ECHO: health care performance assessment in several European health systems

E. Bernal-Delgado1,2, T. Christiansen3, K. Bloor4, C. Mateus5, A.M. Yazbeck6, J. Munck7, J. Bremner8*, on behalf of ECHO Consortium

1 Institute for Health Sciences in Aragon, IIS Aragon, Zaragoza, Spain
2 Red de Investigación en Servicios de Salud en Enfermedades Crónicas (REDISSEC), Spain
3 Centre of Health Economics Research, University of Southern Denmark, Denmark
4 Department of Health Sciences, University of York, England, UK
5 Escola Nacional de Saúde Pública, Universidade Nova de Lisboa, Portugal
6 Institute of Public Health, Ljubljana, Slovenia
7 UMIT, Innsbruck, Austria
8 European Health Management Association, Brussels, Belgium

Correspondence: Enrique Bernal-Delgado, Institute for Health Sciences in Aragon, IIS Aragon, San Juan Bosco 13, 50009 Zaragoza, Spain, Tel: +34976716792, e-mail: ebernal.iacs@aragon.es.

*The members of the ECHO Consortium are listed in the Acknowledgement.

Background: Strengthening health-care effectiveness, increasing accessibility and improving resilience are key goals in the upcoming European Union health-care agenda. European Collaboration for Health-Care Optimization (ECHO), an international research project on health-care performance assessment funded by the seventh framework programme, has provided evidence and methodology to allow the attainment of those goals. This article aims at describing ECHO, analysing its main instruments and discussing some of the ECHO policy implications. Methods: Using patient-level administrative data, a series of observational studies (ecological and cross-section with associated time-series analyses) were conducted to analyze population and patients’ exposure to health care. Operationally, several performance dimensions such as health-care inequalities, quality, safety and efficiency were analyzed using a set of validated indicators. The main instruments in ECHO were: (i) building a homogeneous data infrastructure; (ii) constructing coding crosswalks to allow comparisons between countries; (iii) making geographical units of analysis comparable; and (iv) allowing comparisons through the use of common benchmarks. Conclusion: ECHO has provided some innovations in international comparisons of health-care performance, mainly derived from the massive pooling of patient-level data and thus: (i) has expanded the usual approach based on average figures, providing insight into within and across country variation at various meaningful policy levels, (ii) the important effort made on data homogenization has increased comparability, increasing stakeholders’ reliance on data and improving the acceptance of findings and (iii) has been able to provide more flexible and reliable benchmarking, allowing stakeholders to make critical use of the evidence.

Introduction

In a recent communication, the European Commission set three pillars upon which the European Union (EU) health-care agenda should be built. Those cornerstones are: strengthening health-care effectiveness, increasing accessibility and improving resilience. The communication bestows a critical role, among other strategies, on health care. Operationally, several performance dimensions such as health-care inequalities, quality, safety and efficiency were analyzed using a set of validated indicators. The main instruments in ECHO were: (i) building a homogeneous data infrastructure; (ii) constructing coding crosswalks to allow comparisons between countries; (iii) making geographical units of analysis comparable; and (iv) allowing comparisons through the use of common benchmarks. Conclusion: ECHO has provided some innovations in international comparisons of health-care performance, mainly derived from the massive pooling of patient-level data and thus: (i) has expanded the usual approach based on average figures, providing insight into within and across country variation at various meaningful policy levels, (ii) the important effort made on data homogenization has increased comparability, increasing stakeholders’ reliance on data and improving the acceptance of findings and (iii) has been able to provide more flexible and reliable benchmarking, allowing stakeholders to make critical use of the evidence.

Methodology

ECHO has used two complementary methodological approaches, geographical and hospital specific. The first one, stemming from the tradition of Small Area Analysis, tries to delimit whether the place of residence determines the population’s exposure to health systems performance (an extensive review of those research projects can be found at http://www.euroreach.net/compendium). The European Collaboration of Health-Care Optimization (ECHO) project was among those research initiatives.

ECHO is a demonstration project aimed at analysing healthcare performance within and across several European countries. Stemming from the numerous scientific achievements of the Dartmouth Atlas of Healthcare and the Spanish Atlas of Variations in Medical Practice,11,12 ECHO has used the well-accepted OECD performance framework,13 studying several of its performance dimensions (i.e. utilization of and equitable access to effective care, quality and safety, and efficiency in terms both of opportunity costs and provider-level efficiency) and using indicators developed or validated by various of the aforementioned initiatives.4,12 This scientific orientation, rather than focusing on ‘average differences’, aims to analyze systematic variations in performance. This article aims at framing ECHO’s methodology, analysing its main instruments and discussing some of the ECHO policy implications.
care. The second one, in turn, is interested in patients’ hospital care experience and seeks to attribute patients’ outcomes to the hospital where they were treated or operated on.

In terms of the study design, ECHO has been conducted as an observational study. In the case of the geographical approach, as an ecological study using either cross-section or time series analyses, whereas in the case of the hospital-specific approach as a cross-sectional study with or without time series analyses. Table 1 provides further detail on both approaches.

The study population at ECHO has virtually been all public hospital discharges produced in the ECHO countries (Austria, Denmark, England, Portugal, Slovenia and Spain) from 2002 to 2009, which has entailed the analysis of 191,136,051 discharges.

Each one of those discharges was allocated to either a health-care unit (geographical perspective) or to a hospital (hospital-specific approach). Therefore, considering the smallest unit of analysis, ECHO has represented the experience of the population living in 922 health care areas (9 ‘Länder’ in Austria, 98 ‘Kommuner’ in Denmark, 326 Health Authorities in England, 278 ‘Concelhos’ in Portugal, 12 Statistical Regions in Slovenia and 199 Health Care Areas in Spain) as well as patients’ exposure to 850 hospitals (169 in Austria, 66 in Denmark, 220 in England, 59 in Portugal, 19 in Slovenia and 317 in Spain).

**ECHO instruments**

Conceived as a demonstration project, ECHO set about the task of bringing together routinely collected patient-level administrative data from Austria, Denmark, England, Portugal, Slovenia and Spain, making them comparable. Key instrumental issues in ECHO have been: (i) building a homogeneous data infrastructure; (ii) constructing coding crosswalks to allow comparisons between countries; (iii) making units of analysis comparable in the geographical analysis and (iv) allowing comparisons through the use of common benchmarks.

**Building a homogeneous data infrastructure**

A major concern when building a single data infrastructure from different country datasets is figuring out how to offer comparable information. In ECHO, three main sources of heterogeneity in the original datasets had to be considered beforehand: (i) whether the elemental unit of information (EUI) (e.g. discharge) represents the same construct; (ii) whether the information contained within each one of those elemental units is similar and (iii) whether each EUI could be allocated to the different units of analysis (geographical areas or hospitals) of interest for each country.

With regard to the EUI, Denmark and England record more disaggregated information than the other countries; accordingly, in Denmark several contacts compose an episode, while in England spells may be composed of several episodes. In the other countries this differentiation does not exist because they just record discharges. A valid common construct for all the countries was to consider ‘episodes’ as the elemental unit for Denmark, ‘spells’ as the elemental unit for England, and ‘discharges’ as the elemental unit for the remaining countries.

With regard to the second concern, whether the EUI contained similar information across datasets, this might affect the construction of those performance indicators of interest. Each original dataset was analyzed to check feasibility by using the common information metatile in Supplementary Appendix S1. Except for some information gaps affecting Slovenia in the first years of the series (2002–2005), a change in the ‘transfers out’ definition in Denmark in 2006, and the inexistence of new-born records (different from mothers’) in all countries, ECHO was able to study the vast majority of those projected performance indicators.14,15

Finally, geographical analyses require each EUI to be allocated to spatial units. The concern with regard to this point is that different effectiveness in the allocation across countries could hamper comparability. Fortunately, numbers ensured quite fair comparability, with figures of effective allocation ranging between 100% in Portugal and 97.7% in England.

**Construction of coding crosswalks**

ECHO countries use different coding languages in the description of main diagnoses, secondary diagnoses and procedures within a EUI. So, Spain and Portugal uses the ICD 9th for both diagnoses and procedures. All the remaining countries use the ICD 10th in the description of diagnoses. More variety appears in the description of procedures where Denmark uses NOMESCO, England OPCS, Slovenia ACHI and Austria the Leistung-Kataloge-BMG.

The process was not a code-to-code translation but a conceptual approach where the rationale and the ICD 9th definition of those previously validated indicators were taken as a reference.4,12,13 Once a first version was developed, a face and empirical validation process was run before accepting the final definitions. For that purpose, a two-round validation was carried out by a local team providing feedback on local coding practices, misused or missing codes, and on whether the rough figures obtained using that definition (i.e. rates or incidence) matched with local knowledge. Detail on the crosswalks can be found in http://www.echo-health.eu/handbook/getting-indicators.html.15

**Developing comparable units in the geographical analysis**

ECHO represents populations’ exposure to health care by using geographical units of analysis (areas with some meaning in terms of policy making, planning or health-care provision). However, the original units of analysis are extraordinarily heterogeneous. Austria has nine ‘Länder’ for a population of 8.5 million inhabitants, Denmark is divided into 98 ‘Kommuner’ for a population of 5.6 million individuals, England is composed of 326 Health Authorities for a population of 53 million inhabitants, Portugal is composed of 234 ‘Concelhos’ for a population of 10.5 million individuals, Slovenia has 12 Statistical Regions for a population of barely 2 million and Spain is composed of 199 Health Care Areas for a

<table>
<thead>
<tr>
<th>Table 1 Methodological approaches in ECHO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geographical approach</strong></td>
</tr>
<tr>
<td>Research question</td>
</tr>
<tr>
<td>Main endpoint</td>
</tr>
<tr>
<td>Denominator</td>
</tr>
<tr>
<td>Analysis</td>
</tr>
</tbody>
</table>
population of 46.5 million inhabitants. This heterogeneity could lead to the misinterpretation of variations in a country’s performance, attributing to providers what is in fact noise sensitive to differences in population size.

Consequently, a third instrumental key in ECHO has been the reconstruction of map boundaries in those cases where population size was more heterogeneous—Denmark, Portugal and England. Specifically, cross-country comparisons of geographically based performance indicators have used those new, more homogeneous, areas. Details of the methodology can be found in http://www.echo-health.eu/handbook/unit_analysis.html.17

Proper benchmarking
Proper benchmarking depends on whether it is possible to estimate the true fraction of variance that would remain significant once population or patients’ differences were controlled. If that were the case, it would be safe to flag those providers that behave differently to the expectation (i.e. benchmark). Two general methodologies are used in ECHO to allow sound comparisons between providers: in the case of geographical analysis, rate standardizations; and in the case of hospital-specific analysis, multilevel modelling using several risk-adjusters.

With respect to geographical analyses, age and sex are generally used as predictors of risk at population level by standardizing rates using either direct or indirect methods.18 Nevertheless, in some specific indicators other predictors can also be used to improve risk adjustment validity. For example, the burden of ischaemic disease could be used when studying the differences in utilization of coronary interventions, or the use of quintiles of socioeconomic deprivation could be used as a proxy for need in, for example, access to mastectomy or colectomy in cancer.

With regard to hospital-specific analyses, risk difference across patients is adjusted using multilevel modelling—linear when studying continuous variables (e.g. length of stay) or ‘logit’ when studying dichotomic events (e.g. in-hospital mortality).19 Table 2 details those risk factors included in the models.

The massive pooling of data permits the building of ad hoc benchmarks, allowing comparison between any provider (i.e. area, hospital, region or country) with a reference built within the same country, a benchmark estimated on the experience of several providers, or an aspirational target using the best provider in the whole sample as a reference.

Yet in some hospital-specific indicators, the limited nature of the clinical information registered in administrative datasets may jeopardize risk adjustment or might be insufficient to control phenomena such as the patient selection bias. Depending on the indicator of interest, several strategies could be deployed to reduce both threats to validity. For example, when studying percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG) mortality, confining comparisons to only high-volume hospitals, given the use of all available discharges in several years, would mitigate risk adjustment failures or, using PCI/CABG rates in the same hospital in the preceding years would act as a good instrument to palliate selection bias.

As a final comment with regard to risk adjustment quality, a differential coding intensity (i.e. the number of codes in each EUI) might challenge comparability between hospitals across countries. As an example, the number of secondary diagnoses—those used in the Elixhauser comorbidity index in table 2—in patients undergoing CABG in ECHO is very similar in England, Portugal, Slovenia and Spain, but fairly different in Denmark—in the former, 45–50% of cases had two or three comorbidities vs. 24% in Denmark. In those cases where differences are important, ECHO performs sensitivity analyses in order to figure out whether this factor affects the relative position of hospitals with regard to benchmarks.

Implications on policy making
ECHO provides numerous policy-oriented findings that could be used to strengthen health systems’ effectiveness while improving their sustainability. As examples from among those findings, ECHO has highlighted the following: (i) at population level, the burden of coronary ischaemic disease barely explains variation in exposure to revascularization procedures, with some areas likely to be over-exposed and others likely to be underserved; (ii) orthopaedic procedures are performed differently across income quintiles, not always coherent with social gradient, causing concern about effective access to those procedures; (iii) case-fatality rates vary dramatically across high-volume hospitals, irrespective of the differences in patient case-mix, pointing to health-care institutions’ responsibility for those results; (iv) variation remains high and stable over time in potentially avoidable admissions in chronic diseases regardless of the increasing number of chronic care policies across Europe; (v) variation in lower-value procedures is huge, within and across countries, implying unacceptable opportunity costs in those areas performing above the reference pattern.

Although these (and more) messages are yielded using a robust methodology, ECHO is an observational study built upon

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elixhauser Comorbidities</td>
<td>Cardiac arrhythmias; valve disease; congestive heart failure; chronic lung disease; hypertension; uncomplicated; hypertension, complicated; pulmonary circulation disorders; renal failure; pre-existing hypertension complicating pregnancy; other hypertension in pregnancy; diabetes; without chronic complications; diabetes, with chronic complication; hypothyroidism; liver disease; obesity; alcohol abuse; drugs abuse; lymphoma; solid tumour without metastasis; metastatic cancer; weight loss; psychoses; depression; AIDS/HIV; fluid and electrolyte disorders; peptic ulcer disease excluding bleeding; deficiency anaemia; blood loss anaemia; coagulopathy; rheumatoid arthritis/collagen vascular diseases; peripheral vascular disorders; paralysis; other neurological disorders.</td>
<td>Depending on the case, age is used as a continuous variable or included as age-groups strata</td>
</tr>
<tr>
<td>Severity</td>
<td>Primary diagnosis of Myocardial Infarction; ST elevation vs. Non ST elevation; Concomitant structural cardiac surgery.</td>
<td>Variables that are used either as a number of comorbidities or as individual dichotomic factors. Usually not all comorbidities are considered in the model but those that theoretically might be related to the exposure and to the outcome of interest.</td>
</tr>
<tr>
<td>Severity</td>
<td>Emergent vs. non-emergency admission; number of secondary diagnoses.</td>
<td>Specific for cardiovascular events</td>
</tr>
</tbody>
</table>

Unspecific, not used primarily, but as part of sensitivity analyses.
administrative data. Therefore, most of the time ECHO exhibits serious limitations when making inferences about a provider’s responsibility for performance outcomes. In other words, rather than a firm diagnosis, ECHO should be used as a screening tool, meant to monitor health-care providers’ behaviour with regard to a benchmark, and over time. In any case, providers beyond the expectation will always require further examination.

Moreover, because there is no general causal model explaining performance variation, and underlying factors may be different or affect variation across procedures and contexts differently, ECHO monitoring results have to be interpreted in country-specific institutional contexts, particularly local policies, organizational conditions and incentive structures.

Nevertheless, beyond the required wariness in the interpretation of ECHO results, a main purpose of ECHO is to report findings in a way that stakeholders can easily understand and use. Thus, ECHO reports systematically: the magnitude of the phenomenon (e.g. rate maps and dotplots), its persistence over time (e.g. evolution graphs), to what extent the observation is systematic (e.g. maps exhibiting standardized utilization ratios), if the phenomenon could reasonably be attributed to a particular provider (e.g. rho statistic, or median odds ratio), the impact of the need (i.e. evolution of rates by social gradient quintiles) and whether a finding deserves further investigation at local level (e.g. excess cases maps and funnel plots). The panoply of ECHO statistics, maps and graphs can be found at http://www.echo-health.eu/handbook/metrics.html.\textsuperscript{21,22}

Conclusion

ECHO has produced some innovations in international comparisons of health-care performance, mainly derived from the massive pooling of patient-level data; thus: (i) ECHO has expanded the usual approach based on average figures (i.e. scarcely informative) providing insight into within and across-country variation at various meaningful policy levels, (ii) the emphasis on data homogenization has increased comparability, increasing a stakeholder’s reliance on data and improving the acceptance of findings and, (iii) ECHO has been able to provide more flexible and reliable benchmarking, allowing stakeholders critical use of the evidence.

Supplementary data

Supplementary data are available at EURPUB online.

Acknowledgements

ECHO Consortium: IACS’s team (E. Bernal-Delgado, S. García-Armesto, N. Martínez, M. Seral, F. Estupiñán, M. Comendeiro, E. Angulo-Pueyo, M. Ridao and C. Baixaulí). Libero J (as affiliated researchers), University of Southern Denmark’s team (T. Christiansen, L.C. Thygesen), University of York’s team (K. Bloor, R. Cookson and N. Gutacker), University Nova de Lisboa’s team (C. Mateus, C. Nunes and L. Joaquim), Institute of Public Health of Iñigo’s team (A.M. Yazbeck, M. Galsworth and T. Albrecht), UMIT’s team (J. Munck and B. Güntert) and EHMA’s team (J. Bremer, P. Giepmans and O. Dix).

Funding

This work was supported by the FP-7 Programme of the EU with Grant HEALTH-F3-2010-242189 for ECHO’ project.\textsuperscript{23}

Conflicts of interest: None declared.

Key points

- European Collaboration for Healthcare Optimization has provided methodological foundation for the use of routinely collected data in the international comparison of health-care equity, quality, safety and efficiency.
- In order to increase comparability and stakeholders’ reliance on data and findings, harmonizing performance concepts and datasets is key.
- Pooling patient-level data allocated to geographic areas or care providers, into a single dataset, allows the development of comparable performance indicators, as well as more flexible and reliable benchmarking.
- International health care performance assessment should replace comparisons based on average figures, with the insight provided by within and across country variation, at various meaningful policy levels.

References

3 Institute for Clinical Evaluative Sciences (ICES), Canada. Available at: http://www.ices.on.ca/Publications/Atlases-and-Reports (10 January 2015, date last accessed).
5 NHS Right Care, UK. Available at: http://www.rightcare.nhs.uk/index.php/nhs-atlas/ (10 January 2015, date last accessed).


