Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals?

Hospitals in most European countries are paid on the basis of diagnosis related groups. Reinhard Busse and colleagues find much variation within and between systems and argue that they could be improved if countries learnt from each other.

Reinhard Busse professor1, Alexander Geissler research fellow1, Ain Aaviksoo director of health policy programme2, Francesc Cots director of management control3, Unto Häkkinen research professor5, Conrad Kobel research fellow, Céu Mateus assistant professor, Zeynep Or research director7, Jacqueline O’Reilly research analyst8, Lisbeth Serdén researcher9, Andrew Street professor of health economics10, Siok Swan Tan researcher11, Wilm Quentin research fellow1

1Department of Health Care Management, Straße des 17 Juni 135, 10623 Berlin, Germany; 2PRAXIS Centre for Policy Study, Tallinn, Estonia; 3Parc de Salut Mar, Barcelona, Spain; 4Centre for Health and Social Economics, National Institute for Health and Welfare, Helsinki, Finland; 5Department of Medical Statistics, Informatics and Health Economics, Innsbruck Medical University, Innsbruck, Austria; 6Escola Nacional de Saúde Pública, CMDT.LA, Universidade Nova de Lisboa, Lisbon, Portugal; 7Institut de Recherche et de Documentation en Economie de la Santé, Paris, France; 8Health Research and Information Division, Economic and Social Research Institute, Dublin, Ireland; 9Department of Statistics, Monitoring and Evaluation, National Board of Health and Welfare, Stockholm, Sweden; 10Centre for Health Economics, University of York, York, UK; 11Institute for Medical Technology Assessment, Erasmus Universiteit Rotterdam, Rotterdam, Netherlands

Diagnosis related groups (DRGs) were originally developed in the 1970s by researchers at Yale University. Their aim was to define “hospital products” and to measure what hospitals actually do. Medicare in the United States soon realised the potential of DRGs for paying hospitals for their work and introduced the first DRG based payment system in 1983. Since then, DRGs have spread around the world, gradually becoming the basis for paying hospitals in most industrialised countries, particularly in Europe. Even though hospital doctors in Europe are salaried rather than paid by DRGs, their clinical decisions largely determine the income of their hospitals and their specialist societies can influence the way DRG systems work. We recently reviewed the experience with DRG systems in 12 European countries, and here we compare the different systems and discuss current trends and the potential for improvement.

Why use diagnosis related groups?

European countries introduced hospital payments based on DRGs either to increase transparency or to improve efficiency. DRGs increase transparency because they condense the confusingly large number of individual patients treated by hospitals into a manageable number of clinically meaningful and economically homogeneous groups (for example, primary hip replacement in elective patients or transient ischaemic attack in patients under 70 without complications). By providing a definition for “hospital products” DRGs enable comparisons that would otherwise not be possible. For example, examining the proportion of cases in each hospital falling into more costly DRGs can help show whether one hospital treats more complex cases than another. Use of resources can be compared by assessing whether patients in the same DRGs are staying significantly longer (or are more costly) in one hospital than in another. Similarly, quality can be assessed by comparing the care processes or rate of complications for patients in a particular DRG.

DRGs are thought to improve efficiency because they provide incentives for hospitals to limit the services per patient and to treat more patients. Previously health systems had used fee for service or global budgets. Fee for service encourages hospitals to provide many services for each patient but may result in unnecessary or inappropriate treatment and a disregard of treatment costs. By contrast, global budgets contain costs by
Intended and unintended consequences

Because most countries in Europe moved to DRG based payment from global budgets, their experience differs from that in the US, where DRG based payments succeeded for service. In the US the introduction of DRG based payments initially led to a reduction in hospital activity; the effect (as intended) in Europe was an increase in activity, particularly in day care. For example, in England between 2003 and 2007, day case activity increased by about 15% while total NHS inpatient activity increased by 10%. While in the US DRG based payment helped contain costs, the increase in activity in Europe mostly led to higher hospital costs. For example, in Austria between 1997 and 2007, total hospital costs increased by 3.8% annually, while costs per case increased by 1.7%. Whether increased activity and reduced costs per case led to improved efficiency in European hospitals and whether this was the direct effect of DRG based payment is difficult to say.

DRG based hospital payment may have unintended consequences if the effects are too strong. For example, DRG based payments are intended to reduce length of stay and cost of treatment, but an excessive reduction in length of stay may reduce the quality of care. Research in Europe found little change in death rates and readmissions when DRG based hospital payments were introduced. However, in France 30 day readmission rates after discharge seem to have increased since the introduction of DRG based payment, and a study from Sweden showed that patient perceived quality of care decreased.

Other potential unintended consequences of DRG based hospital payment include cherry picking, dumping, upcoding, overtreatment, and frequent readmissions. Cherry picking occurs if certain patients within one group are systematically more costly than others, leading to incentives for hospitals to select the less costly, more profitable cases and to transfer or avoid the unprofitable ones (“dumping”). Upcoding refers to hospitals increasing their revenue by coding additional diagnoses to move patients into a higher paying group. Furthermore, hospitals may even change their practice patterns, providing procedures that place patients in higher paying groups (such as treating patients with acute myocardial infarction with drug eluting stents instead of bare metal stents); or they could admit or readmit patients to hospital for unnecessary services or for conditions that could be treated in outpatients (currently few countries use DRG based payments for outpatient activity and inpatient treatment may be more profitable).

Our research in Europe suggests that these unintended consequences are relatively rare—or at least that they have been detected rarely despite (or because of) regular audits by monitoring and review bodies. Evidence of cherry picking is available only for England and France, where private providers have been found to treat less complex patients than public hospitals. Intentional upcoding and overtreatment are substantial problems in France and Germany (box 1) but seem to be uncommon elsewhere.

Several countries have introduced policies to try to prevent frequent readmissions and overtreatment. For example, in Germany and England, hospitals are—under certain conditions—not paid for readmissions within 30 days from the initial admission or discharge. Furthermore, to avoid an excessive increase in the number of admissions, hospitals in England in 2013 receive only 30% of the full tariff for emergency admissions that exceed the number of admissions in the financial year 2008-09. Similarly, in Germany, total hospital activity is limited by negotiated target budgets and hospitals can keep only 35% of the revenue earned for activity provided in excess of the budget.

Similar problems, different solutions

Most unintended consequences can be avoided or reduced if the DRG system uses high quality data on costs and is good at creating groups of patients that have similar costs and cannot be easily manipulated. If differences between patients within the same groups are not adequately controlled for, the payment is either too high or too low for many patients, creating strong incentives to game the system. In theory, the more groups created, the more homogenous they will be, but a balance has to be struck between resource homogeneity and having a manageable number of groups for comparative and payment purposes.

Table 1 shows that in most countries the number of groups has increased considerably in recent years. Between 2005 and 2011 the number of groups in the German system increased by almost 40%, the number in England more than doubled, and the number in France more than tripled. By contrast, the number of groups in the Netherlands, which introduced its first classification system in 2005, fell from an extremely high number of about 100 000 groups to about 4000 in the latest version of the system. Under the previous versions of the system, identical patients were sometimes put into different groups, depending on the specialty that treated them. The smaller number of groups makes it easier to keep the system up to date and to reliably calculate average costs.

Large differences remain between countries in the number of groups. Interestingly, countries with a high number of groups for one condition do not necessarily have a high number for all conditions. For example, England has two groups for stroke but 14 for hip replacement. In Germany, where patients with stroke are classified into 10 groups, stroke unit care, systemic thrombolysis, intracranial haemorrhage, and death, among other things, are taken into account, whereas in 2008 England distinguished between only cerebral infarction and intracranial haemorrhage, although stroke unit care has since been incorporated into the standard tariff. For acute myocardial infarction, eight countries differentiate between patients with and without certain complications and comorbidities (such as diabetes and heart failure) but three countries (Austria, England, and the Netherlands) do not.

The reasons why DRG systems were established in the way they were—and the reasons for revising them—are rarely made public, and the processes are not completely objective. Usually, the process relies on input from medical specialist associations or expert consultants and the decisions are made by a national authority. Some have been developed from previous systems that classified groups of procedures (Austria, England, and Poland), and the systems in these countries still have a different grouping logic from that used in other countries. Other
We analysed the ability of DRG systems to form cost homogeneous groups of patients into homogeneous groups and in pushing for revisions to the DRG system than those in other countries or specialties. Usually countries claim that their national system is tailored to their needs, implying that their DRG system is good at explaining national practice and resource consumption patterns while keeping the number of groups manageable. But we could not identify any relation between differences in clinical practice and DRG system design. For example, England does not use comorbidities as a classification variable for patients with acute myocardial infarction, even though the number of patients with comorbidities is similar to that in countries that do split for comorbidity. European treatment guidelines exist for acute myocardial infarction and stroke, implying a consensus about what constitutes good quality and cost effective care; more uniformity should therefore be possible.

Countries also differ in the relative prices paid for different groups (figure 2). For example, in Germany, payments for the most complex stroke group were 10 times higher than for the least complex stroke group (price index range 0.25-2.55). In England, reimbursements for most stroke patients were similar irrespective of complications and procedures performed (price index range 0.88-1.00), implying that the system is not good at distinguishing between complex and less complex patients. For acute myocardial infarction the price differed almost 13-fold between the four groups in Estonia, whereas there was a less than fourfold difference between the 10 groups in Germany.

Quality of DRG systems

We analysed the ability of DRG systems to form cost homogeneous groups of patients using the adjusted (deviance) R² as an indicator of the proportion of variation in costs (or length of stay) explained by the groups. This measure should be used with caution for comparisons because of differences in the underlying data. The ability of systems to form homogeneous groups of patients is not the same for all conditions (table 1). Finland, France, and Germany do best for acute myocardial infarction and England does best for hip replacement, suggesting that countries have the potential to learn from each other. Other possible measures of DRG systems could include how resistant the system is to manipulation, how well it is accepted among the relevant stakeholders, and whether it improves transparency.

Current trends and potential for improvement

Three major trends in the development of DRG systems in Europe are observable. Firstly, countries are trying to integrate incentives for improving quality into their payment systems (table 2). For example, as with the Commissioning for Quality and Innovation (CQUIN) framework in the UK, DRG based payments can be adjusted at the hospital level by increasing payments for all patients if hospital quality indicators show good quality. Alternatively, payments to a hospital for all patients falling into one DRG could be increased if the hospital scores above average on DRG specific quality indicators. For example, in Germany a pay for quality contract between a large insurer and a hospital in Karlsruhe awards higher payments for bypass surgery if in-hospital mortality and post-surgical mediastinitis is below the German average.

Secondly, because more hospital activity is taking place in day care and outpatient settings, DRG systems are being extended to include these areas. Several European countries, such as England, Germany, and Sweden, are transferring the concept of DRGs to psychiatric and rehabilitation hospitals, following the example of the US.

Thirdly, there is great interest in the development of bundled payment systems to better align incentives across multiple providers, including hospitals and ambulatory care, and to improve coordination and quality. However, so far, it remains difficult to define products (similar to DRGs) for these extended care bundles. Existing definitions, such as episode treatment groups, which classify patients for entire treatment episodes, remain too imprecise to be used for determining payment. One thing is certain: as long as there are hospitals, measuring hospital activity and using this information for management and payment purposes will continue to be important. Therefore, DRGs will remain popular around the world as they are the best available measure for what hospitals actually do. Episode based bundled payments are likely to complement DRGs rather than replace them. It is therefore important that DRG systems in Europe learn from each other.

This paper is part of an occasional series prepared in conjunction with the European Observatory on Health Systems and Policies (www.healthobservatory.eu)

Contributors and sources: The article is based on results of the EuroDRG-diagnosis related groups in Europe: towards efficiency and quality project, which was funded through the seventh framework programme (FP7) by the European Commission under Grant Agreement No 22330 between 2009 and 2011. The project was led and coordinated by RB, who drafted this article, together with AG and WQ. The project on which this article is based was conceptualised jointly by all authors. AA, FC, UH, CK, CM, ZO, JOR, LS, AS, and SST were the main contributors.
contributors from the other 11 countries; they all reviewed the draft article, contributed to its development, and approved of the final version. WQ is the guarantor.

Competing interests: All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare financial support from the European Commission under Grant Agreement No 223300. The funding source was not involved at any stage of the research. FC received grant funding from the Instituto de Salud Carlos III; AG, RB, and WQ received honoraria from the Korean Health Insurance Review and Assessment for lectures held at the international DRG symposium 2012 in Seoul.

Provenance and peer review: Commissioned; externally peer reviewed.

2 Goldfield N. The evolution of diagnosis-related groups (DRGs): from its beginnings in case-mix and resource use theory, to its implementation for payment and now for its current utilization for quality within and outside the hospital. Qual Manage Health Care 2010;19:3-16.

Accepted: 24 April 2013

Cite this as: BMJ 2013;346:f3197

© BMJ Publishing Group Ltd 2013
<table>
<thead>
<tr>
<th></th>
<th>Austria</th>
<th>England</th>
<th>Estonia</th>
<th>Finland</th>
<th>France</th>
<th>Germany</th>
<th>Ireland</th>
<th>Netherlands</th>
<th>Poland (Catalonia)</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of diagnosis related groups†</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>16 545</td>
<td>73 857</td>
<td>3 409</td>
<td>12 007</td>
<td>69 054</td>
<td>202 758</td>
<td>6 192</td>
<td>31 341</td>
<td>81 634</td>
<td>7 721</td>
</tr>
<tr>
<td>Stroke</td>
<td>18 092</td>
<td>70 256</td>
<td>5 729</td>
<td>13 095</td>
<td>96 920</td>
<td>267 592</td>
<td>5 380</td>
<td>31 663</td>
<td>87 397</td>
<td>11 089</td>
</tr>
<tr>
<td>Hip replacement</td>
<td>19 363</td>
<td>86 090</td>
<td>1 737</td>
<td>21 762</td>
<td>140 313</td>
<td>222 261</td>
<td>5 231</td>
<td>22 935</td>
<td>45 312</td>
<td>2 902</td>
</tr>
<tr>
<td><strong>Relative price range for diagnosis related groups‡</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>0.87-1.52</td>
<td>0.51-1.38</td>
<td>0.81-11.05</td>
<td>0.10-2.79</td>
<td>0.37-3.32</td>
<td>1.00-3.69</td>
<td>0.83-2.03</td>
<td>0.80-2.15</td>
<td>1.00-8.84</td>
<td>0.97-2.64</td>
</tr>
<tr>
<td>Stroke</td>
<td>1.00-2.27</td>
<td>0.88-1.00</td>
<td>1.00</td>
<td>0.06-3.80</td>
<td>0.21-3.01</td>
<td>0.25-2.55</td>
<td>0.28-8.41</td>
<td>0.12-1.17</td>
<td>0.48-2.10</td>
<td>0.76-2.44</td>
</tr>
<tr>
<td>Hip replacement</td>
<td>0.78-2.64</td>
<td>0.53-1.70</td>
<td>1.00-2.14</td>
<td>0.65-1.37</td>
<td>1.00-2.03</td>
<td>1.00-1.87</td>
<td>1.00-2.18</td>
<td>0.38-1.00</td>
<td>0.43-2.10</td>
<td>1.00-1.66</td>
</tr>
<tr>
<td><strong>Adjusted (deviance) R²§</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>0.09</td>
<td>0.39</td>
<td>NA</td>
<td>0.50</td>
<td>0.53</td>
<td>0.49</td>
<td>0.18</td>
<td>NA</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td>Stroke</td>
<td>0.14</td>
<td>0.25</td>
<td>0.33</td>
<td>0.26</td>
<td>0.53</td>
<td>0.49</td>
<td>0.40</td>
<td>NA</td>
<td>0.30</td>
<td>0.14</td>
</tr>
<tr>
<td>Hip replacement</td>
<td>0.26</td>
<td>0.50</td>
<td>0.22</td>
<td>0.22</td>
<td>0.37</td>
<td>0.32</td>
<td>0.37</td>
<td>NA</td>
<td>0.39</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*Definitions are available at: www.eurodrg.eu/EuroDRG_EoCdefinitions.pdf.
†Number of groups into which at least 1% of patients are classified; Figures are for 2008 for all countries except France and Poland, for which 2009 data were used.
§See fig 1 for explanation.
¶Measure of the ability of a DRG system to explain variation in costs or length of stay. Length of stay data were used for analyses in Austria, Ireland and Poland, where cost data were unavailable.
Table 2: Options for integrating quality adjustments into DRG based hospital payment systems and examples from selected European countries and the United States

<table>
<thead>
<tr>
<th>Type of payment adjustment/calculation</th>
<th>Mechanism</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital based</td>
<td>Payment for entire hospital activity is adjusted upwards or downwards by a certain percentage</td>
<td>Predefined quality results are met or not met—e.g., the Commissioning for Quality and Innovations Framework in England</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall hospital readmission rate is below/above average or below/above agreed target (for example, in the US)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hospitals invest in quality improvement programmes</td>
</tr>
<tr>
<td>DRG/disease based</td>
<td>Payment for all patients within a certain DRG (or a disease entity) is adjusted upwards or downwards by a certain percentage</td>
<td>Insurers negotiate with hospitals that DRG payment is higher/lower if certain quality standards are met/not met. For example, in Germany, a large insurer awards higher payments to a hospital for bypass surgery if the hospital scores better than the German average on in-hospital mortality and post-surgical mediastinitis for these patients.</td>
</tr>
<tr>
<td></td>
<td>DRG payment is not based on average costs but on costs at hospitals delivering good quality</td>
<td>DRG payment for some treatment episodes (such as cholecystectomy, stroke care, or primary hip replacement) is based on a &quot;best practice&quot; tariff, which could be higher or lower than national average costs. For example, in England a payment is made for compliance with all seven best practice criteria for fragility hip fracture in addition to the base tariff, which can vary depending on the level of trauma, complexity of the procedure, and presence of complications</td>
</tr>
<tr>
<td>Patient based</td>
<td>No payment is made for a case</td>
<td>Readmissions within 30 days after the initial admission are not paid separately but as part of the original admission (for example, in England and Germany)</td>
</tr>
<tr>
<td></td>
<td>Payment for an individual patient is adjusted upwards or downwards</td>
<td>Complications (that is, certain conditions that were not present on admission) cannot be used to classify patients into DRGs that are weighted more heavily (for example, in the US)</td>
</tr>
</tbody>
</table>
Number of DRGs and relative price range for three episodes of care in 11 countries. The length of the bars indicates the range of the price index, which compares country specific DRG weights (relative weights, tariffs, or scores) with the weight of an index DRG (price index = 1) for the episode of care (that into which a standard case without complications would be classified). The size of the circles represents the number of DRGs used to classify patients.