

Epidemiologia & Saúde


Fundamentos, Métodos, Aplicações



Naomar de Almeida Filho
Mauricio L. Barreto

Epidemiologia & Saúde

Fundamentos, Métodos, Aplicações

- Os autores deste livro e a EDITORA GUANABARA KOOGAN LTDA. empenharam seus melhores esforços para assegurar que as informações e os procedimentos apresentados no texto estejam em acordo com os padrões aceitos à época da publicação, e todos os dados foram atualizados pelos autores até a data da entrega dos originais à editora. Entretanto, tendo em conta a evolução das ciências da saúde, as mudanças regulamentares governamentais e o constante fluxo de novas informações sobre terapêutica medicamentosa e reações adversas a fármacos, recomendamos enfaticamente que os leitores consultem sempre outras fontes fidedignas, de modo a se certificarem de que as informações contidas neste livro estão corretas e de que não houve alterações nas dosagens recomendadas ou na legislação regulamentadora. Adicionalmente, os leitores podem buscar por possíveis atualizações da obra em <http://gen-io.grupogen.com.br>.
- Os autores e a editora se empenharam para citar adequadamente e dar o devido crédito a todos os detentores de direitos autorais de qualquer material utilizado neste livro, dispondo-se a possíveis acertos posteriores caso, inadvertida e involuntariamente, a identificação de algum deles tenha sido omitida.
- Direitos exclusivos para a língua portuguesa
Copyright © 2012 by
EDITORA GUANABARA KOOGAN LTDA.
Uma editora integrante do GEN | Grupo Editorial Nacional
Travessa do Ouvidor, 11
Rio de Janeiro – RJ – CEP 20040-040
Tels.: (21) 3543-0770/(11) 5080-0770 | Fax: (21) 3543-0896
www.editoraguanabara.com.br | www.grupogen.com.br | editorial.saude@grupogen.com.br
- Reservados todos os direitos. É proibida a duplicação ou reprodução deste volume, no todo ou em parte, em quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação, fotocópia, distribuição pela Internet ou outros), sem permissão, por escrito, da EDITORA GUANABARA KOOGAN LTDA.
- Edição eletrônica:  Diagrama Ação - Produção Editorial

■ Ficha catalográfica

A446e

Almeida Filho, Naomar de, 1952-

Epidemiologia & saúde : fundamentos, métodos, aplicações / Naomar de Almeida Filho, Mauricio Lima Barreto. - Rio de Janeiro : Guanabara Koogan, 2011.

il.

ISBN 978-85-277-1619-2

1. Epidemiologia. 2. Saúde pública. I. Barreto, Mauricio Lima. II. Título.

11-6186.

CDD: 614.4

CDU: 616-036.22

21

Modelos Básicos de Análise Epidemiológica

Vilma Sousa Santana, Inês Dourado, Ricardo Ximenes e Sandhi Barreto

A escolha do método de análise mais adequado para estudos epidemiológicos baseia-se na pergunta de investigação, nos objetivos, desenho, tipo e características das bases de dados e na natureza das variáveis do estudo, dentre outros aspectos discutidos na Parte 2 deste volume. Não existe uma receita a ser seguida, pois mesmo resultados preliminares podem indicar direções valiosas para os procedimentos de análise a serem adotados. Os métodos disponíveis de análise são múltiplos e relativamente fáceis para o pesquisador utilizá-los, porque muitos deles fazem parte de pacotes estatísticos que permitem facilmente o manuseio dos dados e a obtenção de medidas interpretáveis.

Neste capítulo, pretendemos abordar as técnicas de análise mais comuns em Epidemiologia, na perspectiva de apresentar um guia prático, de nível intermediário, para familiarização do leitor com a análise epidemiológica. Mostraremos os principais recursos e instrumentos básicos de análise, incluindo análise descritiva, elaboração de gráficos, tabelas de frequências simples e relativas, convergindo para as medidas de associação desenvolvidas especialmente para a pesquisa epidemiológica. Enfatiza-se o uso apropriado de tais métodos, bem como a necessidade de sua contextualização teórica, com interpretação, explicação e compreensão adequadas, concluindo-se com a exposição dos limites desses métodos, especialmente da inferência estatística como base para a interpretação dos resultados.

► Análise de estudos descritivos

Na Epidemiologia, estudos descritivos têm como objetivo o conhecimento e o registro sistemático da extensão e da gravidade dos problemas de saúde na população e seus padrões de distribuição de enfermidades, agravos, exposições, programas, custos de intervenções, dentre outros aspectos. Por essas características, são de grande utilidade para os serviços de saúde e para o conhecimento geral da magnitude, repercussões e tendências das doenças e agravos. Quando se destinam a realizar testes iniciais, preliminares, sobre hipóteses causais, alguns os chamam de estudos exploratórios, especialmente em comparação aos estudos realizados com o propósito de identificar relações causais entre exposição e desfecho. Os estudos exploratórios se justificam nas situações em que ainda é limitado o

conhecimento existente sobre a relação entre variáveis de exposição que são possíveis fatores de risco, ou variáveis descritoras de interesse, e o desfecho, comumente um evento relacionado com saúde, estado de saúde, ou nível de desenvolvimento. Portanto, contribuem com informações preliminares, não conclusivas, sobre causalidade. Ou seja, não permitem a completa avaliação de hipóteses sobre fatores de risco e/ou causas, mas permitem identificar hipóteses para teste em outros estudos que utilizem abordagens analíticas mais específicas e adequadas.

Não obstante, tais estudos são muito empregados no planejamento e na programação de ações de saúde, por serem de fácil e rápida execução, além do baixo custo, especialmente quando conduzidos com dados secundários. Nesses casos, estudos exploratórios de fato podem ser os mais adequados às perguntas que fundamentam a pesquisa, permitindo a obtenção de resultados válidos e conclusivos.

Análises exploratórias são valiosas nas descrições de situações de saúde, ao permitir que gestores conheçam a situação de saúde de áreas, regiões, distritos, grupos vulneráveis como idosos, gestantes etc., ao representarem um diagnóstico populacional, delineando as necessidades de serviços de saúde, a exemplo do número de casos graves, e assim inferir sobre os custos das intervenções, ao identificar tendências passadas e prospectivas, como nos estudos de tendências no tempo, para assim delinear respostas adequadas à oferta dos serviços de saúde. Medidas de morbidade e mortalidade são apresentadas em estimativas específicas para cada um dos grupos ou categorias de variáveis do tipo, pessoa, tempo ou lugar, conforme apresentado na Parte 2 deste volume.

■ Análises gráficas

Na análise epidemiológica descritiva, os dispositivos gráficos contribuem para a demonstração de evidências sobre relações entre quantidades, volumes, formas, que também podem ajudar na compreensão dos fenômenos relativos ao processo saúde/doença. Às vezes esta é a maneira de mais rápida e fácil compreensão. Além das medidas de morbimortalidade, podem representar medidas de associação, e também a forma da relação entre dimensões. Para responder à pergunta de investigação, é sempre útil, e algumas vezes imprescindível, a verificação preliminar da distribuição das variáveis de interesse. Isso pode ser

realizado por meio de apresentação gráfica, como curvas, histogramas, gráficos em torta, ou também mediante a distribuição espacial na forma de mapas. Tais formatos de apresentação vêm se tornando populares com a facilidade de uso de pacotes estatísticos que disponibilizam ferramentas amigáveis para a elaboração de gráficos ou mapas de modo simples e prático.

Na construção de gráficos, Cleveland (1985) aponta alguns princípios que devem ser considerados:

- Visualização – os dados mais importantes devem estar destacados dos menos importantes, empregando-se adequados contrastes de cores ou formas;
- Clareza na compreensão – legendas devem ser compreensíveis e informativas;
- Adequada escolha das escalas – devem incluir a totalidade da faixa de distribuição dos dados.

Parcimônia e simplicidade são fundamentais para alcançar esses princípios, devendo-se evitar número excessivo de informações, legendas, barras de referência etc. Assim, disposições comuns, como os conhecidos gráficos em curvas ou em barras (histogramas), podem ser empregadas tanto para apresentações de distribuições de morbimortalidade ou exposições isoladas, como também para comparações em relação a uma variável.

Gráficos de curvas, como o da Figura 21.1, são muito utilizados. Nela, mostram-se coeficientes padronizados, por sexo, da mortalidade por doenças respiratórias selecionadas na população de 60 anos e mais de idade, segundo grupos de idade (Campagna *et al.*, 2009). Constatou-se que o coeficiente de mortalidade médio anual pelas causas selecionadas vem aumentando ao longo da série histórica, de 1992 a 2005, para todas as faixas de idade estudadas, exceto para os indivíduos de 60 a 69 anos de idade. Nesse grupo, o coeficiente médio de mortalidade anual, padronizado por sexo, estabilizou-se em torno de 11 óbitos para cada 10 mil idosos. As estimativas para a faixa etária de 70 a 79 anos, no mesmo período, apresentaram leve tendência de incremento, partindo de 32,97 óbitos por 10 mil

em 1992 para 37,41 óbitos por 10 mil no ano de 2005. Entre os de 80 anos e mais de idade foi possível constatar maior variação da mortalidade pelas causas selecionadas, *i.e.*, após o ano 2000, houve marcada tendência de crescimento desse indicador, atingindo 132,53 óbitos por mil pessoas em 2005.

Um outro tipo de histograma pode ser visto na Figura 21.2, que apresenta resultados de um estudo sobre ocorrência de casos de sarampo no município de São Paulo. Veja que, neste exemplo, são três as variáveis que estão sendo comparadas: percentual de casos de sarampo (1), por faixa de idade (2), e o ano da ocorrência (3) categorizada em blocos de 5 anos. Com essa disposição pode-se comparar a distribuição etária dos casos de sarampo nas décadas de 1970 a 90, observando-se uma modificação da participação relativa dos diferentes grupos etários. Especificamente, observa-se um contínuo decréscimo proporcional dos casos entre as crianças de 1 a 4 anos e simultânea elevação de casos entre os menores de 1 ano e os maiores de 15 anos.

Para um dado conjunto de medidas de proporções que totalizam 100%, pode-se empregar o gráfico de torta, que permite uma compreensão rápida das relações entre quantidades com essas características. Esses gráficos também podem ser dispostos lado a lado para melhor compreensão das diferenças e semelhanças entre grupos em comparação. Em um estudo sobre perfil dos idosos, vítimas de acidentes de trânsito, atendidos em um hospital governamental de Ribeirão Preto, São Paulo, os dados de indivíduos com mais de 60 anos de idade, de ambos os sexos, que sofreram algum tipo de acidente de trânsito, foram coletados pelo Serviço de Vigilância Epidemiológica do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, de janeiro a dezembro de 1998 (Silveira *et al.*, 2002). Entre 112 vítimas, atendidas na Unidade de Emergência desse hospital, a idade variou de 60 a 85 anos. Para avaliar a condição de alta após o atendimento hospitalar, os autores apresentaram esses resultados em gráficos de tortas. Observa-se que a maioria (50,9%) apresentou algum tipo de seqüela no momento da alta,

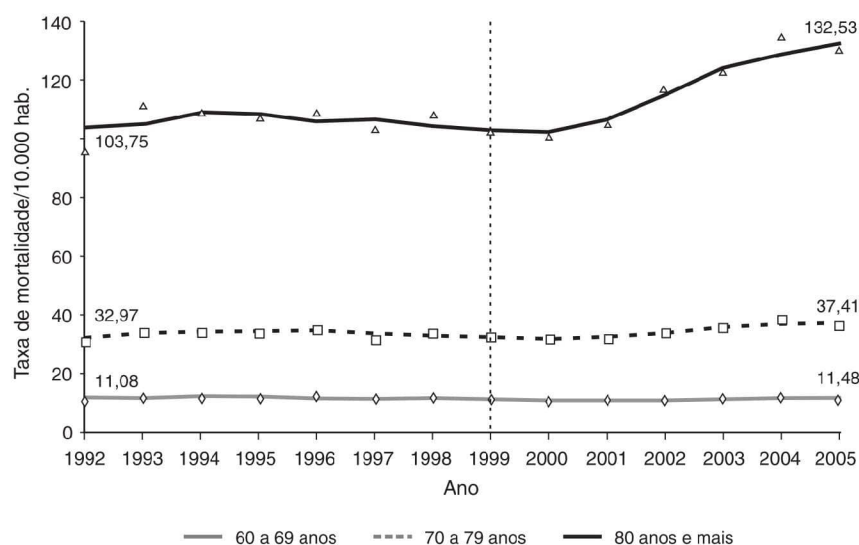


Figura 21.1 Coeficientes de mortalidade, de doenças respiratórias selecionadas, por 10 mil habitantes na população de 60 anos e mais, padronizados por sexo, segundo grupos de idade. Brasil, 1992-2005. Doenças respiratórias selecionadas incluíram: pneumonias, *influenza*, bronquites e obstrução das vias respiratórias. Para o período de 1992 a 1999, as estimativas da população idosa foram calculadas pelo método de interpolação populacional (Arriaga *et al.*, 1994). As taxas padronizadas foram calculadas pelo método direto adotando como padrão a população brasileira do Censo em 2000 (IBGE). Fonte: Campagna AS, Dourado I, Duarte EC, Daufenbach LZ. Mortalidade por causas relacionadas à *influenza* em idosos no Brasil, 1992-2005. *Revista Epidemiologia e Saúde* 18(3):209-218, 2009.

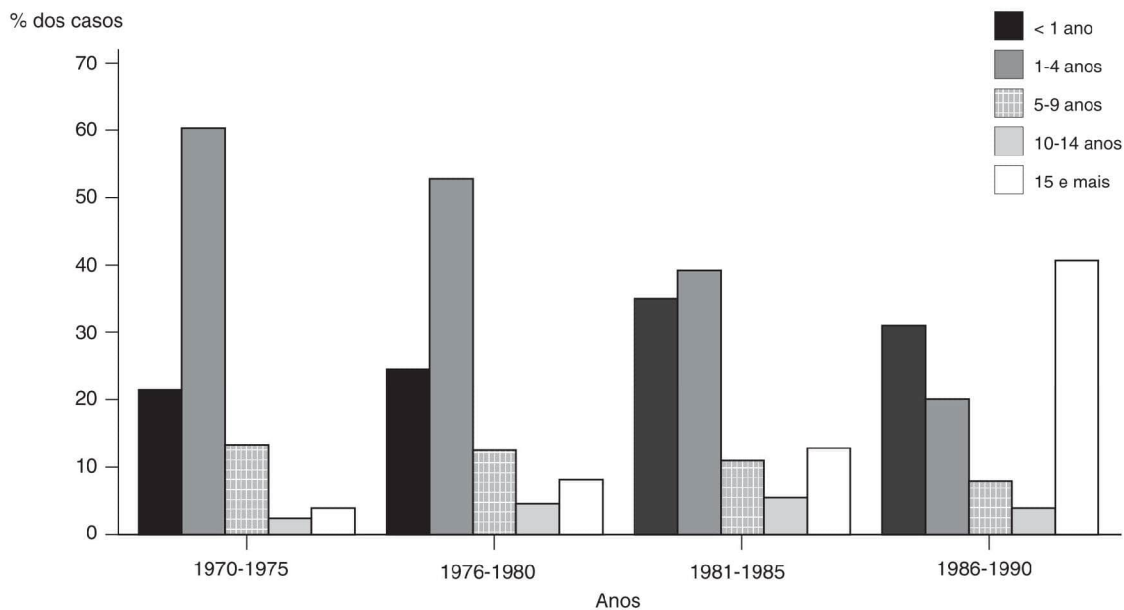


Figura 21.2 Distribuição percentual dos casos de sarampo por faixa etária no município de São Paulo, 1970-1990. Fonte: Museu Emílio Ribas (1950-1980); Centro de Informações de Saúde/Centro de Vigilância Sanitária Alexandre Vranjac (1981-1990).

enquanto 10,7% evoluiu para óbito (Figura 21.3A). Em relação à condição do acidente, observou-se que os pedestres são os mais afetados (41%), seguido pelos acidentes em que o idoso estava no veículo a motor (32%) (Figura 21.3B).

Quando a variável é contínua, conforme visto anteriormente, pode-se utilizar os valores médios de cada uma das categorias da correspondente outra variável categórica, comparando-as com uma de referência. Todavia, é possível que o investigador tenha interesse de apresentar outros detalhes descritivos, que algumas vezes trazem outras informações importantes. Por exemplo, além da média aritmética, que é uma medida pontual de uma distribuição, o investigador pode querer apresentar como os valores encontrados na população se comportam em torno da média, se há assimetria, se existem valores aberrantes, sejam eles muito elevados ou muito baixos etc. Existe um tipo

de gráfico chamado *box plot* (Figura 21.4), porque a sua forma se assemelha a uma caixa, que pode ser facilmente elaborado com rotinas de muitos pacotes gráficos ou estatísticos, de uso amigável. O *box plot* apresenta, além da média, a mediana (valor que delimita 50% da distribuição da variável) e a dispersão dos valores em torno da média, tal qual medida pelos quartis, valores que delimitam cada uma das quartas partes da distribuição, 25%, 50% e 75%. No exemplo, note que a média se encontra acima da mediana, denotando uma assimetria da distribuição em torno dos valores mais elevados.

Estes gráficos são usados para descrições de variáveis contínuas, quando se quer mostrar mais informações descritivas do que simplesmente a média e o desvio padrão. O exemplo na Figura 21.5 ilustra como *box plots* podem ser utilizados para estudos de avaliação de tendência histórica do impacto de um

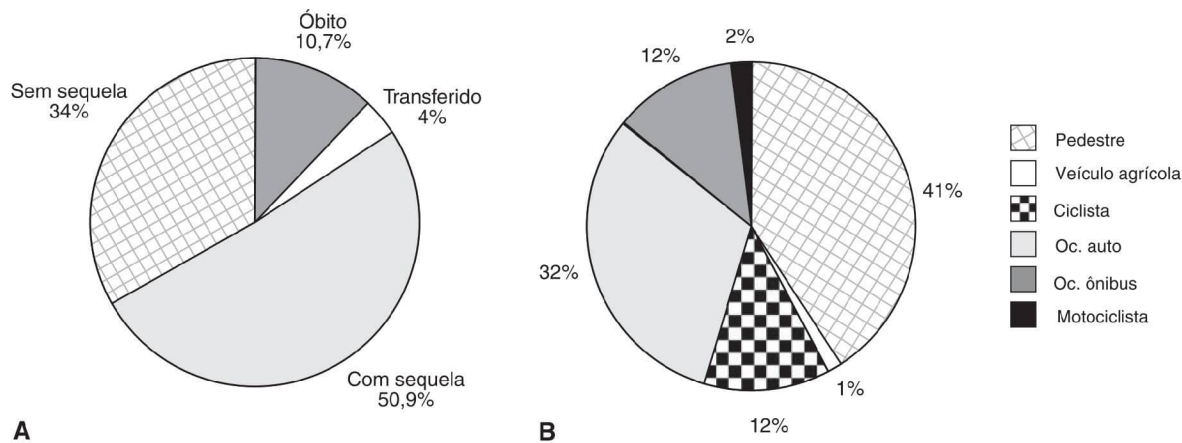


Figura 21.3 A. Distribuição dos idosos vítimas de acidentes segundo condição de alta. Ribeirão Preto-SP, 1998. **B.** Distribuição dos idosos acidentados segundo a classificação da vítima. Ribeirão Preto-SP, 1998. Fonte: Silveira R, Rodrigues RAP, Costa Júnior ML. Idosos que foram vítimas de acidentes de trânsito no município de Ribeirão Preto-SP, em 1998. *Rev Latino-Americana de Enfermagem* 10(6):765-771, 2002.

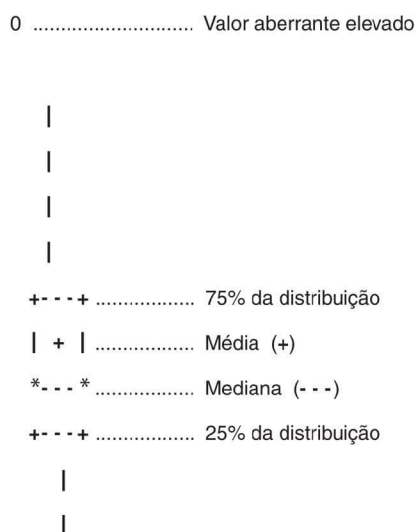
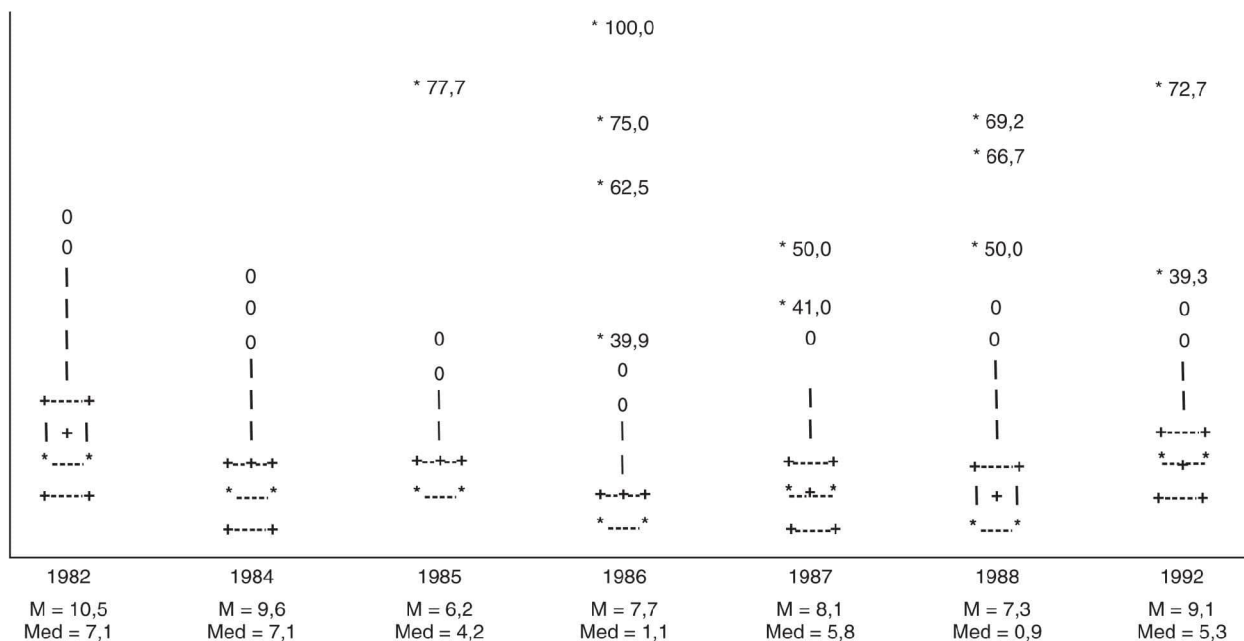


Figura 21.4 Distribuição esquemática de um gráfico do tipo box plot.

programa de controle da esquistossomose (Santana *et al.*, 1996). Esse programa era desenvolvido por órgãos do Estado da Bahia e consistia em tratamento para a infecção e atividades educativas, com análise parasitológica das fezes a cada ano, com a meta de se alcançar uma prevalência de infecção de 5%. Analisaram-se dados históricos de prevalência em 115 vilarejos e aglomerados da região da Bacia do Paraguaçu na Bahia, entre 1982 e

1992. As distribuições das prevalências de infecção por esquistossomose para cada ano eram bastante assimétricas, com muitos valores atípicos ou aberrantes (chamados *outliers*) para a esquerda, ou seja, para estimativas muito elevadas. Com isso as distribuições não eram normais, e as médias pouco informavam sobre a evolução da infecção e do efeito da prevenção. No entanto, os *box plots* permitiram a compreensão de que em alguns locais as medidas não estavam sendo efetivas, com prevalências que persistiam ou aumentavam apesar das medidas, que eram os valores atípicos elevados. Estes, nesse caso, não representavam problemas que limitavam a análise ou a interpretação, como são comumente tratados, mas eram evidências de desarranjos nas ações de saúde que reduziam o impacto, merecendo ajustes nas estratégias adotadas.

O uso de **mapas** de base geográfica, onde distribuições de medidas epidemiológicas são dispostas de acordo com uma base espacial, seja de natureza geopolítica, como os estados de uma federação, ou as regiões administrativas de um estado ou de uma cidade, também podem ser empregados para a descrição e comparação de dados. Nas figuras abaixo, apresentam-se resultados de um estudo de corte transversal, realizado em 1999 no município de Olinda-PE, com a finalidade de estimar a prevalência de filariose e verificar sua distribuição nos diferentes estratos urbanos segundo o indicador socioambiental de risco de transmissão (Braga *et al.*, 2001). A rede de drenagem em Olinda é constituída pelas bacias dos rios Beberibe e Fragoso, cujos pequenos afluentes foram retificados em muitas localidades que recebem o escoamento da drenagem urbana. No primeiro mapa (Figura 21.6) observa-se a distribuição de casos de filariose por setor censitário. No segundo mapa (Figura 21.7),



M = média aritmética; Med = mediana.

Localidades com: 1985 – Desterro 77,%; 1986 – Altamira 100%, Engenho da Praia 75%, Engenho 62,5% e Cachoeira e 33,9%; 1987 – Engenho da Praia 50%, Campinas B 41%.

Valores extremos: 1988 – Pedra Funda 69,2%, Ponta do Paraguaçu 66,7%, Rosário 50%; 1992 – Engenho Novo 72,7%, Pedra Funda 39,3%.

Figura 21.5 Box plots das prevalências de infecção esquistossomótica em localidades do município de Cachoeira-Ba, para cada um dos anos 1982-1992. Fonte: Santana VS, Teixeira MGLC, Santos CP. Avaliação das Ações de Controle da Infecção Esquistossomótica nas localidades de Cachoeira-Ba, Bacia do Paraguaçu. *Revista da Sociedade de Medicina Tropical*, 29(2):185-195, 1996.

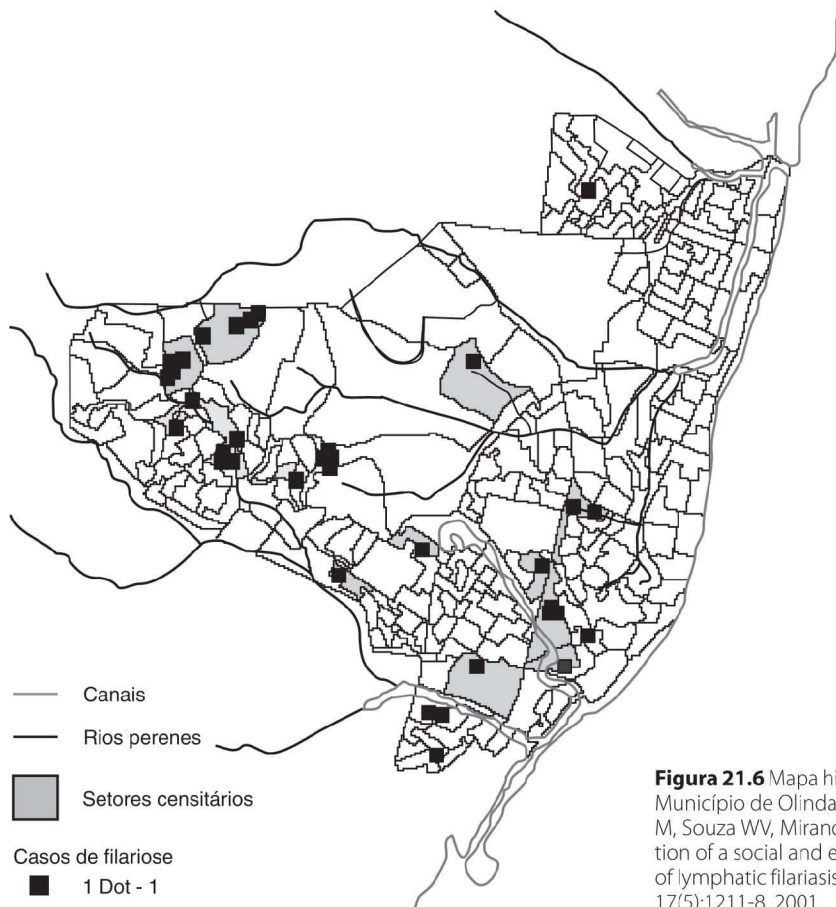


Figura 21.6 Mapa hidrográfico e casos de filariose por setor censitário. Município de Olinda, 1999. Fonte: Braga C, Ximenes RA, Albuquerque M, Souza WV, Miranda J, Brayner F, Alves L, Silva L, Dourado I. Evaluation of a social and environmental indicator used in the identification of lymphatic filariasis transmission in urban centers. *Cad Saúde Pública* 17(5):1211-8, 2001.

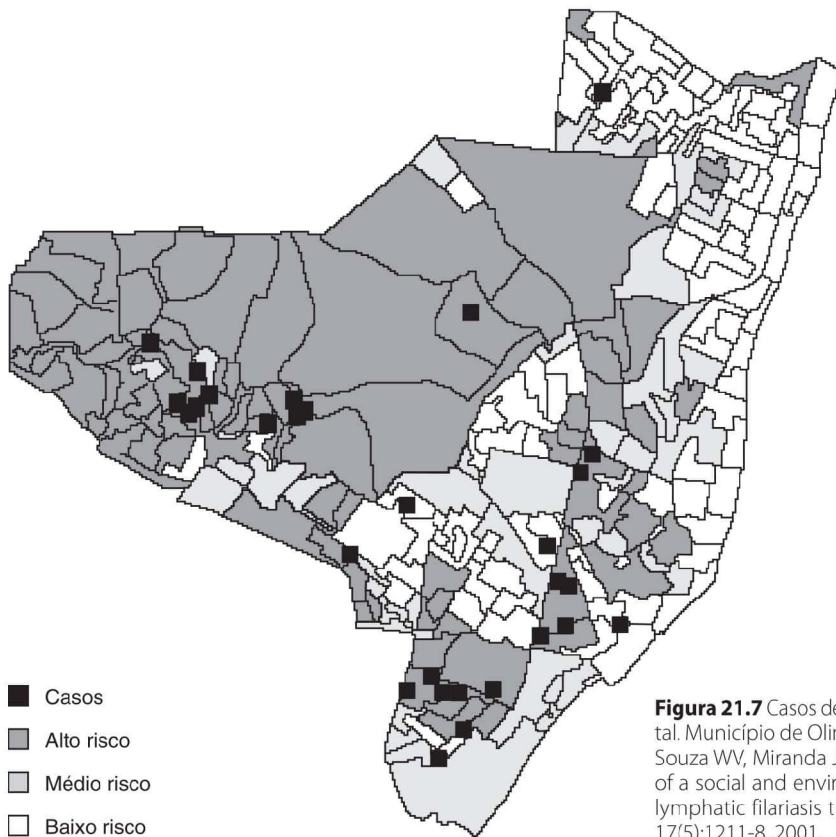


Figura 21.7 Casos de microfilaremia por estratos de risco socioambiental. Município de Olinda. Fonte: Braga C, Ximenes RA, Albuquerque M, Souza WV, Miranda J, Brayner F, Alves L, Silva L, Dourado I. Evaluation of a social and environmental indicator used in the identification of lymphatic filariasis transmission in urban centers. *Cad Saúde Pública* 17(5):1211-8, 2001.

observa-se que, dos 42 casos de microfilaremia registrados no estudo, 36 (85,7%) foram identificados nos dois estratos de maior risco de transmissão. Não se observaram casos de microfilaremia em 386 pessoas residentes no primeiro estrato, ou de menor risco de transmissão, que realizaram o exame da gota espessa.

Apesar de importantes para alguns tipos de dados, os gráficos e mapas não devem ser utilizados exageradamente, dado que são limitados na quantidade de informação que podem adequadamente representar. Por condensarem maior volume de informação, dispositivos analíticos chamados de tabelas mostram-se mais eficientes e adequados aos objetivos da pesquisa epidemiológica; por isso, a análise de dados categoriais em sua expressão tabular constitui objeto da seção seguinte.

■ Análise tabular

Os exemplos apresentados a seguir mostram a importância da análise tabular de dados categoriais em estudos descritivos, pois permitem demonstrar relações quantitativas entre variáveis, comumente pela estimativa de medidas de frequências absolutas e relativas de uma variável de interesse, em relação às categorias de outras variáveis, chamadas de descritoras. Essa análise é, particularmente, importante em Epidemiologia por permitir a visualização das relações entre as variáveis, suas dimensões e a checagem dos resultados para as medidas pontuais, as proporções. É um método, portanto, transparente e de fácil compreensão. Quando as variáveis são contínuas, a medida pontual é a média, comumente a média aritmética, e a sua respectiva medida de dispersão, o desvio padrão. Nesse caso, é preciso considerar a forma da distribuição da variável, se é normal, para que assim a média seja um parâmetro interpretável. Em situações de grande assimetria pode-se empregar a mediana, ponto de corte que corta a distribuição em 50%. Notar que frequências relativas como proporções e médias constituem medidas pontuais de um conjunto de valores adequadas para descrever populações, o que as torna medidas muito úteis na análise epidemiológica.

Variáveis descritoras representam as que serão empregadas para caracterizar a distribuição dos principais desfechos de interesse do estudo. Resultados da distribuição desses desfechos em relação às descritoras podem indicar o papel que têm essas variáveis descritoras como possíveis fatores de risco para o desfecho. Evidências de associação entre um fator de risco potencial e uma variável de resposta nos estudos exploratórios fortalecem a plausibilidade de existirem relações causais que podem ser analisadas com o teste de hipóteses causais, com a realização de estudos mais específicos, confirmatórios ou inferenciais, para verificar se o fator associado, de acordo com os resultados do estudo descritivo, se confirma como um fator de risco. Conforme assinalado no Capítulo 20, fatores de risco são fatores associados, independentemente das variáveis de confundimento, considerando-se as variáveis modificadoras de efeito, e que atendem, dentre outros, os aspectos apontados por Hill (1965) para evidenciar relações causais em Epidemiologia.

Um exemplo interessante de análise descritiva é mostrado na Tabela 21.1, que apresenta o número de casos de hanseníase notificados no Amazonas e em Manaus, em pessoas abaixo de 15 anos, e o coeficiente de detecção anual por 10.000 habitantes. Os dados provêm do Sistema Nacional dos Agravos de Notificação, SINAN, Ministério da Saúde, e estão disponíveis publicamente. Este sistema também divulga os dados populacionais, permitindo facilmente a produção de estimativas epidemiológicas (Imbiriba *et al.*, 2008). Os resultados evidenciam que a morbidade por hanseníase está se reduzindo ao longo do período

de tempo analisado, tanto no estado quanto na capital, informação importante para a definição de estratégias de intervenção sobre esta enfermidade, e também para prever tendências futuras de necessidades específicas de saúde. Notem que, nessa tabela, não são mostrados testes estatísticos ou estratégias de inferência estatística. A razão é que os dados são censitários, para toda a população, não sendo limitados a amostras, e não se está testando hipóteses causais. Neste último caso, poder-se-ia utilizar testes estatísticos considerando que mesmo um dado censitário compõe uma parte do universo do estudo. Por tudo isso, fica fácil interpretar os resultados e possivelmente utilizá-los na tomada de decisão dos serviços, pois a pergunta da investigação operacional estava claramente delineada.

Estudos descritivos também podem ser realizados com dados populacionais primários. A Tabela 21.2 mostra os resultados de um estudo que teve como objetivo estimar a magnitude da não realização do exame clínico das mamas e da mamografia, que indica quanto uma dada população não está aderindo a recomendações de programas de prevenção. Trata-se, portanto, de informação importante para gestores que precisam realizar ajustes nos modos como se estão oferecendo tais serviços para a população. Como se trata de um estudo exploratório, cujo objetivo é descritivo, os autores analisam as duas variáveis de desfecho, não realização do exame de mamas e não realização de mamografia, segundo seus descritores, fatores socioeconômicos, demográficos e de comportamentos relacionados com a saúde. A população do estudo foi composta por mulheres com 40 anos ou mais, residentes na cidade de Campinas, São Paulo, Brasil (Amorim *et al.*, 2008). O estudo foi do tipo transversal, de base populacional, conduzido em uma amostra de 290 mulheres. Entre elas, 38,2% não realizaram o exame clínico das mamas no ano prévio à entrevista e 50,8% não haviam realizado mamografia nos 2 anos anteriores. Observou-se que a não realização do exame clínico das mamas e a não realização da mamografia foram mais prevalentes em mulheres com 70 anos ou mais de idade, ou que tinham cor negra/parda. As diferenças nas prevalências de acordo com idade ou raça/cor foram estatisticamente significantes, dado que os valores de *p* foram menores do que o alfa definido para o estudo, de 0,05. Portanto, as diferenças entre as categorias das variáveis foram maiores do que as esperadas apenas por erro randômico, ou seja, devido ao uso de amostra, podendo ser diferente caso se utilizasse a população inteira.

■ **Tabela 21.1** Casos de hanseníase em menores de 15 anos e coeficiente de detecção ($\times 10.000$) no Estado do Amazonas e em Manaus, 1998-2005

Ano	Casos no Amazonas	Coefficiente de detecção	Casos em Manaus	Coefficiente de detecção
1998	200	2,00	90	2,14
1999	183	1,78	81	1,88
2000	142	1,30	58	1,24
2001	140	1,24	52	1,07
2002	163	1,42	65	1,31
2003	133	1,13	61	1,19
2004	114	0,95	38	0,73
2005	83	0,66	29	0,53
Total	1.075		474	

Fonte: Sistema Nacional de Agravos de Notificação (SINAN), Banco de Dados Estadual. Imbiriba *et al.* Perfil epidemiológico da hanseníase em menores de 15 anos de idade, Manaus (AM) 1998-2005. *Revista de Saúde Pública* 42(6):1021-6, 2008.

■ **Tabela 21.2** Prevalência (%) de não realização da mamografia nos últimos dois anos e do exame clínico das mamas no último ano, segundo variáveis socioeconômicas e demográficas em mulheres de 40 anos ou mais. Campinas, São Paulo, Brasil, 2001-2002

Variáveis	Exame clínico			Mamografia	
	N	Prevalência %	Valor P	Prevalência %	Valor P
Idade em anos			0,0474		0,0564
40-59	82	34,1		48,1	
60-69	107	38,3		46,5	
70 ou mais	101	56,5		67,7	
Total	290	38,2		50,8	
Cor da pele autorreferida			0,0064		0,0039
Branca	236	32,3		45,8	
Preta/parda	52	58,6		71,7	

Fonte: Adaptada de Amorim VMSL, Barros MBA, Chester Galvão LC, Carandina L, Goldbaum M. Fatores associados a não realização da mamografia e do exame clínico das mamas: um estudo de base populacional em Campinas, São Paulo, Brasil. *Cadernos de Saúde Pública* 24(11): 2623-2632, 2008.

Valor de **p** ou probabilidade de significância é um valor usado para expressar a conclusão final de um teste de hipóteses. Na análise tabular em Epidemiologia, comumente trata-se de comparação para verificar diferenças entre dois valores, por exemplo, duas medidas de prevalência. Compara-se o valor obtido no teste realizado com o valor obtido a partir da distribuição teórica, específica para o teste, para um valor prefixado do nível de significância. De um modo geral, na área de saúde, considera-se que o valor **p** menor ou igual a 0,05 indica que há diferenças estatisticamente significantes entre os grupos comparados.

Algumas vezes, a necessidade de informação se origina dentro dos próprios serviços, como uma outra forma de conhecer aspectos da utilização dos serviços, da clientela etc., para assim também se ajustar o atendimento, melhorando a satisfação da população atendida, por exemplo. Isso ocorreu no estudo conduzido com os clientes de um serviço para portadores de HIV/AIDS, no qual o propósito era conhecer como o uso de certos medicamentos alteravam ou não a ocorrência de problemas de saúde bucal, verificando a necessidade de ampliar o atendimento especializado (Miziara *et al.*, 2004). No exemplo seguinte, comparam-se lesões orais em pacientes com HIV/AIDS, com o número de células T CD4⁺ e a carga viral. Os dados foram coletados de 124 pacientes atendidos em um hospital, sendo incluídos apenas os que não estavam em uso de nenhum tipo de terapia antirretroviral. Após a primeira consulta, os pacientes foram divididos em dois grupos, um que apresentava candidíase oral e/ou leucoplasia pilosa e outro sem lesões orais. Todos foram submetidos a testes laboratoriais para contagem

■ **Tabela 21.3** Comparação entre contagem média de células CD4 e nível médio de carga viral (CV) entre aqueles com e sem lesão oral na linha de base

Variáveis	Lesão oral		Valor P
	Sím	Não	
CD4 (média)	164,9 céls./mm ³	542,6 céls./mm ³	0,001
CV (média)	84.745 cópias/ml	2.504 cópias/ml	0,001

Fonte: Miziara ID, Lima AS, Cortina RAC. Candidíase oral e leucoplasia pilosa como marcadores de progressão da infecção pelo HIV em pacientes brasileiros. *Rev Bras Otorrinolaringol* 70(3):310-314, 2004.

de células CD4⁺ e o nível de carga viral do HIV, indicadores empregados para o monitoramento clínico de pacientes com HIV/AIDS. Os resultados mostrados na Tabela 21.3 revelam que o grupo com lesões orais apresentou contagem de células CD4⁺ diminuída e carga viral elevada em comparação com o grupo não afetado, diferenças estatisticamente significantes.

Um outro ponto a destacar é que estudos descritivos, mesmo não testando hipóteses causais e com limitada capacidade conclusiva, podem ser usados para investigar exploratoriamente se existem associações entre alguns fatores e certos desfechos, identificando pistas para possíveis associações causais. Essas associações poderão ser então testadas como hipóteses causais em estudos posteriores, com procedimentos analíticos adequados, incorporando covariáveis modificadoras de efeito ou com papel potencial de confundimento. Notem que não é ainda a identificação de fatores causais, mas sim de possíveis fatores de risco, ou fatores de risco potenciais. Como já mencionado, fatores de risco requerem a consideração das variáveis confundidoras e modificadoras de efeito, levando em conta elementos teóricos consistentes, como pontos a serem considerados para identificação, dentre outros, dos critérios de causalidade de Hill (1965). Todavia, é muito comum encontrar estudos de fatores associados, de natureza exploratória, cujas conclusões remetem equivocadamente a fatores de risco ou associações causais. Análises de estudos para teste de hipóteses serão vistas mais adiante.

Até agora, os exemplos apresentados não mostram medidas de associação, que podem ser empregadas também em análises exploratórias, especialmente para medidas brutas, *i.e.*, não ajustadas para outras variáveis. Considerando-se ainda a análise tabular, pode-se verificar se existem fatores associados dentre um conjunto de variáveis descritoras, em relação a um ou mais desfechos, as variáveis de resposta. Nesse caso, cada variável descritora é um possível ou potencial preditor do desfecho. Portanto, para cada uma das variáveis descritoras, será necessário definir uma categoria “não exposta”, que servirá de referente, ou valor de referência para comparação com as demais categorias. Por exemplo, se a variável é sexo, será necessária a escolha do grupo, masculino ou feminino, que servirá de categoria de comparação.

Para ilustrar, apresentamos dados de um estudo do tipo caso-controle no qual se avaliaram fatores de risco para hospitalização por doença respiratória aguda de crianças até 1 ano de idade (Macedo *et al.*, 2007). A coleta de dados foi realizada em quatro hospitais da cidade de Pelotas, RS, no período de 1.º de

agosto de 1997 a 31 de julho de 1998. O grupo de casos incluiu crianças hospitalizadas (UTI pediátricas e enfermarias), de zero a 1 ano de idade, com diagnóstico de doença respiratória aguda na alta hospitalar. Os controles foram crianças selecionadas por amostragem aleatória (por sorteio), da comunidade onde os casos moravam, com até 1 ano de idade e que não haviam sido hospitalizadas por doença respiratória aguda até aquele momento. Informações sobre condições sociais, de nascimento e história médica pregressa da criança (antecedentes de sintomas respiratórios, hospitalizações anteriores e outros) foram obtidas por meio de questionário respondido pela mãe ou responsável pela criança caso e controle.

Na "análise bivariada", terminologia usada por alguns autores para determinar comparações entre duas variáveis, sendo uma única variável de exposição (p. ex., escolaridade materna) e um único desfecho (nesse exemplo, doença respiratória aguda), observou-se que a doença respiratória aguda era mais comum no sexo masculino, na faixa etária menor de 6 meses, entre os filhos de mães com baixa escolaridade e fumantes. Essas diferenças correspondem a associações entre variáveis e foram estimadas por uma medida de associação conhecida como razão de chances ou *odds ratio*. Aqui, as categorias de comparação ou referência são do sexo feminino, faixa etária de 6,1 a 12 meses, escolaridade materna ≥ 9 anos, e ter mãe não fumante (Macedo *et al.*, 2007). Nessas categorias de comparação, como a medida se iguala, a *odds ratio* é igual a 1, que corresponde à não diferença para razões ($a/a = 1$). Essas associações atingiram níveis de significância estatística, pois o intervalo de confiança de 95% não engloba o valor nulo (1,0) de não associação, ou não diferença. Intervalo de confiança é um conjunto de valores calculados com base nos dados do estudo. Pressupõe-se que o parâmetro de interesse se encontra nesse intervalo com um dado grau de confiança (95% é o nível de confiança mais utilizado na pesquisa em saúde).

Notem que os resultados apresentados nas Tabelas desta seção são brutos, ou seja, não foram ajustados ou padronizados para outras variáveis. Conforme já mencionado, os ajustes neutralizam o efeito de outras variáveis, permitindo maior compa-

rabilidade das medidas de associação em análise, em relação a fatores estranhos ao foco da pesquisa. Alguns estudos, mesmo sendo exploratórios e não objetivando a identificação de fatores de risco, mas apenas fatores associados, podem apresentar medidas de associação ajustadas para algumas variáveis. Com isso, é possível comparar medidas, sabendo-se que o fator para o qual o ajuste foi feito não atrapalha a comparação, a interpretação das diferenças. Deve-se notar que, nesses casos, não há um tratamento semelhante ao adotado para o ajuste para variáveis de confusão, que requer fundamentação teórica, pressupostos e procedimentos empíricos específicos para sua identificação. Na análise exploratória, o ajustamento de medidas de associação é feito independentemente desses critérios, com base apenas no conhecimento subjetivo da importância das variáveis a serem ajustadas, comumente, idade e sexo. Ou seja, o ajustamento é realizado para permitir a comparação de medidas pontuais. Reiterando as diferenças, a razão principal para a distinção entre fator associado e fator de risco é que este último resulta de estudos cujas análises levaram em conta confundimento e possíveis modificações de efeito, e fundamentos teóricos específicos para a associação causal em questão, que constituem os estudos de testes de hipóteses. Notem que, para os pesquisadores, praticamente não é factível o desenvolvimento de fundamentação teórica de várias associações causais a serem testadas em um mesmo estudo.

Na seção seguinte, apresentamos as principais técnicas analíticas de padronização de indicadores que visam ao ajustamento de medidas de associação para o efeito de fatores externos.

► Padronização de medidas

Comumente o propósito do investigador ou dos gestores de saúde é comparar medidas de morbidade ou mortalidade entre diferentes grupos. Por exemplo, deseja-se saber se a mortalidade por tuberculose pulmonar é maior em uma cidade A em relação a outra cidade B, ou maior do que a média nacional, o que seria sugestivo de pior qualidade do cuidado para esta en-

■ **Tabela 21.4** Comparação de casos de hospitalização por doença respiratória aguda e controles em relação às características socioeconômicas e maternas, em análise bivariada. Pelotas, RS, 1998

Variável	Casos N (%)	Controles N (%)	OR	Intervalo de confiança de 95%
Sexo				
Feminino	264 (42,2)	81 (57,8)	1,0	---
Masculino	361 (53,3)	71 (46,7)	1,5	1,1 - 2,2
Faixa etária (meses)				
1 a 6	413 (66,1)	80 (52,6)	1,7	1,2 - 2,5
6,1 a 12	212 (33,9)	72 (47,4)	1,0	---
Escolaridade materna (anos)				
≥ 9	68 (10,9)	23 (15,9)	1,0	---
5-8	319 (51,3)	44 (30,3)	4,0	2,6 - 6,2
1-4	197 (31,7)	51 (35,2)	9,2	5,1 - 16,6
Sem escolaridade	38 (6,1)	27 (18,6)	16,7	3,8 - 74,0
Hábito de fumar da mãe				
Não fuma	341 (54,6)	110 (32,4)	1,0	---
Até 10	235 (37,6)	229(72,4)	1,9	1,3 - 2,9
> 10	49 (7,8)	339 (25,7)	5,1	1,6 - 16,5

Adaptada de: Macedo SEC, Menezes AMB, Albernaz E, Post P, Knorst M. Fatores de risco para internação por doença respiratória aguda em crianças até um ano de idade. *Rev Saúde Pública* 41(3):351-358, 2007.