Empirical comparison of DRG variants using cardiovascular surgery data: initial results of a project at 18 German hospitals

Norbert Roeder, Bernhard Rochell, Christian Juhra, and Marcel Mueller

Norbert Roeder is head of the Centre of Medical Controlling at the University of Muenster, Germany. Bernhard Rochell is a member of the German Medical Association (BÄK). Christian Juhra is a member of the Centre of Medical Controlling at the University of Muenster, Germany. Marcel Lukas Mueller is a member of the Institute of Medical Informatics, University of Muenster, Germany. All authors are members of the DRG Research Group at the University of Muenster.

Abstract

In 2000, the responsibility for selecting a DRG variant for use in Germany was assigned to a body comprising representatives of hospitals and insurers called the Self-Administration Board (or Board in this paper). To help the Board, we applied cardiac surgery data from 18 German hospitals to eight different DRG variants. The error caused by bad coding quality could be minimized this way, since all diagnoses and procedures in cardiac surgery must be recorded for quality assurance purposes. To match the German code to the appropriate code required by the DRG variant, we created mapping tables whenever needed.

As far as cardiac surgery is concerned, the Australian AR-DRG and the French GHM variants provided the best medical relevance, while the AR-DRG variant considered the level of severity better. Other variants would have to be updated to better reflect the level of medical complexity. Three main causes for wrong grouping could be identified for all systems: incomplete mapping, not enough reference to multidisciplinary treatments, and system construction problems.

Reasons for this study

According to section §17b of the Hospital Financing Law (KHG) of May 2000, the German Self-Administration Board must select a classification system for use in reimbursing all hospital inpatient care from January 2003. The classification system must be a variant of Diagnosis Related Groups currently in use in at least one country for payment purposes. This is a unique challenge, since in all other countries DRGs so far are used either only for partial reimbursement or as only one of several factors used in budget determination.

The University of Muenster, the German Society for Thoracic and Cardiovascular Surgery (DGTHG) and the German Hospital Association (DKG) initiated a project to aid the selection process. Data for 1999 would be analysed for cardiac surgery in 18 different German hospitals in order to evaluate eight variants of DRGs, as follows:

1. Health Care Financing Administration DRGs, version 17.0 INTERNOVA (HCFA-DRGs)
2. All-patient DRGs, version 12.0 3M (AP-DRG)

3. Group homogènes de malades, France (GHM)
4. Refined DRGs, INTERNOVA (R-DRGs)
5. All Patient Refined DRGs, version 15.0 3M (APR-DRGs)
6. Australian Refined DRGs, version 4.1, Australia (AR-DRGs)
7. International All Patient DRGs, 3M (IAP-DRGs)
8. Leistungsgerechte Diagnosefallgruppen, Austria (LDF)

The selection cannot be made only on cardiovascular surgery data. The decision to start by analysing cardiovascular surgery data was influenced by the high quality of coding as a result of a well established quality assurance program in place throughout Germany. It was anticipated that results developed from the cardiovascular surgery data would serve as a template to investigate other specialties in due course.

Evaluation criteria

We established a set of clinical and logistical criteria for use in evaluating the eight variants, which may be summarised as follows.

- Is every case group unambiguous and reproducible?
- How do the existing German diagnosis and procedure coding systems (ICD10-SGB V and OPS 301) influence grouping and how do they need to be altered when the DRG variant is introduced?
- Are comorbidities and complications taken adequately into consideration?
- Are main and expensive procedures taken adequately into consideration?
- How sensitive is the system to 'extreme' cases and what is the level of residual cases? Do they remain infrequent, and are the residual groups adequately cost homogeneous?
- Is the system resistant to gaming (such as that undertaken through clinical documentation)?
- Are modern treatment procedures covered?
- Is the administrative health care system in Germany covered?
- Are the groups homogeneous in regard to costs?
- Are paediatric groups formed correctly?
- What changes to the DRG variant would be necessary?
- Is the system easy to use and easily computerised?

An early step involved study of classification logic, documentation standards, and any adaptations that had been made for the German market. If a system had never been used in Germany, completely new mapping tables were developed to match the German codes.

Detailed comparisons of the underlying theories and constructions of the eight variants have been published elsewhere (Fischer 2000; Rochell & Roeder 2000). Here we merely present an overview of the main attributes of each variant. Table 1 shows various performance indicators, and Table 2 shows the methods used to discriminate among patients according to severity.

Table 1: important characteristics of the investigated DRG variants

Variant	HCFA	AP	GHM	R	APR	AR	IAP	LDF
Version	17.0	12.0	5.6	17.0	12.0	4.1	N/A	LKF 2000
No. of MDCs	25	25	26 + 3	25	25	23	?	
No. of case groups	499	641	600	1.198	1.530	661	1.040	867
Cardiac groups	45	41	44	117	152	64	?	72
Edit (error) groups	5	5	8	5	5	7	?	?
Diagnosis classification	ICD9CM Vol. 1	ICD9 ICD10	ICD10	ICD9CM	ICD9CM	ICD10-AM	ICD10	ICD9 BMSG
Procedure classification	ICD9CM Vol 3	OPS-301	CdAM	ICD9CM Vol 3	ICD9CM Vol 3	ICD10 AM-3	OPS-301	BMSG-procedure catalog
	Map: OPS301			Map: OPS301	Map: OPS301			Map: OPS301
Multiple procedures lead to additive reimbursement	_	_	_	_	_	_	_	✓
Consideration of outpatient procedures / day care	_	_	✓	_	_	(✓)	_	(✓)
Used for reimbursement	✓	✓	✓	(✓)	(✓)	✓	_	✓
Benchmarking	✓	✓	✓	✓	(✓)	(✓)	_	(✓)
Quality assurance	(✓)	(✓)	(✓)	✓	✓	✓	_	✓

Why cardiac surgery?

An important requirement for evaluating DRG variants is a high quality of diagnostic and procedural coding and a solid cost base. Coding quality is an especially crucial criterion since incorrect coding leads to inaccurate clinical complexity levels. Missing codes usually result in allocation of the case into a lower group. Thus a systematic error would arise when looking if varying costs of different cases (with a range of severity) were taken into correct consideration. The systems are not responsible for incorrectly grouping badly documented cases.

Table 2: criteria for differentiation of severity levels

	HCFA	AP	GHM	R	APR	AR	IAP	LDF
Version	17.0	12.0	5.6	17.0	12.0	4.1	N/A	LKF 2000
Severity levels triggered by additional diagnoses	114 DRGs with/ without CC 8 other 60	127 DRGs with/ without CC MCC DRGs 2 other	106 GHM with/ without CMA 22/2 GHM with/ without CMAS	341 ADRG with 4 surg. and 3 med. CC levels	382 ADRG with 4 surg. or med. CC levels	409 ADRG with 1-4 surg., medical or other CC levels	348 ADRG with 3 surg. or med. CC levels	173 MEL 249 HDG with up to 4 LDF
Dimensionality (case groups/ reimbursement groups per case)	one	one	one	one	one	one	one	one to more
Special Groups for long term respiration, intensive care	✓	✓	_	✓	✓	✓	?	Intensive-day rates
Age split	Neonat. Pediatric Geriatric	Neonat. Pediatric	Neonat. Pediatric Geriatric	Neonat.	Neonat.	Neonat. Pediatric Geriatric	?	Pediatric Geriatric
Severity differentiation based on systematic case groups	_	_	_	✓	✓	✓	✓	✓
Consideration of status of discharge	✓	✓	✓	✓	✓	✓	✓	✓

Initial experiences with AP-DRG data in Germany also revealed an undercoding problem with respect to complications and comorbidity (CC) diagnoses. Every participating hospital coded an average of only 1.5 to 2.5 diagnoses per case (some only coded a single diagnosis per case). Therefore fewer patients were assigned to higher groups, like CC or MCC (major complication and comorbidity) AP-DRGs. The allocation into a higher-cost and payment group would have been triggered by additional diagnoses that increase the level of severity.

To avoid the effect of bad coding on the results of this evaluation, data from the cardiac surgery area was chosen. This is because in 1992 a nationwide quality assurance system was established that led to a higher quality of documentation of each case in comparison to general data from national or international hospitals.

Most diagnoses from cardiac surgery are coded by QUADRA or ICD-9-codes. QUADRA coding is a special German classification for heart surgery used in most of the German heart surgery units.

QUADRA codes can be mapped directly to ICD-9 or ICD-10 codes. For this reason, it was possible to provide the same data for various systems that required different diagnosis codes (ICD-9 or ICD-10). Procedural codes can be recorded at the same level of differentiation as diagnosis codes. Furthermore, additional diagnoses and post-operative complications are coded, so procedures or diagnoses that may have not been coded with QUADRA or ICD-9 could easily be generated.

Recent calculations on patient costs were also available. Finally, the field of cardiac surgery has the highest rate of per case payment. Thus, evidence could be derived for cost homogeneity of the cases that were assigned to the various groups.

Method of study

After drafting a description of the interfaces for transferring the data, 40 different cardiac surgery departments were contacted. Of these a total of eighteen universities and private and community hospitals agreed to cooperate and were able to fulfil the prerequisites for participation.

Data collection and processing

The first stage in data processing was done by the hospitals. Depending on the local IT infrastructure, the data was either simply downloaded from the IT system without modification or separately retrieved from the patient data management and quality assurance systems and then concatenated. The dataset was then imported into a specially developed program that checked for completeness and plausibility. Finally, the dataset was processed by each of the DRG groupers.

After importing the initial data, additional diagnosis and procedure data were generated from information derived from quality assurance data following a standardized protocol. In entering new codes, care was taken to avoid duplicates. The basis for this recoding process was the documentation of the quality assurance sheets. In most cases only unspecific diagnoses could be generated (.8 or .9 codes). In some cases coding was not possible at all, since the documented diagnosis could be coded as more than one code. The impact of this data derived from quality assurance will need to be investigated further. The occasional redundant diagnoses or procedures were marked and not considered during the later grouping process.

Mapping German ICD-9 and OPS-301 codes

To group German data using foreign grouper software, the codes for diagnoses and procedures had to be mapped to the codes used by the specific foreign grouper.

The diagnosis data were collected in 1999 and was therefore only in ICD-9 format. As ICD-10 codes are required by AR-DRG and GHM, it was necessary to map the ICD-9 codes to ICD-10 using a table provided by the German Institute for Medical Documentation and Information (DIMDI), version 3.0.

A direct, one to one, mapping was not possible for all cases because ICD-10 is more detailed than ICD-9. It was therefore necessary to use judgement based on clinical experience. The results may be debated, but this was preferable to simply mapping problematic cases into non-specific ICD-10 codes. We used the latter approach only for cases that could not be covered by the existing codes.

The mapping task is summarised in Table 3. Note that, even where the DRG variant used ICD-9, it was often necessary to define mappings across ICD-9 versions.

Table 3: Overview of coding systems used for diagnosis coding

Sample data	DRG variant	Diagnosis coding system
ICD-9	HCFA	ICD-9-CM-1
	AP	ICD-9 (D)
	GHM	CIM-10
	R	ICD-9-CM-1
	APR	ICD-9-CM-1
	AR	ICD-10-CM-1
	IAP	ICD-9 (D)
	LDF	ICD-9-BMSG

None of the DRG variants under test were able to accept German procedure coding (OPS-301 together with ICPM v1.1 in a few circumstances). Various mappings had therefore to be developed as summarised in Table 4.

Table 4: Overview of coding systems used for procedural coding

Sample data	DRG variant	Procedure coding system
OPS-301 / ICPM	HCFA	ICD-9-CM-3
	AP	ICD-9-CM-3, German adoption
	GHM	CdAM
	R	ICD-9-CM-3
	APR	ICD-9-CM-3
	AR	ICD-10-AM
	IAP	ICD-9-CM-3, German adoption
	LDF	BMSG-Catalogue

Some of the DRG variants had already been implemented in Germany, and near-complete mappings of procedure codes already existed. However, we had to check them for accuracy.

Problems were experienced where mapping tables had to be developed from scratch by the project team. This was the case for the GHM, AR-DRG and LDF variants. The OPS catalogue is sometimes very unspecific and no 1:1 matching was possible. We faced this problem in all coding systems (ICD-9, CdAM, ICD-10-AM-3, etc).

Where only 1-to-many or many-to-1 mapping was possible, we identified two other circumstances: where this affected assignment to a DRG, and where it did not. We devoted most effort to the former category.

During this project 128,586 OPS-301 procedures were documented, which could be translated into 1,251 different codes. Nearly all cases which were placed into the wrong group could be explained by mapping errors. Besides ICD-9 and OPS-301 codes, admission and discharge status had also to be mapped for use in the various DRG variants.

Grouping and analysis

After the source data files had been mapped and checked in other ways, the records had to be processed through each of the groupers.

We asked 3M Germany to group using its AP-DRG, APR-DRG, and IAP-DRG software. In the end, they were able only to provide results for AP-DRGs.

INTERNOVA was asked to assist by grouping to R-DRGs and HCFA-DRGs. We ourselves did the grouping for GHM, AR-DRGs and LDF. For the French variant, we used the grouper software 'Programme de Groupage' version 5.6 for GHM. Two Australian software packages (Flexdata and Dremsel-Visasys DrGroup) were used to group the data into AR-DRGs. Both are certified by the Australian government.

All the sample records were included in a single file that was input to each of the grouper software packages in more or less the same way. All available manuals as well as other sources of advice were used to resolve the problems we found in the various grouping packages. For further analysis, commercial software products (MS Excel 97, SPSS v9.0, UNISTAT v5) were used as well as software we developed.

Initial comparative results

As noted above, the research group had problems in accessing two of the software packages (APR- and IAP-DRG) within the timeframe for completion of the analysis. Therefore the following summary of findings is mostly restricted to the remaining six variants.

The grouper software

We noted earlier that 3M grouped our data using AP-DRGs, and INTERNOVA did the same for HCFA-DRGs and R-DRGs. Since then, we have obtained a copy of the 3M software and are currently testing it. No software product for R-DRGs is available thus far.

There are currently seven companies that can provide AR-DRG groupers, having met the quality standards of the Australian government (the Commonwealth Department of Health and Aged Care). During this project, we made use of two of the groupers which differed considerably in their performance.

The Flexdata software is sparse, and is simply a grouper module that is included in the company's battery of hospital information system products. In this respect, it is similar to the GHM grouper.

In contrast, the DrGroup product from Visasys is extremely fast (it is capable of processing 41,000 records in 8 seconds). It is easily adapted to different input files, and retrieves a lot of useful information via the user interface.

The patient data software in France is coordinated in a manner similar to the Australian system. The French government is responsible for developing the core of the grouper software which it then provides to commercial suppliers, who then can integrate it into their own products. Additionally, French software companies can develop proprietary grouper software, but the product has to be certified by the government (ie, it must give exactly the same results as the official grouper) before it can be distributed.

In summary, the software packages we used had been produced by independent companies working in different countries. They varied in ease of use and functionality, but performed similarly in terms of the core task of DRG assignment.

The marketplace has provided choice and good value for money at the high end. It seems wise to use a similar approach in Germany, whereby there is competition but users are protected by having an official agency to test and certify products before they can be made available in the marketplace.

Processing of the transferred diagnoses and procedures

As noted earlier, it was necessary to add diagnosis and procedure data to the basic set available for all types of patients in the discharge databases. This was mainly with respect to post-operative complications that are rarely documented for other purposes than quality assurance. The two sources in combination gave a satisfactory data set.

Unfortunately, only the Australian groupers had no limits on the number diagnoses and procedures. All other products, and especially the Austrian LDF software, restricted the number and any additional codes could not be taken into consideration when grouping.

Information provided by the grouper software

The different products tested revealed their underlying decision process in varying degrees. The HCFA-DRG, R-DRG, and AR-DRG groupers gave a large amount of information about each case as it was grouped. This included not only details about the DRG and MDC assignment, but also about the relevant additional diagnoses, procedures, age and so on.

The DrGroup package for AR-DRGs provided the most extensive feedback. It added details such as the PCCL and CCL values, cost weights, and important information on how each code triggered the grouping process. The LDF software also provides a wide range of useful additional information.

3M's AP-DRG grouper provided little more than the core items such as DRG, MDC and the New York State cost weight. More information, especially on how a group was found and how the various codes triggered the process, is not revealed. The results are consequently harder to interpret.

There was an incomplete German manual for the German adaptation of AP-DRG (with German procedure codes) which did not allow us to answer many questions about the grouping process. Furthermore, the most important appendices (CC diagnoses, MCC diagnoses, and hierarchy of diagnoses) were not available at all. To interpret the results, we had to refer to the American manual with ICD-9-CM diagnoses and procedure codes.

The feedback given by the GHM software is also not very extensive. It includes only MDC, DRG and error code for the grouping process.

Manual testing for the correctness of the grouping process is possible in theory - there is a clear algorithm in place, and therefore we can understand how diagnoses and procedures influence the severity of the case - but it is extremely time-consuming. Computer feedback on such matters as the cause of assignment to an error (or edit) DRG is quite critical for daily use.

We believe that grouper systems need not return many other kinds of information (such as cost weights) during interactive grouping as long as the fields are available in generated tables. However, we find that access to this information is a critical feature of any software package.

Mapping German codes: mapping versus native grouping

The 3M software for German AP-DRGs makes direct use of German OPS-301 codes. This is called native grouping, since no further mapping is necessary. 3M changed ICD-9-CM-1 diagnosis and ICD-9-CM-3 procedure codes, that are used in the USA, to the German ICD-9 / ICD-10 / OPS-301 codes.

However, this advantage to the mapping process claimed by 3M is disputable as a consequence of our experiences in this project. AP-DRG 117 may serve as an example. Table 6 contrasts the original ICD-9-CM-3 codes with the codes generated by native grouping.

Table 6: Mapping ICD-9-CM-3 to German OPS-Code (? = no 1:1 mapping possible)

AP-DRG 117 (cardiac pacemaker revision except device replacement)			
ICD9-CM-3	Original (AP V12.0)	OPS-301	Mapping OPS-301
3774	Int or repl lead epicar	5-378.3	PM + AICD lead correction (?)
3775	Revision of lead	5-378.3	PM + AICD lead correction
3776	Repl tv atri-vent lead	5-378.3	PM + AICD lead correction (?)
3777	Removal of lead w/o repl	5-378.1	PM + AICD lead removal
3779	Revis or relocate pocket	5-378.4	PM + AICD device relocation
3789	Revis or remove pacemacer	5-387.0	PM + AICD device removal (?)
		5-387.2	PM + AICD device + lead removal (?)

Changing of a probe is missing in OPS-301, as well as the differentiation between transvenous and epicardial insertion. OPS-301 is also not able to describe the revision of an aggregate. Codes marked with (?) are compromises that match ICD-9-CM-3 as closely as possible (but how closely remains to be investigated). The result is that OPS-301-native codes are unable to match codes 1:1 with any other mapping table. Negative results will occur especially if the native grouping or mapping is wrong. Every DRG grouping process is based on a correct set of diagnoses and procedures.

If it is not possible to map German ICD and OPS classifications correctly to the DRG variant, then precisely coded cases may be allocated to a totally inappropriate group. The possible effects on the deviation of cost weights in Germany can be easily imagined. 3M native grouping is no more resistant to mistakes than mapping solutions. For example, most of the patients with congenital heart disease were not allocated to the appropriate AP-DRG 809, but spread into other groups. Analysing this problem, it became clear that the German version of the grouper was not identical to its American original. The most likely reason for this is an inadequate native grouping. By entering the correct American codes directly, those patients were allocated to the correct DRG.

Thus the native grouping method is neither better nor worse than any other mapping method, as long as a test with a representative dataset shows how the codes generated by native grouping fit the original codes. Only then native grouping will bring an advantage for acceptance and precision of classification.

3M specifically marked their German manuals as drafts. Nevertheless, it is already safe to say that experts will have to check the complete mapping procedure of OPS-301 codes before its introduction in Germany. This means that the same effort will be needed as for other variants where no native grouping is possible.

Grouping results

The entire dataset from one hospital was discarded from final analysis due to a software error. Other occasional implausible data was excluded as well. From the original 41,444 cases 40,061 were used for final analysis.

Nearly all the groupers sorted 90% of all data into 20 DRG classes. The only exceptions were R-DRGs (30 groups) and LDF (23 groups), as they are providing a higher number of total groups and thus a wider distribution.

By restricting analysis to cardiac surgery cases only, the distribution is even smaller (Table 7). Nearly 99% of all cases were included in these 20 groups. Again, the outlier R-DRG variant divided only 95% among the 20 most common groups.

Table 7: grouping results in cardiac surgery

	HCFA-DRG	AP-DRG	GHM	R-DRG	AR-DRG	LDF
Maximum number of groups	499	641	600	1.198	661	867
Groups in sample (including error groups)	53	66	52	106	86	55
Sample as proportion of total (%)	10,6	10,3	8,7	8,8	13,0	6,3
Number of cases in the 20 strongest groups	31333	31233	31307	29849	31234	31337
Proportion of cases (%)	99,8	99,5	99,7	95,0	99,4	99,8
Number of groups with 90% of cases	5	7	5	15	8	3

The number of groups that contain 90% of all cases may serve as a hint to effective granularity (ie, the degree of differentiation) of a variant. Due to the higher differentiation with respect to level of severity, the R-DRG variant was dominant here with 15 groups, followed by the AR-DRG variant with 8 groups (out of 661 possible groups). The lowest level of differentiation can be observed in the GHM variant with only 5 groups. It is worth noting that most of the variants needed no more than two groups to sort more than 50% of all cardiac surgery cases. These groups are bypass and valve surgery. Even the R-DRG variant with its high granularity only needed three groups.

Characteristics of effective granularity are also indicated by the gradient of reduction of cases in each group. Figure 4 shows the distribution of all cases within the 11 largest groups for each of the investigated variants.

The first few groups generated by the HCFA-DRG, AP-DRG, GHM and LDF variants contain a large proportion of the codes and therefore have a narrow distribution. We found the distribution to be wider when using R-DRG and AR-DRG.

Using R-DRG or AR-DRG, each of the ten largest groups contained at least 500 cases, whereas the use of the other variants resulted in more groups with a smaller number of cases within those groups. As an example, the HCFA-DRGs created only six groups with more than 500 cases.

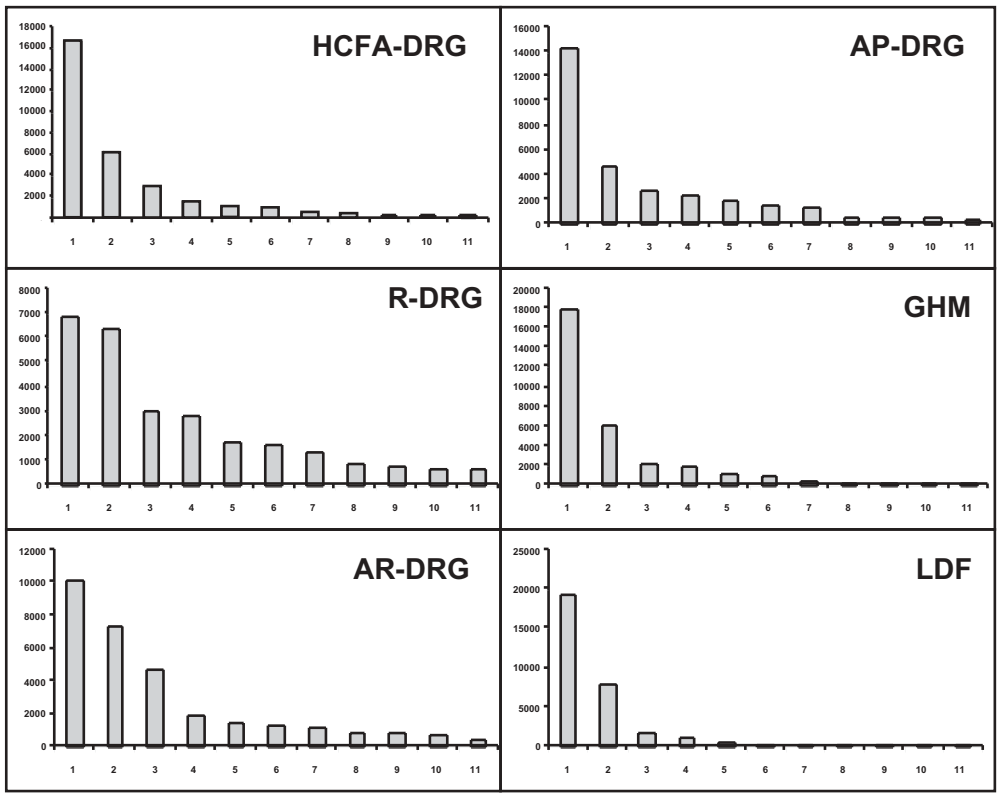
More details are shown in Tables 8 to 13. Within the top ten cardiac surgery groups, the highest similarity exists between HCFA-DRG, AP-DRG and GHM. We found no difference between the top six groups of the HCFA and GHM variants. This is due to the fact that the GHM variant was built on the HCFA-DRG model. The slight difference in numbers between these two variants may reflect further independent software development or mapping effects.

The AP-DRG and the HCFA-DRGs resemble each other as they also have similar origins. Since the AP-DRG variant subdivides the codes further (based on the MCC severity concept), it differs more from HCFA than GHM does. This is because GHM has not adopted the MCC concepts for the major group of cardiovascular diseases (CMD5).

These three variants differ in three other areas. The HCFA-DRG and the AP-DRG variants provide a case group to which cases with long term respiration are allocated via Pre-MDC logic (cases which are allocated directly to a DRG due to special circumstances). This is defined by tracheostomy (except for ENT and oral-facial surgery), and it is not differentiated further according to main diagnoses and procedures.

The GHM variant does not have a similar group. This implies that within the French variant all cases with long term respiration are allocated to groups according to the underlying disease. In contrast, all other variants use indicators such as tracheostomy to allocate those cases directly to a special group that is not directly related to the principal diagnosis. Thus this group is very heterogeneous in accordance with diagnosis, but includes all cases with considerable intensive care.

Figure 4: distribution of cardiac cases in the 11 most frequent DRGs in descending order



The various DRG variants also differ by special groups for paediatric cardiac surgery. While such a group does not exist within the HCFA-DRG variant, there are two GHM classes for children under two years of age with and without usage of the pump (ie, extra-corporeal circulation during surgery).

AP-DRGs are unique in how they allocate newborns under 29 days of age: according to this age the case is allocated to MDC15 (diseases of the newborn), and not further differentiated by cardiology / cardiac surgery characteristics. It is differentiated by birthweight and status of discharge. Additionally, the AP-DRG variant is the only one that contains a special group for congenital heart diseases being surgically treated. All other systems place these cases in common groups with cardiac surgery, although sometimes with an additional split by age.

The R-DRG variant resembles the HCFA-DRG variant closely, since the basic DRGs are derived from that variant. Nevertheless, the R-DRG variant has a higher level of differentiation by severity. Compared to the HCFA-DRG system, the top ten most frequent R-DRGs that are derived from HCFA-DRGs are ranked as 1 to 3 and 4 to 6 (Table 11). The variation in these groups is significantly lower than in all other DRG variants as a consequence of the larger number of cases.

Within the AR-DRG variant, the cases are distributed evenly among the levels of severity within one adjacent DRG (eg, DRGs F06A and F06B). Long-term respiration is detected not only by tracheostomy but also by duration of respiration. A patient has to be temporarily tracheotomised or be on mechanical respiration for more than 95 hours (but may not belong to MDC22, burns, or be younger than 16 years). 444 cases were allocated to this Pre-MDC DRG A06Z. In Australia, a further very important group exists: A04Z (ECMO without cardiac surgery).

Sometimes problems arose when the LDF system was confronted with multiple procedures. The procedures are sorted by predefined scores, according to the manual, but this cannot always be confirmed by checking the results manually. Why this happens has to be further investigated with the support of Austrian experts.

The Austrian system groups are similar to the Refined DRG variants, starting with common case groups (MEL and HDG groups), then assigning a definite LDF by variable criteria for the levels of severity. The assignment of procedures for additional diagnosis to the LDF groups could not be evaluated, so all that is said about the granularity of this system only refers to the assignment of LDF groups (Table 13).

Table 8: top 10 HCFA-DRG classes (cardiac cases)

	HCFA		Cases	%	Cum %
1	109	CORONARY BYPASS W/O CARDIAC CATH	16592	52,83	52,83
2	105	CARDIAC VALVE PROCEDURES W/O CARDIAC CATH	6205	19,76	72,58
3	108	OTHER CARDIOTHORACIC PROCEDURES	2903	9,24	81,82
4	107	CORONARY BYPASS W CARDIAC CATH	1689	5,38	87,20
5	116	OTH PERM CARDIAC PACEMAKER IMPLANT OR PTCA W CORONARY ARTERY STENT IMPLNT	1179	3,75	90,95
6	104	CARDIAC VALVE PROCEDURES W CARDIAC CATH	952	3,03	93,99
7	118	CARDIAC PACEMAKER DEVICE REPLACEMENT	494	1,57	95,56
8	483	TRACHEOSTOMY EXCEPT FOR FACE,MOUTH & NECK DIAGNOSES	428	1,36	96,92
9	103	HEART TRANSPLANT	233	0,74	97,66
10	110	MAJOR CARDIOVASCULAR PROCEDURES W CC	194	0,62	98,28

Table 9: top 10 AP-DRG classes (cardiac cases)

	AP-DRG		Cases	%	Cum %
1	107	CORONARY BYPASS W/O CARDIAC CATH	14222	45,28	45,28
2	105	CARDIAC VALVE PROCEDURES W/O CARDIAC CATH	4646	14,79	60,07
3	546	CORONARY BYPASS W CC	2753	8,77	68,84
4	108	OTHER CARDIAC PROCEDURES W/O CONGENITAL ANOMALIES	2287	7,28	76,12
5	545	CARDIAC VALVE PROCEDURES W CC	1910	6,08	82,20
6	106	CORONARY BYPASS W CARDIAC CATH	1427	4,54	86,74
7	116	OTH PERM CARDIAC PACEMAKER IMPLANT	1283	4,08	90,83
8	104	CARDIAC VALVE PROCEDURES W CARDIAC CATH	600	1,91	92,74
9	547	OTHER CARDIAC PROCEDURES W CC	519	1,65	94,39
10	483	TRACHEOSTOMY EXCEPT FOR FACE,MOUTH & NECK DIAGNOSES	421	1,34	95,73

Table 10: top 10 GHM classes (cardiac cases)

	GHM		Cases	%	Cum %
1	155	CORONARY BYPASS W/O CARDIAC CATH	17857	56,85	56,85
2	153	CARDIAC VALVE PROCEDURES W/O CARDIAC CATH W PUMP	6156	19,60	76,45
3	154	CORONARY BYPASS W CARDIAC CATH	2059	6,56	83,01
4	164	PERM CARDIAC PACEMAKER IMPLANT	1900	6,05	89,06
5	170	OTHER CARDIAC PROCEDURES W PUMP	1189	3,79	92,84
6	152	CARDIAC VALVE PROCEDURES W CARDIAC CATH W PUMP	792	2,52	95,36
7	172	OTHER CARDIAC PROCEDURES W/O PUMP	519	1,65	97,02
8	151	HEARTTRANSPLANTATION	236	0,75	97,77
9	166	PERM CARDIAC PACEMAKER REPLACEMENT	210	0,67	98,44
10	901	PROCEDURE W NO RELEVANCE TO PDX	116	0,37	98,81

Table 11: top 10 R-DRG classes (cardiac cases)

	R-DRG		Cases	%	Cum %
1	1093	CORONARY BYPASS W/O PTCA OR CARD CATH WITH CLASS A CC	6833	21,75	21,75
2	1090	CORONARY BYPASS W/O PTCA OR CARD CATH WITH NO CC	6378	20,31	42,06
3	1092	CORONARY BYPASS W/O PTCA OR CARD CATH WITH CLASS B CC	2981	9,49	51,55
4	1053	CARDIAC VALVE PROC W/O CARD CATH WITH CLASS A CC	2780	8,85	60,40
5	1050	CARDIAC VALVE PROC W/O CARD CATH WITH NO CC	1773	5,64	66,05
6	1052	CARDIAC VALVE PROC W/O CARD CATH WITH CLASS B CC	1602	5,10	71,15
7	1083	OTHER CARDIOTHORACIC PROCEDURES WITH CLASS A CC	1376	4,38	75,53
8	1080	OTHER CARDIOTHORACIC PROCEDURES WITH NO CC	836	2,66	78,19
9	1160	OTH PERM CARD PACE OR PTCA W STENT WITH NO CC	766	2,44	80,63
10	1073	CORONARY BYPASS W CARD CATH WITH CLASS A CC	678	2,16	82,79

Table 12: top 10 AR-DRG classes (cardiac cases)

	AR-DRG	AR-DRG	Cases	%	Cum %
1	F06A	CORONARY BYPASS W/O INVASIVE CARDIAC INVES PROCEDURE W CATASTR OR SEVERE CC	10189	32,44	32,44
2	F06B	CORONARY BYPASS W/O INVASIVE CARDIAC INVES PROCEDURE W/O CATASTR OR SEVERE CC	7313	23,28	55,72
3	F04A	CARDIAC VALVE PROC W PUMP W/O INVASIVE CARDIAC INVES PROC W CAT OR SEV CC	4666	14,86	70,58
4	F04B	CARDIAC VALVE PROC W PUMP W/O INVASIVE CARDIAC INVES PR W/O CAT OR SEV CC	1815	5,78	76,36
5	F17Z	CARDIAC PACEMAKER REPLACEMENT	1437	4,58	80,93
6	F05B	CORONARY BYPASS W INVASIVE CARDIAC INVES PROCEDURE W/O CATASTROPHIC CC	1222	3,89	84,82
7	F01Z	IMPLANTATION OR REPLACEMENT OF AICD, TOTAL SYSTEM	1097	3,49	88,32
8	F07Z	OTHER CARDIOTHORACIC/VASCULAR PROCEDURES W PUMP	838	2,67	90,98
9	F03Z	CARDIAC VALVE PROC W PUMP W INVASIVE CARDIAC INVES PROCEDURE	792	2,52	93,51
10	F05A	CORONARY BYPASS W INVASIVE CARDIAC INVES PROCEDURE W CATASTROPHIC CC	700	2,23	95,73

Table 13: top 10 LDF classes (cardiac cases)

	LDF		Cases	%	Cum %
1	MEL08.02	CORONARY BYPASS, OTHER PROCEDURES W PUMP	19351	61,61	61,61
2	MEL08.03	SINGLE VALVE PROCEDURE	7926	25,23	86,84
3	MEL10.01	PACEMAKER IMPL	1586	5,05	91,89
4	MEL10.02	DEFIBR IMPL	990	3,15	95,05
5	MEL08.01	HEART PROCEDURES W/O PUMP	406	1,29	96,34
6	MEL08.04	DOUBLE VALVE PROCEDURE	250	0,80	97,13
7	MEL18.05	HEART TRANSPLANTATION	214	0,68	97,82
8	MEL07.04	PLASTIC SURGERY ON CHEST WALL	192	0,61	98,43
9	MEL08.07	REKONSTR OF AORTA	138	0,44	98,87
10	MEL09.01	REKONSTR OF PERIPHER VESSELS	63	0,20	99,07

Covering multiple treatments

Multiple procedures are not adequately covered by the evaluated DRG variants. Mostly they are not taken into account at all. Thus a patient with surgery on both carotid arteries is allocated to the same group as a patient with surgery on only one side. The same applies to additional procedures of different nature.

The way the procedure influences the grouping process is heavily dependent on the principal diagnosis. With the exception of special cases like organ transplantation, a pre-selection of the cases is done on the basis of principal diagnosis. This leads to an allocation to a specific, organ-related, group (eg, diseases of the cardiovascular system). Thus far, the order of the documented procedures is of no importance. Only now the documented procedures are sorted to find out the relevant main procedure.

In most DRG classifications, only those procedures related to the principal diagnosis and being in specific lists for every MDC are taken into consideration. The case presented in Table14, a stenosis of the carotid artery (which is very common in cardiac surgery) is placed in MDC 1 (diseases of the nervous system, which also includes diseases of the associated blood vessels) by all DRG variants.

Because only surgery on the carotid artery is listed for this group, and not the bypass surgery which belongs to the procedure list for MDC 5 (diseases of the cardiovascular system), the DRG variant only considers the surgery on the carotid artery regardless of marking the bypass surgery as main procedure.

The bypass surgery does not even indirectly lead to an increase in complexity level. It is completely ignored. If sclerosis of the coronary arteria were the principal diagnosis, this would lead to the identification of bypass surgery as main procedure, and the case would be allocated to the MDC 5.

Table 14: an example of grouping without considering multiple procedures

Age: 64, Gender: female
Length of stay: 48d, ICU: 7d

Diagnoses

No	ICD 9	Description
1	433.1	Stenosis of carotid artery
2	414.0	Sclerosis of coronary artery
3	V15.1	History of surgical procedure on heart or vessels
4	272.0	Elevated cholesterol levels
5	401.9	Essential hypertonia
6	997.0	Complications of the CNS
ICD9CM-1:		43310, 2720, 4019, 41400, 99700, V151

Procedures

No.	OPS-301	Description
1	5-361.24	Triple ACB-Bypass with autog. Vein + Art.
2	1-275.0	Transart. Cath.
3	5-382.0	Resect. of vessels w reanastomosis head / neck, extracranial
4	5-381.54	Resect. of external iliac artery
5	8-851.0	Usage of pump

Reimbursement

	Code	Description
SE	9.08	ACB (artery and vein)
SE	10.08	Proc. on artery w/o prothesis
SE	10.01	Proc. on carotid artery
SE	21.01	Cardiac cath.

DRG Groupings

AP-DRG	Proc. On extracranial vessels
AR-DRG	B04B S : extracranial Vascular Procedures W/O Catastrophic
GHM	Proc. on pre cerebral vessels
LDF	MEL08.02 Bypass procedure with pump
HCFA-DRG	005 EXTRACRANIAL VASCULAR PROCEDURES
R-DRG	0052 EXTRACRANIAL VASCULAR PROCEDURES WITH CLASS B CC
TD:	99700 NERVOUS SYST COMPLC NOS

This dominance of the principal diagnosis over procedures can lead to a disadvantage of cases with multiple procedures or multidisciplinary cases, especially when the additional effort is considered adequately by additional diagnosis, as the former example showed. This implies that the same case can be differently grouped depending on documentation. This may be a wrong stimulus for case management.

The LDF variant is the only variant that includes additional components to the individual case groups, by which the additional resource consumption can be reimbursed. It is also the only system recognizing bypass surgery as main procedure, leading to a clear advantage of the variants regarding multiple procedures.

When implementing a German solution in the future, it may be better to classify and sort procedures regardless of the MDCs. Thus all performed procedures and not only those belonging to the specific principal diagnosis would be taken into consideration for grouping. Combinations of procedures that occur frequently could be assigned to a special group to avoid systematically wrong weights. Procedures which can be performed in combination with different other procedures and are not considered in definition of case groups must be identified by the system. Then it will be possible to allocate the case by using a modular approach.

Treatment of multiple conditions during one hospitalisation

Depending on the size of a hospital, a single patient with multiple diseases can be treated by more than just one department in the same hospital. Again, an example is used to explain this. Table 15 shows one patient who was treated for a total of 83 days.

Table 15: Long term stay with multiple diseases

Age: 75, Gender: female
Admission: Nov 19 1998 Discharge: Feb 2, 1999
Length of stay: 83 days, includes ICU = 2 days, cardiac surgery =10d

Diagnoses

No	ICD 9	Description
1	250.5	Diabetes with neurologic compl.
2	414.0	Sclerosis of coronary artery
3	433.0	Stenosis of basilar artery
4	433.1	Stenosis of carotid artery
5	707.1	Ulcer of the lower extremity
6	785.4	Gangrene
7	V15.1	History of surgical procedures on heart or vessels
8	V13.8	Other history of diseases
9	401.9	Essential Hypertonus
10	250.0	Diabetes mellitus w/o compl.

Procedures

No	OPS-301	Description
1	1-275.2	Cardiac cath.
2	5-395.0	Extracranial patch of vessels (head / neck)
3	5-381.0	Resect. of vessels head / neck, extracranial
4	5-361.21	Triple bypass with autog. Vein
5	8-851.0	Usage of pump

Reimbursement

	Code	Description
SE	9.09	ACB (only vein)
SE	10.01	Procedure on carotid artery

DRG Groupings

AP-DRG	Proc on extracranial vessels
AR-DRG	K09Z Other Endocrine, Nutritional and Metabolic O.R. Procedures
GHM	411 Other Endocrine, Nutritional and Metabolic Procedures, Age >69y
LDF	MEL08.02 Bypass, other procedures with pump
HCFA-DRG	005 EXTRACRANIAL VASCULAR PROCEDURES
R-DRG	0053 EXTRACRANIAL VASCULAR PROCEDURES WITH CLASS A CC

The diagnosis using most of the resources was severe diabetes, but there were extensive additional diseases. After being treated for more than two months in the diabetes department, the patient developed symptoms of coronary heart disease and was thus referred to cardiac surgery in the same hospital. After bypass surgery and surgery on the carotid arteriae, showing signs of stenosis as well, she was referred back after 10 days and was discharged from the hospital two weeks later.

So the same patient was treated for two (or three, including the stenosis of the carotid arteriae) diagnoses in the same hospital. The DRG grouping is either done by stenosis of the carotid arteria (AP-DRG, HCFA-DRG, and R-DRGs) or by diabetes (AR-DRGs, GHM). LDF at least recognised the procedure using the most resources (bypass surgery) as being the main procedure.

This patient cannot be grouped adequately by any of the variants. Most likely, the case would be assigned to an exception group due to the long duration of stay. Nevertheless, this case shows the problems that arise when a patient is treated by more than one department for more than one disease during the same hospital stay.

It is likely that the probability of these cases correlates directly with the number of departments in the hospital. A hospital with few departments will refer the patient to the necessary department of another hospital, which means a new DRG. A hospital with more departments will refer to a department within the same hospital.

To avoid discharge and new admission of the patient to create a new DRG case, a solution has to be found. Maybe more than just one DRG can be reimbursed, if different episodes of treatments are clearly recognizable. With our example, two DRGs could be reimbursed. Nevertheless, this is only a solution if the different episodes are clearly separable.

Specificity of the systems when identifying extensive modern treatments: implantation of defibrillator

Table 16 shows the grouping of 1,021 cases in which only an automatic defibrillator was implanted or changed. The total number of DRGs to which these patients are allocated is an indicator of the precision of the grouping process, since the same information was given to every variant. It also serves as an indicator on how up-to-date the variant is from a medical point of view.

Table 16: implant or replace defibrillator - number of DRGs used in the different systems

AICD	HCFA	AP	GHM	R-DRG	AR	LDF
# of groups	10	12	15	19	5	9
# of groups that include 90%	3	4	2	8	1	1
# of cases in the two most frequently used groups	79.82	66.60	90.72	50.5	99.71	97.94

The fewer groups were used, the better the grouping that had been done. Only the LDF- and AR-DRG variants grouped most of the patients (97.94% for LDF, 99.71% for AR-DRG) in two of the most frequently used groups, closely followed by the GHM-system with 90.72%. This good result is due the special consideration of procedures in the LDF variant, and to a special group for defibrillator patients with procedures in the Australian system.

In the HCFA-DRG, R-DRG, and AP-DRG variants, the cases are allocated to various DRGs (mostly heart valve surgery). As the results show, this leads to a splitting into different groups depending on the principal diagnosis.

The GHM variant allocates these cases to the pacemaker group, which seems to be more adequate than valve surgery. Depending on principal diagnosis, the cases were allocated to different groups in the HCFA-DRG, R-DRG and AP-DRG systems.

In the AP variant, the largest part of AICD cases is allocated to the group with heart valve failure. Assuming a base rate of 2,500 (= Euros), the cost weights were multiplied by this rate to calculate the final rate. Figure

5 shows the results.

Figure 5: an example of same therapy and different reimbursement AP-DRG handling of implantation or replacement of cardiac defibrillator

Primary diagnosis: 428.1 (insuffiency of the left heart)	
ICD9CM-3 3795 Implantation cardio defibrillator leads	ICD9CM-3 3794 Implantation/Repl. cardio defibrillator total
3796 Implantation cardio defibrillator generator	3795 Implantation cardio defibrillator leads &
3797 Replacement cardio defibrillator leads	3796 Implantation cardio defibrillator generator
3798 Replacement cardio defibrillator generator	3797 Replacement cardio defibrillator leads
	3798 Replacement cardio defibrillator generator
Device Rpl. ~ 20,000.- leads to AP-DRGs:	Device + Electrode Rpl ~ 21,000.- leads to AP-DRGs:
⇩	⇩
116: Other Permanent Cardiac Pacemaker Implantation Or AICD Lead Or Generator Procedures (9,495.-*)	104: Cardiac Valve Procedures W Cardiac Cath +(25,230.-*)
548 **: Other Cardiac Pacemaker Implantation or Revision	105: Cardiac Valve Procedures W/O Cardiac Cath (15,462.-*)
Or AICD Procedures (17,147.-*)	545 ** Cardiac Valve Procedures W Major CC (35,250.-*)
MDC 5 **+ add. diagnosis: 427.41 Tachycardia	*Base rate = 2,100.-

If insufficiency of the left heart is the principal diagnosis and a defibrillator is implanted or changed due to tachycardia, the case is grouped either to the AP-DRG 116 (changing of generator) or with additional diagnosis tachycardia into AP-DRG 548 (pacemaker revision), which scores a higher rate.

The patient will be grouped either into AP-DRG 104 or 105 (with or without catheter) when a complete AICD system (generator and lead) is implanted or changed. Tachycardia is a dangerous diagnosis in heart valve diseases but not in AICD cases, because it is the reason for the AICD implantation in combination with heart failure. All AICD patients suffer from tachycardia, but tachycardia as a secondary diagnosis switches the patient to the MCC AP-DRG 554. As a result, all patients are MCC cases.

A little variation in coding (additional diagnosis or procedure codes for lead implantation) has great consequences in this case. On the left side of Figure 5, the reimbursement ranges from 9,445 to 17,147, and on the right side from 15,475 to 35,250.

An AICD is not a pacemaker and not a heart valve. The Australian AR-DRG variant considers this and offers a separate group for AICD implantation. 1,016 cases were assigned to this group, which is shows the correct coding. Only five cases (0.3%) had a wrong principal diagnosis or were grouped elsewhere.

The LDF variant also has a special group for AICD implantation (MEL 10.02A). Sometimes the pump had to be used when changing the AICD and to reconstruct the heart valve which had been damaged by it. The variant identified these cases correctly and allocated them to the groups MEL 8.02 (other procedures with pump) and MEL 8.03 (surgery on one valve). Why 28 cases were grouped into MEL 10.01 (implantation of pacemaker devices) cannot be explained. In these cases a pacemaker had to be removed before an AICD was implanted. A further analysis on this will follow.

The groupings of the HCFA- and R-DRG variants were basically the same as in the AP-DRGs, showing the close relation of these variants. The GHM variant allocated most of the cases into the pacemaker GHMs (164 and 166).

Unusual and high-cost procedures

Extensive procedures like organ transplantations are grouped into special Pre-MDCs, regardless of the principal diagnosis. Thus they can be identified solely on the procedure before being grouped into a MDC based on the principal diagnosis. All investigated variants use this general approach. They differ in degree and details of method.

Only the Australian variant and the AP-DRGs use Pre-MDCs for heart transplantation. The other variants group these cases into MDC 5 (diseases of the cardiovascular system) or into a special group for transplantation as the GHM does. This group is tested like other Pre-MDCs prior to other MDCs and can thus be treated as a Pre-MDC. The Australian variant additionally offers a group for multiple organ transplantation (Table 17).

Table 17: pre-MDC classes for transplantation

Variant and version	HCFA 17.0	AP 12.0	GHM 5.6	R-DRG 17.0	APR 12.0	AR 4.1
Lung transplant	495	795	116 CMD 27	4950-03	484	A03Z
Heart transplant	103 MDC 5	103 MDC 5	151 CMD 27	1030-1033 in MDC 5	103 MDC 5	A05Z
Multiple transplant						A02Z

Differences in grouping results may already be caused by other parameters such as age and weight. If a 21-day-old patient is transplanted, he would be grouped into AP-DRG 622. If the patient were 21 years old, he would be grouped into AP-DRG 103. This is caused by splitting for age, which is done by the system to identify neonatology cases under the age of 29 days (MDC 15).

No exception is made to this rule, so that newborns that would better be assigned to another group may be allocated to this group. A newborn who undergoes a heart transplantation will always be grouped into MDC 15, and not into the DRG specifically created for transplantation.

If the age is changed, then the patient will be assigned into the transplantation MDC, because the age splitting is done prior to considering the transplantation procedure, resulting in neglecting transplantations in newborns in AP-DRGs.

The AR-DRGs instead neglected the age and grouped all patients into the transplantation group, which seems to make more sense. The HCFA-DRGs and R-DRGs grouped a combined heart and lung transplantation into the heart transplantation DRG. The AP-DRGs did not group the correctly coded procedures (OPS-301: 5-375.2) into the heart transplantation group, but to the lung transplantation group.

Different New York State cost weights are assigned to these two procedures. This is most likely a mapping problem of the German version of 3M grouper software. Since no manuals were available, this hypothesis could not be tested.

The GHM variant recognized all cases, except two, correctly as transplantations. A wrong discharge date was entered for one of these patients on the day of admission, so this patient was grouped into an outpatient group. The other patient died on the day of transplantation, thus the discharge date was also the same as the transplantation date, this time correctly. The system grouped the patient into GHM 880 (sudden death, stay <24h). The grouping for these cases is questionable.

One case with heart transplantation and principal diagnosis ICD-9 239.8 (unknown neoplasm, other location) was grouped incorrectly by the HCFA- and R-DRG system into a haemato-oncology group with extensive procedures, while all other systems did group this case correctly. This error is due to the missing pre-MDC for heart transplantation.

Within the R-DRG variant, lung transplant patients are grouped into DRG 4830 for patients with long term respiration. The AR-DRG system is able to group nearly all cases correctly. It also offers a special group for multiple organ transplantation. It is interesting to look at the distribution of different levels of severity among heart transplant patients. 146 of 236 cases were assigned to one level, 90 cases even to highest level 4.

The LDF variant had the highest number of incorrect grouping results. As soon as other procedures were done on the patient (eg, due to complications), those seems to be regarded with priority by the system. The reason may be the order of procedures or the incorrect application of the variant by the project team. This problem will have to be discussed, as well as all other questions, with the suppliers. Combined heart-lung transplantation is also recognized by the system and assigned to a separate group.

It is important that problems with cases that cause extensive costs are reliably placed into an error group to be identified and then corrected. If they are placed (as in the LDF-, HCF-/R-DRG variants) into a totally different, but nevertheless valid group, they would not attract attention and be lost to any correction.

New therapies

More than other variants, the AR-DRG system creates special groups for new therapies. For example, the pump is used as an indicator for extensive disease. The special group for AICD has been mentioned above. There is an explicit difference between surgery with and without the pump (also a feature of the French GHM variant). ERCP and PTCA as therapeutic procedures are handled sensibly. The system identifies implants of coronary stents (as do the HCFA-DRGs and R-DRGs).

By formation of such special groups, new therapies can be better controlled. Exclusive to Australia's system is also the ECMO DRG (A04Z), implemented as Pre-MDC to avoid unreasonable biasing of the cost weight of the corresponding cardiac surgery group. This very expensive therapy is performed after heart failure following cardiac surgery with extreme complications, and is available only at a small number of hospitals. It is an indicator of high morbidity or complexity.

Short term therapy

The distinction between patients with short duration of stay and 'normal' patients also causes great problems. If all cases were included in the same DRG, this group would no longer be homogeneous in terms of costs. Thus it seems reasonable to assign patients with short term or day care to special groups, but only two of the investigated variants do this.

The AR-DRG variant has 12 DRGs for day care patients. For the most part, they derive from diagnostic or short-term therapy. Only one case in our sample was assigned to such a day care group.

MDC 24 in the GHM variant (version 5.6) handles short-term therapy in great detail, using a total of 86 classes. The treatment of cancer (chemotherapy, radiotherapy) can be found as well as care after organ transplantation or outpatient procedures. There is a surgical and a medical outpatient GHM for each MDC. Additionally, procedural GHM classes exist for common day care cases. Thus the mixing of day or short term care with regular patients is avoided. This increases the cost homogeneity in the group while incentives for incorrect classification are decreased. Of the project sample, one case was assigned to this group.

The LDF variant uses a LOS-related reduction of payment for the case group. Most Australian health authorities make lower payments for same-day cases, but this is not reflected in the AR-DRG classification itself.

Cost homogeneity

To evaluate cost homogeneity, the total cost for the treatment of all cases in cardiac surgery in 1999 was used. This was available for 763 cases. Additional costs (such as those for pacemaker implantation) were taken into account. Since these rates include costs of a possible continuing treatment in another hospital, they reflect the total costs for the whole DRG.

Tables 18 and 19 compare the highest-volume classes (coronary artery and heart valve surgery) for three of the DRG variants using the above mentioned data from 1999.

Table 18: cases and case costs, coronary artery surgery, three DRG variants

DRG	AP 107	AR F06A	GHM 155
Number	276	231	374
Mean cost per case ( )	12,849	14,090	13,354
Minimum cost per case ( )	6,389	7,044	6,389
Maximum cost per case ( )	39,292	51,637	51,637

Table 18: cases and case costs, heart valve surgery, three DRG variants

DRG	AP 105	AR F04A	GHM 153
Number	197	210	275
Mean cost per case ( )	16,654	18,026	18,593
Minimum cost per case ( )	9,496	9,168	9,168
Maximum cost per case ( )	79,366	98,410	98,410

In coronary artery surgery, the cost distribution is relatively homogeneous and a clear peak could be seen at 11,000, but the distributions of the percentage of case costs are very heterogeneous in all of the evaluated variants. The incongruent curve shows the differences in grouping, resulting in a different cost structure of cases in the distribution. While in GHM (and even better in AR-DRGs) a peak is at least indicated, the distribution of costs within AP-DRG 105 has no clear peak.

Conclusions and recommendations

The investigations have shown that the correct grouping of a case within all variants is dependent on exact coding of diagnoses and procedures, and on the degree to which the coding systems themselves are suited to the task. We found that the German procedure coding system lacks precision in several respects, and that the inadequate attention to detail resulted in many errors of mapping in the native grouping software (AP-DRG, 3M).

We found that the creation of a mapping table from first principles was possible, and resulted in an adequate outcome. During this project, mapping tables were created for CdAM, ICD-10-AM-3, ICD-3-CM-3, and BMSG within four weeks. We achieved at least the equivalent degree of quality as was present in the existing 3M and Internova native grouping products. Thus it may be assumed that mapping to any of these codes could be done by experts in eight to twelve weeks, and this would be satisfactory if there were adequate testing by experts.

The use of the German diagnosis classification system, ICD-10-SGB V, does not present any serious risks. However, work is needed with respect to coding standards - how the codes are used in terms of selection of principal diagnosis, sequencing, determination of relevance of a condition to the episode of care being grouped, and so on.

Attention needs to be paid to the issue of homogeneity of the classes. In our sample, cost homogeneity was questionable in valve surgery, since the reconstruction of one valve is allocated to the same group as multiple valve replacement. Such examples can surely be found for other medical fields. A process will be needed like that applied in Australia, whereby possible splits are identified and subjected to clinical evaluation. It may be

that many proposed splits are not justified if the aim is equity of payments, but a process that tests ideas and provides statistical evidence is essential if care providers are to be increasingly willing to accept the validity of per case payment by DRG.

The undertaking of costing studies is necessary, to ensure that payment relativities reflect German clinical practice and care methods in general. Before calculations can start, a functional grouper must exist that groups the cases correctly according to the agreed objectives.

This study has not provided an authoritative answer to the question of selection of a starting DRG variant, but we have made some progress. In particular, we have shown significant differences between the variants in the following respects:

- Transparency of the grouping process (displaying of important criteria for grouping: diagnoses, procedures, age, etc.)
- Degree and usefulness of manuals
- Availability of coding standards for the system
- Correct consideration of levels of severity
- Consideration of modern treatments.

In cardiac surgery, the French GHM-DRG and the Australian AR-DRG variants had the best medical logic. The AR-DRG variant was the most modern and the most forward-looking. It was also outstanding in terms of the quality and quantity of explanation of its design and operation, and among the best in terms of statistical performance in resource use homogeneity. An added advantage was its application of PCCL logic whereby there are five 5 levels of severity for every adjacent DRG based on additional diagnoses. We concluded it could be recommended from the point of view of cardiac surgery in terms of its suitability for quality assurance, benchmarking and performance control.

We also believe the number of AR-DRG classes (661) is reasonable from a payment system point of view. It is unclear whether this level of categorisation will be satisfactory in future. Even a system with 2000 or more classes can be handled because an adequate IT infrastructure is needed whether the number is 600 or 2000. However, there is an argument that, if there are too many classes, the estimates of average cost will be less robust. There is also an additional administrative overhead even where computers are used.

The GHM variant, like the AP-DRGs which it strongly resembles, has a disadvantage regarding very heterogeneous groups. It would need to be further developed to take account of variations using PCCL or similar logic in order to be competitive.

Both of these variants would need reconfiguration in paediatric cardiac surgery, since these cases are integrated into other groups. This may make sense from the standpoint of cost homogeneity, but not in terms of clinical acceptance. The AP-DRG variant alleviates a few of the weaknesses by use of the additional group 809, but unfortunately the assignment to this group was not always uniform in the German AP-DRG adaptation.

The HCFA-DRG, R-DRG and AP-DRG variants resemble each other very much - as might be expected, given that they were developed by the same team. The HCFA variant does not adequately account for variations in severity and thus would have to be extended. R-DRGs have nearly the same grouping results as HCFA-DRGs, but the more elaborate differentiation of complexity levels matches multiple morbidity more effectively.

In summary, all the variants tested here could be made to work in the German context and might be expected to perform in a similar way in terms of facilitation of resource allocation through per case payment. However, there are practical and conceptual differences that may be far from trivial. Our analyses favour AR-DRGs overall, because of the open nature of the classification design and software implementation, and because of the degree to which clinical logic has been taken into account over several years of ongoing refinement.

References

Fischer W, ‚Diagnosis Related Groups (DRGs) und verwandte Patientenklassifikationssystemen', http://www.fischer-zim.ch/text-pcssa/t-ga-0103.htm

Heller M 2000, 'Krankenhausfinanzierung in Frankreich, Teil 1 und 2', *das Krankenhaus* no. 1 and 2.

http://visasys.com.au

http://www.flexdata.com.au

http://www.le-pmsi.fr/mco/produits/assistan.html

Jaster K et al 1999, 'Qualitätsicherung in der Herzchirurgie', in Scheld, 'Dokumantationsverfahren in der Herzchirurgie', Steinkopff-Verlag.

Mansky J et al 1999, 'AP-DRG project in Germany: first experiences', *Proceedings Manual, 15th PCS/E International Working Conference.*

Neubauer G & Nowy R 2000,'Wege zur Einführung eines leistungsorientierten und pauschalisierten Vergütungssystems für operative und konservative Krankenhausleistungen in Deutschland', *Gutachten im Auftrag der DKG*, Institut für Gesundheitsökonomie München, Februar 2000.

Rochell B & Roeder N 2000, 'Vergleichende Betrachtung und Auswahl eines Patienten-klassifikationssystems auf der Grundlage von Diagnosis Related Groups, Teil 1 und 2', *das Krankenhaus* no. 4 and 5.

Roeder N et al 2000, 'Kalkulation von herzchirurgischen Fallpauschalen' *f&w* no. 1

Roeder N, Rochell B, Polonius MJ, Achner S, Henke M, Schmidt M, Polei G & Scheld HH 2001, 'Empirischer Vergleich von Patientenklassifikationssystemen auf der Grundlage von Diagnostis Related Groups (DRG) in der Herzchirurgie - Zwischenbericht zum Projekt', http://drg.uni-muenster.de/download/DRGZwischenbericht.zip

The Danish Ministry of Health 1999, 'Hospital Funding Casemix', Health Analysis no. 8.