

# Confiabilidade interobservador da Classificação Internacional de Atenção Primária em uma unidade de atenção básica à saúde

## *Interobserver Reliability of the International Classification of Primary Care at a primary health care unit*

**Mariana Miranda Autran Sampaio<sup>I</sup>**

**Cláudia Medina Coeli<sup>II</sup>**

**Márcia Guimarães de Mello Alves<sup>III</sup>**

**Márcia Fernandes Soares<sup>IV</sup>**

**Kenneth Rochel de Camargo Jr.<sup>V</sup>**

**Arlinda Barbosa Moreno<sup>IV</sup>**

<sup>I</sup>Instituto Nacional de Câncer – INCA.

<sup>II</sup>Instituto de Estudos em Saúde Coletiva – UFRJ.

<sup>III</sup>Universidade Federal Fluminense – UFF.

<sup>IV</sup>Escola Politécnica de Saúde Joaquim Venâncio – Fiocruz.

<sup>V</sup>Instituto de Medicina Social – UERJ.

**Financiamento:** Edital MCT- CNPq/MS-SCTIE- DECIT – Nº 23/2006; Projeto Número: 409348/2006-6.

**Parecer do CEP:** Processo: CAAE - 0011.0.314.000-07.

Os autores não apresentam conflito de interesses.

**Correspondência:** Mariana Miranda Autran Sampaio. Departamento de Epidemiologia, Instituto Nacional de Câncer (INCA). Rua dos Inválidos, 212. 3º andar. Centro, Rio de Janeiro, RJ CEP 20231-048. E-mail: m.autran@gmail.com

## **Resumo**

Avaliou-se a confiabilidade interobservador da Classificação Internacional de Atenção Primária na codificação de queixas relatadas em unidade básica de saúde pública, por meio de uma amostra de 300 atendimentos. Os motivos de consulta foram registrados por médicos do serviço, e sua codificação realizada por duas pesquisadoras. A concordância entre as classificadoras em relação ao número de motivos codificados foi calculada por meio do kappa ponderado; para as demais análises (capítulos e códigos completos), foi utilizado o coeficiente kappa simples. Foram codificados 634 motivos pela primeira classificadora e 699 pela segunda. A concordância simples do número de motivos foi moderada, com kappa ponderado de 0,71 (Intervalo de 95% de Confiança [IC95%]: 0,68-0,73). Em relação aos códigos de capítulos e códigos completos, foram encontradas confiabilidade substancial (kappa = 0,89; IC95%: 0,87-0,91) e moderada (kappa = 0,72; IC95%: 0,68-0,76), respectivamente. Os resultados indicam bom desempenho da classificação para codificar motivos de atendimento na atenção básica, mesmo a partir de registros médicos.

**Palavras-chave:** Classificação Internacional de Atenção Primária. Atenção Primária à Saúde. Confiabilidade interobservador.

## Abstract

We evaluated the International Classification of Primary Care interobserver reliability for coding the complaints reported at a primary public health unit. The study sample consisted of 300 appointments. Reasons for appointment were registered by physicians on the medical record and coding was performed by two investigators. Reliability among the classifiers as to the number of coded reasons was calculated by weighted kappa. Crude kappa coefficients were used for the remaining analyses (chapter codes and full codes). A total of 634 reasons were coded by the first classifier and 699 by the second one. Crude reliability of numbers of reasons was moderate, with a 0.71 weighted kappa (95% Confidence Interval [95%CI]: 0.68-0.73). Regarding chapter codes and full codes, reliability was substantial (kappa = 0.89; 95%CI: 0.87-0.91) and moderate (kappa = 0.72; 95% CI: 0.68-0.76), respectively. Our results suggest a good performance of the Classification for coding reasons for visit in primary care, even when data are obtained from physician consultation notes.

**Keywords:** International Classification of Primary Care. Health Primary Care. Interobserver reliability.

## Introdução

A Classificação Internacional de Atenção Primária (CIAP), do inglês *International Classification of Primary Care*<sup>1</sup>, foi desenvolvida para abarcar limitações da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID) no contexto de cuidados primários em saúde. Para isso, incorporou uma lista de sinais e sintomas, procedimentos e diagnósticos mais comuns na atenção básica. Suas principais características inovadoras são a incorporação de queixas expressadas pelos pacientes e problemas de natureza social. A CIAP sofreu uma adaptação, publicada em 1998 como CIAP-2, para permitir comparabilidade com a CID-10, uma vez que seu objetivo não é substituí-la, mas sim complementá-la<sup>2</sup>.

Seu código é composto por uma letra e dois dígitos. A letra representa, em regra, o sistema corporal onde se localiza a queixa. No caso de queixas não localizadas em sistemas corporais, é possível atribuir um dos seguintes capítulos: Geral e Inespecífico, Psicológico, Gravidez e Planejamento Familiar ou Problemas Sociais. Os dois dígitos numéricos representam sinal e sintoma (no caso da numeração entre 01 e 29), procedimento (entre 30 e 59) ou diagnóstico (entre 60 e 99)<sup>3</sup>.

Esta classificação foi experimentada no Brasil, em 1983, em um teste de viabilidade, que também foi realizado em outros países, como Austrália, Barbados, Estados Unidos da América, Hungria, Malásia, Noruega e Países Baixos<sup>4,5</sup>. Apesar de vir sendo amplamente utilizada em vários países da Europa e na Austrália<sup>6,7</sup>, no Brasil, não foram encontradas mais publicações sobre seu uso após essa data, exceto um teste de confiabilidade em questionário autopreenchido<sup>8</sup>.

O objetivo deste estudo é avaliar a confiabilidade interobservador da Classificação Internacional de Atenção Primária na codificação de queixas relatadas em atendimentos de uma unidade primária de saúde pública.

## Métodos

Foi avaliada uma amostra de 300 atendimentos de uma unidade primária de saúde pública localizada na Cidade do Rio de Janeiro, cuja população usuária é predominantemente de baixa renda. O cálculo da amostra foi realizado empregando o módulo *Kapci* desenvolvido por Reichenheim<sup>9</sup> para o programa Stata<sup>10</sup>, empregando-se uma de amostra de diagnósticos provenientes de um estudo anterior<sup>8</sup> e estabelecendo-se uma precisão menor que 0,05.

Os dados foram colhidos de fichas de atendimento de pessoas que procuraram a unidade para uma primeira consulta acerca de um determinado motivo em agosto de 2007 de médicos que se dispuseram a participar da pesquisa, sendo coletados todos os motivos de consulta relatados pelos pacientes que foram registrados na ficha de atendimento da unidade de saúde. Antes da coleta foi explicada aos médicos a lógica da classificação, sendo ressaltada a importância do registro literal do relato do paciente e, em linhas gerais, como funcionava a classificação que seria usada. Foram excluídas as fichas de pacientes menores de 18 anos e de pessoas que tivessem seus motivos de consulta relatados por terceiros; dos que estavam retornando para acompanhamento de alguma outra consulta; ou quando havia ficha referindo que o paciente não aguardou pelo atendimento.

Inicialmente foi realizado um treinamento com as classificadoras a respeito da CIAP. Uma vez que uma delas (MMAS) já havia trabalhado com a classificação previamente, esta ficou responsável por treinar a outra classificadora (MGMA). Em um primeiro momento, foi explicada a lógica da classificação e MGMA leu todo o manual de uso da ICPC<sup>11</sup>. Posteriormente, a mesma classificou 59 relatos de questionários autopreenchidos, e este seu trabalho foi comparado a um consenso de treinamento anterior de MMAS com uma especialista em classificações. Depois de uma reunião acerca de dúvidas, MGMA codificou também 27 motivos de consulta de um ambulatório

primário de saúde utilizado no treinamento anterior e novamente se reuniu com MMAS para comparar o consenso. Após o treinamento, a codificação foi realizada de forma independente pelas duas pesquisadoras, sendo as divergências resolvidas por consenso.

Foram calculados os valores de mediana e intervalo interquartil (IQ) do número de motivos. Como a queixa de uma pessoa poderia ser interpretada com um ou mais códigos, era possível haver divergência no número de motivos codificados por cada classificadora; por isso foi realizada uma avaliação da concordância entre as duas classificadoras em relação ao número de motivos codificados, por meio do *kappa* ponderado<sup>12</sup>. Além desta, foi analisada a concordância da CIAP-2, segundo capítulos e segundo os códigos completos, tanto globalmente quanto nos estratos por sexo e faixa etária. Em todos os casos, a concordância entre as duas classificadoras foi aferida através da estatística *kappa*<sup>13</sup>. Os intervalos de 95% de confiança (IC 95%) do *kappa* foram calculados empregando-se a rotina *kapci*<sup>10</sup>, desenvolvida para o software Stata<sup>10</sup>.

Para a interpretação do *kappa*, foi empregada classificação proposta por Shrout<sup>14</sup>, que considera as seguintes faixas de concordância:  $k < 0,10$  concordância virtualmente ausente;  $k = 0,10$  a  $0,40$  concordância fraca;  $k = 0,41$  a  $0,60$  concordância discreta;  $k = 0,61$  a  $0,80$  concordância moderada;  $k = 0,81$  a  $1,0$  concordância substancial.

Após a atribuição dos motivos de consulta por cada classificadora, chegou-se por consenso a um código único para cada motivo de consulta. Foram apresentados, neste artigo, os cinco motivos mais frequentes por sexo e faixa etária, incluindo-se todos de igual frequência quando houve empate na quinta posição. Para esta análise optou-se por agrupar as duas categorias mais jovens em virtude da similaridade de seu perfil.

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Secretaria Municipal de Saúde do Rio de Janeiro (CAAE 0011.0.314.000-07).

## Resultados

Das 300 fichas de atendimento analisadas, 86 (28,7%) se referiam a homens e 214 (71,3%) a mulheres. A idade dos pacientes variou de 18 a 89 anos, com mediana de 41,5 anos.

Foram codificados 634 motivos pela primeira classificadora e 699 pela segunda, com uma mediana de 3 (IQ: 2-3) para uma e 3 (IQ: 2-4) para outra. A concordância simples do número de motivos codificados foi de 49,3%, com *kappa* simples de 0,36 (IC 95%: 0,30–0,40) e *kappa* ponderado de 0,71; IC 95%: 0,68–0,73, indicando confiabilidade moderada.

Em relação aos códigos completos e de capítulos, foram encontradas confiabilidade moderada e substancial, respectivamente (Tabela 1), tanto globalmente quanto em cada estrato. A exceção foi a faixa etária mais jovem, que também apresentou confiabilidade substancial, mesmo considerando o código completo, embora a

diferença não seja significativa. Apesar de estarem na mesma faixa de interpretação, houve uma pequena diferença em relação à confiabilidade por sexo, sendo significativa quando foi considerado apenas o capítulo. Também foram observadas diferenças entre as faixas etárias, sendo os maiores valores de confiabilidade encontrados para as faixas até 34 anos. Com relação ao perfil baseado nos códigos obtidos após consenso, não se observou diferença significativa no número de motivos relatados segundo sexo ou faixa etária, sendo em todos os estratos a mediana igual a 2. Na Tabela 2 são apresentados os motivos de consulta mais frequentemente relatados segundo sexo e faixa etária. Em relação ao sexo, o perfil é grosso modo semelhante, embora “procedimento administrativo geral e inespecífico” apareça entre os motivos mais frequentemente relatados apenas entre os homens. Outra diferença diz respeito à hipertensão arterial. Entre os homens houve predomínio de “pressão arterial elevada,” enquanto entre as mulheres de

**Tabela 1** - Concordância entre as classificadoras em relação aos códigos completos e aos códigos dos capítulos dos motivos de consulta mencionados, segundo sexo e faixa etária, usando a ICPC-2. Serviço de Pronto Atendimento de uma unidade primária de saúde pública localizada na Cidade do Rio de Janeiro (2007).

**Table 1** - Interobserver agreement between classifiers with respect to full and chapter codes when classifying reported reasons for medical appointments, according to respondents' sex and age, using ICPC-2. Walk-in clinic from a public primary health care unit located in the city of Rio de Janeiro.

Variável	Capítulo		Código completo	
	percentual de concordância simples	kappa (IC 95%)	percentual de concordância simples	kappa (IC 95%)
Global	90,2%	0,89 (0,87 - 0,91)	72,5%	0,72 (0,68 - 0,76)
Sexo				
masculino	85,5%	0,83 (0,78 - 0,90)	71,1%	0,70 (0,61 - 0,75)
feminino	91,9%	0,91 (0,88 - 0,93)	73,1%	0,73 (0,68 - 0,76)
Faixa etária				
<25	95,1%	0,94 (0,87 - 0,97)	81,5%	0,81 (0,70 - 0,89)
25-34	94,8%	0,94 (0,87 - 0,98)	77,6%	0,77 (0,70 - 0,85)
35-44	84,0%	0,82 (0,75 - 0,91)	68,9%	0,68 (0,60 - 0,77)
45-54	87,6%	0,86 (0,79 - 0,92)	70,3%	0,70 (0,63 - 0,77)
55+	91,5%	0,90 (0,81 - 0,95)	68,2%	0,68 (0,59 - 0,74)

CIAP-2: Classificação Internacional de Atenção Primária, segunda versão.

ICPC-2: *International Classification of Primary Care*, second version.

**Tabela 2** - Principais motivos de consulta, classificados com a CIAP-2, segundo idade e sexo. Serviço de Pronto Atendimento de uma unidade primária de saúde pública localizada na Cidade do Rio de Janeiro (2007).

**Table 2** - Main reported reasons for medical appointments, using ICPC-2, according to respondents' sex and age, Walk-in clinic from a public primary health care unit located in the city of Rio de Janeiro(2007)

		motivos de consulta	Freq.	%
Sexo	Masculino	Tosse	9	4,8
		Cefaleia	8	4,2
		Febre	7	3,7
		procedimento administrativa geral e inespecífico	7	3,7
		Náusea	5	2,7
		pressão arterial elevada	5	2,7
Feminino	Tosse	25	5,0	
	Cefaleia	20	4,0	
	hipertensão sem complicações	17	3,4	
	sinais/sintomas da garganta	16	3,2	
	Febre	14	2,8	
<35 anos	Febre	15	6,7	
	Tosse	15	6,7	
	Cefaleia	13	5,8	
	sinais/sintomas de garganta	12	5,4	
	Náusea	8	3,6	
	medo de outras doenças respiratória	8	3,6	
35-44 anos	Tosse	8	5,6	
	cefaléia	7	4,9	
	sinais/sintomas da garganta	6	4,2	
	sinais/sintomas da região dorsal	4	2,8	
	sinais/sintomas da região lombar	4	2,8	
	Gripe	4	2,8	
45-54 anos	dor abdominal/cólica generalizada	6	3,6	
	pressão arterial elevada	6	3,6	
	cefaléia	6	3,6	
	hipertensão sem complicações	4	2,4	
	sinais/sintomas dos ombros	4	2,4	
	vertigens/tonturas	4	2,4	
	Tosse	4	2,4	
55 anos ou +	hipertensão sem complicaçõe	9	5,8	
	Tosse	7	4,6	
	procedimento administrativo geral e inespecífico	6	3,9	
	Refer.médico/especialista/clínica/hospital	5	3,3	
	outras perturbações visuais	4	2,6	
	sinais/sintomas coxa/perna	4	2,6	
	sinais/sintomas Joelho	4	2,6	
	diabetes não-insulino dependentes	4	2,6	

CIAP-2: Classificação Internacional de Atenção Primária, segunda versão.

ICPC-2: *International Classification of Primary Care*, second version.

“hipertensão arterial sistêmica”, sem complicações. Já em relação à faixa etária, para o grupo mais jovem (< 35 anos), observa-se uma maior concentração dos motivos mais frequentes relatados, que correspondem a 31,8% do total de motivos relatados neste grupo. Como esperado, motivos relacionados a doenças crônicas prevalentes em nosso meio (hipertensão arterial e diabetes), assim como o encaminhamento para especialistas são mais relatados nas faixas etárias mais elevadas.

## Discussão

A confiabilidade moderada em relação ao número de motivos codificados expressa a dificuldade das classificadoras em distinguir o relato do paciente do relato do médico, uma vez que os dados foram coletados a partir de fichas do atendimento previamente realizado e não diretamente no atendimento. Em alguns casos era absolutamente evidente que a queixa registrada era uma interpretação dada pelo profissional, especialmente quando foram consideradas a linguagem técnica utilizada no registro e a baixa escolaridade característica da população. Também houve dificuldade de codificação por não haver a possibilidade de questionar o paciente acerca de determinados relatos com interpretação dúbia.

É possível supor que o fato de uma das classificadoras ser médica e a outra não possa ter contribuído para algumas divergências de códigos. Foi observado que, apesar de a proposta de ser uma classificação desenvolvida para o uso de leigos, muitas vezes o desconhecimento do que vinha a ser a queixa do paciente, principalmente no caso de diagnósticos, criava uma dificuldade adicional para quem não é médico. Mesmo que o uso da Internet tenha sido de extrema importância para suprir o desconhecimento técnico e a inexistência de uma nomenclatura em português, ainda assim, alguns casos em que existiam diferenças mais sutis entre doenças provocaram grandes dúvidas. Em outras situações, o desconhecimento técnico foi uma vantagem, pois evitava a

conclusão de diagnósticos, fato que aconteceu algumas vezes com a médica.

Outra limitação da classificadora não médica foi o fato de não usar outras queixas citadas pelo paciente para detalhar alguns motivos de consulta. Neste caso, não se tratava de fechar um diagnóstico, mas de compreender com mais detalhes a queixa do paciente. Por exemplo, no caso de um paciente que afirma ter “lesão em órgão alvo pela hipertensão”, o código atribuído deveria ser hipertensão com complicações e não apenas hipertensão.

A ausência de concordância de códigos completos em relação aos problemas sociais é um ponto delicado, especialmente em uma classificação que traz este atributo como um de seus pontos inovadores. Alguns dos motivos de consulta detectados pelo consenso como problemas sociais não foram classificados por uma das codificadoras, por terem sido considerados como uma justificativa para o problema de saúde que originou a consulta ou parte de sua história e não o motivo em si. Por exemplo, a queixa “Paciente com receita (...) refere não estar conseguindo pegar medicação na sua área e veio solicitar inscrição na unidade” foi codificada por uma das classificadoras apenas como um procedimento administrativo, desconsiderando-se o código “problemas relacionados com o sistema de saúde”, que foi, de forma equivocada, enquadrado como parte da história da outra queixa ao invés de ter sido considerado como um novo código. Esse problema pode ser resolvido e deve ser focalizado em treinamento. Problemas com ex-marido e filhos que não são necessariamente crianças também provocaram divergências, dessa vez por uma limitação da classificação que não inclui espaço para classificar esses tipos de relações, mas apenas parceiro, família, amigos e crianças. Optou-se por classificar problemas com ex-marido como “problemas sociais não especificados” e perda de filho que não necessariamente criança em “perda de ou falecimento de uma criança”.

Não houve consenso entre os autores de uma hipótese para explicar a diferença na confiabilidade percebida entre os sexos.



Já em relação à diferença entre as faixas etárias, é possível supor que a maior confiabilidade nas faixas etárias mais jovens em relação às mais intermediárias se deva à maior homogeneidade das queixas e a um perfil mais objetivo do que o das demais. No entanto, em ambos os casos seria necessário outro estudo para avaliar diretamente este ponto. Foram encontrados alguns artigos que tratam da confiabilidade da CIAP. No estudo conduzido por Letrilliart *et al.*<sup>15</sup> foi estudada a confiabilidade interobservador da classificação entre médicos generalistas, treinados na CIAP-2, que classificaram a partir de seu atendimento, e epidemiologistas, que classificaram a partir das informações registradas em uma base de dados. O *kappa* ponderado relativo ao número de motivos foi 0,65 (IC 95%: 0,52-0,77). A concordância simples entre códigos, considerando apenas o nível do capítulo, foi de 69,2% (83 de 120) e o *kappa* simples, 0,84 (IC 95%: 0,78-0,91). Já no estudo de Sampaio *et al.*<sup>8</sup> foi estudada a confiabilidade interobservador da CIAP para classificar motivos de saúde que levaram à interrupção de atividade habituais, registrados em questionário autopreenchido. O *kappa* ponderado relativo ao número de motivos foi de 0,94 (IC 95%: 0,93-0,94), a concordância simples entre códigos, considerando apenas o nível do capítulo, foi de 90,6% e o *kappa* simples, 0,89 (IC 95%: 0,89-0,90). A comparação entre estatísticas *kappa* em populações de estudo diferentes é questionável, uma vez que a medida é influenciada pela prevalência do fenômeno em estudo. No entanto, é possível supor que a pior confiabilidade observada no estudo de Letrilliart *et al.*<sup>15</sup> em relação ao presente estudo sejam explicadas, ao menos parcialmente, pela utilização de fontes diferentes para a captação dos dados (primária vs secundária) e profissionais distintos (clínicos vs epidemiologistas). Já o melhor desempenho verificado no estudo de Sampaio *et al.*<sup>8</sup> poderia ser explicada pelo fato de que os motivos de saúde foram relatados diretamente por quem se queixava e não sofreram a interpretação de terceiros como no caso deste estudo.

Britt *et al.*<sup>16</sup> fizeram um estudo com 347 vídeos de consultas de clínicos gerais. Esses vídeos foram classificados e codificados pelo clínico durante a consulta e também por dois observadores, também médicos, que assistiram aos vídeos das consultas posteriormente. Cada um poderia atribuir até três motivos de consulta, procedimento diferente do manual da classificação, que orienta o registro de todos os motivos mencionados durante o atendimento. A distribuição dos motivos codificados pelos capítulos foi semelhante entre os três. A concordância em relação ao número de problemas classificados (em um máximo de três) pelo clínico foi de 81,0% com o primeiro observador e 63,4% com o segundo. Os observadores tenderam a classificar um número maior de motivos do que o médico durante a consulta. A média das concordâncias dos observadores com o clínico geral foi de 78,8% quando se considerou o nível do capítulo, de 69,6% quando se considerou o capítulo e o componente, e de 55,9% quando foi considerado o código completo. É possível que o maior percentual de concordância do número de motivos codificados em relação ao encontrado no presente estudo expresse a possibilidade de observação da consulta, o que facilita a percepção dos motivos realmente expressos pelos pacientes, mas também o limite de registro de até três motivos de consulta. Já os menores percentuais nos níveis de capítulos e componentes podem representar o impacto de relatos registrados de maneiras diferentes. No caso do presente estudo, as classificadoras avaliaram os mesmos registros.

Emond *et al.*<sup>17</sup> também realizaram um estudo de concordância em que dez médicos de família deveriam codificar 44 entrevistas, atribuindo um a três problemas para cada. O objetivo do estudo era avaliar se havia diferença entre a codificação manual, realizada por cinco dos médicos, e a codificação informatizada, efetuada pelo restante. O estudo avaliou a reprodutibilidade de três aspectos: os diagnósticos registrados, a codificação dos problemas de

forma manual e a codificação informatizada, encontrando concordâncias de 70,5%, 70,2% e 75,0%, respectivamente, e concluiu que não houve diferenças significativas entre as formas de codificação. Apesar de o presente estudo avaliar apenas o aspecto da codificação manual, é possível observar resultados bastante semelhantes.

## Conclusão

Foram encontradas algumas dificuldades para o uso da CIAP na codificação de motivos de consulta em uma unidade primária de saúde. É possível supor que

parte desses problemas seria resolvida com um treinamento mais extensivo de seu uso, tanto de quem irá registrar, quanto de quem irá codificar, no caso de não ser possível que a mesma pessoa realize ambas as funções, o que seria o ideal. Apesar disso, apresentar confiabilidade de moderada a substancial é um indício de que a classificação pode ter um bom desempenho para a codificação de motivos de atendimento em unidade primária em saúde, mesmo que esta seja realizada em momento posterior à consulta, e, portanto, deveria ser utilizada nesse contexto.

---

## Referências

1. WONCA. ICPC. *International Classification of Primary Care*. Oxford Oxford University Press; 1987.
2. WONCA. ICPC-2. *International Classification of Primary care*, second edition. Oxford: Oxford University Press; 1998.
3. Bentsen BG. International classification of primary care. *Scand J Prim Health Care* 1986 Feb;4(1):43-50.
4. Lamberts H, Meads S, Wood M. Results of the international field trial with the Reason for Encounter Classification. *Soz Praventivmed* 1985; 30(2): 80-7.
5. Lebrão ML. [International classification of reasons for consultation in primary health care: a test in some Brazilian areas]. *Rev Saude Publica* 1985; 19(1): 69-78.
6. de Lusignan S, Minmagh C, Kennedy J, Zeimet M, Bommezijn H, Bryant J. A survey to identify the clinical coding and classification systems currently in use across Europe. *Studies in Health Technology and Informatics* 2001; 84(Pt1): 86-9. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/11604711>. [Acessado em 30 outubro de 2010].
7. Rodgers RPC, Sherwin Z, Lamberts H, Okkes IM. ICPC Multilingual Collaboratory: a Web- and Unicode-based system for distributed editing/translating/viewing of the multilingual International Classification of Primary Care. *Studies in Health Technology and Informatics*. 2004; 107(Pt1): 425-9. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/15360848>. [Acessado em 30 outubro de 2010].
8. Sampaio MM, Coeli CM, de Miranda NN, Faerstein E, Werneck GL, Chor D, et al. [Interobserver reliability of the International Classification of Primary Care]. *Rev Saude Publica* 2008; 42(3): 536-41.
9. Reichenheim ME. Confidence intervals for the kappa statistic. *Stata Journal* 2004; 4: 421-8.
10. StataCorp. Stata Statistical Software: Release 8.0. In: College Station TSC, editor; 2003.
11. WONCA. *Classificação Internacional de Cuidados Primários - Segunda Edição*. Oxford University Press; 1999.
12. Soeken KL, Prescott PA. Issues in the use of kappa to estimate reliability. *Med Care* 1986; 24(8): 733-41.
13. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; 20: 37-46.
14. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998; 7(3): 301-17.
15. Letrilliart L, Guiguet M, Flahault A. Reliability of report coding of hospital referrals in primary care versus practice-based coding. *Eur J Epidemiol* 2000; 16(7): 653-9. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/11078123>. [Acessado em 30 outubro de 2010].
16. Britt H, Angelis M, Harris E. The reliability and validity of doctor-recorded morbidity data in active data collection systems. *Scand J Primary Health Care* 1998; 16(1): 50-5 Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/9612880>. [Acessado em 30 outubro de 2010].
17. Emond JG, Cauchon M, Ouellet J, Beaulieu MD, Brailovsky CA, Leduc Y. Interobserver agreement on diagnoses classified according the International Classification of Primary Care. *Can Fam Physician* 1998; 44: 2128-33. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed/9805167>. [Acessado em 30 outubro de 2010].

Recebido em: 07/11/2010

Versão final apresentada em: 15/07/2011

Aprovado em: 08/09/2011