

## Consistência entre fontes de dados e confiabilidade interobservador do Estudo da Morbi-mortalidade e Atenção Peri e Neonatal no Município do Rio de Janeiro

Consistency between data sources and inter-observer reliability in the Study on Neonatal and Perinatal Morbidity and Mortality and Care in the City of Rio de Janeiro

Mônica Rodrigues Campos <sup>1</sup>  
 Maria do Carmo Leal <sup>2</sup>  
 Paulo Roberto de Souza Jr. <sup>3</sup>  
 Cynthia Braga da Cunha <sup>2</sup>

### Abstract

*The objective of this study was to evaluate the quality of data in the research project entitled "Study on Neonatal and Perinatal Morbidity and Mortality and Care in the City of Rio de Janeiro", analyzing the completeness of patient records, inter-observer reliability, and concordance of collected data. The study interviewed a sample of 10,072 post-partum women, corresponding to 10.0% of the deliveries in the City of Rio de Janeiro. This article analyzed the concordance between data on patient records and the reproducibility of questionnaires by mean of rest/retest, using common and prevalence-adjusted Kappa as well as the intra-class correlation coefficient. Losses totaled 4.5%, and the proportion of unknown data on patient records varied from 3.0 to 90.0%. Lower proportions were concentrated in the neonatal assessment and higher ones in the data on maternal hospital admission. There was a high concordance between the data reported by the mother and that written on the patient record, with the Kappa varying from 0.77 to 0.96. In the test/retest verification, Kappa varied from 0.61 to 0.94. This study demonstrated a high inter-observer reliability as well as reliability between the different data sources (interviewee and patient record).*

*Evaluation; Reproducibility of Results; Maternal and Child Health*

### Introdução

A confiabilidade de um experimento, teste ou medição pode ser definida como a capacidade de obter resultados com variações mínimas, quando realizada por diferentes pessoas ou em diferentes momentos <sup>1</sup>.

A confiabilidade verificada em múltiplas avaliações do mesmo fenômeno é de vital importância nas investigações, constituindo-se em um dos temas mais atuais da epidemiologia moderna <sup>2</sup>. Diversas publicações na área da saúde, envolvendo o estudo das variações entre observadores, têm aparecido nas últimas décadas e podem ser encontradas nos levantamentos bibliográficos feitos por Elmore & Feinstein <sup>3</sup>, ou ainda em Fonseca <sup>4</sup> e Gimeno & Souza <sup>5</sup>. Estudos que avaliam as variações intraobservador, em momentos diferentes, têm motivado a produção acadêmica recente nesta área <sup>6,7</sup>.

Desde a introdução da estatística kappa por Cohen <sup>8</sup>, vêm se medindo a confiabilidade entre avaliadores, corrigindo-se pelo acaso. O teste kappa (k) é o procedimento estatístico adequado para avaliar a confiabilidade de variáveis categóricas e nominais <sup>8</sup>.

Em alguns estudos de confiabilidade é suficiente obter-se o cálculo de uma única medida resumo de concordância. Porém, segundo Silva & Pereira <sup>9</sup>, em outros contextos mais complexos, tem-se demonstrado que a estatística kappa apresenta características indesejáveis.

<sup>1</sup> Departamento de Ciências Sociais, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

<sup>2</sup> Departamento de Epidemiologia e Métodos Quantitativos em Saúde, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

<sup>3</sup> Departamento de Informações em Saúde, Centro de Informação Científica e Tecnológica, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

#### Correspondência

Maria do Carmo Leal  
 Rua Leopoldo Bulhões 1480,  
 Rio de Janeiro, RJ  
 21041-210, Brasil.  
 duca@ensp.fiocruz.br

Nesse sentido, tem-se que a prevalência de uma dada condição, na população, pode afetar o resultado final. Prevalências muito baixas tendem a estar associadas a baixos níveis de reprodutibilidade, ou seja, kappa próximo de zero, levando a conclusão de que o valor de  $k$  depende da concordância devida ao acaso. Assim, é possível encontrar-se pequenos valores de kappa devido à prevalência do evento e não a erros substanciais relacionados ao procedimento de aferição empregado<sup>9</sup>. Outros autores, tais como, Tanner & Young<sup>10</sup> e Maclure & Willett<sup>11</sup>, também têm ilustrado a dependência do kappa em relação à prevalência da característica em estudo. Recentemente, aplicações desses modelos em estudos epidemiológicos de confiabilidade apareceram nos artigos de Graham & Jackson<sup>12</sup> e May<sup>13</sup>.

Quando se trata de obter uma medida de confiabilidade para variáveis discretas, ordinais ou mesmo contínuas, pode-se utilizar a tradicional estimativa do coeficiente de correlação de Pearson<sup>14</sup>. Porém, devido à falta de sensibilidade e robustez desta estatística em detectar a discordância interobservadores e ser influenciada por valores extremos, uma alternativa mais apropriada e, comumente utilizada, é o coeficiente de correlação intraclasse<sup>15,16</sup>.

O presente trabalho tem o objetivo de avaliar a qualidade das informações do projeto de investigação “*Estudo da Morbi-mortalidade e da Atenção Peri e Neonatal no Município do Rio de Janeiro*”, analisando a completude das informações dos prontuários, concordância das informações coletadas de diferentes fontes e a confiabilidade interobservador, utilizando-se a estatística kappa e o coeficiente de correlação intraclasse.

## Metodologia

### Amostra e questionários

Foi feita uma amostra estratificada proporcional, onde os estabelecimentos de saúde foram agrupados segundo a proporção de baixo peso ao nascer (BPN) em três estratos: (1) estabelecimentos municipais e federais; (2) estabelecimentos privados conveniados com o SUS, militares, estaduais e filantrópicos; (3) estabelecimentos privados.

Em cada estrato foi selecionada uma amostra de, aproximadamente, 10,0% de parturientes do número previsto de partos em todos os

hospitais de cada estrato, exceto naqueles com menos de duzentos partos por ano (<http://www.saude.rio.rj.gov.br>).

O tamanho da amostra em cada estrato foi estabelecido com o objetivo de comparar proporções em amostras iguais no nível de significância de 5,0% e detectar diferenças de pelo menos 3,0% com poder do teste de 90,0%, baseando-se na proporção de BPN.

No total, foram selecionadas 47 instituições, 12 compuseram o primeiro estrato amostral, correspondendo a 34,8% dos partos; dez no segundo estrato, correspondendo a 34,4% dos partos; e 25 no terceiro, com 30,8% dos partos.

O tamanho inicial calculado foi de 3.282 puérperas para cada estrato. Considerando-se a possibilidade de perdas, o total de cada estrato foi estabelecido em 3.500 puérperas. Ao final da pesquisa foram realizadas 10.072 entrevistas, no período compreendido entre julho de 1999 e fevereiro de 2001.

Foram utilizados três tipos de questionários para coleta de dados. O primeiro aplicado à mãe no pós-parto imediato, composto por variáveis de identificação materna e paterna, escolaridade e renda, local e condições de moradia, antecedentes obstétricos, dados antropométricos maternos, informações sobre a gestação em questão, assistência pré-natal, enfermidades e uso de medicamentos na gestação, trabalho de parto e avaliação da assistência prestada a ela e ao recém-nascido (RN). O segundo questionário, preenchido com base em dados disponíveis no prontuário médico, contemplava questões acerca do pré-natal, admissão, pré-parto e parto, medicamentos e intervenções realizadas, evolução da puérpera e do RN. O terceiro questionário levantava informações sobre as condições de alta ou óbito do conceito.

Em uma amostra aleatória de 2,5% das mães entrevistadas foi feita, pelos supervisores de campo, uma replicação de um questionário resumido, composto por questões preenchidas previamente pelo bolsista, com o intuito de avaliar a qualidade da entrevista.

Para os dois primeiros estratos amostrais foi realizada observação do pré-parto em uma subamostra de 5,0% dos partos (diurnos e noturnos), para verificação de procedimentos padronizados do atendimento, por um período de até quatro horas, para cada gestante amostrada.

### **Treinamento de entrevistadores e trabalho de campo**

As equipes de entrevistadores deste estudo foram compostas por acadêmicos bolsistas de medicina e enfermagem, selecionados por concurso público e cedidos pela Secretaria Municipal de Saúde do Rio de Janeiro (SMS-RJ), especialmente para atuar na pesquisa. O treinamento constou de duas etapas; a primeira, com a participação do grupo da pesquisa e de especialistas convidados, constou de aulas expositivas sobre temas de abrangência do estudo, tais como os sistemas de informação em saúde, dinâmica de grupo sobre parto humanizado, noções básicas em epidemiologia e ética em pesquisas na área da saúde, terminologia técnica e farmacológica aplicadas à área materno-infantil. A segunda constou de uma aplicação, em campo, dos questionários já elaborados e de um treinamento para observação dos procedimentos na sala de pré-parto.

O estudo piloto (n = 300 questionários) ocorreu em um Centro de Saúde-Escola e em uma Maternidade-Escola, cujos questionários não foram incluídos na base definitiva de dados.

Elaborou-se um *manual para os entrevistadores*, contendo um livro de registros (memória de campo), a escala da equipe, instruções para a observação do pré-parto, realização da entrevista e do resumo de alta. Uma relação dos termos técnicos utilizados nos questionários e dos medicamentos mais utilizados em maternidades e unidades neonatais, também compunha este material.

Esse manual foi sistematicamente atualizado no decorrer da pesquisa, a partir da revisão dos questionários preenchidos em campo. As questões que suscitaram um maior número de correções foram identificadas e corrigidas na revisão.

A *supervisão de campo* foi realizada por três médicos com formação em pesquisa científica, que tinham como tarefa a elaboração de escalas dos bolsistas por instituição e turnos das atividades (diurnos e noturnos). No campo, esses profissionais exerciam a função de esclarecer dúvidas, suprir questionários, quantificar e justificar as perdas ocorridas e aplicar o questionário de reteste.

### **Análise da consistência da revisão e digitação**

Antes da digitação, os questionários eram submetidos a uma *revisão*, realizada por profissionais encarregados especificamente desta tarefa, com o objetivo de identificar inconsistências,

dando retorno imediato aos entrevistadores e supervisores de campo. Os questionários eram compostos, em sua maioria, de questões fechadas, mas algumas perguntas semi-abertas, como por exemplo, o bairro, município, região administrativa, estabelecimento de saúde, dentre outras, demandavam uma *codificação prévia*. Para verificar a *consistência do trabalho dos revisores*, em uma remessa de cerca de seiscentos questionários do primeiro lote de mil, foi realizado um trabalho paralelo de segunda revisão pela coordenação do trabalho de campo.

No processo de digitação, optou-se por uma validação via *dupla digitação*. Dessa forma, todos os questionários da pesquisa foram digitados e redigitados. Caso houvesse discordância, o digitador responsável pelo instrumento fazia a correção utilizando como base o respectivo questionário.

Foi também elaborado um *programa de crítica*, com base no Statistical Package for Social Sciences (SPSS) versão 8.0, para a correção automática de variáveis com codificações associadas ou dependentes.

### **Análise estatística**

Para avaliação da qualidade dos prontuários foram contabilizadas as proporções de variáveis com informações ignoradas, por blocos de variáveis, segundo os estratos amostrais do estudo.

A concordância entre fontes foi testada pela comparação de fontes de informações independentes – a referida pelas mães e a coletada em prontuário médico, sendo medida pelo cálculo da estatística kappa tradicional e do kappa ajustado pela prevalência (IC95%).

A confiabilidade interavaliadores (teste-reteste) foi avaliada pela comparação das aferições de informações independentes – a realizada pelos acadêmicos bolsistas e uma segunda aplicação pelos médicos supervisores de campo, tendo sido medida pelo cálculo do coeficiente de correlação intraclassa (ICC), pela estatística kappa tradicional e o kappa ajustado pela prevalência (IC95%).

A concordância medida pelo kappa seguirá a orientação da literatura especializada: kappa < 0,10 ausente; de 0,11 a 0,40 fraca; de 0,41 a 0,60 discreta; de 0,61 a 0,80 moderada; 0,81 a 0,99 substancial e 1,00 concordância perfeita<sup>17</sup>. Porém, cabe destacar que, em geral, recomenda-se maior ênfase ao valor numérico propriamente, uma vez que tais escalas podem variar entre os autores.

Serão comparados os valores do kappa “tradicional” e o “ajustado pela prevalência”, ao ní-

vel de significância de 5,0%. Para o cálculo do kappa ajustado pela prevalência utilizou-se o programa PEPI<sup>18</sup>, disponível na Internet (<http://www.usd-inc.com/pepi.html>).

## Resultados

Nos três estratos, foram realizadas 10.072 entrevistas, com uma perda total de 4,5%, tendo sido maior (9,3%) no estrato três, que congrega as maternidades privadas. O principal motivo das perdas foi a recusa da mãe em participar do estudo.

Na análise de consistência entre a primeira e a segunda digitação dos questionários da entrevista materna houve 2,0% de itens discordantes entre a dupla digitação nas duas primeiras maternidades, diminuindo para 0,4% nas duas seguintes. A melhoria deste índice se deu com base na identificação de variáveis com maior discordância, concentrando-se 23,0% dos erros em apenas uma delas.

No que diz respeito ao preenchimento de informações nos prontuários, como pode ser visto na Tabela 1, a variável com maior proporção de dados incompletos no item identificação materna foi a “procedência”, com 33,0% de ausência para o conjunto dos prontuários, atingindo 56,2% no estrato 2. Quanto às informações sobre o pré-natal, o número de consultas foi o pior preenchido, chegando a estar ausente em quase 50,0% dos registros e em 78,2% dos prontuários no estrato 3.

Para os dados coletados na admissão da gestante, registrou-se, ao contrário, presença de “partograma” na quase totalidade dos prontuários. As informações de “peso”, “altura” e “frequência cardíaca da mãe” deixaram de ser registradas para a grande maioria, chegando a atingir índices acima de 80,0% para os dois primeiros e de 66,1% para a “frequência cardíaca”. Além disso, para um quarto das puérperas não havia registro da “pressão arterial”. Quanto ao monitoramento fetal, os “batimentos cardíacos” não estiveram anotados em 42,0% do total, ficando ausente em um quarto dos prontuários do primeiro estrato, um terço no segundo e em mais da metade do terceiro estrato. Situação semelhante foi identificada quanto ao registro da hora da “amniorrexe” ou “amniotomia”.

Ainda na Tabela 1, observa-se que o registro da evolução da mãe no pós-parto imediato mostrou que cerca de 90,0% delas foram vistas pelo médico nas “primeiras seis horas”, com exceção do estrato 1 que visitou apenas 70,0% da clientela. A ampliação do tempo de visita

para “doze horas” do pós-parto não modificou a situação descrita.

Quanto à evolução do recém-nascido na sala de parto, notifica-se uma melhora substantiva dos registros, ficando sem informação do “peso ao nascer” apenas 3,3%. O “apgar” no primeiro e quinto minutos e a medida do “comprimento” do recém-nato estiveram com valores não registrados próximos a 10,0%. Em quase todos os prontuários, 93,4%, registrou-se a presença ou ausência de “anomalias congênitas”. O registro do tipo de “aleitamento” do recém-nascido esteve também ausente em 17,2% da amostra, mas em menor proporção, 9,6%, no estrato 3. A variável pior preenchida neste grupo foi a “idade gestacional”, 34,1%, tomando como referência o escore de “capurro” que foi amplamente utilizado apenas no estrato 1, em quase 90,0% dos partos ali ocorridos.

A Tabela 2 apresenta a análise de concordância entre algumas informações referidas pela mãe e os dados colhidos, com base no prontuário, com as estatísticas de kappa bruto e kappa ajustado pela prevalência. Encontrou-se concordância acima de 0,90 (kappa ajustado) e percentual de discordantes abaixo de 6,0%, entre as informações relatadas e os dados de prontuário para as variáveis de realização de “pré-natal”, “fator RH”, “tipo de parto” e “apresentação fetal” em todos os estratos. Foram também elevadas as concordâncias para “episiotomia” e “óbito fetal”, sendo que ambas as variáveis apresentaram melhores resultados no terceiro estrato (kappa acima de 0,90) e piores no segundo (kappa próximo de 0,70). No primeiro e segundo estratos, a proporção de discordância quanto ao “óbito fetal” foi respectivamente de 10,0% e 15,8%, enquanto no estrato 3, sua concordância foi de 100,0%. Além disso, concordâncias fracas foram observadas para a variável “amniorrexe” com 0,34, 0,29 e 0,52 no primeiro, segundo e terceiro estratos, respectivamente, bem como se observou também elevados valores de informações discordantes, 33,0%, 35,0% e 24,0%.

Destacam-se, ainda, baseando-se na da Tabela 2, consideráveis diferenças entre as medidas de comparabilidade do kappa bruto e ajustado pela prevalência, principalmente quando o valor do percentual de discordância é muito pequeno e, mesmo assim, o kappa bruto apresenta-se muito baixo. Um exemplo é o caso da variável “fez pré-natal” (no terceiro estrato), na qual tem-se um percentual de discordância de 2,4% com um kappa bruto de 0,08 (péssimo) e um kappa ajustado de 0,95 (ótima concordância). Porém, quando a proporção de discordantes é razoavelmente alta, como no caso da va-

Tabela 1

Distribuição de dados incompletos no prontuário das puérperas segundo estratos amostrais.  
Município do Rio de Janeiro, Brasil, 1999-2001.

Variáveis do prontuário	Total de ignorados			Total n = 10.072 %
	Estrato 1 n = 3.504 %	Estrato 2 n = 3.468 %	Estrato 3 n = 3.100 %	
<b>Identificação da mãe</b>				
Procedência	18,6	56,2	23,6	33,1
Município	8,0	21,0	3,0	10,9
Bairro	9,1	21,5	2,9	11,5
Idade (anos)	7,7	7,5	4,7	6,7
<b>Pré-natal</b>				
Fez pré-natal	8,2	23,7	42,3	24,0
Cartão do pré-natalno prontuário	3,8	6,5	8,9	6,3
Consultas no pré-natal	23,9	43,4	78,2	47,3
Fator RH	13,9	12,3	28,4	17,8
<b>Dados da admissão</b>				
Partograma no prontuário	3,1	2,2	3,3	2,8
Pressão arterial	26,7	29,5	20,2	25,7
Peso	94,0	94,3	67,8	86,1
Altura	96,5	98,8	70,5	89,3
Frequência cardíaca	72,2	92,5	29,7	66,1
Fundo do útero	23,0	24,3	90,6	44,3
Idade gestacional (data da última menstruação)	28,0	18,2	48,4	30,9
Batimentos cárdio fetais	24,7	32,6	72,4	42,1
Hora da amniorrexe	35,3	26,8	54,5	38,3
Hora da amniotomia	31,6	24,7	50,0	34,9
<b>Evolução da mãe</b>				
Visitas nas primeiras 6h	30,5	10,4	11,4	17,7
Visitas das 6 às 12h	30,9	11,6	11,5	18,3
<b>Evolução do recém-nascido</b>				
Peso ao nascer	4,1	2,9	3,1	3,3
Comprimento	23,1	7,5	6,9	12,8
Apgar no 1º minuto	3,9	19,3	4,3	9,3
Apgar no 5º minuto	3,7	19,4	4,5	9,3
Idade gestacional (capurro)	11,3	41,4	51,7	34,1
Anomalias congênicas	6,7	7,3	5,5	6,6
Tipo de amamentação	22,6	19,2	9,6	17,4

Tabela 2

Comparações entre informações referidas pela mãe e colhidas no prontuário das puérperas segundo estratos amostrais. Município do Rio de Janeiro, Brasil, 1999-2001.

Variáveis comparadas	Discordantes		Medidas de comparabilidade		
	n	%	Kappa	Kappa*	Sign (95%)
<b>Estrato 1</b>					
Fez pré-natal	120	3,7	0,73 (0,69; 0,78)	0,93	0,000
Fator RH (mãe)	143	5,8	0,74 (0,69; 0,78)	0,88	0,000
Amniorrexe	790	32,9	0,36 (0,32; 0,40)	0,34	0,000
Tipo de parto	125	3,7	0,92 (0,90; 0,93)	0,94	0,000
Apresentação fetal	183	6,3	0,52 (0,46; 0,57)	0,92	0,000
Episiotomia	298	9,8	0,80 (0,78; 0,82)	0,80	0,000
Óbito fetal	7	10,0	0,72 (0,53; 0,91)	0,80	0,000
<b>Estrato 2</b>					
Fez pré-natal	77	2,9	0,66 (0,59; 0,73)	0,94	0,000
Fator RH (mãe)	103	3,8	0,81 (0,77; 0,84)	0,92	0,000
Amniorrexe	933	35,3	0,30 (0,26; 0,34)	0,29	0,000
Tipo de parto	71	2,1	0,96 (0,95; 0,97)	0,97	0,000
Apresentação fetal	175	5,8	0,96 (0,38; 0,51)	0,92	0,000
Episiotomia	339	11,0	0,78 (0,76; 0,80)	0,78	0,000
Óbito fetal	6	15,8	0,54 (0,21; 0,87)	0,68	0,000
<b>Estrato 3</b>					
Fez pré-natal	43	2,4	0,08 (-0,05; 0,21)	0,95	0,000
Fator RH (mãe)	35	1,6	0,92 (0,90; 0,95)	0,97	0,000
Amniorrexe	372	23,8	0,47 (0,42; 0,52)	0,52	0,000
Tipo de parto	44	1,5	0,94 (0,92; 0,96)	0,98	0,000
Apresentação fetal	134	7,2	0,55 (0,48; 0,61)	0,90	0,000
Episiotomia	66	2,4	0,87 (0,84; 0,90)	0,95	0,000
Óbito fetal	0	0,0	1,00 (1,00; 1,00)	1,00	0,014
<b>Total</b>					
Fez pré-natal	240	3,1	0,67 (0,64; 0,71)	0,94	0,000
Fator RH (mãe)	283	3,8	0,81 (0,79; 0,83)	0,92	0,000
Amniorrexe	2100	31,7	0,37 (0,35; 0,39)	0,37	0,000
Tipo de parto	241	2,4	0,95 (0,95; 0,96)	0,96	0,000
Apresentação fetal	494	6,3	0,50 (0,47; 0,54)	0,92	0,000
Episiotomia	703	7,8	0,84 (0,83; 0,85)	0,84	0,000
Óbito fetal	13	11,4	0,68 (0,51; 0,84)	0,77	0,000

\* Ajustado pela prevalência.

riável “amniorrexe” (no segundo estrato) com 35,0% de discordantes, tem-se as medidas de comparabilidades com valores bem próximos (kappa bruto 0,30 e ajustado 0,29).

A Tabela 3 mostra a comparação das variáveis discretas e ordinais entre as informações referidas pelas mães no teste-reteste. Observou-se elevados ICC, que variaram de 0,78 para a variável “grau de instrução materno” a 1,00 para a “idade” materna referida.

Na Tabela 4, são apresentadas as comparações entre os entrevistadores e supervisores de campo para as variáveis categóricas. Calculou-se a estatística kappa bruto e a ajustada pela prevalência. Tomando como referência o kappa ajustado pela prevalência, observa-se, na Tabela 4, uma menor concordância para a anotação da “hipertensão arterial sistólica” com valor de 0,61 e uma maior para “a mãe fez aborto”, com kappa de 0,96, sendo todas as estimativas estatisticamente significativas e encontrando-se primordialmente acima de 0,80, configurando uma situação de muito boa concordância.

## Discussão

Neste estudo, pôde-se verificar que o trabalho de supervisão constituiu-se em uma ferramenta fundamental para o bom desempenho do trabalho de campo, evitando perdas, corrigindo falhas e esclarecendo as dúvidas quanto ao preenchimento de questionários. A baixa pro-

porção de perdas, inferior a 5,0%, pode ser considerada um ponto alto desta pesquisa, principalmente por incluir a população de elevado nível sócio-econômico e, mesmo neste grupo, ter conseguido entrevistar mais de 90,0% das mulheres selecionadas. A dupla digitação e o trabalho de revisão dos questionários forneceram um acurado mecanismo de controle de qualidade das informações coletadas e garantiu a consistência dos resultados obtidos. A reprodutibilidade dos questionários, aferida pelo teste-reteste, também mostrou o bom desempenho do trabalho de campo e a concordância dos dados.

A análise da completude dos prontuários denotou uma diferenciação importante no preenchimento das informações relativas ao momento de pré-parto e do parto. Dados vitais indispensáveis para avaliação clínica das mães e recém-nascidos, estiveram muito ausentes nos registros da admissão. Foram, entretanto, anotados em mais de 90,0% nos registros da sala de parto. Se por um lado é louvável que os registros das salas de parto sejam elevados, é grave a ausência de informações na admissão, dado que as maternidades do Município do Rio de Janeiro, na sua grande maioria, funcionam em sistema de plantões, permitindo que o profissional médico trabalhe por um período de 24 horas seguidas, só voltando ao estabelecimento na semana seguinte. Vale lembrar que a ausência de anotações sobre os dados clínicos das puérperas, na admissão e no seu seguimento, impossibilitam o acompanhamento adequado da sua evolução. A ausência de informações em prontuários tem sido notificado por outros pesquisadores nacionais<sup>19,20</sup>.

Outro fato relevante percebido aqui foi que na evolução do recém-nascido, a idade gestacional aferida pelo Capurro apresentou grandes diferenças entre os estratos, devido a pouca utilização desse escore por maternidades privadas e conveniadas com o SUS. Este dado é importante para os pesquisadores da área materno-infantil que pretendam utilizar o indicador de idade gestacional disponível em prontuários hospitalares do Município do Rio de Janeiro.

Um problema na análise de concordância dos dados dos prontuários médicos é que a informação obtida pode não estar refletindo a realidade da assistência, uma vez que pode não ter sido registrada porque o médico não identificou qualquer necessidade de fazê-lo, ficando a escolha a seu critério, pela não existência de uma normatização dos serviços sobre o que deve ou não deve ser referido no prontuário<sup>19</sup>. Frequentemente, a baixa concordância para

Tabela 3

Comparação entre as informações referidas pelas mães no teste-reteste\*, para variáveis discretas/ordinais (n = 247). Município do Rio de Janeiro, Brasil, 1999-2001.

Variáveis contínuas/ordinais	ICC IC95%
Idade da mãe (em anos)	1,00 (0,99; 1,00)
Última série com aprovação da mãe	0,95 (0,93; 0,96)
Gravidez de instrução materno	0,78 (0,72; 0,83)
Renda mensal materna	0,94 (0,92; 0,95)
Última série com aprovação do pai	0,92 (0,89; 0,94)
Gravidez de instrução paterno	0,81 (0,75; 0,86)
Números de moradores em casa	0,86 (0,83; 0,89)
Números de abortos provocados	0,95 (0,92; 0,97)
Mês da última menstruação	0,82 (0,76; 0,86)
Números de consultas pré-natais	0,96 (0,95; 0,97)
Mês de detecção da sífilis	0,97 (0,81; 0,99)

\* 1ª aferição feita pelos bolsistas após treinamento padronizado e 2ª aferição pelos supervisores de campo.

Tabela 4

Comparação entre as informações referidas pelas mães no teste-reteste\*, para variáveis categóricas (n = 247). Município do Rio de Janeiro, Brasil, 1999-2001.

Variáveis categóricas	Kappa (95%)	Kappa ajustado	p-valor
Situação conjugal	0,75 (0,66; 0,85)	0,85	0,00
Emprego paterno	0,82 (0,74; 0,90)	0,91	0,00
Tipo sanguíneo da mãe	0,90 (0,84; 0,96)	0,94	0,00
Vacina antitetânica	0,86 (0,81; 0,91)	0,87	0,00
História de aborto	0,96 (0,93; 0,99)	0,96	0,00
Desejo de engravidar	0,82 (0,76; 0,88)	0,84	0,00
Orientação sobre a importância do pré-natal	0,75 (0,66; 0,84)	0,86	0,00
Orientação sobre os sinais da hora do parto	0,77 (0,69; 0,85)	0,84	0,00
Orientação sobre a alimentação	0,68 (0,59; 0,77)	0,78	0,00
Orientação sobre o aleitamento materno	0,77 (0,69; 0,85)	0,84	0,00
Orientação sobre os cuidados com o RN	0,71 (0,63; 0,80)	0,79	0,00
Presença de sífilis nesta gravidez	0,46 (0,26; 0,66)	0,89	0,00
Aminiorrexe antes da internação	0,82 (0,75; 0,90)	0,87	0,00
Procurou por outro hospital para o parto	0,84 (0,77; 0,91)	0,91	0,00
Satisfação com o atendimento pré-natal	0,77 (0,70; 0,84)	0,82	0,00
Satisfação com o atendimento ao parto	0,75 (0,68; 0,82)	0,82	0,00
Hipertensão arterial sistólica**	0,34 (0,21; 0,47)	0,61	0,00
Óbito fetal ou neonatal**	0,38 (0,14; 0,61)	0,91	0,00

\* 1ª aferição feita pelos bolsistas após treinamento padronizado e 2ª aferição pelos supervisores de campo.

\*\* Dados procurados no prontuário médico.

variáveis de um questionário baseado em um prontuário advém destas dificuldades, às quais se associam a ilegibilidade ou a ambigüidade das anotações médicas<sup>21</sup>.

A concordância entre o relato da mãe e o registro no prontuário médico foi elevada, com exceção para a variável “aminiorrexe”, que demanda informação anterior à entrada no hospital, não decorrente da observação direta e, talvez por isso mesmo, negligenciada pelos profissionais de saúde.

Oliveira<sup>22</sup>, estudando gestantes do Município do Rio de Janeiro, também descreveu uma elevada concordância entre o relato do peso pela gestante e a medida pela pesquisadora.

Veras & Martins<sup>21</sup>, analisando o grau de concordância entre os dados anotados nos formulários da AIH e prontuários médicos, em hospitais do Município do Rio de Janeiro, encontraram uma baixa concordância para a variável “óbito do recém-nascido”, estando mais bem informado nos prontuários médicos.

Foi também muito elevada a confiabilidade para variáveis discretas e ordinais, neste estudo. Na análise das variáveis categóricas verificou-se uma concordância fraca para a informação hipertensão arterial sistólica e óbito fetal ou neonatal, e discreta para a puérpera com

ocorrência de sífilis na gestação, ao utilizar o kappa tradicional. Porém, em todas essas variáveis, o uso da estatística, ajustada pela prevalência, mais que duplicou os valores da concordância, mostrando, dessa forma, o quanto é adequado usá-la para evitar distorções nos resultados.

No reteste, os supervisores, que eram médicos, tiveram maior êxito em encontrar as informações nos prontuários do que os entrevistadores, estudantes de medicina e enfermagem. Houve maior familiaridade dos médicos com a “ilegibilidade” ou a “ambigüidade das anotações médicas”, como referido por Veras & Martins<sup>21</sup>.

Diante dessas constatações, sugere-se que o preenchimento adequado dos prontuários materno-infantis seja considerado um indicador, de fácil acesso, um evento sentinela, da qualidade do atendimento a este grupo.

Por fim, no presente trabalho, pôde-se mostrar que tanto as informações colhidas das diferentes fontes (puérperas e prontuários), quando aquelas obtidas no teste-reteste tiveram, em geral, substancial concordância, sugerindo alta estabilidade das informações coletadas no projeto de investigação “*Estudo da Morbi-mortalidade e da Atenção Peri e Neonatal no Município do Rio de Janeiro*”.



## Resumo

*O objetivo deste estudo é avaliar a qualidade das informações do projeto de investigação "Estudo da Morbi-mortalidade e da Atenção Peri e Neonatal no Município do Rio de Janeiro", analisando a completude dos prontuários, a confiabilidade interobservador e a concordância das informações coletadas. Este estudo entrevistou uma amostra de 10.072 puérperas, correspondente a 10,0% dos partos ocorridos no Município do Rio de Janeiro. Analisou-se a concordância das informações dos prontuários e a reprodutibilidade dos questionários por meio do teste e reteste, utilizando-se o kappa comum e ajustado pela prevalência, bem como o coeficiente de correlação intraclass. As perdas contabilizaram 4,5% e a proporção de dados ignorados no prontuário variou de 3,0 a 90,0%. As menores proporções concentravam-se na avaliação do neonato e as maiores nos dados da admissão materna. Foi alta a concordância entre as informações relatadas pela mãe e as anotadas no prontuário, variando o kappa de 0,77 a 0,96. Nas aferições do teste reteste o kappa variou de 0,61 e 0,94. Este estudo demonstrou uma alta confiabilidade interobservador, bem como entre as diferentes fontes de informação utilizadas; a entrevistada e o prontuário.*

*Avaliação; Reprodutibilidade dos Resultados; Saúde Materno-Infantil*

## Colaboradores

M. R. Campos participou da construção dos bancos de dados e análise dos dados, desenvolvimento da metodologia empregada, análise estatística e redação do artigo. M. C. Leal foi responsável pela definição dos indicadores analisados, análise e discussão dos resultados e redação final do artigo. P. B. Souza Jr. colaborou na construção do banco de dados e análise estatística. C. B. Cunha contribuiu na construção do banco de dados e análise estatística, e na redação da metodologia.

## Referências

1. Almeida Filho N. Epidemiologia sem números: uma introdução crítica à ciência epidemiológica. Rio de Janeiro: Campus; 1989.
2. Rothmann KJ, Greenland S. Modern epidemiology. 2<sup>nd</sup> Ed. Philadelphia: Lippincott-Raven; 1996.
3. Elmore JG, Feinstein AR. Publications on observer variability. *J Clin Epidemiol* 1992; 45:567-80.
4. Fonseca SC. Qualidade da mensuração na pesquisa perinatal: apreciação crítica da literatura sobre baixo peso ao nascer [Dissertação de Mestrado]. Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro; 1997.
5. Gimeno SGA, Souza JMP. Reprodutibilidade de informação em estudo de fatores de risco para o diabetes mellitus insulino-dependente. *Rev Saúde Pública* 1997; 31:508-11.
6. Griep RH, Chor D, Faerstein E, Lopes CS. Confiabilidade teste-reteste de aspectos da rede social no Estudo Pró-Saúde. *Rev Saúde Pública* 2003; 37:379-85.
7. Coutinho ESF, Silva SD, Paiva CHA, Freitas GG, Miranda LP, Mendonça R. Confiabilidade da informação sobre o uso recente de medicamentos em um estudo caso-controle de base hospitalar. *Cad Saúde Pública* 1999; 15:553-8.
8. Cohen JA. Coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20:37-46.
9. Silva EF, Pereira MG. Avaliação das estruturas de concordância e discordância nos estudos de confiabilidade. *Rev Saúde Pública* 1998; 32:383-93.
10. Tanner MA, Young MA. Modeling agreement among rates. *J Am Stat Assoc* 1985; 80:175-80.
11. Maclure M, Willet WC. Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol* 1987; 126:161-9.
12. Graham P, Jackson R. The analysis of ordinal agreement data: beyond weighted kappa. *J Clin Epidemiol* 1993; 46:1055-62.
13. May SM. Modeling observer agreement – an alternative to kappa. *J Clin Epidemiol* 1994; 47:315-24.
14. Snedecor GW, Cochran WG. Statistical methods. 6<sup>th</sup> Ed. Ames: The Iowa University Press; 1967.
15. Fleiss JL. Large sample standard errors of kappa and weighted kappa. *Psychol Bull* 1969; 12:323-7.
16. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86:420-8.
17. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998; 7:301-17.
18. Abramson JH, Gahlinger PM. Computer programs for epidemiologists: PEPI Version 3. London: Brixton Books; 1999.
19. Silva T, Schenkel EP, Mengue SS. Nível de informação a respeito de medicamentos prescritos a pacientes ambulatoriais de hospital universitário. *Cad Saúde Pública* 2000; 16:449-55.
20. Santos SR, Paula AFA, Lima JP. O enfermeiro e sua percepção sobre o sistema manual de registro no prontuário. *Rev Latinoam Enfermagem* 2003; 11: 80-7.
21. Veras CMT, Martins MSA. Confiabilidade dos dados nos formulários de autorização de internação hospitalar (AIH), Rio de Janeiro, Brasil. *Cad Saúde Pública* 1994; 10:339-55.
22. Oliveira AF. Estudo sobre validação das informações de peso e estatura em gestantes no Município do Rio de Janeiro [Dissertação de Mestrado]. Rio de Janeiro: Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz; 2001.

---

Recebido em 29/Set/2003

Versão final reapresentada em 09/Fev/2004

Aprovado em 04/Mar/2004