

DO DIAGNOSIS-RELATED GROUPS APPROPRIATELY EXPLAIN VARIATIONS IN COSTS AND LENGTH OF STAY OF HIP REPLACEMENT? A COMPARATIVE ASSESSMENT OF DRG SYSTEMS ACROSS 10 EUROPEAN COUNTRIES

ALEXANDER GEISSLER*, DAVID SCHELLER-KREINSEN, WILM QUENTIN
ON BEHALF OF THE EURODRG GROUP†

Department of Health Care Management, Berlin University of Technology, Berlin, Germany

ABSTRACT

This paper assesses the variations in costs and length of stay for hip replacement cases in Austria, England, Estonia, Finland, France, Germany, Ireland, Poland, Spain and Sweden and examines the ability of national diagnosis-related group (DRG) systems to explain the variation in resource use against a set of patient characteristic and treatment specific variables. In total, 195 810 cases clustered in 712 hospitals were analyzed using OLS fixed effects models for cost data ($n = 125\,698$) and negative binominal models for length-of-stay data ($n = 70\,112$). The number of DRGs differs widely across the 10 European countries (range: 2–14). Underlying this wide range is a different use of classification variables, especially secondary diagnoses and treatment options are considered to a different extent. In six countries, a standard set of patient characteristics and treatment variables explain the variation in costs or length of stay better than the DRG variables. This raises questions about the adequacy of the countries' DRG system or the lack of specific criteria, which could be used as classification variables. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS: DRG; diagnosis-related groups; cost analyses; length of stay; artificial joint implants; hip replacement; arthroplasty

1. INTRODUCTION

Most European countries adopted diagnosis-related group (DRG) systems to increase the transparency of services, which are delivered in hospitals and as a basis to pay hospitals in relation to their activity (Wiley, 2011; Geissler *et al.*, 2011). DRGs classify patients with similar diagnoses and/or procedures, which are characterized by comparable resource consumption (Kobel *et al.*, 2011). DRG-based payments incentivize hospitals to cut down costs per patient, for example, by reducing the length of stay (LoS) or the intensity of services (Cots *et al.*, 2011). These behavioural responses may be intended (e.g. a reduction of resource intensity via the introduction of efficient clinical pathways) but may also take unintended forms such as inappropriate early discharge, skimping, cream-skimming or dumping (Ellis, 1998; Newhouse and Byrne, 1988; Levaggi and Montefiori, 2003; Martinussen and Hagen, 2009). To avoid these unintended consequences, DRGs need to be defined resource homogeneously to ensure that DRG-based payments reflect treatment costs as precisely as possible for a given patient.

*Correspondence to: Department of Health Care Management, Berlin University of Technology, Straße des 17. Juni 135, H80 10623 Berlin, Germany. E-mail: a.geissler@tu-berlin.de

†www.eurodrgeu/EuroDRG_group.pdf. Funding: The presented results were conducted within the research project 'EuroDRG – Diagnosis Related Groups in Europe: towards efficiency and quality', which was funded by the European Commission under the Seventh Framework Programme. Research area: HEALTH-2007-3.2-8 European System of Diagnosis-Related Groups, Project reference: 223300.

However, beyond application descriptions of DRG systems (Kimberly *et al.*, 2008; Busse *et al.*, 2011), little is known whether these systems can adequately account for treatment costs. Therefore, we assess the ability of national DRGs in 10 European countries (Austria, England, Estonia, Finland, France, Germany, Ireland, Poland, Spain and Sweden) to account for patient-level variation in hospital resource consumption against a standard set of patient characteristics, treatment and quality variables. By way of example, we analysed DRG classification of hip replacement patients as they account for major epidemiological and economic burden in all European health systems (Stargardt, 2008).

Replacement of the hip joint by an artificial implant can provide effective relief for patients with osteoarthritis of the hip (coxarthrosis), which is estimated to become the fourth leading reason for disability in 2020 (Woolf and Pfleger, 2003; Kurtz *et al.*, 2007; United Nations, UN *et al.*, 2007, Bitton, 2009). Hip replacement is provided with increasing frequency but with varying rates throughout Europe. In 2009, the rate ranged from 296 per 100 000 inhabitants in Germany to 44 per 100 000 inhabitants in Poland (OECD Health Data, 2011).

Diagnosis-related group classification and payment for hip replacement differ substantially across Europe. For example, the number of DRGs, which individually account for at least 1% of hip replacement patients, ranges from two in Estonia and Sweden to 10 in France or even 14 in England. Similarly, the number and type of patient and treatment characteristics, which are taken into account by each DRG system to define DRG split variables differ greatly (Table I). For example, all DRG systems distinguish between cases with total hip replacement and the revision of an implant, but only four systems (Austria, England, Germany and Poland) have specific DRGs for partial replacements. Moreover, main and secondary diagnoses are used to classify patients only in England, France, Germany, Ireland and Spain. Age and LoS is only used to group hip replacement patients in France and Germany.

2. DATA AND METHODS

Researchers from 10 European countries performed quantitative analyses on national routine patient level data samples from 2008 as these were the latest available data samples across all countries (Street *et al.*, 2012).

Table I. Number of diagnosis-related groups and split variables used to classify hip replacement cases in 10 European countries

Country (system)	Number of DRGs	Diagnosis-related group-split variables							
		Age	Length of stay	Complications/comorbidities	Main/secondary diagnosis	Total replacement	Partial replacement	Revision of replacement	Multiple/bilateral procedures
Austria (LKF)	8	—	—	—	—	X	X	X	—
England (HRG)	14	—	—	X	—	X	X	X	—
Estonia (NordDRG)	2	—	—	—	—	X	—	X	—
Finland (NordDRG)	3	—	—	—	—	X	—	X	X
France (GHM)	10	X	X	X	X	X	—	X	X*
Germany (G-DRG)	8	X	—	X	X	X	X	X	X*
Ireland (AR-DRG)	3	—	—	X	—	X	—	X	—
Poland (JGP)	6	—	—	—	—	X	X	X	—
Spain (AP-DRG)	3	—	—	X	X	X	—	X	X*
Sweden (NordDRG)	2	—	—	—	—	X	—	X	—

Notes: X variable is used to classify patients; — variable is not used to classify patients; * DRGs for multiple/ bilateral procedures are part of the DRG system but not considered by the analysis because they are populated with less than 1% of hip patients in the available data sources.

We identified hip replacement cases using ICD-9 CM procedure codes 81.51–81.53 (total, partial and revision of hip replacement), 00.7 (other hip procedures), 00.85–00.87 (hip resurfacing) or equivalent codes in national coding systems, which were mostly available from the Hospital Data Project (Magee, 2003). Patients in ambulatory settings and infants aged less than 1 year were excluded from the analysis.

In addition to the standard set of patient-level variables, such as age categories in quintiles, gender and Charlson index co-morbidities (Street *et al.*, 2012), we defined three dummy variables to control severity and case-mix differences specific to hip replacement patients: First, hip replacement is increasingly required for patients with severe hip fractures, and we expect that their costs are higher than replacements in elective settings (Gullberg *et al.*, 1997; Burge *et al.*, 2007). Hence, patients with fractures were identified via a dummy variable that captures recoded ICD-10 codes S72, M84, M960 or S324. Second, there are different treatment options for hip replacement: it can be performed as total replacement, replacing both the acetabulum (hip socket) and the femoral head, or as partial (hemi) replacement, replacing just the femoral head using press or cement-fit implants (Wülker, 2010). As we expect that patients with partial hip replacements have lower costs and lengths of stay because of a less complex surgical intervention, we account for this by a dummy variable defined via ICD-9-CM procedure codes 81.52, 00.86, 00.87 or equivalent. Finally, artificial hip joints have to be revised some years after implementation, for example, because of aseptic loosening, instability or osteolysis (Clohisey *et al.*, 2004), and we expect related procedures to be more costly than the reference case (Braithwaite *et al.*, 2003; Bozic *et al.*, 2005). Hence, we define a dummy variable for revisions via ICD-9-CM procedure codes 81.53, 00.70–00.73. Some surgical techniques such as ‘press-fit’ and ‘cemented’ could not be differentiated across countries as the relevant procedures cannot be consistently identified across countries.

A two-stage multilevel regression approach was adopted to identify the effects of patient and hospital variables. Seven of 10 countries (England, Estonia, Finland, France, Germany, Spain and Sweden) provided cost data. Because of data restrictions, three countries (Austria, Ireland and Poland) had to rely on LoS data, which were used as a proxy for treatment costs (Street *et al.*, 2012). Cost and LoS data were collected from hospital admission to the discharge of the patient to either another institution (e.g. rehabilitation unit) or home (*ibid.*). The costs analyses were performed using OLS fixed effects models with log costs as the dependent variable. For the LoS analysis, negative binomial (NegBin) models were used, as over dispersion was detected during exploratory data analyses.

For the first stage, three models were designed to assess the determinants of costs and LoS variation. First, M_D contains only DRG dummy variables to test the ability of relevant DRGs to reflect resource consumption. Second, M_P uses the standard set of patient characteristics such as age, treatment variables, quality indicators and co-morbidities to predict variation in costs or LoS. Finally, the fully specified model M_F includes both sets of variables, that is, DRGs and the variables considered in M_P . A comparison of the model results will uncover if the set of patient characteristic and treatment variables has a greater ability to predict the costs or LoS of a given patient than the national DRGs or vice versa. In the second stage, we analyse the estimated hospital fixed effects to understand why some hospitals have higher average cost or LoS than others after taking patient characteristics and treatment variables into account. A set of hospital-specific variables, such as teaching and ownership status, volume of activity and specialization, were used for explaining hospital level differences in cost and LoS. The methodological approach and a detailed description of the core variables are presented in Street *et al.* (2012) as part of this freely available supplement.

3. RESULTS

3.1. Descriptive statistics

Table I provides an overview of the DRGs and underlying split variables used to group hip replacement cases across European DRG systems. It illustrates that common origins of DRG systems are reflected in the selection

of classification variables. For example, the Estonian, Finnish and Swedish DRG system are very similar as they are all operated under the framework of the NordDRG system (Linna and Virtanen, 2011).

Table II provides descriptive statistics of the data used for the analysis. The sample size differs widely across countries from 86 090 cases in England to 2941 cases in Spain. Additionally, cases were clustered in a varying number of hospitals - from 277 hospitals in Poland to 5 hospitals in Finland.

In all countries, except Poland, England and France, more than 70% of patients are grouped into one DRG. In England, the country with the highest number of DRGs, just 31% of cases are grouped into the most populated DRG, whereas in Estonia, 92% of cases are grouped into one DRG. In Table II, the DRGs are ordered by ascending DRG weights. It indicates that in seven countries, the DRG with the highest volume is also the least expensive.

The share of male patients varies from 34% (Poland) to 45% (Ireland). The mean age ranges from 67.8 years in Poland to 72.9 years in England.

The LoS in hospitals varies remarkably from 4.7 days in Finland to 15.2 days in Germany. These significant differences can be explained with different treatment patterns and care organization including the arrangements about rehabilitation after replacement (its availability and place). In Finland, patients are kept in the acute care wards just for the hip surgery and transferred very quickly into specialized rehabilitation units, whereas in Germany, a longer observation in hospital is the norm before deciding a proper setting for rehabilitation. In any case rehabilitation is not part of the DRG-based payment in any of the countries studied here and reimbursed separately.

Germany has the highest mean number of different coded diagnoses (6.6 against 1.2 in Poland or Finland). Similarly, the average number of different procedures differs widely from 5.2 in Poland to 1.0 in Estonia.

In Spain, almost 40% of patients have a recorded fracture, whereas in Austria, the share of fracture patients is less than 20%. The proportion of revisions ranges from 6% (Sweden) to 13% (France and Spain). We find the highest rate of partial replacements in Finland (35%) and the lowest in Poland (less than 17%).

Ireland has the highest rate of recorded wound infection during the hospital stay (2%), and Germany has the highest recorded urinary tract infections (7%). Adverse events were most frequent with 2% in France (mostly because of pulmonary embolism and deep vein thrombosis) and Spain (mostly because of sepsis).¹ The highest mortality rate was recorded in England with 3% of patients deceased after being admitted for a hip surgery procedure. The English in-hospital mortality rate is notable as it is at least twice as high as in any other country.

The highest frequency of Charlson co-morbidities was in England and Germany. One third of English patients have at least one coded co-morbidity, which is mostly chronic pulmonary disease, diabetes and dementia. In Germany, 41% of patients are frequently affected by diabetes, congestive heart failure, chronic pulmonary disease and dementia¹.

3.2. Regression results

3.2.1. Stage 1. Tables III and IV present results of the first stage regressions. In model M_D , where DRGs are ordered by ascending DRG weights, most DRG coefficients are greater than one (LoS analysis) or positive (costs analysis) and significant, which shows that patients who are grouped into these DRGs have higher costs or longer LoS than patients grouped to the most populated DRG, which is our reference case. Two exceptions (DRG5 in Poland²; DRG6 in England³) stand out. Patients within these groups have shorter LoS and lower costs than patients within the reference group but higher DRG weights, which means higher payments to hospitals.

M_p estimated the impact of a set of patient characteristics on LoS and costs of hip replacement. We find a significant positive relationship between age and the LoS. In cost data, this relationship is less clear as it is only

¹Data not shown but available on www.eurodr.org.

²Primary total hip replacement with osseous reconstruction, metaphyseal stem, hip resurfacing [H06]

³Intermediate hip procedures for trauma without CC (includes partial hip joint endoprosthesis) [HA13C]

observed for England. In Estonia, Germany and Sweden, patients at age younger than 61 years are significantly more costly than the reference group (aged between 61 and 70 years). In Finland and Estonia, older patients (age category 3, 4 and 5) have significantly lower resource consumption. In France and Spain, we find no significant impact of age on costs.

Gender is significantly affecting LoS: hospital stays tend to be shorter for male patients. The relationship between gender and costs is significant only for three countries (Estonia, Spain and Sweden) at the 5% level, indicating that male patients tend to be more costly.

The number of coded diagnoses and procedures is highly significant across countries, except for Estonia. Regardless of different coding standards, this reflects that patients with more coded diagnoses and procedures have higher costs and longer LoS.

Patients who were transferred to a hospital from another institution tend to have significantly lower costs in England, Estonia, France and Sweden and a shorter LoS in Ireland. A transfer out of the hospital to another institution, such as rehabilitation units, seems to reduce the LoS in Poland and costs in Estonia and Sweden. In contrast, in England, France and Spain, cases with a transfer out of the hospital appear to have higher costs. Interestingly, in Finland, the country with the lowest LoS, where patients receive rehabilitation care at home or in other institutions than hospitals, both the transfer to or out of a hospital had no significant effect on costs.

An emergency admission of patients is associated with significantly longer LoS in Austria, Ireland and Poland and higher costs in Finland and Sweden, whereas in England, France and Germany, emergency patients have significantly lower costs.

Deceased patients have significantly shorter LoS and significantly lower costs in France, Germany and Sweden.

Co-morbidities assessed via the Charlson index (Street *et al.*, 2012) are rarely significant, and where they are, they show inconsistent influences on costs and LoS across countries. The weak explanatory power of the Charlson index variables might be caused by the fact that both the Charlson index and the total number of diagnostics are used to measure patients' morbidity.

Patients with fractures have significantly longer LoS in Poland and higher costs in England, Germany and Sweden. Patients who receive a partial replacement have significantly lower LoS in Austria and Poland and lower costs in Estonia, France, Spain and Sweden. The revision of a hip implant leads to significantly longer LoS and higher costs in all countries.

In all countries, except Austria, Estonia, Finland and Sweden, patients who suffer from urinary tract or wound infection have significantly longer LoS and imply higher costs. Adverse events are not significant in explaining differences in costs and LoS across countries, except in England.

In some countries, certain patient and treatment variables included in M_P are also used to define DRGs (see Table I). If DRGs adequately reflect differences in resource consumption, the patient and treatment variables should have less explanatory power when adding them to the DRG variables as is done in the full model (M_F). Therefore, the coefficients of these variables are expected to be lower in M_F than in M_P or become insignificant in M_F . This applies to the majority of significant coefficients such as age categories, diagnoses, procedures, emergency and revision. In contrast, the coefficients are mostly not significant for the Carlson index, adverse events, urinary tract and wound infection variables. The revision variable stands out because in five countries the coefficient becomes insignificant in M_F .

The degree to which DRGs and patient/treatment variables are able to explain costs differences among patients varies across countries. The adjusted R^2 statistics indicate that for the patient characteristic models (M_P), the explained variance in resource consumption ranges from 17% in Finland up to 61% in Estonia, whereas the remaining countries are in a corridor of 30% to 50%. In most countries, except for England and Finland, M_P are superior to the DRG models in explaining variation of costs. The R^2 difference between M_D and M_P in most countries approximates 10% with the exception of Estonia where the difference is substantially higher at 40%.

Table II. Descriptive patient-level and hospital-level statistics of hip replacement samples in 10 European countries

Country	Length-of-stay data					Cost data				
	Austria	Ireland	Poland	England	Estonia	Finland	France	Germany	Spain	Sweden
No. of patients	19569	5231	45312	86090	1737	1872	17650	5166	2941	10242
No. of hospitals	105	22	277	151	11	5	80	22	8	31
Variables, patients*					Percentage of patients (%)					
DRG1	2.4	71.9	7.2	1.4	91.7	88.7	43.5	70.1	81.4	87.4
DRG2	74.4	23.1	16.9	2.8	7.1	8.1	4.7	4.0	13.2	11.2
DRG3	2.3	2.4	44.1	21.6	—	2.1	16.5	10.6	4.0	—
DRG4	1.6	—	3.3	30.8	—	—	9.7	3.5	—	—
DRG5	13.9	—	24.1	3.4	—	—	6.4	1.5	—	—
DRG6	2.3	—	2.8	2.7	—	—	3.5	1.0	—	—
DRG7	1.1	—	—	10.9	—	—	3.8	3.5	—	—
DRG8	1.0	—	—	1.2	—	—	5.1	2.3	—	—
DRG9	—	—	—	7.7	—	—	1.5	—	—	—
DRG10	—	—	—	1.4	—	—	1.8	—	—	—
DRG11	—	—	—	3.9	—	—	—	—	—	—
DRG12	—	—	—	1.3	—	—	—	—	—	—
DRG13	—	—	—	3.3	—	—	—	—	—	—
DRG14	—	—	—	2.6	—	—	—	—	—	—
DRG other	1.1	2.6	1.6	5.1	1.2	1.1	3.5	3.5	1.3	1.4
Male	40.6	45.0	33.7	34.0	37.0	36.5	39.6	37.3	42.8	38.0
Transfer in	0.8	7.7	NA	3.2	3.3	11.7	2.1	1.3	0.4	4.4
Transfer out	10.1	31.7	4.1	10.9	25.1	17.8	45.4	37.5	3.7	17.3
Emergency	25.5	32.4	31.8	36.1	28.0	33.8	11.1	31.6	37.1	39.7
Deceased	0.9	1.6	0.9	3.3	0.6	0.9	1.1	1.7	0.8	1.4
Charlson index = 0	86.6	82.2	97.5	68.1	84.2	95.7	79.2	58.6	72.5	71.9
Charlson index = 1	8.8	12.0	2.1	22.0	11.4	3.4	14.4	22.5	19.0	20.4
Charlson index = 2	4.7	5.8	0.5	9.9	4.4	0.9	6.4	19.0	8.4	7.7
Adverse event	0.3	1.0	0.1	0.9	0.8	0.2	1.9	1.3	2.1	0.6
Urinary tract infection	2.1	3.4	0.0	3.5	0.2	0.2	3.1	6.7	2.3	2.8
Wound infection	0.1	2.0	0.0	0.9	0.1	0.0	0.3	0.8	0.8	0.4
Fracture	18.5	30.2	22.3	29.0	23.9	36.8	23.0	24.1	39.6	31.6
Partial	17.5	30.0	16.8	27.5	21.3	35.1	2.2	17.8	27.4	24.8
Revision	9.6	8.3	7.1	8.6	7.3	8.1	12.6	7.1	13.1	5.8

(Continues)

Table II. (Continued)

Country	Length-of-stay data					Cost data				
	Austria	Ireland	Poland	England	Estonia	Finland	France	Germany	Spain	Sweden
Log cost	NA	NA	NA	8.6 (0.4)	8.1 (0.3)	9.1 (0.3)	8.9 (0.3)	8.8 (0.3)	8.9 (0.4)	9.0 (0.3)
Cost	NA	NA	NA	6081 £ (2120.0)	3351.8 € (1 017.7)	8854.9 € (2 416.2)	7313.5 € (2 226.1)	6848.1 € (2487.0)	7736.8 € (3 101.7)	84014 SEK (30603.0)
LoS	13.8 (5.5)	12.8 (9.3)	13.3 (6.1)	12.1 (14.1)	8.6 (5.3)	4.7 (2.4)	11.0 (4.9)	15.2 (7.6)	9.7 (5.8)	7.6 (4.7)
Age	68.8 (12.4)	70.6 (12.7)	67.8 (12.0)	72.9 (12.1)	68.7 (12.6)	70.8 (12.9)	71.2 (12.8)	71.4 (11.7)	72.9 (13.1)	72.8 (12.1)
No. of diagnoses	2.6 (2.1)	3.9 (3.2)	1.2 (0.5)	4.0 (2.8)	1.9 (1.3)	1.2 (0.6)	2.9 (2.2)	6.6 (4.2)	3.9 (2.4)	3.0 (2.0)
No. of procedures	2.4 (0.9)	4.0 (1.5)	5.2 (4.7)	2.6 (1.2)	1.0 (0.0)	4.3 (1.2)	4.9 (2.3)	3.6 (2.3)	1.5 (0.8)	1.7 (1.1)
Variables, hospital					Percentage of hospitals (%)					
Teach	2.9	36.4	7.9	0.1	0.1	20.0	11.3	NA	50.0	27.6
Private	19.0	NA	NA	0	0.1	0	62.5	4.5	0	0
Urban	65.7	NA	98.6	NA	0.6	NA	53.8	NA	37.5	37.9
Cases treated annually ('000 s)	20.0 (20.6)	42.0 (29.6)	16.2 (11.5)	84.6 (46.0)	20.3 (15.5)	90.5 (142.4)	30.2 (29.0)	20.1 (13.4)	32.1 (24.7)	26.8 (25.7)
Share of hip replacement cases (%)	1.4 (2.0)	3.2 (6.7)	2.1 (6.4)	0.9 (1.1)	0.7 (0.4)	0.5 (0.1)	1.5 (3.7)	2.2 (3.4)	1.3 (0.5)	2.2 (2.1)
Specialisation index	0.4 (0.2)	0.5 (0.3)	NA	0.2 (0.1)	NA	0.2 (0.0)	0.5 (0.2)	0.4 (0.2)	NA	0.2 (0.2)
adv_event_1	2.00 (1.98)	11.61 (9.88)	10.00 (7.00)	15.07 (5.42)	0.00 (0.00)	17.76 (6.25)	12.90 (8.44)	47.77 (12.17)	0.43 (0.41)	23.50 (7.60)
adv_event_2	2.45 (1.69)	5.35 (3.10)	2.00 (1.00)	6.99 (2.21)	0.00 (0.10)	18.48 (2.30)	7.71 (4.98)	17.61 (5.90)	0.23 (0.11)	11.90 (3.10)

NA = not available.

*DRGs ordered by ascending DRG weights (DRGs vary by country).

Table III. First stage regression results (part 1)

Dependent variable	Length of stay [^]						Log of cost								
	Austria			Ireland			Poland			England			Estonia		
	M _D	M _P	M _F	M _D	M _P	M _F	M _D	M _P	M _F	M _D	M _P	M _F	M _D	M _P	M _F
DRG1	1.089***	—	1.127**	ref	—	ref	1.038***	—	0.991	—0.745***	—	—0.761***	ref	—	ref
DRG2	ref	—	ref	1.561***	—	1.138***	1.066***	—	1.003	—0.550***	—	—0.586***	0.453***	—	0.488***
DRG3	1.042*	—	1.051	1.928***	—	1.232**	ref	—	ref	—0.087***	—	—0.070***	—	—	—
DRG4	1.723***	—	1.448***	—	—	—	1.264***	—	1.046	ref	—	ref	—	—	—
DRG5	1.139***	—	0.937	—	—	—	0.986**	—	0.972***	0.266***	—	0.236***	—	—	—
DRG6	1.032*	—	0.95	—	—	—	1.314***	—	1.071*	—0.026***	—	—0.072***	—	—	—
DRG7	1.341***	—	1.058	—	—	—	—	—	—	0.055***	—	—0.01	—	—	—
DRG8	1.830***	—	1.438***	—	—	—	—	—	—	0.160***	—	0.130***	—	—	—
DRG9	—	—	—	—	—	—	—	—	—	0.317***	—	0.217***	—	—	—
DRG10	—	—	—	—	—	—	—	—	—	0.177***	—	0.070***	—	—	—
DRG11	—	—	—	—	—	—	—	—	—	0.146***	—	0.188***	—	—	—
DRG12	—	—	—	—	—	—	—	—	—	0.140***	—	0.287***	—	—	—
DRG13	—	—	—	—	—	—	—	—	—	0.255***	—	0.142***	—	—	—
DRG14	—	—	—	—	—	—	—	—	—	0.355***	—	0.468***	—	—	—
DRGother	1.728***	—	1.255***	2.249***	—	1.456***	1.608***	—	1.355***	0.050***	—	—0.011	0.163	—	0.211*
Age 1 (1–60)	—	0.943***	0.943***	—	0.958*	0.946**	—	0.961***	0.965***	—	—0.013***	—0.009**	—	0.071***	0.068***
Age 2 (61–70)	—	ref	ref	—	ref	ref	—	ref	ref	—	ref	ref	—	ref	ref
Age 3 (71–75)	—	1.053***	1.048***	—	1.076***	1.075***	—	1.018***	1.018***	—	0.008*	0.003	—	—0.031***	—0.031***
Age 4 (76–80)	—	1.099***	1.093***	—	1.157***	1.151***	—	1.045***	1.045***	—	0.011***	0.006*	—	—0.034*	—0.035***
Age 5 (>80)	—	1.150***	1.142***	—	1.197***	1.195***	—	1.045***	1.048***	—	0.030***	0.023***	—	—0.037	—0.037
Gender	—	0.968***	0.967***	—	0.970*	0.972*	—	0.986***	0.985***	—	0.002	0.002	—	0.018*	0.019*
No. of diagnoses	—	1.036***	1.034***	—	1.071***	1.059***	—	1.069***	1.068***	—	0.017***	0.005***	—	0.027***	0.031***
No. of procedures	—	1.128***	1.107***	—	1.114***	1.107***	—	1.032***	1.034***	—	0.021***	0.020***	—	0.522***	0.520***
Transfer in	—	0.995	0.956	—	0.932*	0.929*	—	NA	NA	—	—0.067***	—0.031***	—	—0.119***	—0.125***
Transfer out	—	0.997	0.996	—	0.987	0.988	—	0.920***	0.922***	—	0.054***	0.032***	—	—0.023*	—0.018
Emergency	—	1.121***	1.112***	—	1.224***	1.198***	—	1.026***	1.022**	—	—0.323***	0.025**	—	0.011	0.026
Deceased	—	0.720***	0.726***	—	0.823*	0.795**	—	0.762***	0.756***	—	0.020**	—0.004	—	0.339	0.346*
Charlson index = 1	—	0.986	0.988	—	1.002	1.006	—	0.975	0.975	—	0.018***	0.005*	—	0.022	0.013
Charlson index = 2	—	0.961*	0.963*	—	0.973	0.951	—	1.065	1.031	—	0.019***	0.002	—	0.001	—0.013
Adverse event	—	1	1.015	—	1.038	0.991	—	1.171**	1.180**	—	0.112***	0.028*	—	0.243	0.253*
Urinary tract infection	—	1.03	1.038*	—	1.142**	1.097*	—	1.251***	1.274***	—	0.106***	0.032***	—	0.025	0.031
Wound infection	—	1.507***	1.381***	—	1.134**	1.108*	—	2.020***	2.055***	—	0.123***	0.068***	—	1.465***	1.474***
Fracture	—	1.007	1.009	—	0.904	0.936	—	1.091***	1.084***	—	0.365***	0.094***	—	—0.036	—0.06
Partial	—	0.910***	0.978	—	0.917	0.926	—	0.927***	0.930***	—	0.016*	—0.084***	—	—0.332***	—0.318***
Revision	—	1.001	0.889**	—	1.139***	1.102	—	1.320***	1.182***	—	0.156***	—0.154***	—	0.408***	0.399***
Intercept	13.613***	9.494***	9.923***	11.241***	5.475***	5.642***	12.651***	9.993***	9.999***	8.624***	8.491***	8.556***	8.047***	7.551***	7.552***
Dispersion parameter	0.034***	0.021***	0.018***	0.151***	0.103***	0.100***	0.044***	0.038***	0.037***	—	—	—	—	—	—
Adjusted deviance R^2	0.263	0.373	0.4	0.372	0.515	0.525	0.387	0.422	0.429	—	—	—	—	—	—
Adjusted R^2	—	—	—	—	—	—	—	—	—	0.504	0.407	0.519	0.222	0.612	0.618

[^]Exponentiated coefficients; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ref = reference variable; NA = not available.

Dependent variable		Log of cost															
Finland						France			Germany			Spain			Sweden		
MD	MP	MF	MD	MP	MF	MD	MP	MF	MD	MP	MF	MD	MP	MF	MD	MP	MF
DRG1	—	ref	ref	—	—	ref	—	—	ref	—	—	ref	—	—	ref	—	ref
DRG2	—	0.181***	0.22	0.016	—	-0.108***	0.134***	—	0.090***	0.321***	—	0.166***	—	—	0.274***	—	0.225***
DRG3	—	0.625***	0.609***	0.050***	—	0.026***	0.199***	—	0.071***	0.274***	—	0.158***	—	—	—	—	—
DRG4	—	—	—	0.026***	—	-0.088***	0.245***	—	0.134***	—	—	—	—	—	—	—	—
DRG5	—	—	—	0.119***	—	-0.049*	0.391***	—	0.174***	—	—	—	—	—	—	—	—
DRG6	—	—	—	0.233***	—	0.022	0.371***	—	0.161***	—	—	—	—	—	—	—	—
DRG7	—	—	—	0.114***	—	0.048***	0.503***	—	0.317***	—	—	—	—	—	—	—	—
DRG8	—	—	—	0.096***	—	-0.055*	0.600***	—	0.304***	—	—	—	—	—	—	—	—
DRG9	—	—	—	0.314***	—	0.042	—	—	—	—	—	—	—	—	—	—	—
DRG10	—	—	—	0.228***	—	0.031	—	—	—	—	—	—	—	—	—	—	—
DRGother	—	0.122	0.159*	0.266***	—	0.070***	0.413***	—	0.160***	0.450***	—	0.304***	—	—	0.150***	—	0.104***
Age 1 (1–60)	0.009	-0.004	—	—	0	-0.001	—	0.034***	0.020**	—	0.016	0.013	—	0.038***	—	—	0.037***
Age 2 (61–70)	—	ref	—	—	ref	ref	—	ref	ref	—	ref	ref	—	ref	—	ref	ref
Age 3 (71–75)	-0.084***	-0.082***	—	—	0.001	0.001	—	-0.005	-0.009	-0.024	-0.024	-0.029	-0.017*	0.017*	—	0.017*	0.017*
Age 4 (76–80)	-0.057**	-0.049**	—	—	0.001	-0.002	—	-0.006	-0.013	-0.018	-0.018	-0.02	-0.052***	0.051***	—	0.051***	0.051***
Age 5 (>80)	-0.059**	-0.051**	—	—	0.002	-0.024***	—	0.004	-0.007	-0.031	-0.031	-0.03	0.071***	0.070***	—	0.071***	0.070***
Gender	0.008	0.003	—	—	-0.005	-0.007*	—	0.014*	0.014*	0.026*	0.026*	0.026*	0.013*	0.009	—	0.013*	0.009
No. of diagnoses	-0.008	-0.013	—	—	0.019***	0.014***	—	0.017***	0.011***	—	0.013***	0.009*	0.036***	0.036***	—	0.036***	0.036***
No. of procedures	0.043***	0.039***	—	—	0.047***	0.046***	—	0.065***	0.055***	—	0.095***	0.082***	0.079***	0.071***	—	0.079***	0.071***
Transfer in	0.014	0.01	—	—	-0.030*	-0.035*	—	0.04	0.009	0.093	0.093	0.114	-0.054*	-0.054*	—	-0.060*	-0.054*
Transfer out	-0.022	-0.025	—	—	0.031***	0.030***	—	0.003	0.015	0.117**	0.117**	0.112**	-0.069***	-0.071***	—	-0.069***	-0.071***
Emergency	0.120*	0.117*	—	—	-0.036***	-0.027**	—	-0.051***	-0.053***	0.009	0.009	0.015	0.044***	0.032**	—	0.044***	0.032**
Deceased	0.048	0.025	—	—	-0.232***	-0.233***	—	-0.241***	-0.233***	—	0.219*	0.157	-0.220***	-0.218***	—	-0.220***	-0.218***
Charlson index = 1	0.006	0.004	—	—	-0.008	-0.008	—	-0.023*	-0.012	-0.017	-0.017	-0.011	-0.043***	-0.030***	—	-0.043***	-0.030***
Charlson index = 2	0.145	0.137*	—	—	0	0.001	—	-0.028*	-0.025*	0.037	0.037	0.031	-0.085***	-0.060***	—	-0.085***	-0.060***
Adverse event	0.445	0.271	—	—	0.023	-0.006											

Copyright © 2012 John Wiley & Sons, Ltd.

3.2.2. Stage 2. After controlling for different patient/treatment characteristics and DRGs, Figure 1 shows the unexplained variance in costs or LoS across hospitals within each country. Hospitals are ranked from left to right by their deviation from the national sample mean costs or LoS in ascending order. Figure 1 illustrates that even after taking different patient characteristics and treatment variables into account, the variation of unexplained LoS and costs from one hospital to another remains large. Countries with a larger hospital sample (e. g. Austria, Poland, England and France) show an S-shaped distribution of hospitals with remarkable outlier hospitals at the extremes. The confidence intervals reflect the distribution of the dependent variable. Wide intervals are mostly because of a low number of cases with highly different costs or LoS within one hospital and because of low or high cost outlier cases at the end of the distribution.

A second-stage analysis of hospital-specific variables and their impact on LoS and costs was conducted for Austria, Ireland, Poland, England, France, Germany and Spain (Street *et al.*, 2012).⁴ Patients in Austrian private hospitals stay significantly longer than in public or non-profit hospitals. In France, private for-profit ownership leads to significantly lower treatment costs. In Poland and England, LoS and costs slightly increase with the size of the hospital, that is, we find modest for support diseconomies of scale. In addition, in England specialized hospitals, which were identified using an adapted Gini index (Street *et al.*, 2012), appear to treat hip patients at significantly higher costs. None of the other explanatory variables (i.e. total volume of hospital cases, share of hip surgery cases on total hospital volume, teach, urbanity and adverse events) was a significant predictor of resource consumption in the second stage analyses.

4. DISCUSSION AND CONCLUSIONS

The ability of European DRG systems to explain the variation in resource use was compared against a standard set of patient characteristic and treatment-specific variables based on patient-level data. However, cross-country comparisons should be made with caution because of the differences in data/hospital samples. First, the coding practices, especially for secondary diagnoses, vary from one country to another based on different national (or even regional) regulations and coding incentives induced by the given DRG-based payment methodology. Therefore, conflicting results across countries for any given explanatory variable may partially be the result of coding differences. Moreover, patients' costs are reported differently in each country and may include a varying number of cost categories because of national specific costing and reporting standards (Tan *et al.*, 2011). Third, the need to define a common set of variables, which reflect patient characteristics and treatments consistently across countries, prevented the inclusion of some potentially important treatment variables such as different surgical techniques (e.g. press-fit versus cemented) as information about these were not available in all countries. In addition, in countries where add-on payments for high cost implants or rehabilitation exist, DRGs might systematically differ from countries where they do not as for example different split-variables are needed to account for resource consumption variation or the shape of the cost distribution differs. Lastly, regulators may be motivated by concerns other than resource homogeneity when defining DRGs that we did not pay attention to in our analysis such as low transaction costs or a fast incorporation of technological innovation.

Nevertheless, in six countries, the set of patient characteristics and treatment variables proposed in our model appear to explain better than the DRG variables the variation in costs or LoS. Given the relatively weak ability of some DRG models to explain cost variation, these systems may benefit from integrating selected patient characteristics in the definition of relevant groups. For example, although all countries differentiate between primary and revision hip replacement, just four countries have DRGs for partial replacements. However, our results suggest that incorporating this variable in classification can improve their capacity to account for resource use.

⁴Detailed results are available at: www.eurodrgeu.

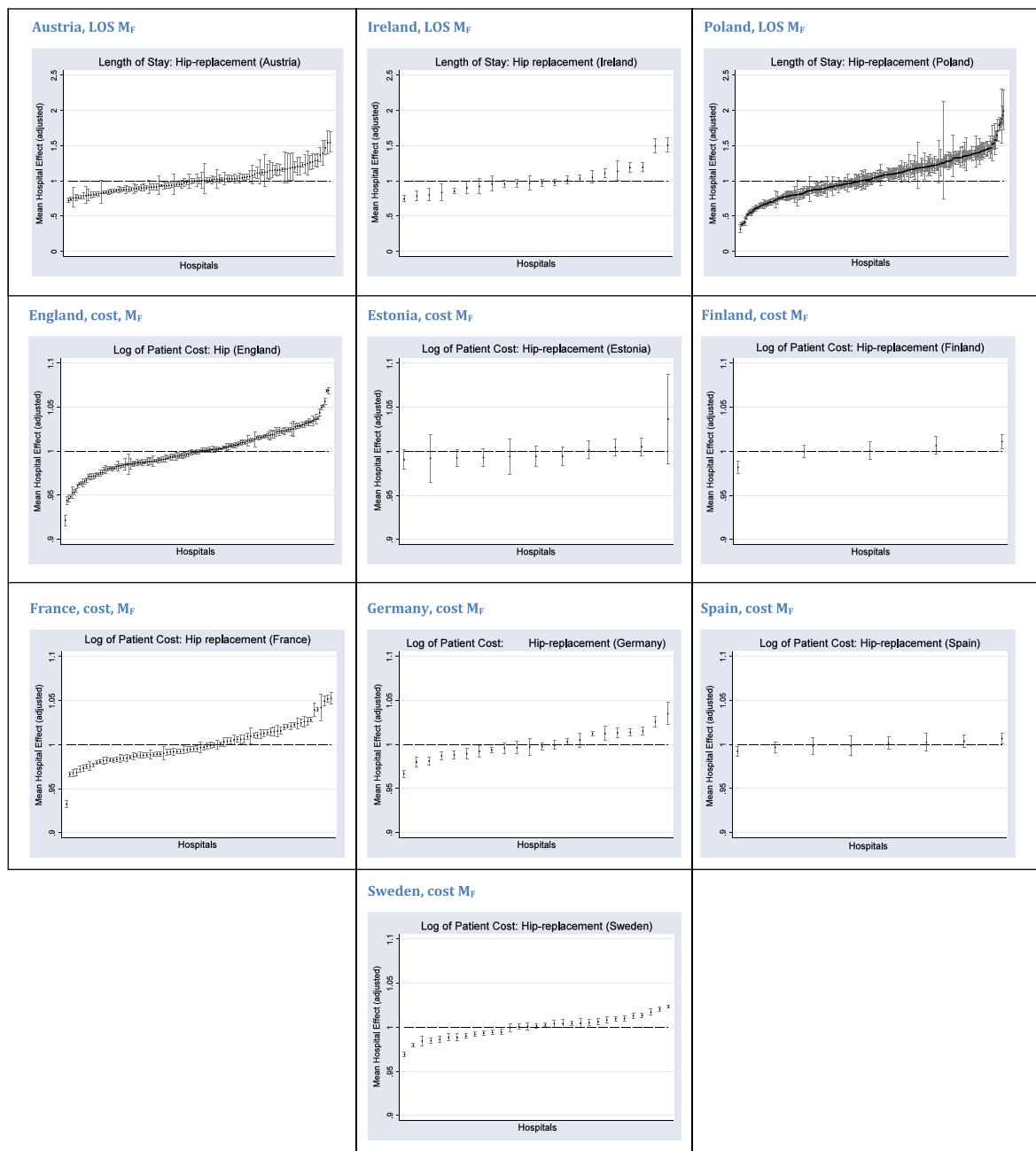


Figure 1. Variation in length of stay ($N=3$) and log of cost ($N=7$) in hospitals within countries

For other classification variables, we find that their operationalization seems to matter less than expected. For example, the mechanism to account for secondary diagnoses differs widely across countries, and much attention is paid to the definition of these measures in many national policy discourses (Busse *et al.*, 2011). However, some countries, such as the NordDRG countries and Austria, do not account for secondary diagnoses and, nevertheless, do not systematically perform worse than the other countries in accounting for resource consumption variation. Our analysis suggests that co-morbidities play only a minor role for resource consumption

variation of hip replacement patients as, for example, Charlson index co-morbidities and adverse events are not significant explanators of hospital costs. Moreover, co-morbidity measures taking into account the number and type of diagnoses, such as the patient cumulative complexity level used in Germany or Ireland, should – theoretically – perform better than simple measures. However, our results do not confirm this hypothesis because the total number of diagnoses remains significant in the full model (M_F) in all countries except Finland.

It is also interesting to note that countries with a higher amount of DRGs to group hip replacement cases do not necessarily perform better than those with fewer groups. Considering the cost analyses, the adjusted R^2 of M_D for Sweden (two DRGs) is in the same range of approximately 30% as the adjusted R^2 of Germany with eight DRGs. Despite this, the English HRG system with 14 DRGs has the greatest ability to explain differences in resource consumption (adjusted R^2 of M_D : 50%). Similarly, the adjusted R^2 of M_D for LoS analyses show that Ireland with three DRGs has an adjusted R^2 similar to Poland with six DRGs. This suggests that the number of DRGs for hip replacement is not per se a good predictor for the ability of DRG systems to explain differences in resource consumption.

In some countries (e.g. Austria, Finland, Ireland and Sweden), DRGs are used for measuring hospital activity or for adjusting global hospital budgets rather than to directly determine payment rates (Geissler *et al.*, 2011). In these contexts, it might be less relevant for stakeholders whether DRGs precisely account for patient-level costs variation. However, especially in countries that perform poorly in our comparative analyses and where hospital revenue is directly linked to DRG-based payments, policymakers should consider revisiting their grouping algorithms for hip replacement patients.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

The authors are grateful to Zeynep Or and Mikko Peltola for their comments on earlier drafts. Moreover, the authors thank the journal's referees for their constructive support in improving this article.

REFERENCES

- Bitton R. 2009. The Economic Burden of Osteoarthritis. *The American Journal of Managed Care* **15**: 230–235.
- Bozic KJ, Katz P, Cisternas M, Ono L, Ries MD, Showstack J. 2005. Hospital resource utilization for primary and revision total hip arthroplasty. *The Journal of Bone and Joint Surgery. American Volume* **87**: 570–576.
- Braithwaite RS, Col NF, Wong JB. 2003. Estimating Hip Fracture Morbidity, Mortality and Costs. *Journal of American Geriatrics Society* **51**(3): 364–370.
- Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. 2007. Incidence and Economic Burden of Osteoporosis-Related Fractures in the United States, 2005–2025. *Journal of Bone and Mineral Research* **22**(3): 465–475.
- Busse R, Geissler A, Quentin W, Wiley M (eds). 2011. *Diagnosis-Related Groups in Europe – Moving towards transparency, efficiency and quality in hospitals*. Maidenhead: Open University Press.
- Clohisey JC, Calvert G, Tull F, McDonald D, Maloney WJ. 2004. Reasons for revision hip surgery: a retrospective review. *Clinical Orthopaedics and Related Research* **429**: 188–92.
- Cots F, Chiarello P, Salvador X, Castells X, Quentin W. 2011. DRG-based hospital payment: Intended und unintended consequences. In *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- Ellis RP. 1998. Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics* **17**: 537–555.

- Geissler A, Quentin W, Scheller-Kreinsen D, Busse R. 2011. Introduction to DRGs in Europe: Common objectives across different hospital systems. In *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- Gullberg B, Johnell O, Kanis JA. 1997. World-wide Projections for Hip Fracture. *Osteoporosis International* 7:407–413.
- Kimberly JR, de Pourville G, D'Aunno T (eds). 2008. *The Globalization of Managerial Innovation in Health Care*. Cambridge: Cambridge University Press.
- Kobel C, Thuilliez J, Bellanger M, Pfeiffer KP. 2011. DRG systems and similar patient classification systems in Europe. In *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- Kurtz S, Ong K, Lau E, Mowat F, Halpern M. 2007. Projections of Primary and Revision Hip and Knee Arthroplasty in the United States from 2005 to 2030. *The Journal of Bone and Joint Surgery. American Volume* 89:780–785.
- Levaggi R, Montefiori M. 2003. Horizontal and Vertical Cream-Skimming in the Health Care Market. DISEFIN Working Paper, 11/2003, Rochester, NY: Social Science Research Network. <http://ssrn.com/abstract=545583>. [5th January 2012]
- Linna M, Virtanen M. 2011. NordDRG: The benefits of coordination. In *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- Magee HF. 2003. The Hospital Data Project. *European Journal of Public Health* 13(3): 73–79.
- Martinussen PE, Hagen TP. 2009. Reimbursement systems, organizational forms and patient selection: evidence from day surgery in Norway. *Health Economics, Policy, and Law* 4: 139–158.
- Newhouse JP, Byrne DJ. 1988. Did Medicare's prospective payment system cause length of stay to fall? *Journal of Health Economics* 7: 413–416.
- OECD Health Data. 2011. http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT [17 September 2011]
- Stargardt T. 2008. Health service costs in Europe: cost and reimbursement of primary hip replacement in nine countries. *Health Economics* 17(S1): S9–S20.
- Street A, Kobel C, Renaud T, Thuilliez J. 2012. How well do Diagnosis Related Groups explain variation in costs and length of stay among patients and across hospitals: Methods for analysing routine patient data. *Health Economics* 21 (Suppl. 2): 6–18.
- Tan S, Serden L, Geissler A, Ineveld M, Redkop K, Heurgren M, Hakkaart-van Roijen L. 2011. DRGs and cost accounting: which is driving which? In *Diagnosis-Related Groups in Europe – Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- United Nations (UN), Department of Economic and Social Affairs, Population Division. 2007. World Population Prospects: The 2006 Revision, Highlights. Working Paper No. ESA/P/WP.202.
- Wiley M. 2011. From the origins of DRGs to their implementation in Europe. In *Diagnosis-Related Groups in Europe – Moving towards transparency, efficiency and quality in hospitals*, Busse R, Geissler A, Quentin W, Wiley M (eds). Maidenhead: Open University Press.
- Woolf AD, Pfleger B. 2003. Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization* 81:646–656.
- Wülker N. 2010. Hüfte, Becken, Oberschenkel. In *Orthopädie und Unfallchirurgie* (2.Edition), Wülker N (ed.). Thieme: Stuttgart, New York.