization of data-requirements from a public information perspective seem to hinder further growth of HSR in this area. If the potential of the EHR for helping monitoring quality of care is to be fulfilled, Health Services Researchers should keep on addressing policy makers with the message that the performance data they want can only be acquired when they make sure the necessary legislation on minimal data requirements and privacy is in place. Given the fact that some countries in Northern Europe seem to have overcome these problems, there is potential for mutual learning.
Patient Safety Indicators

Patients Safety has become a major focus point of health policies over the past ten years. The EU, after the US report To Err is Human, has initiated several activities to coordinate policy development and research in this area. In 2007 a meeting was held in Porto to provide an overview of the ongoing research efforts. At present the EU funded EUNEtPass project tries to coordinate various national efforts. Also WHO, on a global scale has launched programs which also include inventories of ongoing research. Many European countries have executed studies to assess the magnitude of adverse events in their country, mostly based on detailed audit studies on medical records. Also several countries have set up national patient safety agencies that are mostly also involved in running adverse event reporting programs. Furthermore a growing body of knowledge has been created on studying safety culture, the implementation of safety systems and implementation programs for specific safety project on topics such as handovers, medication-errors and reduction of hospital infections. In the area of patient safety indicators, the example of the PSI reporting system of AHRQ in the US has been broadened to 17 other counties, including many European one’s, in the work of OECD’s Health Care Quality Indicator program.

The type of Health Services Research applied in all these efforts is not fundamentally different to the types of questions and methods discussed earlier in this paper and focuses of measurement of risks, adverse events and errors and their contexts as well as implementation research evaluating the effectiveness of interventions.

With respect to data collection similar problems as with data collection for quality indicators on hospital care can be made:

- many research is dependent on the quality of medical records
- Electronic Health Records are as yet often still insufficient as a source for the necessary data
- administrative systems often don’t have sufficient secondary diagnoses coded to calculate Patient Safety Indicators
- administrative data bases often don’t record whether relevant conditions (i.e. infections, bedsores) where present at admission
- linking with other data bases within the hospital (i.e. laboratory, pharmacy) or outside the hospital (data bases in primary care) is often not possible or not allowed.

When these data-availability problems are not solved the possibilities and impact of HSR in the area of patient safety will remain limited with respect to performance measurement and benchmarking.
Performance Indicators in Primary Care

Traditionally the design and functioning of primary care is, alongside hospital care, an important focal point of health policies. A well functioning primary care system is considered to help contain costs and improve the quality of care. HSR on the organization of primary care is discussed in a separate paper elsewhere. As a part of the assessment on HSR related to performance indicators and benchmarking, it suffices to state here that monitoring data on the quality of primary care are still relatively scarce and heterogeneous. This is partly due to the fact that the information infrastructure in primary care is often still patchy and less developed than the administrative data-bases and (electronic) health record availability in hospital care. Although there are notable examples of countries that have some part of their PHC information infrastructure more developed for governance purposes, an assessment made by the OECD in 2007 showed that national data bases were not developed and comparable enough to merit cross national collection and comparison of performance indicators on primary care. As a conclusion the OECD is at present assessing the quality of primary care by looking at the rates of avoidable hospital admissions derived from the hospital administrative data bases. Despite the limitations of the data systems, a lot of HSR is going on in primary care, often focusing on specific diseases (especially chronic diseases) and the related care arrangements (i.e. disease management) or implementation of guidelines. Reported benchmarking studies in primary care seem to be rare.

Patient Experiences

The systematic measurement of patient experiences has become a fundamental element of assessing the performance of health care services and systems. Health Services Research in this area consists of methods to determine the domains and topics that are considered to be important for performance (methods such as focus groups, interviews, concept mapping) and all types of questionnaire development and testing. Instrumental research on valid methods to assess and report on patient experiences seems to be the core of the HSR in this area. After the example of the US (CAHPS) and UK (Picker) a growing number of European countries is standardizing and institutionalizing the systematic measurement of patient experiences. Apart from the validation of patient questionnaires, work has been reported on population based surveys to measure the experience and opinions of citizens on health care (i.e. Eurobarometer, Common Wealth Fund surveys, WHO). Apart from the ongoing validation of sets of questions, work is going on the use of vignettes to capture opinions. Overall this seems a fruitful area for further exchange of instruments and methods on a
European level as this area of performance measurement and management is at the core of the EU values to create transparency in the health care markets and strengthen the position of the health care consumer.

**Research on concepts and performance frameworks**

Performance indicators are never a goal in itself but derive their meaning from the management cycle and context they are part of. As a consequence one should be very careful to assume that an indicator that is useful for one goal (selective contracting, public accountability) can also be useful for another goal (internal quality improvement). As such, performance indicators are not, as in bio-medical research, aiming for universal truths, but for truths within a specific context and setting from a specific goal. This also explains why it is often difficult to get HSR studies published in mainstream journals as they are often considered as less relevant for a broad international audience, or, erroneously so, considered less scientific because of the (intentionally created) limited generalisability.

Given this dependency on context and goals, the work of performance measurement has become embedded in a growing body of research work on concepts and frameworks used for performance measurement and benchmarking. One obvious line of research here is linking specific quality aspects to specific sets of indicators (what is a balanced set or bundle), another line of research looks at how sets of indicators can be grouped in broader conceptual frameworks to keep an overview over the performance on all aspects and being able to derive strategical meaning from developments on performance on various domains and aspects at the same time. Frameworks from industry, such as the Balanced Score Card are taken as examples. This type of frameworks can be found for assessing the performance of whole health care systems and quality of health care (i.e. OECD, WHO) but also for individual services such as hospitals (PATH/WHO). Theoretical and explorative HSR can underpin the policy need to develop and test conceptual frameworks that are consistent with the management and governance goals policy makers have with using performance indicators. Strengthening this line of thought and research remains necessary to enhance appropriate embedding of performance indicators and prevent dysfunctional use and bureaucratization.

A specific area of research seems to be the linkage between quality and costs. As two domains of health services and health systems performance they are most of the time considered separately. The measurement of costs has his own processes of standardization, often materialized in (national) accounting schemes and international efforts such as the
systems of health accounts work of OECD, EU and WHO. Linking quality and costs as part
of overall performance management seems less common although work in this area has
been reported (Ontario). On health services levels quality/cost methods from industry have
been applied (Custers et al) but were in the reported cases not congruent with the external
financial incentives; i.e. in many cases there does not seem to be a business case for quality.
This notion has already been recognized by policy makers and performance indicators on
quality are at present increasingly used in Pay for Performance schemes and (selective)
contracting. Evaluation research on these policy interventions, although quite prevalent in
North America, is in Europe still limited. It should, however, be a promising area for future
HSR given the increased interest by policy makers and health care managers. Thus the
present more theoretical and explorative work might be extended to hypotheses testing
research on experiments with financial incentives and contracting.

Research on the practice of benchmarking and performance improvement

Although HSR specifically focusing on the technique and methods of industrial benchmarking
in health care is still rare, the body of evidence on implementation and innovation in health
care is substantial. It seems wise to take the implementation of strategies to use
performance indicators not as a separate research topic, but consider the use of indicators
as one strategy alongside others to improve the quality of health care. The EU funded
Marquis Project demonstrated in a group of 489 hospitals the synergy between strategies to
use indicators, measure patient experiences, perform audits, clinical guidelines, patient
safety systems and TQM. HSR on benchmarking and the use of performance indicators
could therefore borrow and become complementary to the already existing theories and
study results on quality improvement.

Pathway to strengthen this paper

This paper aimed to give an overview of Health Services Research related to performance
indicators and benchmarking. As stated at the beginning the paper does not aim to be
complete in describing all ongoing research but rather provide a bird’s eye overview on the
various main research lines and the associated methodological and policy debates. During
the planned European HSR conference April 8 and 9 2010 a session will be held on HSR
research related to performance indicators and benchmarking. During this session, based on
this paper, conclusions will be drawn and recommendations formulated for Health Services
Researchers and policy makers. It is anticipated that the following potential recommendations will be discussed.

1. Health Services Research on benchmarking and performance indicators addresses both questions on the development and the use of benchmarks and indicators.

2. Health Services Research on the development of measures focuses on validity, reliability and relevance; HSR on the use focuses on the effective embedding in policy and management on health system and health services level.

3. Given the applied nature of HSR on benchmarking and performance indicators active involvement of the potential users is essential.

4. HSR on benchmarking and performance indicators needs combined research capacities from the bio-medical and the social-sciences.

5. Further progress of HSR on benchmarking and performance indicators is hampered by data-availability. The following issues need addressing:
   - use of Unique Patient Identifiers to facilitate linkages between separate data-bases
   - further standardization of coding
   - use of present-at admission codes in administrative databases
   - recording of secondary diagnoses in administrative data-bases
   - facilitate secondary data use from Electronic Health Records
   - facilitate standardized measurement of experiences by patients and citizens

6. HSR research on benchmarking and performance indicators on European level would benefit from strengthening the clearinghouse function on research findings, training of researchers and appropriate scientific publication media.

7. Results of HSR research on benchmarking and performance indicators should be systematically shared with policy makers and managers of health services and systems.

8. Networking should be stimulated on European level between the research groups involved in this kind of work and the growing number of national/regional institutes involved in quality measurement and reporting.

9. The 8th framework programme should address the topic of research on benchmarking and performance indicators as one of the research items.