

# Inter-rater reliability of the ICPC-2 in a German general practice setting

Thomas Frese<sup>a</sup>, Kristin Herrmann<sup>a</sup>, Peggy Bungert-Kahl<sup>b</sup>, Hagen Sandholzer

<sup>a</sup> Department of Primary Care, Leipzig Medical School, Leipzig, Germany

<sup>b</sup> Department of Neurology, Leipzig University Hospital, Leipzig, Germany

## Summary

**QUESTIONS:** Three- and four-digit International Classification of Diseases (ICD-10) is not a reliable classification system in primary care. The reliability of the International Classification of Primary Care (ICPC-2) as an alternative coding system has not yet been investigated in a German general practice setting.

**METHODS:** Cross-sectional data were collected during a one year period in a general practice setting. Participants: A total of 8,877 patients were randomly selected. Main outcome measures: The first of the reasons for encounter was taken into account on new and chronic managed problems. The ICPC-2 coding of each case was performed by two raters to investigate the inter-rater agreement. The degree of agreement between the raters was assessed by using Cohen's kappa ( $\kappa \geq 0.61$  meaning high or satisfactory and  $\kappa \leq 0.6$  (incl.  $\leq 0.000$ ) meaning low or unsatisfactory).

**RESULTS:** The reliability was good to excellent at the chapter level, at the component level the reliability was moderate though good in the components 1-symptoms and 7-diseases. At single code level the agreement was only fair to moderate in both chapters and components. One third to half of the used codes showed good inter-rater agreement.

**CONCLUSION:** The ICPC-2 is an adequate and feasible instrument for routine use in general practice. The fair to moderate reliability on the single code level should be considered when designing studies and interpreting data that are based on the ICPC-2.

**Key words:** ICPC; ICD; general practice; reliability

## Introduction

Describing general practice epidemiology is an important topic for various stakeholders, such as medical practitioners, epidemiologists and health system managers. As in the

German health care system, general practice epidemiology is recently often described by ICD-10 codes (International Classification of Diseases and Related Health Problems, 10<sup>th</sup> Revision). Earlier results have suggested that the three- and four-digit ICD-10 is not a reliable classification system in primary care [1]. ICD-10 does not contain enough individual categories for many of the common and ill-defined problems encountered in general practice. Neither the patient's reason for encounter nor the primary care interventions are represented as codes. It is also a fact that the ICD-10 at the three-digit level cannot serve as a core classification for an international primary care system [2]. Therefore an alternative clinical classification system especially adapted to primary care, the International Classification of Primary Care (ICPC), was chosen as the measurement device of interest in the current investigation. ICPC and its currently used version ICPC-2 were derived from the International Classification of Health Problems in Primary Care (ICHPPC-2) [3], the Reason for Encounter Classification (RFEC) and the IC-Process-PC Classification [4]. The ICPC became a polyvalent classification system. It allows classification of all elements of the problem-solving process and thereby the complete treatment period [5]. The ICPC is biaxial, each of the 17 chapters (table 1) contains seven components (symptoms and complaints; diagnostic, screening, and preventive procedures; medication, treatment and procedures; test results; administrative; referrals and other reasons for encounter; diseases). The resulting ICPC-codes contain three-digits: The first for the chapter specification and the other two for the component within the chapter [3, 6].

The SESAM 2-study [7] evaluated the reasons for encounter, the performed procedures, and the results of encounter of patients in a German general practice setting (subject population). Being a minor part of the SESAM 2 study, the recent investigation set out to study the reliability of the ICPC-2 if it is applied by different general practitioners (rater population). An earlier investigation towards the reliability of the ICD-10 that was based on the same data and followed the same design [1] enabled comparisons between the reliability of this classification system and the ICPC-2.

### Abbreviations

ICD-10 International Statistical Classification of Diseases and Related Health Problems, 10<sup>th</sup> Revision

ICPC-2 International Classification of Primary Care, 2<sup>nd</sup> Revision

SESAM2 Saxon Epidemiological Study of General Medicine

## Method

**Design:** Cross-sectional data were collected from 1<sup>st</sup> October 1999 to 30<sup>th</sup> September 2000. Ethical approval was stated not to be necessary. **Setting:** The Saxon Society of General Medicine (SGAM) contacted all general practitioners in Saxony. A total of 209 of the 2,510 physicians cooperated. **Selection of participants:** Case recording was carried out one day per week that was chosen randomly (either morning or afternoon consulting hours). Data were collected from one out of ten patients. Each patient was estimated once. House calls were not considered. A total of 8,877 patients were included. **Main outcome measures:** A [standardised data collection](#) (see appendix) form was used. It was developed by general practitioners. Each patient's reasons for encounter, symptoms, diagnostic procedures, recent results of encounter / diagnoses, general morbidity and therapeutic procedures were assessed. Data were documented *verbatim* (according to the study instructions), either as told by the patients (e.g. reasons for encounter) or in the physician's words (e.g., chronic managed problems). Only completely filled forms were considered. **Coding:** ICPC-2 was used to code the patients' first reason for encounter and the first new and the first chronic problem managed. Data was edited by two medical doctoral candidates (comparable to PhD students, specialised general practitioners with their own medical practice, educated at different universities, both female, age 32 and 48 years) of Leipzig Medical School's Department of General Practice. The raters were not specially trained in ICPC-2 coding. The rating was performed independently. The time that was necessary for coding the first 100 reasons for encounter, new and chronic managed problems was estimated and averaged. **Statistical analysis:** The data analysis was performed with SPSS (Version 11.0). The frequency of rating different codes by both raters was recorded using contingency tables. The Cohen's Kappa ( $\kappa$ ) was used to assess the inter-rater reliability on two levels of specificity (chapter or component and single code). The Kappa statistics can range from -1 to 1. 0 indicates only chance agreement and 1 indicates perfect chance corrected agreement. Following the

classification by Landis and Koch [8], interpretations of agreement for different scores are: <0.20 poor, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 good, and 0.81 to 1.00 very good. Additionally, the means of the single code Kappas were calculated.

## Results

Data from 8,877 subjects were returned in completely filled forms and were coded. Of the 1,366 potential ICPC-2 codes, 641 (46.9%) were used for coding the 8,710 reasons for encounter, 468 (34.3%) were used for the 5,573 new managed problems and 395 (29.9%) were used for the 7,050 chronic managed problems. This information was coded by two raters, and there were no replicate observations. Table 1 illustrates the distribution of the encoded data among the ICPC-2 chapters.

Comparisons to unpublished data from other investigations which were performed in the same setting showed that the estimated subjects were representative. The described raters (see methods for details) were not a representative sample of Saxon general practitioners.

### Reasons for encounter

Inter-rater agreement was high on chapter level with a mean  $\kappa$  of 0.83 (table 2). In 11 of 17 chapters,  $\kappa$  was over 0.81. The inter-rater reliability was highest in the chapter H-Ear with a kappa of 0.95. The largest proportion of chapter discrepancies involved chapter A-General and unspecified ( $\kappa = 0.58$ ). On the single code level mean kappa for all used ICPC-2 codes was 0.39, with 237 of 641 used codes (37%) reaching good or very good agreement ( $\kappa > 0.61$ ). The highest level of agreement on single code level was reached in the chapters R-Respiratory, N-Neurological and W-Pregnancy and family planning (mean  $\kappa = 0.48$ ; table 2). Solely in chapter N-Neurological more than half of the single codes had a good or very good inter-rater reliability. At component level, the average  $\kappa$  amounted to 0.58 (table 3). It ranged from 0.16 to  $\kappa = 0.92$ . Regarding means of single code  $\kappa$ , a moderate degree of agreement was found in component 1-Symptoms ( $\kappa = 0.59$ ) and com-

**Table 1:** The average numbers of the reasons for encounter (RFE), the new managed problems (NMP) and the chronic managed problems (CMP) among the chapters of the ICPC-2 according to the coding of two raters.

Chapter of the ICPC-2	RFE (n)	NMP (n)	CMP (n)
A General, unspecified	1,514	355.5	145.5
B Blood, lymphatics, spleen	137	58	64.5
D Digestive	718.5	558.5	414
F Eye	59.5	54	21.5
H Ear	88.5	100.5	38
K Circulatory	1,430.5	464.5	2,844
L Musculoskeletal	1,524.5	1,273	936.5
N Neurological	407.5	172.5	228
P Psychological	187	210	275.5
R Respiratory	1,258	1,349.5	478.5
S Skin	587.5	566	232
T Endocrine, metabolic, nutrition	592.5	215.5	1,160.5
U Urology	144	140.5	98
W Pregnancy, family planning	6	6	4.5
X Female genital system	24.5	25.5	63
Y Male genital system	15.5	16	40.5
Z Social problems	15	8	5

ponent 7-Diagnoses ( $\kappa = 0.54$ ). A high degree of agreement was found in 61.6% of all codes in component 1-Symptoms and in 54.2% of all component 7-Diagnoses codes. Agreement on the components 2 to 6 was poor ( $\kappa < 0.20$ ). The mean time for coding of reasons for encounter was 69 seconds.

### New managed problems

Of all new managed problems, rater 1 and 2 achieved a high agreement on chapter level (mean  $\kappa = 0.80$ ; table 2). Good agreement ( $\kappa > 0.80$ ) was obtained in 12 of the 17 ICPC-2 chapters and low reliability scores were obtained for chapter Z-Social problems ( $\kappa = 0.25$ ). Of all 468 used ICPC-2 codes, 223 (47.7%) had good reliability scores ( $\kappa > 0.61$ ). The mean of all single code  $\kappa$  was 0.48 (range from 0.21 to 0.71). More than half of the used codes of chapters D-Digestive, F-Eye, H-Ear, N-Neurological, S-Skin, U-Urology, W-Family Planning, X-Female and Y-Male genital system showed good reliability scores. At component level, we found on average moderate agreement (mean  $\kappa = 0.44$ ) (table 3). The inter-rater reliability was highest at the component 2-Diagnostic procedures ( $\kappa = 0.91$ ). No agreement was found in the component 4-Test results. On single code level a substantial agreement was only found in component 7-Diagnoses (mean  $\kappa = 0.61$ ) with 60.6% of all codes having a good inter-rater reliability.

The coding of new managed problems took a mean of 33 seconds.

### Chronic managed problems

A very good reliability (average  $\kappa = 0.89$ , range from 0.60 to 0.99) was found on chapter level (table 2). Thus, at chapter level a very good degree of agreement ( $\kappa > 0.80$ ) was reached in 13 of the 17 ICPC-2 chapters. The average inter-rater reliability of all single-codes was moderate (mean  $\kappa = 0.53$ ) ranging from  $\kappa = 0.34$  for the chapter A-General and unspecified to  $\kappa = 0.72$  for the chapter Y-Male genital system. For 202 of the used 395 codes good reliability scores ( $\kappa > 0.61$ ) were calculated. Observer agreement on component level was fair (mean  $\kappa = 0.40$ , range from 0.00 to 0.71; table 3). Means of the single code  $\kappa$  ranged from 0.00 to 0.60. Only in component 7-Diagnoses more than half of the ICPC-2 codes had good inter-rater reliability scores. The coding of known managed problems took a mean of 36 seconds.

## Discussion

We present for the first time data regarding the reliability of the ICPC-2 at different levels of specificity in a German general practice setting. Furthermore, we can draw direct comparisons to our earlier work that elucidated the reliab-

**Table 2:** Inter-rater reliability towards for reasons for encounter, new and chronic problems presented in Cohens Kappa on chapter level ( $\kappa$ ), on single code level mean Kappa over all used codes (mean  $\kappa$ ) and frequency of codes with good or very good Kappa ( $>0.61$ ; %  $\kappa > 0.61$ ).

Chapter of the ICPC-2		Reasons for encounter			New managed problems			Chronic managed problems		
		$\kappa$	Mean $\kappa$	% $\kappa > 0.61$	K	Mean $\kappa$	% $\kappa > 0.61$	$\kappa$	Mean $\kappa$	% $\kappa > 0.61$
A	General, unspecified	0.58	0.29	24.2	0.47	0.35	27.9	0.70	0.34	32.0
B	Blood, lymphatics, spleen	0.80	0.41	33.3	0.79	0.33	41.2	0.88	0.66	64.7
D	Digestive	0.91	0.45	41.2	0.91	0.53	54.6	0.97	0.62	58.5
F	Eye	0.94	0.27	27.6	0.96	0.63	68.8	0.93	0.58	53.9
H	Ear	0.95	0.39	34.8	0.86	0.56	53.9	0.87	0.37	40.0
K	Circulatory	0.85	0.42	38.9	0.83	0.40	36.8	0.99	0.53	52.6
L	Musculoskeletal	0.90	0.46	43.2	0.94	0.50	48.2	0.97	0.56	54.0
N	Neurological	0.90	0.48	51.4	0.79	0.58	59.3	0.79	0.59	53.9
P	Psychological	0.84	0.31	30.0	0.86	0.40	42.9	0.80	0.37	37.5
R	Respiratory	0.92	0.48	43.5	0.90	0.51	47.1	0.95	0.45	40.7
S	Skin	0.85	0.44	40.9	0.90	0.62	59.3	0.92	0.61	55.3
T	Endocrine, metabolic, nutrition	0.83	0.35	34.2	0.82	0.37	34.5	0.98	0.49	50.0
U	Urology	0.92	0.40	44.8	0.97	0.71	73.3	0.97	0.54	57.9
W	Pregnancy, family planning	0.67	0.48	40.0	0.67	0.21	25.0	0.89	0.67	66.7
X	Female genital system	0.78	0.33	30.0	0.98	0.59	61.5	0.98	0.51	53.9
Y	Male genital system	0.71	0.14	21.4	0.81	0.54	50.0	0.96	0.72	75.0
Z	Social problems	0.73	0.22	18.2	0.25	0.21	20.0	0.60	0.62	50.0
	Average	0.83	0.39	37.0	0.80	0.48	47.7	0.89	0.53	51.1

**Table 3:** Inter-rater reliability towards for reasons for encounter, new and chronic problems presented in Cohens Kappa on component level ( $\kappa$ ), on single code level mean Kappa over all used codes (mean  $\kappa$ ) and frequency of codes with good or very good Kappa (%  $\kappa > 0.61$ ).

Component of the ICPC-2		Reasons for encounter			New managed problems			Chronic managed problems		
		$\kappa$	Mean $\kappa$	% $\kappa > 0.61$	K	Mean $\kappa$	% $\kappa > 0.61$	$\kappa$	Mean $\kappa$	% $\kappa > 0.61$
1	Symptoms and complaints	0.92	0.59	61.6	0.63	0.46	45.3	0.64	0.40	37.7
2	Diagnostic, screening and preventive procedures	0.62	0.19	12.4	0.91	0.07	3.5	0.00	0.00	0.0
3	Medication, treatment and procedures	0.46	0.16	8.6	0.38	0.21	25.9	0.71	0.27	25.0
4	Test results	0.89	0.19	3.7	0.00	0.00	0.0	0.00	0.00	0.0
5	Administrative	0.23	0.19	7.1	0.23	0.00	0.0	–	–	–
6	Referrals and other reasons for encounter	0.16	0.11	5.5	0.32	0.13	8.3	–	–	–
7	Diseases	0.76	0.54	54.2	0.59	0.61	60.6	0.64	0.60	59.0
	Average	0.58	0.39	37.0	0.44	0.48	47.7	0.40	0.53	51.1

ility of the ICD-10 in exactly the same setting and using exactly the same basic data [1]: Kappa values were all satisfactory on the ICD-10 chapter level for new and chronic managed problems. However, coding with the ICPC-2 reached higher Kappa values at the chapter level: The number of chapters reaching a Kappa >0.8 was 12 of 17 for new managed problems and 13 of 17 for chronic managed problems. The corresponding numbers when coding with ICD-10 were 5 of 21 and 10 of 21. When coding new managed problems with the ICD-10 the average kappa values were fair at a three-digit level and poor at a four-digit level. Kappa was fair to moderate when the three-digit level was used and poor for terminal codes (four-digit level) for chronic managed problems [1]. In contrast, we found moderate Kappa values when coding chapter-based single codes with the ICPC-2. Therefore this classification system could be assumed to be more reliable. However, it is difficult to compare single ICPC-2 codes to terminal ICD-10 codes since the structures of the two coding systems are very different.

Coding reasons for encounter with the ICPC-2 took a mean of 69 seconds while coding with the ICD-10 needed 96 seconds. The time that was necessary for coding of new (36 vs. 33 seconds) or chronic managed problems (42 vs. 36 seconds) revealed no relevant differences between the two classification systems [9]. The average duration of a general practice consultation is about 10 minutes [10, 11]. The coding of the reasons for encounter in particular consumes a high percentage of that valuable time. However, it should be regarded that the chosen method of assessing the coding time might overestimate the required time because it did not consider a training effect.

Encoding with the ICPC-2 was reliable on chapter-level (table 2). At component level reliability was moderate overall (table 3). For single codes, we found a fair to moderate reliability overall. One third to half of the used codes showed good inter-rater agreement. Overall agreement was highest by coding chronic problems and lowest when coding reasons for encounter. Greater disagreement especially appeared in single codes of the components 2 to 6 and in chapters A-General & unspecified, P-Psychological, T-Endocrine, and Z-Social problems. This may be partially due to a small sample number of cases.

Our findings are in accordance with those of others: The inter-rater reliability of coding morbidity data with the ICPC ranged between 84% and 96% [12–13]. At chapter level, the positive agreement was between 79% and 84% [3, 14]. Thus, 70% agreement was reached in 14 of the 17 ICPC chapters in a study by Britt et al. [14]. In the study of van der Horst et al. [15] Kappa scores for different organ and problem systems ranged from a minimum of 0.61 to a maximum of 0.96. Low reliability scores were obtained for Z-Social problems, problems pertaining to Y-Male genital system and B-Blood, lymphatics and spleen. The highest scores were obtained for circulatory disorders, respiratory disorders, and disorders of the eye and ear. Chapter and code discrepancies in chapter A-General and unspecified was also found by Letrilliart et al. [3] and indicated a lack of specificity of this chapter. The positive main agreement decreased from 60% to 65% for the number of codes in each chapter [3, 14]. At single code level 70% agreement

was found in only one chapter [14]. The reliability based on components was lower than that based on chapters [15]. However, the analyses showed a good degree of agreement between the different components of the ICPC code with Kappas ranging from 0.70 to 0.78 [14–15]. In this study as well as in specialist literature it was reported that the majority of chapters and chapter-component agreement is better than chance. At a specific contact the morbidity recorded is reliable and it is valid at chapter level and, in most cases, at chapter-component level. At single code level, variance between practitioners in labelling the problem calls into question the validity and reliability of the data [14].

Our results suggest that coding with the ICPC-2 is slightly more reliable than coding with the ICD-10. However, both classification systems should be improved to become more reliable and applicable. Solving problems that occur in coding with ICPC-2 [16] might be a first approach. In contrast to the ICD-10, the ICPC-2 enables the coding of reasons for encounter because it covers the complete treatment period [5]. Our earlier results suggested that only 13.6% of the possible ICD-10 codes but 30% to 47% of the possible ICPC-2 codes were used to encode the 8877 consultations [1]. Thereby, the spectrum of the ICD-10 codes is not necessary in a general practice setting. However, it was concluded from a Norwegian study that physicians did not adhere to the ICPC-2 standard due to its incompleteness (i.e. lack of many clinically important diagnoses) [17]. This may also limit the usefulness of the ICPC-2 for comprehensive disease registers, as necessary for diabetic patients for example [18].

It should be kept in mind that epidemiologic and other scientific investigations, for example the Suisse FIRE-project (Family Medicine ICPC-Research using Electronic Medical Records [19]), the German CONTENT project [20], other morbidity studies [21] and electronic patient records [22], are often based on the ICPC-2. The recent investigation is relevant because it illustrates limitations of this classification system. This may exert consequences towards a further development of the ICPC-2 to the ICPC-3. The reported results should also be taken into account when designing studies and interpreting data that are based on the ICPC-2. It can be concluded that the ICD-10 cannot be substituted by the ICPC-2. In accordance to the reported results of Botsis et al. an in depth revision of ICPC-2 [17] or other approaches (e.g. the creation of a general practice thesaurus [23]) are necessary to get a coding system that is appropriate for clinical work in general practice settings.

### Strengths of the study

The study investigates a widely used classification system (ICPC-2).

The topic is relevant for various stakeholders, such as medical practitioners, epidemiologists and health system managers.

This is the first investigation of the reliability of the ICPC-2 in a German general practice setting.

The study included all groups of reasons for encounter, results of encounter or diagnoses.

### Weaknesses of the study

Only about 10% of the general practitioners cooperated.

Data could only be used from randomly selected patients. The raters could not be present at the consultations. The raters were not representative for all Saxon general practitioners. Coders viewing time was unlimited whereas the practitioner works at normal clinical pace.

### Conclusion

The ICPC-2 is an adequate and feasible instrument for routine use in general practice. The fair to moderate reliability on single code level should be considered when designing studies and interpreting data that are based on the ICPC-2.

**Acknowledgements:** The authors would like to thank Dr. Hanno Grethe, honorary president of the SGAM, and Dr. Johannes Dietrich, president of the SGAM, for their kind support.

**Funding / potential competing interests:** This work was supported by the Leipzig Medical School and the Saechsische Gemeinschaft fuer Allgemeinmedizin (SGAM; Saxon Society of General Practice).

**Correspondence:** Thomas Frese, MD, Department of Primary Care of Leipzig Medical School, Philipp-Rosenthal-Straße 55, DE-04103 Leipzig, Germany, [mail\[at\]thomasfrese.de](mailto:mail[at]thomasfrese.de)

### References

- 1 Wockenfuß R, Frese T, Herrmann K, Claussnitzer M, Sandholzer H. Three- and four-digit ICD-10 is not a reliable classification system in primary care. *Scand J Prim Health Care*. 2009;27:131–6.
- 2 Wood M, Lamberts H, Meijer JS, Hofmans-Okkes IM. The conversion between ICPC and ICD-10. Requirements for a family of classification systems in the next decade. *Fam Pract*. 1992;9:340–8.
- 3 Letriliart L, Guiguet M, Flahault A. Reliability of report coding of hospital referrals in primary care versus practice-based coding. *Eur J Epidemiol*. 2000;16:653–9.
- 4 Meyer RL. Von der ICHPPC zur ICPC. *Primary Care*. 2005;5:236–9.
- 5 Lamberts H, Hofmans-Okkes I. The core of computer based patient records in family practice: episodes of care classified with ICPC. *Int J Biomed Comput*. 1996;42:35–41.
- 6 Wonca International Classification Committee. International Classification of Primary Care-2<sup>nd</sup> Ed. [cited 2011 Aug 20] Available from: [www.globalfamilydoctor.com/wicc/pagers/english.pdf](http://www.globalfamilydoctor.com/wicc/pagers/english.pdf)).
- 7 Frese T, Klauss S, Herrmann K, Sandholzer H. Children and adolescents as patients in general practice – The reasons for encounter. *J Clin Med Res*. 2011;3:177–82.
- 8 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
- 9 Wockenfuß R. Überprüfung der Reliabilität der ICD-10 in der Allgemeinmedizin. Leipzig Medical School 2010;33–4.
- 10 Deveugele M, Derese A, de Bacquer D, van den Brink-Muinen A, Bensing J, de Maeseneer J. Consultation in general practice: a standard operating procedure? *Patient Education and Counseling*. 2004;54:227–33.
- 11 Litschgi L, Fehr F, Zeller A. Characteristics of consultations by general practitioners in Switzerland – time, topics and verbal behaviour. *Praxis (Bern 1994)*. 2011;18:1085–95.
- 12 Britt H, Meza RA, Del Mar C. Methodology of morbidity and treatment data collection in general practice in Australia: a comparison of two methods. *Fam Pract*. 1996;13:462–7.
- 13 Bernstein R, Hollingworth G, Viner G, Shearman J, Labelle C, Thomas R. Reliability issues in coding encounters in primary care using an ICPC/ICD-10-based controlled clinical terminology. *J Am Med Inform Assoc*. 1997;21:843.
- 14 Britt H, Angelis M, Harris E. The reliability and validity of doctor-recorded morbidity data in active data collection systems. *Scand J Prim Health Care*. 1998;16:50–5.
- 15 van der Horst F, Metsemakers J, Vissers F, Saenger G, de Geus C. The Reason-for-encounter mode of the ICPC: reliable, adequate, and feasible. *Scand J Prim Health Care*. 1989;7:99–103.
- 16 Boot CR, Meijman FJ. Classifying health questions asked by the public using the ICPC-2 classification and a taxonomy of generic clinical questions: an empirical exploration of the feasibility. *Health Commun*. 2010;25:175–81.
- 17 Botsis T, Bassoe CF, Hartvigsen G. Sixteen years of ICPC use in Norwegian primary care: looking through the facts. *BMC Med Inform Decis Mak*. 2010;10:11.
- 18 Maier M, Knopp A, Pusarnig S, Rurik I, Orozco-Beltran D, Yaman H, van Eygen L. Diabetes in Europe: role and contribution of primary care – position paper of the European Forum for Primary Care. *Qual Prim Care*. 2008;16(3):197–207.
- 19 Chmiel C, Bhend H, Senn O, Zoller M, Rosemann T. The FIRE project: a milestone for research in primary care in Switzerland. *Swiss Med Wkly*. 2011;140:w13142.
- 20 Laux G, Kuehle T, Rosemann T, Szecsenyi J. Co- and multimorbidity patterns in primary care based on episodes of care: results from the German CONTENT project. *BMC Health Serv Res*. 2008;8:14.
- 21 Wolff H, Sebo P, Haller DM, Eytan A, Niveau G, Bertrand D, et al. Health problems among detainees in Switzerland: a study using the ICPC-2 classification. *BMC Public Health*. 2011;11:245.
- 22 Verbeke M, Schrans D, Deroose S, De Maeseneer J. The International Classification of Primary Care (ICPC-2): an essential tool in the EPR of the GP. *Stud Health Technol Inform*. 2006;124:809–14.
- 23 Zentralinstitut für die kassenärztliche Versorgung in Deutschland. Hausarzt-Theasaurus. [cited 2012 Apr 03] Available from: <http://www.zi.de/cms/projekte/kodierung/hausarzt-theasaurus/>

### Appendix

» [Standardised data collection form](#) (PDF-File).