I. On-line data supplement: Appendix

In a 2x2 contingency table (cf. table A1) the proportion A/(A+C) is called recall in retrieval studies. This corresponds to sensitivity as benchmark of diagnostic tests. The proportion A/(A+B) is called precision, which corresponds to the positive predictive value as benchmark of diagnostic tests. Hogan and Wagner[18] defined completeness similar to recall and correctness similar to precision. Logan et al.[13] introduced an additional category for the documentation under consideration (cf. table A2). They use four categories in comparison with a gold standard: 1) items present and correct (n1), 2) items present and incorrect (n2), 3) items absent (n3), 4) new and therefore incorrect items (n4). Completeness is then defined as (n1+n2)/(n1+n2+n3), correctness as n1/(n1+n2+n4). For completeness the denominator is equal in both definitions. It consists of all items present in the gold standard. But Logan et al. shift the present but incorrect items into the numerator, which leads to a higher value for completeness. For correctness numerator and denominator are equal. In case of a missing gold standard we have to be aware of s1 and s2 in table A2 too.

II. On-line data supplement: Tables

Table 1
Inter-coder reliability of the principal diagnosis

Level of ICD-10-SGB-V 2.0	Number of concordant cases	Weighted kappa	95 % Confidence limits	Assessment according to Landis and Koch[13]
Terminal code	120	0.49	0.42-0.55	Moderate
Maximally 4-digits	125	0.51	0.44-0.57	Moderate
3-digits	166	0.67	0.61-0.73	Substantial
Group	202	0.82	0.77-0.87	Almost perfect
Chapter	211	0.84	0.78-0.89	Almost perfect

Table 2
Use of “other” and “unspecified” in both sets before and after validation. Percents are calculated from the total number of diagnoses (909 PPR-set, 959 EPR-set), from the number of principal diagnoses (244 in each set) and from the total number of procedures (765 PPR-set, 940 EPR-set).

	PPR-set		EPR-set	
	Native	Validated	Native	Validated
ICD-10-SGB-V 2.0 (all diagnoses)				
---.8- (“other”)	68 (7.5 %)	59 (6.5 %)	82 (8.6 %)	71 (7.4 %)
---.9- (“unspecified”)	154 (16.9 %)	66 (7.3 %)	217 (22.6 %)	106 (11.1 %)
ICD-10-SGB-V 2.0 (principal diagnoses)				
---.8- (“other”)	11 (4.5 %)	8 (3.3 %)	9 (3.7 %)	5 (2.0 %)
---.9- (“unspecified”)	46 (18.9 %)	13 (5.3 %)	70 (28.7 %)	28 (11.5 %)
OPS-301 2.0				
x (“other”)	11 (1.4 %)	11 (1.4 %)	13 (1.4 %)	13 (1.4 %)
y (“unspecified”)	34 (4.4 %)	34 (4.4 %)	22 (2.3 %)	22 (2.3 %)

Table 3
Procedures with more than 10 occurrences in descending order of the EPR-set. The German text was translated by the authors.

Code OPS-301 2.0	Text OPS-301 2.0	PPR-set	EPR-set
5-572.1	Percutaneous cyststomy	14	23
5-530.1	Repair of inguinal hernia with synthetic substitute	16	19
5-511.11	Laparoscopic cholecystectomy without revision of bile ducts	13	13
		...	
5-893.0g	Cleaning of small wound foot	3	13
5-385.70	Crossektomie and stripping of greater saphenous vein one side	6	12

Table 4
Concordance of the PPR-set and the EPR-set concerning DRGs. The numbers refer to identical results for a case.

Level of AR-DRGs	Number of concordant cases	Percentage of concordant cases
Principal diagnosis and DRG	93	38.1 %
DRG	152	62.3 %
Adjacent DRG	179	73.4 %
MDC	213	87.3 %

Table 5
Economical profit through optimization of coding from the PPR. The results of Mieth et al.[33] were estimated with a CMI of 3,5 before optimization.

	Profit by coding from the PPR	CMI PPR/EPR
This study	1.5 %	2,09/2,06
Langrehr et al.[32]	14.0 %	3,52/3,28
Mieth et al.[33]	16.4 %	
Ingenerf et al.[26]	36.7 %	2.83/2.07

Table A1
Definition of completeness (A/(A+C)) and correctness (A/(A+B)) of Hogan and Wagner[18]

	Gold standard		
Documentation	Present	Absent	
Present	A	B	A/(A+B)
Absent	C	D	
	A/(A+C)		

Table A2
Definition of completeness ((n1+n2)/(n1+n2+n3)) and correctness (n1/(n1+n2+n4)) of Logan et al.[13]

	Gold standard		
Documentation	Present	Absent	
Present correct	n1	s1	
Present incorrect	n2	n4	n1/(n1+n2+n4)
Absent	n3	s2	
	(n1+n2)/(n1+n2+n3)		