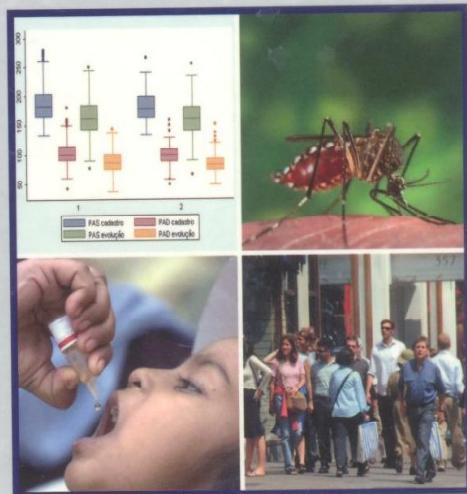


EPIDEMIOLOGIA

2^a Edição

Incluído Caderno de Exercícios



Roberto A. Medronho

Katia Vergetti Bloch | Ronir Raggio Luiz | Guilherme Loureiro Werneck

 Atheneu

Seção 2

PESQUISA EPIDEMIOLÓGICA

Fundamentos da Pesquisa Epidemiológica

Katia Vergetti Bloch
Evandro da Silva Freire Coutinho

INTRODUÇÃO

A epidemiologia tem recebido diversas definições, mas, de acordo com o Dicionário de Epidemiologia¹, trata-se do estudo da distribuição e dos determinantes dos eventos ou padrões de saúde em populações definidas, e a aplicação deste estudo para controlar problemas de saúde.

O uso de populações diferencia a epidemiologia da medicina clínica e de outras ciências biomédicas, as quais, em geral, observam um pequeno número de indivíduos, tecidos ou órgãos.

Existem duas razões para o uso de populações em epidemiologia²:

- embora o nível primário de interesse seja o indivíduo, o objetivo final da epidemiologia é melhorar o perfil de saúde das populações.
- do ponto de vista metodológico, populações são necessárias para se fazer inferências sobre a relação entre determinados fatores e a ocorrência de doenças.

OBJETIVOS DA PESQUISA EPIDEMIOLÓGICA

- Descrever freqüência, distribuição, padrão e tendência temporal de eventos ligados à saúde em populações específicas e/ou subpopulações.
- Explicar a ocorrência de doenças e distribuição de indicadores de saúde, identificando as “causas” e

os determinantes da sua distribuição, tendência e modo de transmissão nas populações.

- Predizer a freqüência de doenças e os padrões de saúde em populações específicas.
- Controlar a ocorrência de doenças e de outros eventos ou estados negativos para a saúde, através da prevenção de novos casos, cura de casos existentes, aumento da sobrevida e melhoria da saúde.

Dos objetivos anteriormente descritos, depreende-se que a epidemiologia tem o seu foco ligado tanto à compreensão ou explicação dos fenômenos relacionados à saúde em populações quanto à intervenção para modificar o padrão de saúde dessas populações.

PESQUISA EPIDEMIOLÓGICA

A pesquisa epidemiológica é empírica, baseada na coleta sistemática de informações sobre eventos ligados à saúde em uma população definida e na quantificação destes eventos. O tratamento numérico dos fatores investigados se dá através de três procedimentos relacionados: mensuração de variáveis aleatórias, estimativa de parâmetros populacionais e testes estatísticos de hipóteses.

Mensuração

É a atribuição de um valor ou qualidade a cada unidade de observação. Por exemplo, 154 mm de Hg

para a pressão arterial sistólica de um indivíduo, ou a sua classificação como “hipertenso”. Ou ainda, dois salários mínimos como renda para um domicílio ou a classificação deste domicílio no estrato de “baixa renda”. O ato de mensurar envolve, portanto, o posicionamento de unidades de observação (indivíduos, domicílios) ao longo de um contínuo (154 mm de Hg, dois salários mínimos) ou sua classificação em categorias (hipertenso/não-hipertenso; renda baixa/média/alta).

Estimação

É um processo matemático pelo qual se obtém um valor numérico a partir de uma amostra (estimativa) para representar o valor numérico desta variável numa população (parâmetro). Isto é, utilizar uma prevalência de 30% de obesos encontrada numa amostra para inferir a proporção de obesidade na população da qual se extraiu a amostra. O parâmetro costuma ser desconhecido pois, em geral, não é possível medir toda a população, mas só uma amostra desta.

Testes Estatísticos de Hipóteses

Avaliam o quanto o acaso pode ser responsável por um resultado encontrado numa amostra.

Os processos para estimação de parâmetros e testes de hipóteses, assim como as limitações desses procedimentos, são apresentados no Capítulo 24 – Associação Estatística em Epidemiologia: Análise Bivariada.

CAUSALIDADE

Vimos anteriormente que um dos objetivos da epidemiologia é explicar a ocorrência de doenças ou de outros eventos ligados à saúde. Num modelo determinístico puro, existe uma ligação única e perfeitamente previsível entre dois eventos. Na pesquisa epidemiológica, esse modelo é substituído por outro, de natureza probabilística. Através da teoria das probabilidades e de técnicas estatísticas avalia-se, de forma empírica, uma possível associação entre duas variáveis. O modelo probabilístico permite predizer quantos irão adoecer numa população, mas não quem irá adoecer. Portanto, através da pesquisa epidemiológica, são estimados parâmetros causais para populações, e não parâmetros determinísticos para indivíduos.

A incerteza advinda do modelo probabilístico fez com que os epidemiologistas passassem a utilizar o termo fator de risco para indicar variáveis que modificam a

probabilidade de um certo evento ocorrer. Para um fator ser denominado de fator de risco é preciso que:

- a doença varie de acordo com o fator, ou seja, esteja estatisticamente associada a este, de modo que a freqüência da doença seja diferente para distintos níveis do fator. Por exemplo, a incidência de hipertensão arterial varie segundo diferentes níveis de ingestão de sal;
- o fator de risco (ou uma mudança relevante deste) preceda a ocorrência da doença. Para que o tabagismo seja considerado fator de risco para enfisema, o hábito de fumar deve ter-se iniciado antes da doença;
- a associação observada não seja totalmente atribuída a qualquer tipo de erro, incluindo o erro amostral (erro aleatório), ao envolvimento de outros fatores de risco (variável de confusão) ou a problemas no desenho de estudo ou na análise (viés).

Erro aleatório será objeto do capítulo Inferência Estatística, enquanto variáveis de confusão e vieses serão tratados no Capítulo 15 – Validade em Estudos Epidemiológicos.

Além dos fatores de risco, a pesquisa epidemiológica interessa-se também pelos fatores de prognóstico, isto é, aqueles capazes de influenciar o curso de uma doença.

MÉTODO CIENTÍFICO

O método científico, no qual a epidemiologia se insere, é um processo pelo qual se busca conectar observações e teorias. Neste processo, hipóteses conceituais, mais amplas, são reescritas sob a forma de hipóteses operacionais, passíveis de serem mensuradas. A teoria que gerou a hipótese conceitual é então confrontada com os dados obtidos na investigação.

Tomemos como exemplo a controversa teoria de Barker³:

Teoria – O ambiente nutricional, hormonal e metabólico propiciado pela mãe durante a gestação, “programa” de modo permanente a estrutura e a fisiologia do bebê.

Hipótese conceitual – A subnutrição intra-uterina é um fator de risco para doença cardiovascular.

Hipótese operacional – Crianças com baixo peso ao nascer (menos de 2.500 g) têm maior probabilidade de sofrer de doença coronariana na vida adulta, em comparação com crianças nascidas com peso normal.

Com base nos dados coletados e analisados, a hipótese operacional é aceita ou rejeitada e, a partir desses

resultados, são feitas inferências causais. A teoria pode ser reformulada gerando novas hipóteses.

Há uma vasta literatura sobre inferência causal, isto é, sobre o modo pelo qual se procura ligar a observação à teoria. Diferentes visões sobre esse mecanismo têm sido propostas, e algumas têm sido objeto de discussão entre os epidemiologistas. Rothman e Greenland⁴ resumem os mecanismos pelos quais se busca estabelecer a inferência causal:

- *inferência indutiva*. Parte-se do específico, de observações, para leis gerais da natureza. Essas observações, chamadas de evidências, levam a generalizações que vão além dessas observações. Nesse processo, observam-se fenômenos, identifica-se uma relação constante entre eles e, finalmente, generaliza-se essa relação para fenômenos que podem ainda não ter sido observados. Por exemplo, a observação de que pessoas que tinham um consumo de leite e carne acima de um certo nível não adoeciam de pelagra levou à conclusão de que esta doença era causada por uma deficiência de substâncias presentes nesses alimentos.

No século XVIII, David Hume chamou a atenção para o fato de que inúmeras repetições de uma sequência particular de eventos observados não eram garantia de uma relação causal entre esses fenômenos, uma vez que a possibilidade de coincidência não pode ser descartada. Essa argumentação de Hume ficou conhecida como “Problema de Hume”.

A falácia da inferência com base na coincidência faz parte de uma falácia mais geral conhecida como “falácia de afirmar o consequente”⁴. Por exemplo, se o uso de computador protege contra doenças infecciosas, as taxas dessas doenças devem ser menores nos países com maior número de computadores *per capita*. Como a incidência de doenças infecciosas é de fato menor nos países com maior razão computador/habitante, conclui-se que o uso desse equipamento protege contra doenças infecciosas. Assim, embora a observação seja correta, a conclusão é falsa.

- *inferência por refutação*. O filósofo Karl Popper buscou superar as críticas ao indutivismo propondo que o conhecimento científico avança não através da comprovação de hipóteses, mas de sua refutação, num processo de eliminação de hipóteses denominado “conjectura e refutação”. A estratégia de Popper consiste em construir hipóteses a partir da intuição e de conjecturas; a partir dessas hipóteses são feitas previsões usando a lógica dedutiva. A inferência dedutiva é um método que, ao contrá-

rio do raciocínio indutivo, parte do geral para o específico. Se as premissas fôrêm verdadeiras, a conclusão também o será. Se temos como premissas verdadeiras que a raiva humana tem uma letalidade de 100% e que João contraiu raiva, a conclusão é de que João morrerá.

No método de inferência proposto por Popper, também se compararam as previsões baseadas nas hipóteses com as observações. Entretanto, em vez de buscar a confirmação das hipóteses, procura-se a sua refutação com base nessas comparações. Enquanto uma hipótese não for refutada, ela permanecerá como uma boa explicação para o fenômeno em estudo. Usando o exemplo anterior, o pesquisador buscaria identificar um caso de raiva humana que não tivesse morrido para então refutar a sua hipótese, em vez de buscar a sua confirmação através de repetições.

As idéias de Popper começaram a ser discutidas pela epidemiologia em meados da década de 1970 e têm profundas implicações para a metodologia da ciência, uma vez que estimulam os cientistas a buscarem múltiplas hipóteses e testes capazes de escolher, entre as hipóteses alternativas, aquela que não pode ser refutada. Essa estratégia se contrapõe à idéia clássica do cientista que busca evidências para fundamentar uma hipótese “de estimação”, considerada propriedade sua e que, uma vez comprovada, é vista como uma lei universal.

- *inferência por consenso*. Para Thomas Khun, refutar uma teoria é uma questão de escó!^{1b-a}. Ao se deparar com uma observação que leva à refutação, o investigador pode escolher entre rejeitar a teoria em teste ou toda a infra-estrutura científica das teorias em estudo. Quando a infra-estrutura científica é rejeitada em lugar da teoria, ocorre o que Khun chama de revolução científica. A decisão de aceitar ou refutar, segundo ele, depende do consenso da comunidade científica.
- *inferência bayesiana*. De acordo com esse método de inferência, a lógica dedutiva não permite saber com 100% de certeza se as premissas sobre as quais se apóia a conclusão são ou não verdadeiras, uma vez que todas as observações são sujeitas a erros. A inferência bayesiana lida com essa questão atribuindo um grau de certeza às premissas (probabilidade a priori). Então, através da teoria probabilística, calcula-se um grau de certeza para a conclusão (probabilidade a posteriori). Portanto, as conclusões obtidas dependem da probabilidade atribuída às premissas. Ao contrário do que ocorre na lógica dedutiva, permite lidar com fenômenos científicos em que essa probabilidade não é de 100%.

Rothman e Greenland⁴ concluem que do debate filosófico emerge o desafio deixado por Hume de que “provar é impossível” para as ciências empíricas. Segundo os autores, essa crítica atinge especialmente o epidemiologista, uma vez que leva à suposição de que experimentos são a fonte definitiva do conhecimento científico. Essa suposição é equivocada porque:

1º – ciência não-experimental não implica necessariamente em descobertas imprecisas, como por exemplo: a evolução das espécies, planetas orbitando outras estrelas e os efeitos do cigarro à saúde humana;

2º – mesmo quando possíveis, experimentos (incluindo ensaios randomizados) não produzem nada que se aproxime de uma prova e podem até mesmo ser controversos, contraditórios ou irreprodutíveis.

DESENHOS DE ESTUDOS

O processo pelo qual as hipóteses conceituais são transformadas em hipóteses operacionais é intermediado pelos desenhos de estudos. Esses desenhos têm em comum a observação sistemática dos fenômenos de interesse, o uso da teoria e dos métodos estatísticos para analisar os dados e interpretar os achados, e a comparação entre grupos com o objetivo de identificar associações estatísticas entre variáveis.

Como dissemos no início desse capítulo, a epidemiologia utiliza populações e, portanto, é importante que se discrimine as populações relacionadas com a investigação.

População do estudo ou amostra – é o grupo de indivíduos sobre os quais se fazem as observações e coletam-se os dados.

Ex.: 100 crianças nascidas em 2000, no município X.

População-fonte (universo amostral) – é o conjunto de indivíduos elegíveis para constituírem a população do estudo, isto é, o conjunto de indivíduos que está sendo representado pela população do estudo.

Ex.: Todas as crianças nascidas em 2000 no município X.

População-alvo – é o grupo de indivíduos para o qual se deseja fazer as inferências estatísticas relacionadas com o objetivo do estudo.

Ex.: Crianças nascidas em 2000, no município X.

Embora a população do estudo deva ser representativa da população-alvo, isto pode não ocorrer, caracterizando um erro sistemático ou viés. Por exemplo, se

usarmos os nascimentos em 2000 registrados no cartório do município X como população-fonte para representar os nascimentos ocorridos em 2000 naquele município (população-alvo) a população do estudo não será representativa da população alvo se houver um importante sub-registro de nascimentos.

Base populacional – é o conjunto de indivíduos de onde surgem os casos da doença que farão parte da investigação.

Ex.: Para investigar casos de sarampo em menores de um ano no município X, a base populacional será constituída por todos os menores de um ano residentes naquele município. Para investigar os casos de sarampo em menores de um ano atendidos no posto de saúde X, a base populacional será constituída por todos os menores de um ano que irão ao posto de saúde caso adoecam.

Os desenhos de estudos epidemiológicos diferem entre si no modo pelo qual selecionam as unidades de observação, mensuram os fatores de risco ou de prognóstico, identificam as variáveis de desfecho e garantem a comparabilidade entre os grupos que fazem parte do estudo.

Quanto à comparabilidade, os desenhos de estudos podem ser agrupados em três amplas categorias, com base no modo pelo qual os indivíduos são alocados nos grupos a serem comparados, e no controle que o investigador tem sobre a exposição ao fator de risco ou de prognóstico:

- experimentais – os indivíduos são alocados de modo aleatório em diferentes grupos de exposição aos fatores que se julga serem de risco ou de prognóstico. Esse processo de alocação aleatória (com base no acaso) garante a todos os indivíduos a mesma probabilidade de fazer parte de qualquer um dos grupos que estão sendo comparados. Outra característica dos estudos experimentais é que o investigador controla a exposição ao fator de interesse. Por esta razão, questões éticas fazem com que tais estudos se restrinjam a fatores que tenham uma influência positiva (protetores) sobre a saúde.

Os estudos experimentais mais comuns em epidemiologia são o ensaio clínico (terapêutico ou profilático), e o estudo de intervenção em comunidades. No caso do ensaio clínico terapêutico, busca-se avaliar a efetividade e a segurança de uma intervenção terapêutica sobre pacientes (ex.: combinação de drogas para o tratamento da tuberculose). Nos ensaios profiláticos, a intervenção se dá sobre indivíduos não-doentes, com um objetivo

preventivo (ex.: aspirina para prevenir infarto do miocárdio). As intervenções em comunidades são uma extensão dos ensaios profiláticos, onde a intervenção se dá sobre a comunidade como um todo (ex.: tratamento da água que abastece uma comunidade);

- *quasi-experimentais* – há controle do fator em estudo pelo investigador, mas a alocação dos indivíduos nos grupos a serem comparados não é aleatória. Esses estudos costumam ser direcionados a uma ou mais populações em vez de a indivíduos. Um exemplo é a comparação da ocorrência de cárie dentária entre duas comunidades, sendo que em uma foi aprovada a fluorização da água e em outra não;
- *observacionais* – o investigador não controla nem a exposição nem a alocação dos indivíduos, lançando mão de uma situação dada, e observando os resultados. Como os indivíduos estão expostos ou não a uma causa potencial de doença independentemente da interferência do observador, esses estudos não apresentam problemas de natureza ética para investigar fatores de risco.

Tipos de Estudos Observacionais

Pelo menos três dimensões devem ser consideradas para a taxonomia dos estudos epidemiológicos observacionais: a estratégia de observação da população, o esquema de seleção e a unidade de observação.

- *estratégia de observação*. Dissemos anteriormente que o uso de populações distingue a epidemiologia de outras ciências biomédicas e da clínica. O modo como se observa as populações na investigação epidemiológica distingue os diferentes desenhos de estudos observacionais. Cada população apresenta uma dinâmica própria, na qual seus elementos adoecem, curam-se, morrem, expõem-se a fatores diversos. Diz-se que a estratégia de observação de uma população é seccional quando existe uma só observação dessa população, isto é, as informações coletadas referem-se a um só momento. Portanto, na observação seccional de uma população geralmente não é possível estabelecer a relação temporal entre as características presentes no grupo. É como uma fotografia, onde o momento é congelado. A figura de um homem com um cigarro próximo à boca não permite dizer se ele está se preparando para uma tragada ou se acabou de fazê-lo. Ou ainda, a imagem de duas crianças sentadas não permite concluir qual

delas o fez primeiro. Entretanto, existem algumas exceções em que se pode estabelecer uma relação temporal numa observação seccional como, por exemplo, em associações que envolvem variáveis constitucionais como sexo ou grupo sanguíneo.

A estratégia de observação é dita longitudinal quando pelo menos duas observações são realizadas, em momentos diferentes. Portanto, pode-se captar algum grau da dinâmica dessa população. É como observar duas ou mais fotos obtidas numa seqüência temporal conhecida. A população que está sendo acompanhada através de duas ou mais observações ao longo do tempo será chamada de coorte fixa, caso não se permita a entrada de novos indivíduos no grupo após o início do estudo. Se essa entrada for aceita, teremos então uma população dinâmica. Por fim, esta última será considerada estável quando o seu tamanho e a distribuição das características de interesse (sexo, idade, por exemplo) se mantiverem constantes durante o estudo.

- *esquema de seleção*. Outra característica considerada na classificação dos estudos observacionais é o esquema de seleção. Quando todos os indivíduos da base populacional da qual surgem os casos são selecionados, fala-se em esquema completo ou censo. Quando apenas uma fração dos não-casos é selecionada, fala-se em esquema incompleto. Essa característica é amplamente discutida no Capítulo 11 – Estudos Caso-controle;
- *unidade de observação/análise*. Por fim, temos a unidade de análise, que é o nível no qual as informações são coletadas e analisadas. No nível individual as informações são obtidas para os indivíduos da população do estudo. Já no nível agregado (também chamado ecológico) as informações são obtidas para um conglomerado de indivíduos (domicílio, área geográfica), e não para o indivíduo isoladamente. Por exemplo, ao coletarmos informações sobre o hábito de fumar e doença isquêmica coronariana (DIC) no nível individual, saberemos, para cada indivíduo, se ele fuma ou não e se ele tem DIC ou não. A mesma investigação, se conduzida no nível agregado, poderia levantar a proporção de fumantes e portadores de DIC em diferentes cidades, sem saber se essas duas características estão presentes ou não em um mesmo indivíduo.

A maioria dos estudos observacionais, se não todos, tem elementos descritivos e analíticos, sendo estas características consideradas como dois extremos de um contínuo.

Com base nas três dimensões descritas teremos, os seguintes desenhos de estudos observacionais:

- *estudo de coorte.* É um estudo no qual um conjunto de indivíduos sem a doença de interesse é classificado em grupos segundo o grau de exposição a um possível fator de risco ou de prognóstico, sendo então acompanhados para se comparar a ocorrência da doença em cada um desses grupos. Trata-se, portanto, de um desenho de observação longitudinal da base populacional (fixa ou dinâmica), isto é, da população de onde surgirão os casos de doenças ao longo do seguimento. Toda a base populacional predefinida é observada (esquema de seleção completo), sendo as informações obtidas para o nível individual.

Suponhamos um estudo de coorte em que se pretenda investigar a associação entre tabagismo e infarto agudo do miocárdio (IAM) em funcionários públicos universitários em atividade. A base populacional seria formada pelo conjunto desses funcionários, os quais seriam divididos, por exemplo, em dois grupos, segundo a exposição ao tabaco: fumantes e não-fumantes. Esses grupos seriam então acompanhados por um certo período, sendo contabilizados os casos de IAM ocorridos em cada grupo durante o seguimento.

- *estudo caso-controle.* É um estudo onde casos da doença de interesse são identificados e classificados segundo o grau de exposição ao fator de risco ou de prognóstico de interesse. É um estudo longitudinal, pois a informação sobre a exposição e a doença se refere a momentos diferentes. Uma amostra da base populacional é selecionada para formar o grupo-controle (esquema de seleção incompleto), o qual também é classificado segundo o grau de exposição. A informação sobre a exposição nos casos e nos controles é obtida no nível individual.

Consideremos agora que se queira investigar a relação entre o tabagismo e a ocorrência de IAM em funcionários públicos universitários, através de um desenho caso-controle. Nesse estudo a base populacional seria ainda o mesmo conjunto de funcionários públicos universitários. Todos os casos de IAM ocorridos num certo período de tempo seriam identificados, levantando-se a informação sobre a presença ou não de tabagismo nesses indivíduos. Uma amostra dos funcionários públicos, que não tiveram IAM, seria então selecionada para constituir o grupo controle, coletando-se também a história de tabagismo. A exposição ao tabaco seria comparada em ambos os grupos.

Deve-se notar que em ambos os desenhos, a base populacional era formada pelos funcionários públicos universitários, pois foi desse grupo que surgiram os casos de IAM detectados por ambos estudos. Entretanto, enquanto o estudo de coorte coletou a informação em toda base, isto é, em toda a população de onde surgiram os casos de IAM (esquema completo), o estudo caso-controle coletou informação apenas de uma parte da base populacional quando selecionou os controles (esquema incompleto).

- *Estudo seccional.* Aqui a estratégia de observação da população é seccional, isto é, as informações obtidas referem-se ao mesmo momento, e os indivíduos são selecionados da população fonte inteira (censo) ou de uma fração desta (amostra). Esse desenho faz uso de dados individuais, e, como será visto no Capítulo 10 – Estudos Seccionais, apresenta algumas limitações enquanto ferramenta para investigações causais, sendo mais adequado para a obtenção de estimativas populacionais como prevalências e médias, ou para levantar hipóteses etiológicas.

Caso se utilizasse o estudo seccional para investigar a associação entre tabagismo e IAM, a população fonte seriam os funcionários públicos universitários, dos quais se estudaria todo o grupo ou uma amostra representativa (população do estudo). As informações sobre doença isquêmica do coração e sobre tabagismo seriam coletadas em relação a um mesmo momento no tempo.

- *Estudo ecológico.* É um estudo seccional com a informação obtida e analisada no nível agregado. Nesse desenho não se conhece a distribuição conjunta da exposição e da doença no nível individual. No nosso exemplo, não se conheceria a situação de cada indivíduo com relação ao tabagismo e à presença de IAM. Ao trabalhar no nível agregado, o estudo ecológico buscara observar se populações onde a proporção de tabagistas é mais elevada apresentam maiores taxas de doença isquêmica do coração, do que populações onde essa proporção é menor.

Os termos prospectivo e retrospectivo referem-se à relação cronológica entre a mensuração da exposição e da doença e a ocorrência de ambos os fenômenos. Num estudo prospectivo, também chamado de concorrente, exposição e doença são mensuradas no curso da investigação. Assim, a mensuração se dá simultaneamente à ocorrência desses fenômenos. Já no estudo retrospectivo, também referido como histórico ou não-concorren-

te, exposição e doença são mensuradas após já terem ocorrido, através de relato ou registros.

A importância desta característica (temporalidade, timing) está na avaliação da qualidade do dado colhido, que tende a ser melhor quando feita de forma concorrente.

Os desenhos aqui apresentados não esgotam as estratégias de estudo utilizadas pela epidemiologia. Nos capítulos referentes a cada um desses delineamentos, veremos que outros elementos devem ser considerados numa taxonomia mais fina do que a que apresentamos até aqui. Assim; dentro de cada um destes desenhos existem subcategorias, e diferentes estratégias podem ser combinadas nos chamados desenhos híbridos.

A escolha de um tipo de desenho para investigar uma hipótese de exposição-doença depende da natureza da doença, do tipo de exposição, do conhecimento existente sobre essa relação e dos recursos disponíveis.

Estudos preliminares, de validade e precisão limitadas e de menor custo, podem ser utilizados em um estágio inicial do conhecimento para testar e selecionar hipóteses. Sendo detectada uma associação, esta deverá ser submetida a testes mais rigorosos, geralmente de maior custo em estudos cujo desenho proporcione maior validade.

O julgamento da existência de uma relação causal baseia-se na avaliação da totalidade dos dados disponí-

veis (resultados consistentes de estudos com diferentes populações), e não apenas da existência de uma associação estatística válida. Algumas estratégias e critérios utilizados nesse julgamento são discutidos no Capítulo 7 – Causalidade em Saúde.

CONCLUSÃO

Neste capítulo introduzimos alguns conceitos básicos da investigação epidemiológica, com ênfase nos aspectos gerais que caracterizam os diferentes delineamentos de estudo. Nos próximos capítulos os conceitos básicos aqui apresentados serão aprofundados a partir da discussão das características de cada desenho, das implicações dessas características na validade do estudo e das estratégias de análise mais utilizadas.

BIBLIOGRAFIA

1. Last JM. A dictionary of epidemiology. 3 ed. New York: Oxford University Press, 1995.
2. Kleinbaum D, Kupper LL, Morgenstern H. Epidemiologic Research: principles and quantitative methods. New York: Van Nostrand Reinhold Company. 1982.
3. Barker DJP. Mothers, babies and health in later life. 2 ed. Edinburgh, Churchill Livingstone. 1998.
4. Rothman KJ, Greenland S. Modern Epidemiology. 2 ed. Philadelphia: Lippincott-Raven, 1998.

Estudos Seccionais

Carlos Henrique Klein
Katia Vergetti Bloch

INTRODUÇÃO

O estudo seccional é aquela estratégia de estudo epidemiológico que se caracteriza pela observação direta de determinada quantidade planejada de indivíduos em uma única oportunidade. Os indivíduos são geralmente pessoas, porém também podem ser agregados ou até mesmo animais, no caso de estudos de epidemiologia veterinária. Doravante, para todos os efeitos, vamos também denominar os indivíduos, objetos de estudo, de unidades de observação.

As unidades de observação de um estudo seccional costumam ser selecionadas aleatoriamente, isto é, por algum método orientado apenas pelo acaso, entre todos os indivíduos que compõem uma população. População é o conjunto completo de indivíduos que pode ser descrito de acordo com uma característica comum a todos. Mais freqüentemente, a característica que define uma população, alvo de um estudo seccional, está relacionada com critérios geográficos, políticos e administrativos, que a limitam em termos espaciais. Exemplos desta conjugação de critérios são populações demarcadas por bairros, municípios, regiões geoeconômicas, estados ou países. Entretanto, uma população também pode ser descrita em função de outras características tais como o sexo, a faixa etária, ou a ocupação. De qualquer modo, estas outras características costumam ser complementadas por um critério de demarcação espacial. Assim, a população-alvo pode ser a dos farmacêuticos, formados antes de 2000, residentes no Estado do Rio de Janeiro.

Não é obrigatória a escolha de apenas uma parte dos indivíduos que compõem a população, para for-

mar uma amostra. Se o pesquisador possui recursos para examinar todas as unidades de observação, que constituem a população, poderá optar pela realização de um censo. Entretanto, como na maioria dos estudos seccionais a população-alvo é muito numerosa, opta-se pela seleção de tão-somente uma amostra para exame, e os motivos que justificam tal atitude não são apenas de ordem econômica, pela redução de custos. A distribuição de recursos materiais para fazer observações em uma menor quantidade de indivíduos permitirá exames individuais de melhor qualidade. Ainda assim, mesmo que as observações sejam feitas em indivíduos de uma amostra, esta costuma ser formada pelo menos por centenas ou milhares de indivíduos, na maioria dos estudos seccionais.

Portanto, outra característica fundamental da maioria dos estudos seccionais, que se utilizam de amostras, diz respeito à inferência. Neste contexto a inferência será compreendida pela sua conotação estatística. Inferência é “o conjunto de métodos que permitem formular, em termos probabilísticos, um julgamento sobre uma população a partir dos resultados observados em uma amostra extraída ao acaso dessa população”²⁶. Uma vez que a característica essencial de um estudo seccional é a observação de cada indivíduo em uma única oportunidade, a inferência é feita pela referência dos resultados a uma população definida em um local e em determinada época. Desta forma, deve ficar muito claro que os objetivos de um estudo seccional estarão sempre relacionados com indivíduos em local e época demarcados.

A estratégia de delineamento de estudo até aqui chamada de seccional também recebe outras denominações alternativas. Já há alguns anos, em toda a América Latina, vem sendo muito utilizada a expressão estudo transversal⁴, quando até os anos 1970 do século XX eram utilizados os termos inquérito, em português, e *encuesta*, em espanhol. A palavra inquérito lembra o recurso de medida mais comum utilizado nestes estudos: o questionário, com um conjunto variável de perguntas. Entretanto, o estudo pode prescindir do questionário se a única medida de interesse puder ser feita por meio de uma observação direta, com ou sem aparelhos. E, ainda, com grande freqüência, medidas diretas são acompanhadas de um questionário. Exemplificando: em um inquérito sorológico, que visava determinar freqüências relativas de cobertura de imunização em subconjuntos da população, os indivíduos participantes foram inquiridos sobre idade, local de moradia, e deles foram coletadas alíquotas de sangue para titulação de anticorpos.

Os termos seccional e transversal estão relacionados com a temporalidade, isto é, com a época da coleta de dados do estudo. Se, como quase sempre ocorre na prática, as informações individuais são coletadas num determinado prazo, o mais curto possível, decorrido entre as observações do primeiro e último indivíduo, a lógica da análise dos dados ignora este intervalo de tempo e considera que todas as observações foram feitas num mesmo instante. Portanto, se a coleta pode ser considerada transversal, por ser oblíqua ou diagonal em relação ao eixo do tempo, a análise dos dados coletados é feita como se houvesse sido feita uma secção perpendicular de observações na população-alvo, quanto ao eixo temporal. Por isto, o termo seccional é mais abrangente, pois tanto pode ser perpendicular ou instantâneo para a análise, como transversal ou oblíquo para a coleta dos dados. De todo modo, é preciso sempre lembrar que todas as informações são coletadas na mesma oportunidade, ou seja, em um único instante, de cada indivíduo. A Figura 10.1 apresenta o esquema de coleta de dados de um grupo selecionado de seis indivíduos, em que as linhas horizontais representam segmentos de tempo de suas vidas próximos à época do estudo seccional.

Grande parte da literatura existente sobre os estudos seccionais está em língua inglesa, na qual são utilizadas as expressões *cross-sectional*, especialmente na área da saúde, e *survey*, esta de sentido ainda mais extenso e a preferida na área das ciências sociais. Por sinal, os primórdios do uso desta estratégia de estudo podem ser creditados a investigadores da área de ciências sociais, especialmente da área de economia política e sociologia². Para exemplificar, podemos citar um inquérito feito

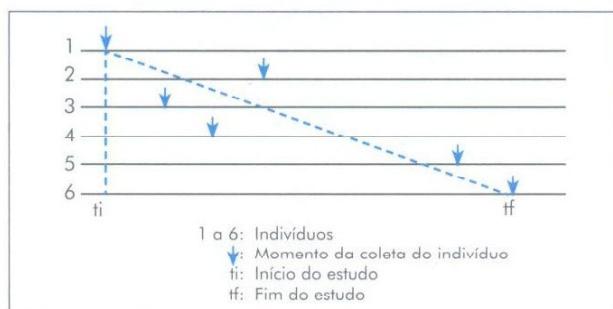


Fig. 10.1 — Esquema de coleta de dados individuais e o tempo em estudo seccional.

por Karl Marx, em 1880, que enviou, pelo correio, a cerca de 25 mil trabalhadores franceses, um questionário com perguntas sobre condições e relações existentes nos locais de trabalho.

Também é relativamente comum o uso da expressão estudo de prevalência como sinônimo de estudo seccional. Isto porque é muito freqüente que o resultado mais geral que se procura alcançar neste estudo é uma prevalência. Neste caso está subentendido que se trata de uma prevalência instantânea devido ao método de coleta de dados. A prevalência instantânea mede a proporção de indivíduos com uma determinada característica, que pode ser uma doença estabelecida, um sintoma, um sinal, uma seqüela, ou outro agravo qualquer da saúde, relacionando os afetados no numerador com todos os examinados no denominador. Entretanto, é também comum que nos estudos seccionais se obtenham várias informações de cada indivíduo examinado, com o intuito de estabelecer relações de associação entre as características investigadas. Por exemplo, um estudo seccional com adultos da Ilha do Governador, uma região administrativa do município do Rio de Janeiro, tinha por objetivo principal estimar a prevalência de hipertensão arterial, para o que foi necessário medir a pressão arterial e obter informações suplementares sobre o tratamento e controle da pressão. Além disto, neste mesmo estudo ainda foram coletadas outras informações sobre características que poderiam estar associadas com a hipertensão arterial.

USOS E LIMITAÇÕES

Estudos seccionais têm sido utilizados para investigar uma ampla gama de problemas de saúde pública, com os mais variados propósitos, desde administrativos até analíticos. De um modo geral, a aplicação mais comum de um estudo seccional está ligada à necessidade de conhecer de que maneira uma ou mais características,

tanto individuais como coletivas, distribuem-se em uma determinada população. Estas características são também chamadas de variáveis. Um estudo seccional pode ser a melhor estratégia de investigação para se conhecer estimativas de parâmetros como médias (de altura, peso, pressão sistólica), proporções (de categorias de obesidade, de grupos etários, de hipertensos) e outras razões ou índices (moradores por domicílio), assim como também as dispersões (variância, desvio-padrão) das características desejadas em uma população definida no tempo e no espaço, e também em suas subdivisões, subconjuntos ou grupamentos. Portanto, em princípio, um estudo seccional é um excelente método para descrever características de uma população, em uma determinada época. A descrição da distribuição de um agravo de saúde em uma população é uma das fontes imprescindíveis para o planejamento e a administração de ações voltadas para prevenção, tratamento e reabilitação, tanto em nível coletivo como individual.

Para exemplificar, o estudo seccional da Ilha do Governador, já citado, tinha por justificativa para seu objetivo principal obter subsídios para o planejamento de ações de prevenção e controle da hipertensão arterial²³. A escolha do local de estudo se deveu a sua proximidade a um hospital universitário, de alta complexidade, que daria suporte às ações de saúde, incluindo sua integração com unidades de diferentes níveis de atenção.

Os estudos seccionais são também muito úteis para determinar a presença de抗ígenos, anticorpos e outros marcadores biológicos no sangue coletado de indivíduos nos chamados estudos soroepidemiológicos. A presença isolada de抗ígenos no sangue, não acompanhada de sintomatologia, indica infecção sem doença. A prevalência de anticorpos específicos pode ser muito importante para avaliar a necessidade de promover campanhas ou outras estratégias de imunização, em populações estudadas com inquéritos soroepidemiológicos. Índices de concentração de colesterol sérico obtidos em estudos seccionais podem ser correlacionados com outros indicadores de risco cardiovascular.

Algumas propriedades dos instrumentos de coleta de dados utilizados em um estudo seccional podem também estender seus propósitos. Se o questionário contiver perguntas classificadas como abertas, isto é, cujas possíveis respostas não são conhecidas ou previstas antes de sua aplicação, então o estudo seccional poderá ter propósitos exploratórios, de busca de hipóteses que expliquem inter-relações de fenômenos ainda não conhecidas ou, pelo menos, elaboradas de forma consistente pelos pesquisadores. O levantamento destas informações poderá levar o pesquisador a novas percepções sobre as conexões entre eventos, dando origem ao planejamento

de novos estudos, com metodologias apropriadas para investigar as hipóteses derivadas do estudo seccional de caráter exploratório.

Os estudos seccionais algumas vezes não são considerados como estratégias de investigação analítica, isto é, que possam testar hipóteses sobre vínculos causais entre eventos. Ainda assim, sempre é possível testar a existência de associações de freqüência, ou estatísticas, entre pelo menos dois eventos classificados como doença (agravo) e exposição (suposta causa), na amostra de indivíduos examinados. A Figura 10.2 mostra como em uma amostra única, obtida de uma população, podem ser identificados pelo menos quatro grupos de indivíduos com a finalidade de analisar associações entre exposição e doença, em um estudo seccional.

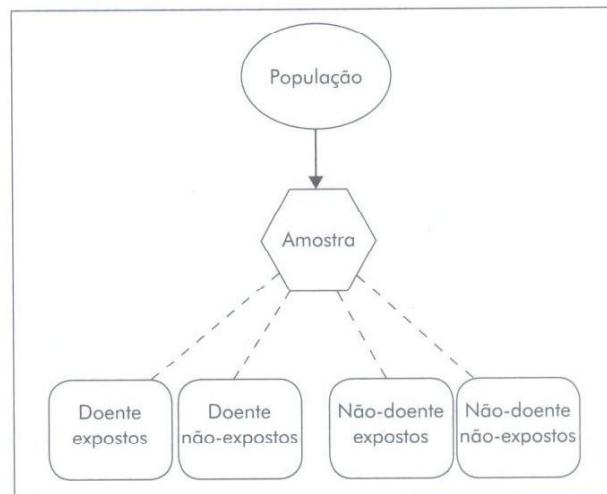


Fig. 10.2 – Grupos constituídos para estudo de associação em estudo seccional.

As quantidades de indivíduos nos subconjuntos, identificados na amostra, também podem ser representados como na Tabela 10.1.

Em uma amostra de n indivíduos, selecionados aleatoriamente da população, a prevalência da doença é dada por $(a+c)/n$. Uma das alternativas de análise da

Tabela 10.1
Freqüências de Doença e Exposição Observadas em um Estudo Seccional

Freqüências	Doentes	Não-doentes	Total
Expostos	a	b	$a+b$
Não-expostos	c	d	$c+d$
Total	$a+c$	$b+d$	n

associação entre a exposição e a doença é comparar a prevalência de doença nos expostos com a prevalência de doença nos não-expostos, ou seja, $a/(a+b)$ com $c/(c+d)$. Se a primeira, $a/(a+b)$, for significativamente diferente da segunda, $c/(c+d)$, de acordo com um teste estatístico apropriado, pode-se concluir pela alta probabilidade de existência de associação estatística entre exposição e doença. Se a prevalência de doença nos expostos for maior do que nos não-expostos, pode-se concluir que a exposição está associada de forma positiva com a doença ou, se ao contrário, a prevalência de doença nos expostos for menor do que nos não-expostos, conclui-se que a associação entre exposição e doença é negativa. Entretanto, é preciso tomar cuidado para não confundir uma associação positiva, observada em um estudo seccional, com uma associação de fator de risco para o desenvolvimento da doença. Do mesmo modo, a associação negativa, revelada em um estudo seccional, não é necessariamente representativa de uma associação de fator de proteção contra o desenvolvimento da doença.

A relação temporal entre causa e consequência é unidirecional, da primeira à última. Nos estudos seccionais, todas as observações são feitas em cada indivíduo em uma única oportunidade, sejam aquelas relacionadas com as exposições (supostas causas) ou com as consequências (doenças). Por isto, ainda que a construção do questionário procure revelar dados sobre momentos diferentes, as informações relativas a tempos passados são obtidas de forma indireta, isto porque dependem da memória e dos interesses peculiares dos indivíduos em relação aos temas de investigação. Isto torna muito difícil a padronização, ou seja, a uniformização de procedimentos de coleta, o que pode distorcer a análise das relações entre eventos estudados em subconjuntos da amostra. Como exemplo, pode-se citar a dificuldade de se obter informações recordatórias sobre consumo alimentar na tentativa de relacionar hábitos passados de ingestão de alimentos gordurosos com neoplasia de mama ou de cólon. No momento da coleta de dados em um estudo seccional, aqueles hábitos podem ter se modificado justamente devido ao diagnóstico estabelecido. Por isto, quando se supõe que o estado de exposição durante o estudo não representa a exposição à época relevante para a determinação da doença, é necessário avaliar retrospectivamente as mudanças qualitativas e quantitativas do estado da exposição.

Outro problema na investigação de uma relação hipotética entre um fenômeno definido como doença e outro especificado como exposição, que se espera estar relacionado com a causalidade da doença, é que

nem sempre é possível estabelecer com segurança a precedência temporal da exposição sobre a doença em estudos seccionais. Pode-se citar como exemplo desta dificuldade a aparente relação percebida entre o consumo de produtos dietéticos e a obesidade, uma vez que é mais provável que a obesidade é que preceda o uso dos dietéticos. Quando não é possível estabelecer com segurança a seqüência temporal dos eventos, pode ser impossível decidir se determinada exposição é causa ou consequência do aparecimento da doença.

Ainda assim, estes estudos podem ser úteis para auxiliar a esclarecer hipóteses etiológicas relacionadas com doenças crônicas, de início insidioso, quando seu início não é reconhecido com facilidade pelos indivíduos ou mesmo por seus médicos assistentes. Mais ainda, os estudos seccionais são úteis para propósitos analíticos quando a exposição é alguma característica permanente ou invariável dos indivíduos, tais como sexo, grupo sanguíneo, etnia ou local de origem, ou que tenha sido adquirida em fase precoce da vida, anterior ao estabelecimento da doença nos adultos, como geralmente é o caso da classe social e do grau de educação formal.

Outro problema que pode perturbar as propriedades analíticas de um estudo seccional é a duração da exposição nos doentes. A probabilidade de participação de doentes expostos em um estudo seccional é dependente do tempo de duração da exposição, de modo que os doentes com longos períodos de exposição poderão estar super-representados na amostra de estudo em relação àqueles em que a duração de exposição é mais curta. O resultado disto é que uma determinada exposição que prolonga a sobrevida dos doentes, inclusive por tornar a doença mais branda, pode ser confundida com um fator de risco, uma vez que a prevalência de doentes entre os expostos pode ser até maior do que entre os não-expostos.

Também pode ocorrer que a exposição, mesmo não tendo relação causal com a doença, leve os doentes a morrer em prazos mais curtos do que os não-expostos, assim que a comparação das proporções de doença entre expostos e não-expostos, observados em um estudo seccional, leve à conclusão errônea de que a exposição é um fator de proteção contra a doença. Portanto, casos prevalentes, reconhecidos em um estudo seccional, podem refletir não apenas os determinantes de risco de adoecer como também os de sobrevivência. Com isto, torna-se difícil a distinção entre fatores de risco para o aparecimento da doença e fatores prognósticos para sua evolução³⁵. Um fator de risco serve para antecipar a probabilidade da doença vir a ocorrer, enquanto um fator prognóstico serve para predizer sua provável evolução

e seu desfecho. E, mais ainda, um determinado fator pode funcionar como fator de risco e de prognóstico, em épocas diferentes durante a evolução da doença. Um exemplo disto é a associação existente entre grupos étnicos e hipertensão arterial. Um determinado grupo étnico pode representar uma certa carga genética que é fator de risco para a hipertensão arterial, mas também pode estar ligado a exposições ambientais que são fatores prognósticos.

Apesar dos problemas apontados, os estudos seccionais oferecem pelo menos uma vantagem evidente sobre outras estratégias de estudo com finalidades analíticas, sejam as observacionais, como os estudos longitudinais de coortes e de casos e controles, ou as experimentais, como os ensaios clínicos aleatorizados. Esta grande vantagem é a capacidade de inferência dos resultados observados para uma população definida no tempo e no espaço, o que pode não ser permitido em outros estudos cujas amostras não são representativas da população. E, como já foi assinalado antes, os estudos seccionais podem ser bastante adequados para investigar relações de doenças crônicas, de início indeterminado e de longa duração, e ainda, de freqüências relativamente elevadas, com exposições de caráter permanente, fixo, ou que pelo menos sofram pequenas variações de intensidade e presença ao longo da vida, com inícios precoces. A magnitude da freqüência relativa do agravo estudado é importante na medida em que determina a quantidade de observações a serem feitas, de acordo com o grau mínimo de precisão desejado na inferência. Assim, se a doença é muito rara, será necessário fazer um número muito grande de observações em um estudo seccional, de modo que um estudo de casos e controles pode ser mais indicado nos casos de doenças incomuns, de baixa freqüência.

ESTUDOS HÍBRIDOS

Como já foi visto, os estudos seccionais costumam se distinguir de forma nítida dos demais estudos epidemiológicos, pela exigência da coleta de informações em uma amostra representativa de indivíduos, de uma população definida segundo local e época. Entretanto, também podem ser delineados estudos chamados de híbridos, que envolvem outras estratégias inerentes a outros estudos epidemiológicos, além do seccional.

Quando os estados de exposição são detectados nos indivíduos de uma coorte original, a presença de doença pode vir a ser identificada com um estudo seccional somente depois de algum tempo já decorrido desde a detecção inicial dos estados de exposição dos indivíduos.

Esse é o momento final de um estudo híbrido chamado de estudo de prevalência em coorte. Essa estratégia pode ser utilizada quando na fase de planejamento de um estudo de coortes não se considera a doença investigada como possível efeito da exposição.

Se a partir de uma população-fonte original são seguidos seus indivíduos participantes para selecionar apenas aqueles considerados elegíveis para um estudo seccional no momento final da investigação, este híbrido pode ser chamado de estudo de prevalência seletivo. Por exemplo, nos indivíduos que constituem a demanda de um hospital, podem ser selecionados somente aqueles que têm doenças do colágeno para determinação da prevalência de coronariopatias.

Os indivíduos examinados durante um estudo seccional, e que permanecem sob o risco de adoecer, podem ser acompanhados durante um período de tempo adicional, permitindo obter estimativas da prevalência inicial e da incidência neste período que se segue ao estudo seccional inicial. Este tipo de estratégia se constitui em um híbrido chamado de estudo de seguimento seccional. Por exemplo, um estudo seccional²⁵ determina a prevalência de carcinoma de colo de útero nas mulheres de uma população, por meio de exames de massa, sendo posteriormente acompanhadas aquelas sem diagnóstico positivo no primeiro exame, por um período de tempo suficiente para estimar a incidência do câncer.

E, mais ainda, uma metodologia muito utilizada para avaliar o impacto de mudanças induzidas sobre uma população dinâmica, isto é, aquela que sofre modificações de composição ao longo do tempo, é um híbrido chamado de estudo de painéis repetidos. Os painéis repetidos consistem na realização de estudos seccionais em momentos variados, em uma mesma população, entretanto, sem repetir necessariamente as observações sobre os mesmos indivíduos selecionados no estudo inicial. Se fossem examinados sempre os mesmos indivíduos, o estudo seria classificado como longitudinal de coortes. Em cada painel, que corresponde a um estudo seccional, é selecionada uma amostra de indivíduos da população-alvo, independentemente das anteriores. Assim, quanto menor for a fração amostral, que é a razão entre o número de participantes na amostra e o número de indivíduos na população, menor é a probabilidade de repetição de um indivíduo em mais de uma amostra. Nestes estudos, o interesse principal costuma ser a avaliação de impactos ocasionados em uma população dinâmica, submetida a mudanças nos intervalos entre os painéis, e não aqueles que recaem sobre um grupo de indivíduos que viesse a representar apenas a população singular existente na época do

primeiro painel. As mudanças podem ser intervenções planejadas, como campanhas de vacinação, ou eventos potencialmente prejudiciais à saúde das pessoas, como a exposição a poluentes ambientais.

FASES DE UM ESTUDO SECCIONAL

A realização de um estudo seccional pode ser dividida em três fases principais: planejamento, execução, e análise e divulgação de resultados, como é comum a todos os métodos de investigação científica. As atividades e os procedimentos desenvolvidos em cada uma destas fases serão descritos a seguir.

Planejamento

O planejamento é a fase inicial de um estudo seccional que antecede a execução. As atividades principais desenvolvidas durante esta fase são a construção do protocolo, a preparação dos instrumentos para tomada de informações, o processo de amostragem, e a seleção e o treinamento dos pesquisadores de campo.

Protocolo

O planejamento de um estudo seccional se inicia pela construção de um protocolo, também chamado de projeto de pesquisa. O protocolo deve indicar claramente porque é necessário realizar o estudo, em que estágio se encontra o conhecimento atual sobre o tema a ser investigado, e a sua importância para a população-alvo como problema de saúde pública. Portanto, o protocolo deve apresentar os objetivos e as justificativas do estudo seccional.

De uma definição precisa dos objetivos gerais e específicos depende o sucesso do estudo. Pelo que já se viu, em um estudo seccional é fundamental delimitar a população-alvo e explicar os critérios de sua eleição. Além disso, é necessário explicitar objetivamente os indicadores quantitativos que expressarão os resultados principais e a época de referência da coleta de dados. Desta maneira, o objetivo mais geral do estudo seccional geralmente orienta o seu título. Como exemplo, não basta afirmar que o objetivo geral é estudar a tuberculose nos escolares. É preciso demarcar melhor o objeto de estudo com um objetivo tal como: estudo da prevalência de infecção por *Mycobacterium tuberculosis* nos escolares do primeiro grau da rede pública, do Município do Rio de Janeiro, em julho de 2006. Os objetivos específicos poderão detalhar os métodos diagnósticos a

serem utilizados e os subconjuntos ou grupamentos da população a terem seus indicadores particulares revelados. A apresentação de resultados para subconjuntos da população denota, em geral, a intenção de estabelecer comparações ou testar hipóteses de associação entre as características de formação dos grupamentos e o agravo de interesse.

O protocolo deverá conter uma descrição dos métodos propostos para alcançar os objetivos e testar as hipóteses formuladas, se assim for desejado. Entre estes métodos, que serão mais bem detalhados adiante, estão os instrumentos de coleta de dados que serão utilizados, os critérios de classificação e diagnóstico, o registro e o armazenamento de informações, a seleção de indivíduos para compor a amostra, sempre que necessário, a seleção e o treinamento dos observadores, chamados de pesquisadores de campo, o controle de qualidade da coleta, e os procedimentos de análise quantitativa das estimativas desejadas e das hipóteses operacionais. De forma adicional, porém imprescindível, deve ser justificado o tamanho da amostra, ou seja, o número de indivíduos a serem observados, em função da precisão desejada das estimativas e do poder dos testes estatísticos. Da mesma maneira, deverão ser incluídos no protocolo modelos de instrumentos, tais como questionários, instruções de preenchimento e de uso de equipamentos, e fichas de laboratório e de supervisão.

O protocolo deverá também demonstrar a possibilidade de realização do estudo com os métodos propostos, com os recursos solicitados, e no tempo previsto. Detalhes programáticos deverão ser descritos no protocolo a respeito do cronograma de atividades, a guarda e o destino dos materiais utilizados, e os recursos financeiros necessários para pagamento de pessoal, laboratórios, deslocamentos no campo e viagens, aquisição de material permanente e de consumo, e processamento de dados.

Ainda mais, além de demonstrar que o pesquisador é hábil e capaz de conduzir o estudo, o protocolo deverá conter os modos de conduta previstos para resolução de problemas imediatos que ocorram durante o trabalho de coleta de dados (trabalho de campo), inclusive aqueles referentes à identificação de problemas de saúde nos indivíduos selecionados para participar do estudo. Por fim, o protocolo deverá se submeter às diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Para tanto, além de todos os aspectos já descritos, deverá conter o modelo do termo de consentimento a ser apresentado aos indivíduos participantes.

Questionários

Os questionários são os instrumentos de tomada de informações mais frequentemente utilizados nos estudos seccionais. Em geral as informações coletadas por questionários são medidas indiretas, porque são oriundas de testemunhos prestados pelos indivíduos, dependendo de certas condições, tais como memória, capacidade de compreensão e sua própria elaboração das perguntas, além de serem mediadas por interesses particulares dos entrevistados.

Um questionário é um instrumento de coleta de informações que poderá ter seu conteúdo e formato determinados pelas características culturais específicas da população-alvo. Um questionário também pode conter questões que serão preenchidas de acordo com observações feitas pelos próprios pesquisadores de campo, tais como atributos físicos dos entrevistados (cor dos olhos, tipo de cabelos, ausência de membros, e outros), e, nestes casos, diz-se que foram feitas medidas diretas. Os pesquisadores de campo são os observadores especialmente treinados para atuarem como entrevistadores ou examinadores. Algumas informações podem oferecer alternativas de observações diretas ou indiretas. Um exemplo é a idade, em que perguntar a idade é uma manobra mais simples, e indireta, do que solicitar um documento que registre a data de nascimento. Se a primeira alternativa é mais rápida e abrangente, a segunda pode ser mais precisa, ainda que também sujeita a erro.

Há duas formas básicas de aplicação de questionários. Os questionários auto-aplicados são preenchidos pelos próprios entrevistados, que podem recebê-los por via postal ou por intermédio de um portador treinado para complementar esclarecimentos. Os questionários auto-aplicados devem se limitar ao uso da linguagem corrente na população-alvo, e podem conter desenhos ou croquis da anatomia humana para que os entrevistados apontem locais de dor ou outros sinais. São também úteis para obter respostas a perguntas embaralhadas, de foro íntimo, desde que se assegure o sigilo, ou seja, a não-identificação dos entrevistados.

Ainda que o uso de questionários auto-aplicados seja a alternativa mais barata, sua taxa de retorno pode ser frustrante, e, além disto, suas limitações quanto às possibilidades de medições são evidentes. Por isto, é mais freqüente que os estudos epidemiológicos seccionais façam uso de questionários aplicados por entrevistadores. Os entrevistadores, ou pesquisadores de campo, podem ser treinados também para realizar as medidas diretas, seja por observação simples ou com o uso de instrumentos específicos. Além disso, os pesquisadores de campo

ainda podem coletar material biológico dos indivíduos para exames complementares, em laboratório. A discreta coerção exercida pela presença do entrevistador é um fator para obtenção de taxas de respostas mais elevadas do que com o uso de questionários auto-aplicados, inclusive pela aparência de um maior grau de interesse na obtenção de respostas.

O contato entre o entrevistador e o indivíduo pode ser feito até à distância por meio de telefone, por exemplo. Um estudo de monitoramento de fatores de risco para doenças crônicas não-transmissíveis, realizado no Município de São Paulo, em 2003, mostrou que a estratégia de entrevistas telefônicas teve bom desempenho em relação à representatividade da amostra, ao custo e à agilidade do sistema de coleta de informações. É importante salientar que a proporção de domicílios com telefone naquele município foi estimada em 85%, e, portanto, esta boa *performance* só pode ser esperada em localidades de características semelhantes. Contudo, deve-se considerar que o contato telefônico apresenta sérias limitações, pois não permite observações diretas, e nem a conferência de óbvias incoerências entre as respostas dadas e a situação real do entrevistado. Mais ainda, só podem participar de um estudo por telefone aqueles indivíduos que possuem uma linha telefônica ou que se encontram em local servido por uma linha, restringindo muitas vezes, de forma severa, a população de referência. Por isto são preferidos os contatos diretos entre pesquisadores de campo e indivíduos, se possível nos domicílios destes.

Os questionários podem apresentar questões fechadas, abertas ou mistas. Questões fechadas são aquelas cujas alternativas de resposta já estão explicitadas no questionário. Questões abertas são aquelas em que não há alternativas preestabelecidas de respostas, em que, portanto, as respostas deverão ser registradas de forma literal, com as expressões usadas pelos entrevistados. As questões mistas contêm algumas alternativas preestabelecidas, mas admitem respostas não-previstas. É comum que as questões mistas contenham uma última alternativa de resposta que tem o formato: *Outra, qual? _____*. As questões fechadas, e em menor grau as mistas, são as que oferecem as menores dificuldades para os tratamentos necessários para as análises quantitativas. As questões abertas devem ser reservadas para aquelas situações em que as alternativas de resposta são muito numerosas, como no caso da informação sobre os medicamentos em uso pelos entrevistados, ou para quando se deseja fazer uma análise exploratória inicial de um

problema ainda pouco conhecido. A construção de um questionário adequado é uma das tarefas primordiais do planejamento de um estudo seccional.

Para fins de comparação de resultados e de perfis epidemiológicos de instrumentos elaborados e validados em outra língua, é necessário avaliar e garantir a equivalência entre o original e suas traduções. A adaptação transcultural está baseada na apreciação de uma seqüência de diferentes tipos de equivalências: 1. *equivalência conceitual*, relacionada à existência dos conceitos e dimensões apreendidos pelo instrumento original na cultura-alvo da nova versão; 2. *equivalência de itens*, que se destina a avaliar cada item do instrumento original e a correspondência de seus conceitos na população onde será utilizado; 3. *equivalência semântica*, que avalia a transferência de significados das palavras entre os idiomas, com intuito de obter efeitos similares nos respondentes da mesma escala; 4. *equivalência operacional*, na qual se observa a adequação e pertinência de aspectos operacionais no novo contexto cultural, tais como forma de administração e número de opções de resposta; e, por fim, 5. *equivalência de mensuração* para avaliar as propriedades psicométricas do instrumento original e a sua nova versão.

Essas etapas utilizam diferentes estratégias para avaliação, entre as quais as principais são: revisão bibliográfica envolvendo publicações da cultura do instrumento original e da população-alvo, discussão com especialistas, discussão com população-alvo, tradução, retradução, avaliação da equivalência semântica desta retradução com o original, pré-teste da nova versão, e avaliação pelo grupo de pesquisa da pertinência e adequação da forma, e do modo de aplicação.

Medidas Objetivas e Escalas

São também chamadas de medidas objetivas aquelas feitas por observações diretas dos pesquisadores de campo (observadores) sobre os indivíduos (observados), muitas vezes com o uso de algum instrumento específico, algumas vezes com coleta de material, e que não dependem da memória do entrevistado ou de uma resposta dada a um questionário.

São inumeráveis as medidas objetivas possíveis, que muitas vezes se referem a variáveis relacionadas com a fisiologia ou os processos biológicos dos indivíduos, tais como concentrações de substâncias no sangue ou dejetos, pressão arterial, freqüências cardíaca e respiratória, entre tantas. Também podem ser relacionadas com características ou sinais externos dos indivíduos,

tais como peso, altura, coloração dos olhos, tipo de cabelos, presença ou ausência de membros ou sinais patognomônicos de doença, por exemplo.

Os métodos de observação destas variáveis, assim como os tratamentos analíticos, dependem de sua categorização. De um modo geral, as variáveis são divididas em discretas e contínuas, de acordo com a possibilidade da existência de outras categorias no intervalo entre duas categorias quaisquer. As variáveis discretas simples ou qualitativas são aquelas que possuem um número limitado de categorias, como é o caso das medidas em escalas nominais ou ordinais. O exemplo mais comum de uma variável discreta qualitativa em estudos seccionais é o atributo doença. Este atributo, doença, é uma variável nominal dicotômica, ou binária, quando só possui duas categorias, a dos doentes e a dos não-doentes. As variáveis discretas quantitativas são aquelas em que há um grande número de categorias, ordenadas, porém em número finito entre dois limites fixos, e sem valores intermediários nos intervalos entre as categorias. Contagens de células ou indivíduos são exemplos de variáveis discretas quantitativas. Já as variáveis contínuas possuem um número infinito de categorias, mesmo que entre limites mínimo e máximo, de modo que em cada intervalo entre duas categorias há sempre outras possíveis, ainda que não detectáveis, e podem ser medidas em escalas de intervalos ou de razões. A temperatura, medida em graus Celsius ou Fahrenheit, é uma variável contínua de escala de intervalos porque possui infinitas categorias, ordenadas, e com diferenças mensuráveis entre quaisquer duas categorias, mas o valor zero é arbitrário, de modo que não é possível interpretar de modo útil a razão entre dois valores. Já o peso é uma variável contínua de razões, porque possui todas as propriedades das variáveis contínuas e, ainda, é possível calcular razões entre dois valores quaisquer. Esta razão representa quantas vezes o peso do numerador é maior que o do denominador. Na prática dos estudos seccionais, a identificação das categorias das variáveis contínuas está limitada pela capacidade de discriminação dos instrumentos e observadores. Mais detalhes sobre este assunto podem ser vistos no Capítulo 18 – Análise Exploratória de Dados.

Adequação, Validade e Confiabilidade de Medidas

É muito frequente que para cada variável que se pretenda medir existam diversos instrumentos e regras de utilização alternativas, cujas combinações são chamadas genericamente de técnicas de medida. Portanto, é neces-

sário decidir quais técnicas de medida serão empregadas no estudo seccional. De início, é preciso considerar a portabilidade dos instrumentos e a aceitação do seu uso por parte da população-alvo, critérios pragmáticos de escolha. Isto porque a maioria dos estudos seccionais exige razoáveis deslocamentos no espaço em que se encontram as unidades de observação selecionadas, sendo as medições feitas nos domicílios de residência dos indivíduos. Além disto, é preciso levar em consideração que a maioria dos indivíduos geralmente não é afetada pelo agravo em investigação e, principalmente, não é demanda passiva de serviços de saúde. O pesquisador precisa obter a adesão voluntária dos indivíduos selecionados aleatoriamente para participar do estudo, demonstrando que os procedimentos de medida utilizados são inócuos, indolores, e, ainda assim, de utilidade para a pesquisa⁸.

Por outro lado, a decisão de escolha de uma técnica de medida deve ser também norteada por critérios técnicos. Assim, as melhores técnicas de medida são aquelas que fornecem medidas com as menores probabilidades de erro, no caso de classificações dicotômicas, ou os menores desvios (vieses de medida) em relação às verdadeiras. Uma técnica de medida que simplesmente classifique os indivíduos nas duas categorias de uma variável dicotômica, tal como doença, poderá ser analisada em função da probabilidade de erro de classificação em relação a cada uma das categorias, de doentes e não-doentes, por exemplo. Estes erros estão relacionados de forma complementar com os acertos. Os indicadores clássicos de validade ou de acertos são chamados de sensibilidade e especificidade.

A sensibilidade de uma técnica é a proporção de indivíduos corretamente classificados pela técnica como positivos, isto é, com o atributo investigado, entre todos aqueles que são realmente positivos. A especificidade de uma técnica é a proporção de indivíduos corretamente classificados pela técnica como negativos, isto é, sem o atributo investigado, entre todos aqueles que são realmente negativos. Como se percebe, estes indicadores são construídos pela confrontação de medições obtidas por uma técnica de medida com as reais (verdadeiras), em uma amostra de indivíduos selecionados para participar de um estudo com a finalidade de estimar sensibilidade e especificidade, ou seja, a validade da técnica. As medições verdadeiras podem ser por técnicas diagnósticas consideradas como “padrão-ouro”. A técnica de medida que será usada no campo, na prática do estudo seccional, raras vezes tem as qualidades de um “padrão-ouro”, de modo que são esperados erros de classificação, ou seja,

sua sensibilidade e sua especificidade não são perfeitas, quer dizer, não são iguais a 1 (ou 100%). Ainda assim, é possível estimar a prevalência de doença real ou corrigida mesmo tendo sido usada uma técnica de medida imperfeita no estudo seccional, com a fórmula:

$$p_r = \frac{p_a + Esp - 1}{Sen + Esp - 1}$$

em que p_r é a prevalência real, p_a é a prevalência aparente obtida no estudo usando a técnica de medida imperfeita, Esp é a probabilidade de especificidade e Sen é a probabilidade de sensibilidade. Esta manobra algébrica permite corrigir o valor da estimativa da prevalência que se refere ao conjunto de indivíduos participantes do estudo, porém não permite corrigir os diagnósticos individuais²².

Exemplificando: podemos supor que um estudo específico tenha demonstrado que um certo instrumento de diagnóstico de alcoolismo apresenta sensibilidade de 0,9, ou seja de 90%, e especificidade de 0,8, ou seja, de 80%. Estas informações irão servir para estimar a prevalência real de alcoolismo em uma amostra na qual a prevalência obtida com aquele instrumento foi de 0,235, ou seja, de 23,5%. A utilização da fórmula apresentada resulta em:

$$Pr = \frac{0,235 + 0,8 - 1}{0,9 + 0,8 - 1} = 0,05$$

o que significa que a prevalência real de alcoolismo naquela amostra era de apenas 5%.

Técnicas de medida de variáveis contínuas podem ter sua validade de uso avaliada por meio de indicadores como o viés absoluto padronizado e o viés das médias. O viés absoluto padronizado é dado pela média das diferenças absolutas (sem os sinais) dos valores obtidos pela técnica e os reais, dividida pelo desvio-padrão dos valores reais. O viés das médias é dado pela diferença entre a média dos valores obtidos pela técnica e a média dos valores reais¹⁹, dividida pelo desvio-padrão dos valores reais. O primeiro indicador é sensível aos vieses de medidas individuais, ao passo que o segundo só serve para avaliar os vieses entre as médias, referindo-se, portanto, a conjuntos de indivíduos.

Os critérios técnicos descritos até aqui exigem a confrontação dos valores obtidos por uma técnica de

medida com os verdadeiros ou reais. Entretanto, na prática isto não só nem sempre é possível, como algumas vezes nem mesmo há um “padrão-ouro”, ou mesmo, é discutível o conceito de “verdade”, como no caso de diagnósticos na área psiquiátrica. Por isto, um outro critério técnico de escolha de uma técnica de medida está relacionado tão-somente com sua capacidade de concordância interna, ou seja, de replicar os mesmos resultados em indivíduos invariantes. Esta capacidade é chamada de confiabilidade. A construção de indicadores de confiabilidade exige que as réplicas de medidas sejam feitas num mesmo conjunto de indivíduos, pelo mesmo observador ou por observadores diferentes. As comparações de confiabilidade também podem ser feitas entre os resultados obtidos por técnicas diferentes, mas que têm o mesmo propósito, sempre no mesmo conjunto de indivíduos.

A avaliação quantitativa da confiabilidade também depende da escala de medidas, discreta ou contínua. Para medir a confiabilidade de técnicas de medidas discretas, com pequeno número de categorias, um dos indicadores mais utilizados é o kappa (κ), calculado com:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

em que p_o é a proporção total de concordâncias observadas, e p_e é a proporção total de concordâncias que seriam esperadas por pura casualidade. Por exemplo, no caso de uma técnica classificatória simples que detecta apenas duas classes, chamadas de positivos e negativos, os resultados de uma prova de confiabilidade, com n indivíduos, podem ser dispostos como na Tabela 10.2, em que a mesma técnica replicou medidas, em duas oportunidades, da seguinte maneira:

Tabela 10.2 Resultados de uma Prova de Confiabilidade			
Freqüências		Oportunidade B	Total
Oportunidade A	+	n_{11}	n_{12}
	-	n_{21}	n_{22}
Total		$n_{\cdot 1}$	$n_{\cdot 2}$
			n

A proporção total de concordâncias observadas é calculada pela soma das freqüências observadas nas células da diagonal de concordâncias, internas à tabe-

la, n_{11} e n_{22} , dividida pelo número total de indivíduos examinados com:

$$p_o = \frac{n_{11} + n_{12}}{n} = \sum_{i=1}^{\kappa} \frac{n_{ii}}{n}$$

sendo κ o número de categorias da variável, enquanto a proporção de concordâncias esperadas por puro acaso é dada pela soma das freqüências esperadas nas células da mesma diagonal de concordâncias, também dividida pelo mesmo número total de indivíduos (total geral). As freqüências esperadas são calculadas pelo produto das marginais (totais) divididas pelo total geral (n). Portanto, a proporção total de concordâncias esperadas por acaso é dada por:

$$p_e = \frac{\sum_{i=1}^{\kappa} (n_i \cdot n_{\cdot i})}{n^2}$$

A interpretação do índice de kappa deve ser feita com cautela. Apesar de este índice relacionar proporções, observadas e esperadas, kappa não é uma proporção e não possui as propriedades das proporções. O valor do índice kappa pode variar entre um valor negativo até +1, dependendo das distribuições dos valores das marginais (totais de linhas e de colunas) e do número de categorias (κ) da variável. Por exemplo, em uma tabela 2×2 cujas células internas tivessem os seguintes valores: $n_{11}=14$, $n_{12}=7$, $n_{21}=25$ e $n_{22}=171$, o valor calculado de Kappa é 0,39, porém, seu valor mínimo, nesta tabela, seria -0,14 e o máximo, 0,66.

A confiabilidade de medidas em escala ordinal pode ser avaliada pelo kappa ponderado. Os pesos atribuídos a cada categoria são arbitrários, porém relacionados com as distâncias relativas entre as mesmas.

Para medir a confiabilidade de técnicas de medidas contínuas, pode ser utilizada uma estimativa do erro experimental, que consiste na variância existente entre réplicas, combinada para n indivíduos, que participam do experimento. No caso de um experimento com duas réplicas, isto é, duas medidas feitas em cada indivíduo, o algoritmo final de cálculo do erro experimental é bastante simples, consistindo na razão entre a soma das diferenças quadráticas individuais das réplicas, dividida pelo dobro de indivíduos participantes. Como a confiabilidade é inversamente relacionada com o erro experimental, a melhor técnica de medida é aquela que oferece o menor erro experimental. Existem ainda outros

indicadores de confiabilidade para variáveis contínuas, tais como o coeficiente de correlação intraclasse.

Amostragem

Uma das questões fundamentais em um estudo seccional é a sua base de referência de resultados. Uma das grandes vantagens dos estudos seccionais em relação às demais estratégias de estudos epidemiológicos é a generalização de seus resultados para uma população definida, como já foi visto antes. Se todos os elementos da população tiverem sido examinados, a possibilidade de generalização de resultados é evidente. Entretanto, é muito mais frequente que se examine apenas uma amostra de indivíduos que compõem a população, e neste caso, para que a generalização seja válida em termos probabilísticos, é necessário que os processos de seleção das unidades de observação sejam rigorosamente conduzidos apenas pelo acaso. Em outras palavras, a inferência dos resultados da amostra para a população requer o uso de processos de amostragem probabilística. Somente amostras formadas por indivíduos escolhidos aleatoriamente, ou seja, por acaso, podem fornecer estimativas de parâmetros representativas da população. A representatividade de uma amostra significa que ela é um “espelho”, uma “miniatura” da população, que representa com distribuições similares às da população todas as estatísticas desejadas pelo pesquisador.

Amostragem probabilística é aquela na qual todos os elementos da população têm probabilidade conhecida e diferente de zero de participar da amostra. As questões gerais sobre amostragem em estudos epidemiológicos podem ser vistas no Capítulo 22 – Amostragem. Aqui vamos fazer apenas algumas considerações referentes a processos particulares de amostragem, muito utilizados em estudos seccionais.

Como já mencionamos anteriormente, com frequência os estudos seccionais procuram representar, com amostras, populações muito grandes, da ordem de centenas de milhares ou até de milhões de indivíduos, espalhados em áreas territoriais bastante amplas. Assim, nestes casos, os processos de amostragem básicos, como a amostragem aleatória simples ou a sistemática, podem tornar os estudos muito caros pelo grande número de deslocamentos a serem feitos entre os indivíduos selecionados, que muito provavelmente se encontrarão bastante dispersos por todo o espaço. E, ainda mais, raras vezes existem listas completas de todos os indivíduos que compõem a população-alvo, em qualquer recorte que se faça da população brasileira. E, como

aqueles processos de amostragem mencionados exigem listas completas de indivíduos da população, os mesmos são inviáveis. Entretanto, pode ser possível obter listas de agregados das unidades básicas de observação, os indivíduos. Estes agregados ou aglomerados, de graus variados de complexidade, costumam ser chamados de conglomerados.

Os conglomerados podem se apresentar em diversos níveis de organização, sendo os mais elementares geralmente os domicílios ou locais de moradia das pessoas, identificados por endereços. Outros níveis mais complexos de organização de conglomerados podem ser exemplificados com bairros, setores censitários, Municípios, microrregiões homogêneas e Estados. Os setores censitários são conglomerados definidos pelo IBGE (Instituto Brasileiro de Geografia e Estatística), e costumam se constituir em conglomerados de cerca de 200 domicílios, em média, nas áreas urbanas, e em menor número nas áreas rurais.

A existência de listas de conglomerados permite a execução de esquemas de amostragem destes conglomerados, em seus diversos níveis hierárquicos. Estes processos de amostragem são chamados genericamente de amostragem de conglomerados, e utilizam-se de esquemas mais básicos como a amostragem aleatória simples ou sistemática de conglomerados, em um único nível de agregação, até processos mais complexos realizados em estratos da população, divididos em vários estágios ou hierarquias de conglomerados. Em cada nível de seleção dos processos de amostragem de conglomerados, só é necessário conhecer a lista completa de elementos que compõem o estágio imediatamente inferior em hierarquia de complexidade estrutural, que pode ser composto por outro nível de conglomerados, naqueles conglomerados que tiverem sido selecionados. Deste modo, o último estágio de seleção costuma ser o dos indivíduos contidos nos conglomerados mais elementares, que podem ser os domicílios, ou as seções de trabalho, por exemplo. Se for desejada a subseleção destes elementos, seu conhecimento se dará durante a fase de coleta de dados. Por todas estas características é que estes esquemas de seleção mais intrincados são também chamados de processos de amostragem complexa.

O estadiamento da seleção de conglomerados pode ser exemplificado com o caso de uma amostra de pessoas que pretendesse representar a população brasileira. O Brasil é organizado administrativamente de uma maneira que pode ser utilizada para orientar um processo de amostragem de conglomerados em múltiplos estágios de seleção. Estes estágios, citados em ordem decrescente de hierarquia administrativa e espacial, poderiam ser o

dos Estados, dos Municípios, dos setores censitários, dos domicílios e, por fim, dos indivíduos. Em cada um destes cinco estágios pode-se fazer uma seleção aleatória de conglomerados, desde o nível mais elevado, dos Estados, até o dos domicílios, e por fim, escolher, nesses, os indivíduos. A subseleção no último nível de aglomeração é opcional, pois se pode optar pelo exame de todos os indivíduos que se encontram nos domicílios selecionados.

A utilização de processos de amostragem complexa implica no uso de estimadores adequados, durante a análise dos dados, para a obtenção das estimativas desejadas, e principalmente de seus intervalos de confiança. Observações feitas em conglomerados de indivíduos costumam reduzir a heterogeneidade da amostra, em relação às variáveis de interesse. Portanto, é preciso ter cuidado com as medidas de variância de indicadores, tais como prevalências, médias, razões de prevalência ou de chances, derivados de dados obtidos em estudos cuja amostra é de conglomerados, de qualquer grau de complexidade. Os algoritmos de cálculo utilizados na maioria dos aplicativos de computador que geram estatísticas, a partir de bancos de dados, supõem, *a priori*, que os dados foram coletados de uma amostra aleatória com reposição das unidades finais de observação, os indivíduos. Portanto, em uma análise rigorosa dos dados derivados de um estudo seccional de amostragem complexa, é preciso utilizar rotinas especiais de estimação e de testes de hipóteses que levem em consideração o delineamento amostral efetivamente praticado para coletar os dados do estudo. Não são comuns os aplicativos que possuem rotinas adequadas, já preestabelecidas, para resolver alguns destes problemas, tais como *Epi-Info* e *Stata*^{9,37}. As estatísticas avançadas do *Epi-Info (complex sample)* têm por finalidade analisar dados oriundos de estudos seccionais complexos, com amostras simples de conglomerados ou em dois estágios⁹, de conglomerados no primeiro estágio e de indivíduos no segundo estágio. Discussões amplas sobre o efeito de delineamento, a razão entre a variância calculada de forma adequada no estudo de amostragem complexa e a que seria obtida com uma amostra aleatória simples do mesmo tamanho, podem ser vistas na literatura especializada.

Uma adaptação simplificada de um processo de amostragem de conglomerados em dois estágios vem sendo preconizada pela Organização Mundial de Saúde, por meio do Programa Ampliado de Imunizações (*PAI* ou *EPI, Expanded Programme on Immunization*), para uso em condições precárias de campo, quando as informações sobre a distribuição da população, os recursos

financeiros e o tempo disponível para o estudo são escassos. No início, os estudos seccionais que adotaram o método proposto, e que vinham sendo realizados desde 1978, tinham por finalidade estimar coberturas vacinais em crianças nas mais variadas populações, especialmente nas mais pobres. O método original recomendava a seleção de 30 conglomerados no primeiro estágio, com probabilidades proporcionais aos tamanhos dos conglomerados, com um esquema de amostragem sistemática, e a posterior seleção de sete indivíduos, em cada conglomerado, no segundo estágio. Nos primórdios da aplicação do método, também conhecido por método de inquéritos rápidos, apenas um dos indivíduos do segundo estágio era selecionado ao acaso, sendo os restantes seis indivíduos escolhidos entre aqueles mais próximos do primeiro. Esta combinação 30×7 era recomendada para avaliação de atributos comuns, quando a prevalência esperada fosse maior do que 50%. Já para atributos raros, com prevalência menor do que 5%, era sugerida a seleção de 40 conglomerados com 100 indivíduos em cada um. Várias modificações foram sendo propostas nos últimos anos, tais como a seleção aleatória nos dois estágios, tanto de conglomerados como de indivíduos, aumento no uso de conglomerados e uso de computadores de mão, entre outras.

A perda de informações sobre parte da amostra, independente do motivo, resulta em modificação da população de referência original. A não ser que estas perdas estejam, por sua vez, representadas naquela parte da amostra que foi efetivamente observada, os indivíduos perdidos em uma amostra representam uma fração da população sobre a qual se perde a capacidade de inferência. A substituição de indivíduos, originalmente selecionados para participar da amostra, pode parecer uma alternativa razoável para remediar perdas ou até mesmo para acelerar o processo de coleta de dados. Entretanto, mesmo que estas substituições sejam feitas de acordo com regras definidas *a priori*, tais como escolher como substituto o mais próximo à direita, no sentido horário, e assim por diante, é impossível assegurar que os substitutos representem os substituídos em relação às variáveis de interesse. Tais manobras de substituição, com regras preestabelecidas, talvez até possam obter unidades de observação alternativas que se assemelhem às planejadas quanto a algumas escassas características, como o local de residência, sexo ou grupo etário. Entretanto, este tipo de técnica se assemelha com aquela utilizada em um processo de amostragem não-probabilística, por quotas, no qual a conveniência dos pesquisadores de campo, e não o acaso, é um dos critérios mais importantes na escolha das unidades finais de observação. Enfim,

a generalização da técnica de substituição de perdas acaba por descharacterizar a aleatoriedade da seleção dos indivíduos, fazendo com que a amostra também deixe de ser probabilística, prejudicando ou inviabilizando a inferência dos resultados.

Seleção e Treinamento de Pesquisadores de Campo

A quantidade de observações a serem feitas em um estudo seccional é geralmente muito grande, como já vimos anteriormente. Além disto, o ideal seria que todos os indivíduos fossem observados em um único momento para evitar distorções provocadas por mudanças ocorridas durante um intervalo longo de estudo. Entretanto, para isto seriam necessários tantos observadores, ou pesquisadores de campo, quantos fossem os observados, os indivíduos. Portanto, é recomendável utilizar a maior quantidade possível de pesquisadores de campo que forem permitidos pelos recursos técnicos e financeiros disponíveis.

Os pesquisadores de campo devem ser selecionados em função de características pessoais tais como gênero, escolaridade, inteligência, experiência, personalidade, cordialidade, capacidade de persuasão, adaptabilidade a ambientes diferentes, pertinácia e honestidade. Em geral, estas características são inerentes às pessoas. As mulheres costumam ser aquelas que menos rejeição provocam nos entrevistados de qualquer sexo, mas algumas vezes o trabalho de coleta de dados é mais eficiente quando são formadas duplas de pesquisadores de campo, de ambos os gêneros. Nem sempre aqueles com grau mais elevado de escolaridade, com maior capacidade intelectiva, ou de maior experiência no uso de equipamentos, são os melhores pesquisadores de campo. Estas aparentes qualidades podem conflitar com a necessidade de pesquisadores de campo que sejam bem aceitos pela comunidade em estudo. Além disto, a experiência prévia pode se contradizer com uma das características essenciais que se deseja em um pesquisador de campo, a neutralidade. O pesquisador de campo deve estar livre de preconceitos sobre os objetos de estudo e os entrevistados. A presença do pesquisador de campo não deve afetar a medida do sujeito observado e suas percepções sobre o que é inquirido. Por isto, é bastante conveniente que o pesquisador de campo se assemelhe aos entrevistados em relação a várias características pessoais, inclusive gênero, grupo étnico e aparência geral.

O treinamento dos observadores visa obter a padronização de procedimentos na coleta de dados. A padronização é a uniformidade da aplicação das práticas corretas de medida ou entrevista por todos os pesquisadores de campo, e por todo o período de tempo em

que transcorre a coleta dos dados. Em outras palavras, estarão padronizados os pesquisadores de campo quando todos se comportarem como se fossem um único, e que é invariante ao longo de toda a fase de coleta de dados da pesquisa.

No início do treinamento, os pesquisadores de campo deverão receber instruções detalhadas sobre os procedimentos, inclusive por escrito. Como é evidente, cada informação desejada terá sua coleta treinada de acordo com o método preconizado pelo pesquisador principal do estudo. A padronização pode ser alcançada por meio de técnicas de simulação ou de experimentos que reproduzam as situações reais de medida no campo. Para medidas objetivas ou diretas, com ou sem utilização de instrumentos, em sessões de simulação de medições, os observadores deverão fazer medidas replicadas em indivíduos que tenham perfis semelhantes aos que compõem a população-alvo do estudo. Na aplicação de questionários, os entrevistadores deverão ser acompanhados pelo pesquisador principal e seus supervisores, durante o treinamento, de início em “classe” e, posteriormente, no ambiente real de trabalho, o “campo”.

Os pesquisadores de campo deverão ser treinados também em relação aos métodos de abordagem dos indivíduos a serem entrevistados. Estes métodos serão fundamentais na conquista da adesão dos entrevistados ao estudo, e serão em grande parte dependentes de características peculiares da população-alvo. Na abordagem inicial dos indivíduos, deverão ser fornecidas as informações básicas sobre o estudo, sem entretanto revelar detalhes sobre eventuais hipóteses que serão avaliadas na análise, para evitar a contaminação dos resultados por respostas “interessadas”. Cuidado semelhante deve ser tomado no treinamento dos próprios pesquisadores de campo para que a neutralidade inicial dos mesmos não seja alterada.

É recomendável treinar um número maior do que o necessário de pesquisadores de campo, se possível o dobro. Inicialmente, como é óbvio, serão escalados como pesquisadores de campo aqueles que demonstrarem o melhor desempenho durante as provas das sessões de treinamento. Os demais observadores treinados poderão ser admitidos em casos de desistência ou de outro tipo de demissão dos que iniciaram o estudo.

Execução

Estudo-piloto

O estudo-piloto consiste em um ensaio que reproduz todos as estratégias e os métodos utilizados na coleta

de dados, em uma amostra da população-alvo especialmente selecionada para este fim. Esta amostra-piloto deve ser de tamanho bem menor do que a amostra de estudo propriamente dita, suficiente apenas para um ou dois dias de trabalho. A intenção do estudo-piloto é testar de forma definitiva os instrumentos, inclusive os questionários, quanto a sua viabilidade e, ainda, os pesquisadores de campo. Além disso, alguns dos resultados obtidos no estudo-piloto poderão servir para embasar os cálculos relativos aos tamanhos mínimos de amostra necessários. Maiores detalhes sobre este assunto poderão ser vistos no capítulo 23 – Tamanho da Amostra em Investigações Epidemiológicas. Porém, os resultados obtidos no estudo-piloto não servirão para compor as estimativas e os testes de hipóteses finais desejados. Portanto, o estudo-piloto é um ensaio do que virá a ser a execução da coleta de dados, permitindo a correção de erros e a obtenção de informações complementares para o planejamento amostral.

Pode ser oportuno que os pesquisadores de campo não tenham conhecimento dos limites do estudo-piloto. É preciso cuidar para que eventuais inovações e alterações nas rotinas de coleta de dados resultantes do estudo-piloto sejam incorporadas de maneira uniforme por todos os pesquisadores de campo.

Coleta de Dados

A coleta de dados é a aplicação prática de todas as estratégias de abordagem e métodos de tomada de informação que foram planejados em todos os seus detalhes pelo pesquisador principal e associados. Em um estudo seccional, esta fase deve ser a mais curta e a de concentração mais intensa de trabalho.

Esta atividade costuma ser precedida e acompanhada pela divulgação dos objetivos gerais do estudo. A divulgação pode ser de caráter coletivo, pelo uso de meios de comunicação de massa como rádio, televisão e imprensa escrita, e de âmbito individual, por meio de cartas enviadas aos domicílios sorteados para a amostra. O ideal é a utilização de ambos, porque se reforçam e se complementam. A divulgação de massa confere uma legitimidade aos objetivos e aos executores e seus patrocinadores.

O contato prévio individual permite o emprego de manobras que podem resultar em maior eficiência do trabalho de campo. Por um lado, garante uma melhor uniformidade na explanação dos objetivos gerais do estudo, e por outro, ainda pode aumentar a eficiência do trabalho de localização e a abordagem dos entrevis-

tados. Para isto é necessário anexar uma agenda, a ser preenchida pelos indivíduos, apontando os melhores dias e horários para entrevistas e exames. Estas agendas devem ser retornadas para o escritório central da pesquisa em envelopes pré-postados nos correios. O termo de consentimento, uma imposição ética, também pode acompanhar esta agenda ou ser apresentado no contato inicial, efetivo, entre pesquisador de campo e indivíduo. No termo de consentimento, além da garantia de uso de métodos inócuos para a integridade dos participantes, deve ser assegurado o total sigilo dos dados individuais.

É comum que a cada pesquisador de campo seja entregue o trabalho de exame dos indivíduos que se encontram em determinada área geográfica. Isto reduz a quantidade de deslocamentos espaciais do observador, e, consequentemente, o custo e o tempo de pesquisa. Porém, para evitar ou minimizar a influência do pesquisador do campo nos resultados do estudo, o pesquisador principal poderá optar por distribuir as unidades de observação, os indivíduos, aleatoriamente entre os observadores, o que também significa que não só os indivíduos são escolhidos ao acaso para participar do estudo, como, ainda, são distribuídos para observação pelos pesquisadores de campo de maneira aleatória. Apesar de este procedimento distribuir ao acaso uma fonte de erro de medida, o pesquisador de campo aumenta a carga de trabalho desses observadores, e, com isto, eleva os custos financeiros e o tempo de estudo. Esta observação reforça a grande importância do trabalho de treinamento para padronização dos pesquisadores de campo.

Algumas compensações podem ser oferecidas aos indivíduos participantes do estudo, tais como transporte de sua residência ou local de trabalho para o centro de pesquisa, oferta de serviços para resolução de problemas de saúde, inclusive aqueles identificados durante a pesquisa, e resarcimento por perdas ou indenização por danos. Os sujeitos de uma pesquisa não devem ser pagos por entrevista ou procedimento porque isto pode se constituir em uma coerção para sua participação, especialmente para os mais pobres. Do mesmo modo, a oferta de brindes, na forma de bens de pequeno valor, que possam ser obtidos sem ônus adicional para o patrocinador oficial da pesquisa, também deve ser evitada porque se ressente do mesmo problema ético apontado em relação ao pagamento. Entretanto, o resarcimento por perdas, tais como horas de trabalho, deve ser feito, assim como também são imperativas as indenizações por eventuais danos causados aos indivíduos.

Controle de Qualidade

Durante toda a fase da coleta de dados, o pesquisador principal e seus associados devem estar atentos e preocupados para que se cumpram todos os preceitos previstos no protocolo. De maneira especial, deverão ser entrevistados e examinados todos os indivíduos planejados no delineamento amostral, e somente aqueles, e todas as observações deverão ser feitas de acordo com as instruções dadas durante o treinamento. Estes cuidados é que são os principais alvos do controle de qualidade, desenvolvido principalmente por meio da atividade de supervisão.

A supervisão deverá ser exercida pelo pesquisador principal e por seus associados, e também por pesquisadores de campo especiais, denominados de supervisores. Estes pesquisadores de campo supervisores devem ser pessoas de confiança do pesquisador principal, especialmente treinados para exercer atividades de vigilância e replicação de observações. Assim, uma das tarefas primordiais da supervisão é evitar a perda de informações. Estas perdas podem ser inevitáveis, quando indivíduos planejados para participar do estudo não pertencem mais à população no momento da coleta de dados. A probabilidade de ocorrência deste tipo de perda aumenta quando o período decorrido entre a confecção da listagem utilizada no delineamento amostral e a tomada de informações efetiva se amplia. De qualquer maneira, as perdas inevitáveis refletem mudanças na composição da população de referência original. O pesquisador poderá ajustar a população final de referência em função destas perdas, se forem conhecidas algumas características destes indivíduos perdidos. Entretanto, se a seleção tiver sido de conglomerados tais como domicílios, e não havendo subtração ou adição de novos domicílios no intervalo entre a listagem e a coleta, a representatividade da amostra pode ser referida à população total existente na época da tomada de informações, desde que não ocorram perdas de indivíduos nos conglomerados selecionados.

Não podem ser toleradas as perdas evitáveis, especialmente aquelas por recusa ou não-localização dos entrevistados, ou ainda por perda do questionário ou fraude do pesquisador de campo. As recusas poderão ser evitadas pela persuasão eficiente exercida desde a comunicação de massa, o envio de cartas personalizadas, e pela correta abordagem feita pelo observador de campo. A não-localização de um indivíduo pode se dever à visitação feita em ocasião em que o mesmo não se encontrava no domicílio ou local de trabalho. Neste caso, uma nova visita poderá resolver o problema. Entretanto, alguns indivíduos poderão estar hospitalizados

ou internados, e mesmo quando localizados podem não estar em condições de prestar as informações desejadas. Problema semelhante pode ocorrer com indivíduos muito idosos ou com deficiências físicas, ou mentais, severas. Algumas das informações desejadas poderão ser tomadas de informantes que acompanhem habitualmente estes indivíduos, em geral um familiar mais próximo. De qualquer maneira, as informações sobre estes indivíduos poderão ser muito relevantes, devendo ser registradas, especialmente se o objeto de estudo estiver relacionado com a doença incapacitante desses indivíduos. O ideal é que tenham sido desenvolvidos instrumentos de coleta adequados para estas situações e que estas informações sejam tratadas em separado ou em grupo especial.

As fraudes mais comuns cometidas pelos pesquisadores de campo consistem na omissão de entrevistas com indivíduos disponíveis para estudo ou na “criação” de pessoas inexistentes. Geralmente a origem destas fraudes está baseada no sistema de remuneração dos pesquisadores de campo. Quando o pagamento é estipulado somente de acordo com o tempo de trabalho, pode haver um estímulo para a fraude por “omissão”. Já quando o pagamento é feito tão-somente por entrevista realizada, ocorre estímulo à fraude por “criação”. Alternativamente, pode ser instituído um sistema misto que premie os pesquisadores de campo que atinjam as metas prefixadas, consideradas convenientes para o pesquisador principal, de um determinado número de entrevistas por períodos de trabalho.

Nenhum esquema de remuneração e premiação é imune a fraudes. É o trabalho de replicação de observações dos supervisores que inibe as fraudes e detecta erros de procedimento no campo. Estas replicações consistem em visitas a uma subamostra dos domicílios ou indivíduos selecionados para a amostra completa do estudo. Esta subamostra deve ser de pelo menos 10% da amostra total, e distribuída de forma homogênea de acordo com o trabalho de cada pesquisador de campo. Isto quer dizer que, da produção apresentada por cada pesquisador de campo, serão novamente visitados, num intervalo de tempo o mais curto possível desde a primeira entrevista, pelo menos 10% dos indivíduos para exame por um supervisor. A escolha dos indivíduos para subamostra de supervisão deve ser feita, de modo aleatório, somente depois que um certo número de visitas já tenham sido feitas pelos pesquisadores de campo. Com estas visitas de supervisão, poderão ser detectadas fraudes por “omissão” e por “criação”.

Se ainda persistirem perdas, após a supervisão, o pesquisador principal poderá optar entre considerá-las

na estimativa das estatísticas desejadas ou proceder a estudo especial das perdas. A correção das estimativas desejadas, tendo ocorrido perdas, pode ser feita considerando que o estado dos perdidos é extremo em relação às medidas das variáveis de interesse. Assim, se a intenção é estimar a prevalência de uma doença, pode-se supor que todos os perdidos são sadios ou todos são doentes, e a partir daí calcular uma prevalência mínima e outra máxima. Foram desenvolvidos alguns algoritmos para o cálculo dos intervalos de confiança destas prevalências corrigidas, sendo que os interessados neste assunto podem consultar o livro clássico de Cochran⁷. Um estudo de perdas consiste em fazer um esforço adicional ao final da fase de coleta de dados para obter as observações desejadas nos indivíduos perdidos. A melhor opção é realizar este trabalho de recuperação em uma amostra aleatória das perdas, para concentrar os recursos disponíveis e evitar novas distorções nos resultados finais, causadas por vieses de seleção dos perdidos na primeira fase do estudo. Um estudo de perdas pode se desdobrar em fases, de tal modo que em cada uma se coletam informações de uma amostra das perdas da fase anterior.

O trabalho de supervisão desenvolvido no campo, isto é, no ambiente em que se encontram os indivíduos participantes do estudo, também irá permitir o monitoramento constante da qualidade das medidas realizadas. O confronto entre as medidas obtidas pelo pesquisador de campo e o supervisor poderá ser analisado do ponto de vista da confiabilidade, isto é, utilizando índices tais como o kappa, já visto antes.

Em adição, existe outro trabalho de supervisão que pode ser feito no escritório central da pesquisa. Este trabalho também consiste em comparações, porém pode prescindir da necessidade de replicação de medidas durante a coleta de dados. As comparações podem ser feitas entre observado e esperado, de acordo com algum conhecimento prévio ou expectativa lógica. Este tipo de controle é também chamado de controle estatístico de qualidade.

Informações recentes sobre a população de estudo podem servir para comparar com o que é obtido durante a coleta pelos pesquisadores de campo. Assim, por hipótese, se o último censo realizado pelo IBGE¹⁴ revelou que no bairro em que será coletada a amostra do estudo 60% dos adultos eram mulheres, espera-se proporções semelhantes de entrevistas com mulheres por cada um dos pesquisadores de campo. Desvios poderão ser tolerados de acordo com testes de hipóteses específicos, tais como os que se baseiam na distribuição da estatística qui-quadrado³³. Entretanto, alguns desvios

consistentes, do observado em relação ao esperado, podem ser admitidos. A consistência pode se dever ao fato de que todos os pesquisadores de campo apontem o mesmo desvio, em sentido e magnitude. Isto pode ser fruto de mudanças reais ocorridas na população ou devido aos instrumentos e regras utilizados no estudo. Se as médias de peso no estudo atual diferem daquelas obtidas em estudo semelhante, feito recentemente na mesma população, isto pode se dever a balanças diferentes ou métodos de pesagem diversos empregados nas duas pesquisas.

As comparações podem ser feitas também entre os valores observados pelos vários pesquisadores de campo no mesmo estudo. Entretanto, deve-se lembrar que nestas comparações a hipótese nula, de homogeneidade de resultados esperados entre os observadores, só faz sentido se a alocação de indivíduos entre os pesquisadores de campo tiver sido aleatória. Se a cada pesquisador de campo for designado um grupo, com atributos distintos dos demais, não é possível afastar, por razão de lógica, a hipótese de que eventuais diferenças entre os resultados obtidos pelos observadores se devam a diferenças reais existentes entre os grupos de indivíduos. Este aspecto reforça a recomendação de que o ideal é a distribuição aleatória de indivíduos pelos observadores, ainda que os custos sejam mais elevados.

Outra vigilância que se faz no controle estatístico de qualidade é sobre os valores absurdos ou extremos. Valores absurdos são aqueles que estão aquém ou além dos limites de registro possíveis com os instrumentos utilizados ou pela própria natureza dos fenômenos biológicos. Valores extremos são aqueles muito baixos ou muito altos, que apesar de possíveis são incomuns e, portanto, devem ser avaliados com cuidado porque podem representar corretamente as observações, não podendo ser descartados, *a priori*, sem verificações. Estas verificações podem consistir no confronto com outras medidas relacionadas àquela em que se verificou o valor suspeito ou na replicação da medida. Por exemplo, seria altamente improvável, porém possível, que a altura de um menino de 15 anos pudesse ser de 1,90 m.

Por fim, é possível ainda exercer outra modalidade de controle estatístico de qualidade sobre medidas feitas em escala contínua que não depende do conhecimento prévio da distribuição da característica investigada, e nem de seus limites, que é a chamada análise dos dígitos terminais. O dígito terminal é o valor da unidade em cada medição. Por exemplo, em uma medida de pressão arterial sistólica cujo valor registrado em milímetros tenha sido 120, o dígito terminal é zero. Quando a amplitude

total dos valores permitidos pela escala de leitura do instrumento utilizado é razoavelmente ampla, como no caso das medidas de peso, altura ou pressão arterial, podemos supor que não há nenhuma razão lógica para que determinado dígito terminal dos valores registrados seja mais freqüente do que os demais. Portanto, a hipótese nula do teste estatístico empregado na análise dos dígitos terminais é a de que as freqüências relativas observadas dos dígitos terminais são iguais, quer dizer, a distribuição destas freqüências é homogênea. Para isso, um teste bastante simples é aquele que confronta as freqüências absolutas observadas com as esperadas na hipótese de homogeneidade, também chamado de qui-quadrado (χ^2) por comparar seu resultado com o valor daquela distribuição. Maiores detalhes deste tipo de teste podem ser vistos em livros de estatística como o de Noether. Assim, por exemplo, se num inquérito sobre hipertensão arterial os pesquisadores utilizaram um esfigmomanômetro de coluna de mercúrio que possui traços que indicam todos os valores pares, espera-se que em uma série de valores observados, cada dígito terminal (0, 2, 4, 6 ou 8) tenha ocorrido em 20% das medidas. O teste estatístico de χ^2 poderá revelar qual é a probabilidade do desvio da distribuição observada em relação à distribuição esperada, homogênea, de 20% para cada dígito terminal, não se dever ao acaso.

Um dos achados mais comuns em distribuições de dígitos terminais que se desviam significativamente da distribuição homogênea é a concentração dos mesmos em torno dos valores terminados em 0 ou 5, quando este último também é permitido. Uma das explicações mais comuns para estas ocorrências é que nas escalas de leitura dos instrumentos convencionais são gravados apenas os valores correspondentes àqueles terminados em 0 ou 5, enquanto os demais valores são assinalados apenas por traços, o que induz os observadores a “aproximarem” a maioria das leituras para valores terminados em 0 ou 5. Instrumentos com escalas analógicas, de comparação, apresentam este problema. Por isto, deve-se dar preferência a instrumentos com painel de leitura digital que apresentam todos os valores possíveis, sem dar mais destaque aos terminados em 0 ou 5.

Além disto, não são apenas as formas de apresentação das escalas de leitura dos instrumentos que levam à concentração das medidas naquelas terminadas em 0 ou 5. Avaliações preconcebidas dos indivíduos, por parte dos observadores, assim como hábitos culturais de registros mais valorizados de efemérides relativas à vida das pessoas, também reforçam as preferências pelos dígitos terminais 0 ou 5. Assim, por parecerem mais importantes os aniversários múltiplos de 10 é que

se pode verificar que as idades, em anos completos, mais freqüentemente declaradas pelos entrevistados em estudos seccionais são aquelas terminadas em 0.

Um dos maiores problemas causados pela preferência por determinados dígitos terminais está relacionado com a redução escalar, que consiste no registro de valores que, apesar de poderem ser feitos em uma determinada escala de medida, são aproximados para outros em uma escala com menor número de categorias. Um exemplo de redução escalar ocorre quando o observador, ao utilizar um esfigmomanômetro de coluna de mercúrio, que permite registro de valores em milímetros de mercúrio (mmHg), ainda que em intervalos de 2 milímetros, faz todos os registros em centímetros de mercúrio (cmHg), por ter aproximado todos os valores lidos para aqueles terminados em 0.

O efeito de uma redução escalar pode se refletir tanto sobre a precisão de uma medida, que é relacionada inversamente com seu desvio-padrão (a raiz quadrada da variância), como sobre os indicadores de tendência central da distribuição da medida, tais como a média ou a proporção. Média e proporção são viesados pela redução escalar, enquanto a precisão costuma ser diminuída por este tipo de redução. Os dados do estudo seccional sobre hipertensão arterial em adultos da Ilha do Governador podem ser utilizados para exemplificar a comparação de resultados segundo dois métodos de registro de leituras de pressão arterial. Naquela pesquisa foi utilizado um esfigmomanômetro de coluna de mercúrio que permitia leituras com discriminações de até 2 mmHg, o menor intervalo entre dois traços da escala do aparelho. Na Tabela 9.3 podem ser observados as médias e os desvios-padrão das pressões sistólica e diastólica, de acordo com dois métodos de registro de valores individuais, o original em mmHg, e o reduzido em cmHg × 10, pela aproximação para os valores terminados em 0, obtidas naquele estudo, já levando em consideração seu delineamento amostral complexo.

Na Tabela 10.3 pode-se observar os efeitos mais comuns da redução escalar: as médias estão viesadas e ocorre perda de precisão, ainda que de pequena monta. Entretanto, os efeitos mais drásticos da redução escalar podem ser percebidos quando se utilizam pontos de corte em escalas originalmente contínuas, ou discretas com muitas categorias, para classificar os indivíduos em apenas duas categorias. Isto pode ser exemplificado com os dados do estudo citado quando se apresentam as prevalências de hipertensão segundo o critério clínico que classifica como hipertensos aqueles com sistólica maior ou igual a 140 mmHg, ou diastólica maior ou igual a 90 mmHg. Levando-se em conta o delineamento amostral

complexo, que levou ao exame de 1.270 indivíduos, a prevalência de hipertensão foi de 33,86% nos adultos da Ilha do Governador, utilizando-se toda a discriminação permitida pelos esfigmomanômetros utilizados, de 2 mmHg. Porém, se o critério de registro tivesse sido mais frouxo, com aproximações para os valores terminados em 0, aquela prevalência teria sido estimada em 46,31%. A redução escalar teria implicado, portanto, em uma “inflação” de cerca de 37% no número de hipertensos adultos estimados para a Ilha do Governador. Estas observações podem servir para uma reflexão das práticas mais comuns utilizadas tanto na clínica como em estudos de escasso rigor metodológico.

Tabela 10.3
Médias e desvios-padrão das Pressões Sistólica e Diastólica, em Adultos da Ilha do Governador, de Acordo com Dois Métodos de Registro das Leituras

Método de Registro das Leituras de Pressão	Sistólica		Diastólica	
	Média	Desvio-padrão	Média	Desvio-padrão
Original (mmHg)	131,64	21,53	82,01	11,76
Reduzido (cmHg x 10)	131,71	21,79	82,20	12,08

Fonte dos dados: Klein et al.²³

Análise e Divulgação de Resultados

Nos estudos seccionais, a análise costuma exigir a construção de bancos de dados que possam ser processados com computador. Durante a fase de coleta, pelo menos parte dos dados já foi analisada com a finalidade do exercício de controle estatístico de qualidade. O objetivo daquela análise era voltado para detectar problemas na produção dos pesquisadores de campo, porém nada impede que algumas informações básicas sobre o conjunto dos indivíduos também sejam processadas, durante o controle de qualidade, para fornecer informes parciais ou preliminares de resultados básicos. Entretanto, as análises definitivas deverão ser feitas em bancos de dados editados, isto é, conferidos. Um método de conferência simples consiste na chamada digitação dupla, isto é, na montagem de dois bancos de dados completos, com o mesmo conteúdo, por digitadores diferentes. Programas especiais, tais como um dos utilitários do *Epi-Info (Data Compare)*, podem ser usados para comparar os dois bancos e apontar as incoerências, que deverão ser corrigidas pela consulta aos documentos de coleta originais de “campo”. Este processo se repete até que não existam mais diferenças entre os bancos.

A seguir, vamos descrever sucintamente algumas estratégias de análise utilizadas com dados de estudos seccionais, de acordo com o número de variáveis envolvidas.

Análises Univariadas

As análises univariadas são aquelas em que apenas o comportamento de uma variável é objeto de atenção. De acordo com o que foi visto na introdução, este tipo de análise procura descrever a distribuição de cada variável na amostra examinada, visando à referência para a população. De uma maneira geral, o comportamento de uma distribuição pode ser analisado por meio de gráficos e de estatísticas. A escolha das representações gráficas e das estatísticas está relacionada com a escala de medida da variável, se discreta ou contínua.

As distribuições de variáveis discretas com pequeno número de categorias, geralmente até dez, podem ser apresentadas em gráficos de barras ou de setores, de freqüências absolutas ou relativas. As freqüências relativas, ou proporções, das categorias, podem ser apresentadas em tabelas, como estatísticas, com seus respectivos desvios-padrão. Nos estudos seccionais é muito comum que a variável de interesse principal seja discreta, com apenas duas categorias, tais como presença e ausência de doença. Este tipo de variável é também chamado de dicotômica ou binária, sendo os resultados apresentados em função de uma das categorias, por meio de proporções que, no caso de um estudo seccional, são as prevalências. Cada prevalência pode ser apresentada acompanhada de seu erro-padrão, que corresponde ao desvio-padrão de sua distribuição amostral, necessário para estimar os limites de um intervalo no qual se espera que esteja contido o verdadeiro valor da prevalência na população, com um determinado grau de confiança, freqüentemente de 95%. Exemplificando: no estudo seccional da Ilha do Governador estimou-se que 24,9% dos adultos eram hipertensos²⁴. Levando em consideração o delineamento complexo deste estudo, o cálculo do erro-padrão desta prevalência resultou em 1,6%, portanto, esperava-se que a verdadeira prevalência de hipertensão arterial daquela população adulta estivesse, com uma probabilidade de 95%, entre os valores 21,7 e 28,1% (ver capítulo 24 – Associação Estatística em Epidemiologia: Análise Bivariada). O critério utilizado neste estudo para classificar um indivíduo como hipertenso foi misto: sua pressão sistólica deveria ser igual ou maior do que 160 mmHg ou sua diastólica deveria ser igual ou superior a 95 mmHg, ou ainda se estivesse em tratamento específico para hipertensão arterial.

As variáveis discretas com muitas categorias e as contínuas podem ter suas distribuições apresentadas com gráficos como histogramas e polígonos de frequência. As estatísticas vão depender ainda de outras características destas variáveis. Por exemplo, para uma variável contínua poderá ser apresentada a média, e o desvio-padrão ou o erro-padrão. Para variáveis discretas ordinais poderão ser mais adequadas estatísticas como a mediana, e para discretas nominais, a moda.

Estas análises são as mais simples e devem obrigatoriamente preceder aquelas mais complexas que envolvam duas ou mais variáveis. Maiores detalhes podem ser vistos no Capítulo 18 – Análise Exploratória de Dados.

Análises Bivariadas

Análises bivariadas são aquelas que se ocupam das relações entre duas variáveis ou das distribuições conjuntas de duas variáveis. Um exemplo de relacionamento ou associação entre duas variáveis dicotômicas, uma chamada de exposição e outra de doença, foi disposto simbolicamente na Tabela 10.1. Ao discutir os usos e as limitações dos estudos seccionais, chamamos a atenção de que a análise de associação poderia ser feita pela comparação entre as prevalências de doença em expostos e não-expostos. Tradicionalmente, para testar a associação estatística, isto é, de frequências, entre duas variáveis dicotômicas ou mesmo categóricas, podem ser utilizados testes de hipóteses baseados na estatística de qui-quadrado (χ^2).

Também podem ser criados indicadores resumidos daquela relação por meio da diferença ou da razão entre as prevalências. Assim, podem ser testadas as hipóteses de que a diferença entre as prevalências é igual a 0 ou de que a razão entre as prevalências é igual a 1. Qualquer uma destas hipóteses corresponde à de que não há associação estatística entre as variáveis exposição e doença. De qualquer modo, é sempre importante lembrar que a razão de prevalências não é necessariamente um estimador adequado de uma razão de riscos, sejam estes de incidências cumulativas ou de densidades de incidência, como se procura obter nos estudos longitudinais de coortes. Uma razão de riscos quantifica quantas vezes é maior o risco dos expostos em desenvolverem a doença quando comparados com os não-expostos, ao passo que uma razão de prevalências tão-somente estima quantas vezes mais os expostos estão doentes, quando comparados aos não-expostos, na época do estudo seccional. Os motivos para estas distinções foram apresentados na

discussão sobre as limitações dos estudos seccionais. A razão de prevalências (RP) para os dados simbólicos da Tabela 10.1 é obtida com:

$$RP = \frac{\frac{a}{c}}{\frac{c+d}{c}} = \frac{a \cdot (c+d)}{c \cdot (a+b)}$$

Para exemplificar, vamos utilizar os dados de um estudo sobre o consumo de psicofármacos na Ilha do Governador, feito em 1994. A Tabela 10.4 relaciona o consumo de psicofármacos com o gênero dos indivíduos entrevistados.

A prevalência de consumo de psicofármacos entre as mulheres foi igual a 6,7% ($= 58/871 \times 100$), enquanto que entre os homens foi de 3,1% ($= 18/588 \times 100$). A razão de prevalências foi de 2,18 ($= (58/871)/(18/588)$). Os limites de um intervalo de confiança de 95% desta razão de prevalências são 1,30 e 3,65, portanto, é altamente improvável que o verdadeiro valor desta razão, na população, seja igual a 1, o que significaria a não-existência de associação entre gênero e consumo de psicofármacos. Um teste alternativo para a hipótese nula, de que esta razão de prevalência é igual a 1 na população revelou que o valor de χ^2 calculado para a tabela é de 9,20 que está associado com uma probabilidade de ocorrência de 0,2% (também chamada de p-valor). Portanto, se o critério de rejeição daquela hipótese nula havia sido fixado previamente num máximo de 5%, a probabilidade de erro nesta rejeição é muito baixa.

Contudo, não é possível inferir diretamente que o risco de uma mulher vir a fazer uso de psicofármacos é 2,18 vezes maior do que o de um homem. Isto porque, como os dados foram coletados em um estudo seccional, não se obtiveram informações sobre diferenciais

Tabela 10.4
Frequências de Consumo de Psicofármacos Segundo o Gênero, Obtidas em Estudo Seccional na Ilha do Governador, em 1994

Gênero	Consumo de Psicofármacos		Total
	Sim	Não	
Mulheres	58	813	871
Homens	18	570	588
Total	76	1.383	1.459

Fonte dos dados: Almeida e Coutinho¹¹

Tabelas 10.5
Freqüências de Doença e Exposição para o Cálculo de Chances

Chances de Exposição			Chances de Doença		
	Doentes	Não-doentes		Doentes	Não-doentes
Expostos	a	b	Expostos	a	b
Não-expostos	c	d	Não-expostos	c	d

de sobrevivência em coortes de mulheres e homens que fazem uso de psicofármacos. Sempre é possível supor que os homens usuários de psicofármacos sejam retirados da população em idades mais precoces do que as mulheres, fazendo com que seus períodos de tempo de uso de psicofármacos sejam mais curtos do que os das mulheres, reduzindo suas probabilidades de serem detectados como usuários em um momento qualquer em que se realiza um estudo seccional.

Outro indicador resumido da relação entre variáveis dicotômicas depende das chamadas chances (*odds*) de exposição entre doentes e não-doentes, ou de doentes entre expostos e não-expostos. Uma chance de exposição é obtida pela razão entre os expostos e os não-expostos, ao passo que uma chance de doença é obtida pela razão entre doentes e não-doentes. Utilizando os símbolos da Tabela 10.5, podemos calcular a chance de exposição entre os doentes com a/c e a chance de exposição entre os não-doentes com b/d . Da mesma maneira, a chance de doença entre os expostos é calculada com a/b e a chance de doença entre os não-expostos, com c/d . As razões entre os pares de chances fornece um outro indicador singular resumido, da relação entre as variáveis da tabela, chamado de razão de chances (*odds ratio*).

Com os dados de um estudo seccional, como os simbolizados na Tabela 10.5, este indicador é chamado de razão de chances prevalentes (RCP), e pode ser calculado alternativamente com:

$$RCP = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{\frac{a}{c}}{\frac{b}{c}} \cdot \frac{d}{c} = \frac{a \cdot d}{b \cdot c}$$

(a última alternativa de cálculo esclarece porque a razão de chances é também chamada de razão dos produtos cruzados).

Utilizando ainda como exemplo os dados do estudo sobre consumo de psicofármacos¹, dispostos na Tabela

10.4, vamos substituir “doença” pelo uso de psicofármacos. A chance de ser mulher entre os consumidores de psicofármacos era de 3,22 (= 58/18) e entre os não-consumidores era de 1,43 (= 813/570). Portanto, a razão de chances prevalentes foi de 2,26 (= (58/18)/(813/570)). Os limites de um intervalo de confiança de 95% desta razão de chances prevalentes são 1,32 e 3,85, portanto é altamente improvável que o verdadeiro valor desta razão, na população, seja igual a 1, o que significaria a não existência de associação entre gênero e consumo de psicofármacos. O mesmo teste visto antes para testar a hipótese nula, de que a razão de prevalências é igual a 1 na população, serve para testar a hipótese nula de que a razão de chances prevalentes é igual a 1, na população de onde provém a amostra.

A interpretação desta razão de chances prevalentes é mais complexa do que a da razão de prevalências, que deve ser a preferida para expressão de resultados de estudos seccionais, quando for desejada uma razão. O uso tão frequente da razão de chances se deve ao fato de que costuma se constituir na única expressão de resultado permitida em estudos de casos e controles sem base populacional, isto é, que se utilizam de amostras de casos e controles cujas frações amostrais são desconhecidas e, além do mais, não-probabilísticas. Acresce ainda que vários algoritmos de cálculo relacionados com a razão de chances foram desenvolvidos e incorporados aos programas de computador que tratam bancos de dados coligidos em estudos epidemiológicos, principalmente, os de casos e controles.

Em estudos seccionais é possível estimar a razão de chances prevalentes, porém é preciso tomar cuidado com sua interpretação, pois ela não necessariamente estima de forma não viesada a razão de chances que seria obtida em um estudo que incluisse os casos incidentes, como geralmente ocorre em estudos de casos e controles. É possível que a doença afete o estado de exposição, de modo que uma parte dos casos incidentes expostos migre para a categoria dos não-expostos, distorcendo a chance prevalente de exposição entre os doentes

e consequentemente a razão de chances prevalente. Exemplificando, os indivíduos que tiveram infarto do miocárdio são aconselhados a parar de fumar, de modo que se a exposição considerada é o tabagismo, em um determinado momento, em que se realiza um estudo seccional, aqueles que já se tornaram casos de infarto podem ter uma chance menor de exposição ao fumo do que a que seria obtida pelo registro da exposição próxima ou anterior ao evento da doença.

Entretanto, a razão de chances prevalentes, obtida em um estudo seccional, pode ser um estimador razoável da razão de densidades de incidência, obtida em estudos de coortes, se a duração média da doença é semelhante em expostos e não-expostos e o número de casos novos e terminados se equivale em períodos considerados de tempo, em uma população estável²⁵. A razão de densidades de incidência é um bom indicador para testar associações etiológicas de doenças com fatores de risco que permanecem atuando por períodos de tempo muito prolongados. Maiores detalhes a respeito destes indicadores podem ser vistos no Capítulo 11 – Estudos de Casos e Controles e no Capítulo 12 – Estudos de Coortes.

A associação entre variáveis contínuas pode ser analisada com indicadores tais como coeficientes de correlação e de regressão. Assim, por exemplo, em um estudo seccional sobre hipertensão arterial, verificou-se que a correlação linear entre a pressão sistólica e a circunferência de braço era igual 0,2486, o que significa que cerca de 6,18% ($(0,2486)^2 \times 100$) da variação da sistólica eram explicados pela variação da circunferência de braço. Além disso, o coeficiente de regressão da sistólica sobre a circunferência de braço foi estimado em 0,15. Como a circunferência de braço havia sido medida em milímetros e a pressão sistólica em mmHg, aquele valor da regressão pode ser interpretado como a expressão de que o aumento de 1 centímetro na circunferência de braço era acompanhado do aumento de 1,5 mmHg na pressão sistólica.

Análises Multivariadas

As análises em que mais de duas variáveis estão envolvidas são chamadas de multivariadas. Existem várias técnicas de análises multivariadas, adequadas segundo as estratégias e a escala de medida das variáveis envolvidas. Detalhes sobre algumas destas técnicas podem ser vistos em outros capítulos deste livro, como o Capítulo 25 – Associação Estatística em Epidemiologia: Análise com Múltiplas Variáveis. Aqui vamos nos

ocupar apenas de algumas, apropriadas para dados de estudos seccionais, que são também utilizadas na análise de estudos de casos e controles, e de coortes com tempos de acompanhamento semelhantes, procurando ressaltar suas peculiaridades. Também vamos nos limitar àquelas análises que têm por estratégia elucidar a relação entre variáveis, em que uma delas é chamada de dependente e as demais, de independentes. Supõe-se que as variações da dependente são causadas ou influenciadas pelas variações das demais, as independentes. Nos estudos seccionais, a variável dependente costuma ser o agravo ou doença, e as independentes, os fatores de exposição. Entretanto, existem técnicas de análise multivariada que suportam mais de uma variável dependente e outras que até mesmo prescindem da pressuposição da existência de uma variável dependente, sendo todas tratadas inicialmente como independentes.

É também muito freqüente que numa análise multivariada de dados de um estudo seccional, o interesse principal esteja na elucidação da associação de um agravo e de um fator de exposição específico, em que é necessário levar em consideração a influência exercida pelas demais variáveis independentes ou fatores de exposição sobre o agravo e o fator de exposição escolhido. Estas influências podem se traduzir por confusão na relação de interesse principal ou em interação. Maiores detalhes sobre confusão podem ser vistos no capítulo 15 – Validade em Estudos Epidemiológicos, mas, por enquanto, o que se pode dizer é que se trata da alteração causada por outra variável na percepção do efeito da variável de exposição sobre a dependente, o agravo ou doença, e que é constante ao longo de toda a faixa de variação desta terceira variável, chamada de variável de confusão. Numa relação entre duas variáveis podem existir outras de confusão. A interação se refere à modificação de efeito produzida por uma terceira variável no efeito percebido da variável de exposição sobre a dependente, na qual este efeito varia ao longo da faixa de variação da terceira variável, a modificadora¹⁶.

As técnicas multivariadas adequadas para elucidar uma associação de interesse principal entre um fator de exposição e um agravo, que pode estar sendo confundida ou modificada por outras variáveis, podem fazer uso de duas estratégias básicas de análise de dados: a estratificação ou a modelagem matemática.

ANÁLISE MULTIVARIADA POR ESTRATIFICAÇÃO

A estratificação consiste na análise da relação de interesse principal nas categorias de outras variáveis.

Os estratos serão formados pelas categorias de outras variáveis.

Para exemplificar uma análise multivariada por estratificação vamos utilizar os dados do estudo sobre hipertensão arterial na Ilha do Governador. Naquele estudo, a amostra foi planejada para ser obtida em três grupos de renda média dos domicílios, detectada algum tempo antes, durante a realização de um recenseamento na área. Isto porque um dos objetivos principais daquele estudo era saber se a prevalência de hipertensão arterial (HA) estava associada com a renda média domiciliar (RMD). A RMD foi classificada em baixa, média e alta, mas, para simplificar, vamos apresentar análises que compararam o grupo de RMD baixa ao conjunto formado pelos grupos de RMD média e alta. E, como é sabido que a HA está associada com o gênero e a idade dos indivíduos, para analisar a relação entre RMD e HA é preciso considerar as distribuições de gênero e idade nos grupos de RMD. A Tabela 10.6 apresenta os resultados de uma estratégia de estratificação, de maneira resumida. Na célula inferior, à direita, que corresponde ao total geral, pode-se observar que a prevalência de HA na RMD baixa era mais elevada do que na RMD média + alta, porém a probabilidade de se rejeitar falsamente a hipótese nula, de que a diferença entre as prevalências é igual a zero na população de onde provêm os dados, é elevada, maior do que 10% (p-valor). Portanto, o resultado global indicaria que, aparentemente, a prevalência de HA não está associada com RMD.

Ao estratificar a relação entre HA e RMD, segundo dois estratos de idade (de menores de 55 anos, e de 55 anos ou mais) nos dois gêneros, percebe-se, nas células

Tabela 10.6 Prevalências de Hipertensão Arterial (%) em Adultos da Ilha do Governador, Segundo Renda Média Domiciliar (RMD), o Gênero e Dois Estratos de Idade, Corrigidas de Acordo com o Delineamento Amostral					
Gênero	Renda Média Domiciliar (RMD)	Grupo Etário		Total	
		< 55 Anos	≥ 55 Anos		
Mulheres	Baixa	20,78	72,97	31,30	ns
	Média + alta	13,98	49,57	24,05	
Homens	Baixa	14,55	38,19	20,16	ns
	Média + alta	20,72	32,61	24,07	
Total	Baixa	18,14	56,43	26,45	ns
	Média + alta	17,01	41,98	24,06	

ns: p-valor > 10%; *: 10% > p-valor > 5%

Fonte dos dados: Klein et al., 1995

internas em destaque na Tabela 10.6, que aquela associação só foi detectada nas mulheres. Nestas, os p-valores dos testes são menores do que 10%, ainda que maiores do que 5%, o que significa que as hipóteses de igualdade das prevalências de HA segundo RMD, nos dois grupos etários das mulheres, só poderiam ser rejeitadas caso o critério de rejeição fosse de 10%. Nos homens, as diferenças entre as prevalências de HA nos grupos de RMD não são estatisticamente significativas, portanto as hipóteses nulas também não podem ser rejeitadas, nem com um critério mais liberal de 10%.

A comparação do sentido e da magnitude das diferenças de prevalências nos homens e nas mulheres, de acordo com os dois grupos de idade, permite também levantar a suspeita de que a associação entre HA e RMD pode estar sendo modificada pela idade, e de forma diferente nos dois gêneros. É possível testar hipóteses sobre modificação de efeito, ou seja, de interação, por meio da estratificação e estatísticas adequadas. Entretanto, para poder estratificar relações segundo uma variável contínua, como é o caso da idade, só é possível criar um número limitado de categorias que compreendem faixas de valores desta variável. No exemplo, os adultos tinham idades que variavam de 20 a mais de 80 anos, porém na análise por estratificação que apresentamos só foram criados dois estratos, com um ponto de corte situado aos 55 anos. Esta redução escalar da idade redonda em perda da sensibilidade da detecção da modificação de efeito provocada por esta variável na relação entre HA e RMD, em cada um dos gêneros, como se verá a seguir.

ANÁLISE MULTIVARIADA POR MODELAGEM

A modelagem matemática consiste em submeter o conjunto dos dados a algoritmos de cálculo que estimam os chamados coeficientes de regressão, de acordo com modelos lineares ou não-lineares. O modelo não-linear mais popular é o logístico, de largo emprego em estudos de casos e controles, mas que também pode ser usado com os dados de estudos seccionais ou de coortes com tempos de acompanhamento semelhantes, apesar de ser pouco adequado nestes estudos para avaliar a interação.

Um dos modelos matemáticos de análise multivariada é conhecido como regressão linear múltipla, que pode ser descrita simbolicamente com a equação:

$$y = C + \sum_{i=1}^k (\beta_i \cdot x_i)$$

em que y representa a variável dependente, e x_i a i -ésima variável independente, de um conjunto de k variáveis.

Os termos C , uma constante, e β_i , os coeficientes de regressão de um conjunto de k coeficientes, são calculados pelo algoritmo do modelo linear múltiplo. O valor de C estima o valor de y quando todas as variáveis independentes têm valor igual a zero, e cada β_i representa o incremento provocado em y pelo aumento de uma unidade de x_i . Os incrementos podem ser positivos ou negativos. Tradicionalmente, este modelo linear múltiplo é utilizado para estimar variáveis dependentes contínuas, podendo as independentes ser contínuas, categóricas ou binárias. Entretanto, com certas restrições, é possível também construir modelos em que a variável dependente, y , é binária, sendo que nestes casos se os valores das categorias de y são iguais a 1 e 0, então cada β_i representa o incremento proporcional da categoria representada por 1 de y , provocado por x_i . O valor de y estimado pela substituição dos valores dos coeficientes e dos valores das variáveis, corresponde à proporção esperada da categoria 1 de y . Estes modelos servem também para controlar confusão e, especialmente para nosso propósito aqui, que é testar a existência de interação, ou modificação de efeito. As interações são testadas pela inclusão no modelo de termos adicionais que consistem em β_i multiplicados pelo produto das variáveis isoladas que formam a interação.

Retornando ao exemplo dos dados obtidos no estudo sobre hipertensão arterial na Ilha do Governador, vamos construir dois modelos lineares múltiplos para estimar as prevalências de HA, para qualquer combinação de RMD e idade, e para testar a existência de interação entre RMD e idade em cada um dos gêneros. Os modelos serão os seguintes, um para as mulheres e outro para os homens:

$$p_Hipertensão = C + \beta_1 \cdot RMD + \beta_2 \cdot Idade + \beta_3 \cdot RMD \cdot Idade$$

Neste modelo, o último termo, β_3 tem por finalidade testar a existência de interação ou modificação de efeito. A hipótese nula é a de que o valor deste termo é igual a zero, que só será rejeitada se o p-valor associado com a estimativa de β_3 for menor do que um determinado critério de rejeição, que pode ser arbitrado em 5%. Na Tabela 10.7 estão os resultados da aplicação de um algoritmo³⁷ que obtém as estimativas dos coeficientes e seus p-valores associados. Quando a variável dependente é binária, como a de nosso exemplo, é importante escolher algoritmos adequados para a estimativa dos coeficientes de uma equação de regressão linear, especialmente de seus erros amostrais. Estes algoritmos devem ser do tipo “robusto”, ou indicar que a distribuição da variá-

vel dependente é binomial. Somente alguns programas dedicados a análises estatísticas são capazes de permitir a escolha dos algoritmos adequados, o que não é o caso de pacotes populares como o Epi-Info.

Na Tabela 10.7 observamos que só podemos rejeitar a hipótese nula ($\beta_3 = 0$) no modelo das mulheres, para um critério de rejeição de p-valor menor do que 5%. Portanto, podemos assumir que a idade é modificadora do efeito da RMD sobre a HA nas mulheres, mas não nos homens. A Tabela 10.6 já nos fornece pistas sobre os tipos de modificação de efeito provocada pela idade que poderiam estar ocorrendo em mulheres e homens, porém isto fica mais claro nas Figuras 10.3 e 10.4.

Tabela 10.7 Coeficientes de Regressão Linear de Hipertensão Arterial, Segundo Renda Média Domiciliar (RMD), Idade e Interação (RMD x Idade), de Acordo com Dois Modelos Independentes, para Mulheres e Homens (Ilha do Governador)				
Gênero Variável	Mulheres		Homens	
	Coef.	p-valor (%)	Coef.	p-valor (%)
RMD*	-0,06486	34,3	-0,11457	19,7
Idade	0,01276	0,0	0,00565	0,0
RMD x Idade	0,00426	1,4	0,00211	36,1
Constante	-0,33014	0,0	-0,00511	92,8

*Grupo de renda média baixa = 1; Grupo de renda média + alta = 0

Fonte dos dados: Klein et al.²³.

As Figuras 10.3 e 10.4 mostram, graficamente, as prevalências que seriam esperadas, de acordo com modelos lineares, nos dois grupos de RMD ao longo das idades observadas no estudo, em homens e mulheres, respectivamente. Assim como se havia percebido na Tabela 10.6, os homens do grupo de RMD baixa tinham prevalências menores do que os de RMD média + alta, até por volta dos 55 anos, invertendo-se a relação nos mais velhos. Este tipo de modificação de efeito é chamado de interação qualitativa, em que há inversão da relação entre agravo e fator de exposição de interesse segundo as categorias da variável modificadora. Porém, é preciso ressaltar que os resultados indicam que esta interação não é estatisticamente significativa nos homens e, portanto, podemos concluir que sua existência na população de origem dos dados é bastante improvável. Por isto, para os homens, com a finalidade de testar a associação entre HA e RMD, podemos construir, em seqüência, um modelo mais simples, que exclui o termo da interação, que testa se há diferença entre as estimativas médias de HA segundo grupos de RMD, uma vez

que o β_i correspondente representa esta diferença. Se este modelo demonstrar que também esta diferença não é significativa, podemos encerrar a análise com o modelo mais simples, que só inclui a idade como variável independente. Efetivamente, no caso exemplificado, foi este último modelo que demonstrou que somente a idade estava associada com a HA entre os homens.

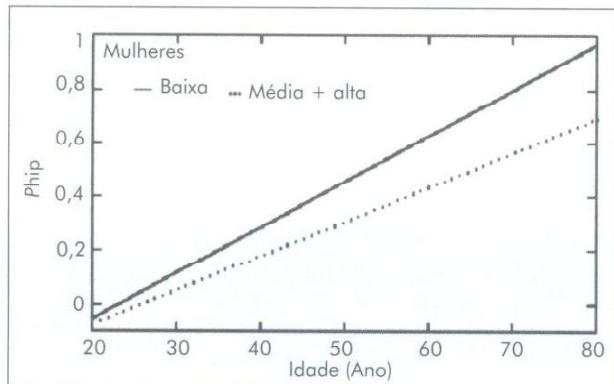


Fig. 10.3 — Prevalências estimadas de hipertensão arterial em mulheres em dois grupos de renda média domiciliar, segundo a idade, de acordo com um modelo de regressão linear (Ilha do Governador).

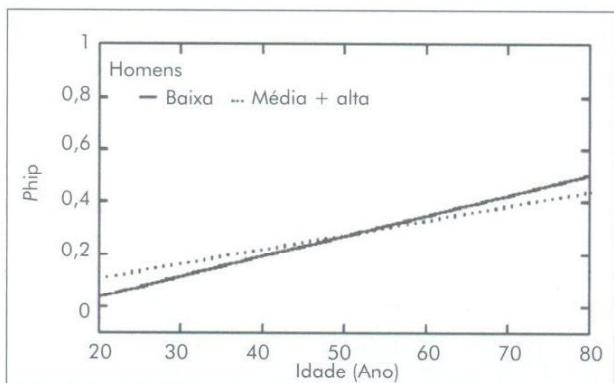


Fig. 10.4 — Prevalências estimadas de hipertensão arterial em homens em dois grupos de renda média domiciliar, segundo a idade, de acordo com um modelo de regressão linear (Ilha do Governador).

A idade também não causa confusão na relação entre RMD e HA nos homens, porque não foi possível rejeitar a semelhança do coeficiente de regressão correspondente a RMD com 0 (zero) no modelo intermediário, que inclui apenas os termos isolados, RMD e idade, sem aquele que representa a interação. Isto significa que não há diferença estatisticamente significativa entre as prevalências ajustadas de HA, segundo RMD nos homens, controlado o efeito da idade por regressão, e, como se observou na Tabela 10.6, a diferença entre as prevalências brutas, não

ajustadas, de HA, segundo RMD nos homens, também é estatisticamente semelhante a 0.

Na Tabela 10.7 também verificamos que podemos aceitar a existência de interação entre RMD e idade nas mulheres (p -valor < 5%). Entre estas, a modificação de efeito provocada pela idade é do tipo interação quantitativa, pois a relação entre HA e RMD tem sua magnitude modificada ao longo do espectro da idade, ainda que não haja alteração de sentido. Portanto, o modelo final para as mulheres é aquele que contém o termo de interação, e que não deve ser reduzido para modelos com menor número de variáveis. De qualquer maneira, tanto para homens como para mulheres, o melhor modelo para estimar as prevalências de HA é o chamado modelo completo, aquele que contém todas as variáveis independentes consideradas e os termos de interação correspondentes. Para exemplificar o uso preditivo da regressão linear para as mulheres, podemos calcular a prevalência de hipertensão arterial esperada de uma mulher, de RMD baixa, de 60 anos de idade, por meio de:

$$Hipertensão = -0,33014 - 0,06486 \cdot 1 +$$

$$0,01276 \cdot 60 + 0,00426 \cdot 1 \cdot 60 = 0,6262 \text{ ou } 62,62\%$$

Entretanto, é neste aspecto que podemos verificar, nas mulheres, que o modelo linear múltiplo para uma variável dependente binária pode apresentar problemas sérios. Na Figura 10.3 é possível verificar que as estimativas de prevalência de HA nas mulheres de RMD média + alta são negativas abaixo de 27 anos de idade, de forma aproximada, o que é evidentemente absurdo. Uma prevalência, que tem as propriedades das proporções, só pode variar entre 0 e 1. Na verdade, é comum que um modelo linear múltiplo com variável dependente binária e variáveis independentes contínuas apresente este tipo de predição impossível próximo aos limites da variável contínua. Este problema só é evitado em modelos lineares em que as variáveis independentes também são binárias, e que são completos, isto é, possuem todos os termos possíveis de interação entre as independentes. De qualquer maneira, para resolver problemas de estimativas de prevalências, que só podem recuar entre os limites 0 e 1, é que se pode preferir um modelo multivariado não-linear, como o da regressão logística.

A expressão algébrica da regressão logística múltipla é a seguinte:

$$y = \frac{1}{1 + e^{-(C + \sum_{i=1}^k (\beta_i \cdot x_i))}}$$

sendo que nesta equação y é obrigatoriamente uma variável binária, enquanto que as independentes, X_i , podem ser binárias, de poucas categorias, ou contínuas. Portanto, em princípio, este modelo é adequado para estimativa de proporções ou prevalências. Estes modelos são muito utilizados em estudos de casos e controles para estimar também razões de chances (*odds ratios*) e testar hipóteses sobre estas. Aqui vamos tão-somente comparar seus resultados com aqueles obtidos pela regressão linear múltipla. Na Tabela 10.8 estão dispostos os coeficientes, cada um com seu p-valor associado para testar a hipótese nula de igualdade do coeficiente com 0 (zero), de duas equações de regressão logística, uma para as mulheres e outra para os homens.

Tabela 10.8
Coeficientes de Regressão Logística de Hipertensão Arterial, Segundo Renda Média Domiciliar (RMD), Idade e Interação (RMD x Idade), de acordo com Dois Modelos Independentes, para Mulheres e Homens (Ilha do Governador)

Gênero Variável	Mulheres		Homens	
	Coef.	p-valor (%)	Coef.	p-valor (%)
RMD*	-0,02489	97,3	-0,94061	14,7
IDADE	0,07965	0,0	0,03077	0,0
RMD x Idade	0,01836	21,7	0,01662	21,1
Constante	-5,09185	0,0	-2,54450	0,0

*Grupo de renda média baixa = 1; Grupo de renda média + alta = 0
Fonte dos dados: Klein et al., 1995.

As Figuras 10.5 e 10.6 apresentam as estimativas das prevalências de HA nos grupos de RMD, ao longo das idades, segundo dois modelos de regressão logística, um para as mulheres e outro para os homens. Nestas figuras se percebe que os limites 0 e 1 são respeitados, portanto, estes modelos de regressão são adequados para as previsões de prevalência, em todo o espectro de amplitude da variável contínua. Para exemplificar o uso preditivo da regressão logística nas mulheres, podemos calcular a prevalência de hipertensão arterial esperada de uma mulher, de RMD média + alta, de 30 anos de idade, por meio de:

$$P_{\text{Hipertensão}} = \frac{1}{1 + e^{-(5,09185 - 0,02489 - (0) + 0,07965 - (30) + 0,01836 - (0) - (30))}}$$

$$\equiv 0,649 \text{ ou } 6,49\%$$

(em que e é a base especial dos logaritmos naturais).

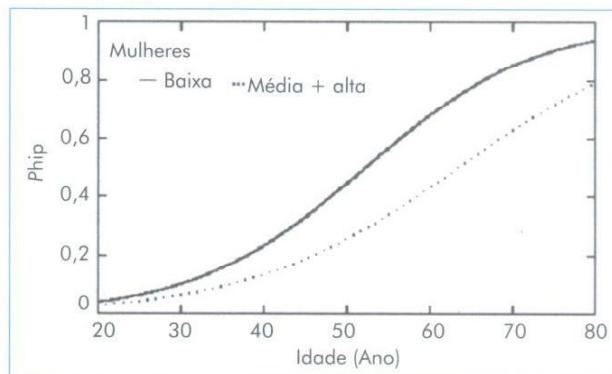


Fig. 10.5 — Prevalências estimadas de hipertensão arterial em mulheres, em dois grupos de renda média domiciliar, segundo a idade, de acordo com um modelo de regressão logística (Ilha do Governador).

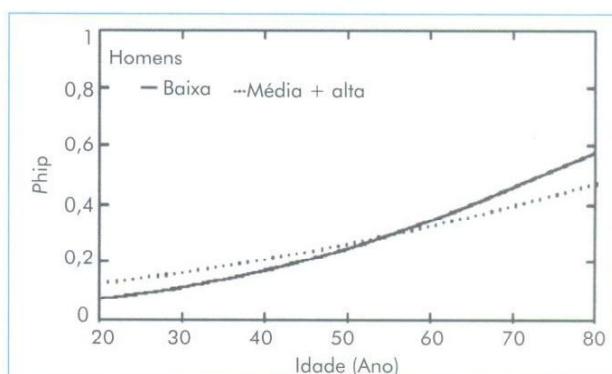


Fig. 10.6 — Prevalências estimadas de hipertensão arterial em homens em dois grupos de renda média domiciliar, segundo a idade, de acordo com um modelo de regressão logística (Ilha do Governador).

Entretanto, na Tabela 10.8 pode-se verificar que a existência de interação entre RMD e idade deve ser rejeitada não só para os homens como também para as mulheres, diferente do que aconteceu com a regressão linear, pelo menos com relação às mulheres. Nos modelos mais simples, que se seguem na análise excluindo o termo da interação, a RMD é reconhecida como fator associado com a HA entre as mulheres, mas não entre os homens. De qualquer maneira, o conflito entre as conclusões a respeito da existência de interação ocorre freqüentemente ao se confrontarem os resultados de regressões lineares e logísticas.

A regressão linear deve ser a preferida para testar hipóteses relacionadas com a detecção de interações relevantes para propósitos de prevenção da combinação de fatores de exposição, mesmo quando a variável dependente é binária, desde que se escolha o algoritmo adequado para variável dependente de distribuição bi-

nomial. A regressão linear múltipla é capaz de detectar a existência de interações segundo modelos aditivos, enquanto a regressão logística detecta interações segundo modelos multiplicativos, que são mais restritivos que os aditivos. A detecção de interação segundo um modelo aditivo com modelos logísticos exige manobras adicionais, não-triviais, e nem sempre disponíveis nos programas mais utilizados.

A existência de interações aditivas é utilizada para planejar a prevenção da concomitância de fatores de exposição. Por exemplo, se for detectada uma interação aditiva entre fatores tais como hipertensão arterial e uso de anticoncepcional oral em relação a uma consequência como acidentes vasculares cerebrais, ainda que a interação multiplicativa tenha sido rejeitada, a suspensão do uso de anticoncepcional terá um impacto maior entre as hipertensas do que entre as normotensas. Uma discussão mais detalhada sobre modelos de interação está além do escopo deste capítulo, mas pode ser vista na literatura.

A interpretação dos valores dos coeficientes das equações de regressão também varia segundo os modelos linear ou logístico. Na linear os valores dos coeficientes, sem termos de regressão, representam diretamente as diferenças entre as prevalências da categoria de cada variável com aquela que foi escolhida como referência. Já na logística, os coeficientes devem ser operados como expoentes da constante e (base especial dos logaritmos naturais) para se obter as razões de chance em modelos sem termos de interação. Para se obter diretamente as razões de prevalências podem ser utilizados outros modelos como regressões logarítmicas, de Poisson ou do risco proporcional de Cox. Entretanto, também estas regressões, em princípio, somente são adequadas para testar hipóteses sobre existência de interações segundo modelos multiplicativos.

Divulgação de Resultados

Os resultados de um estudo seccional devem ser divulgados nos meios especializados em que são relatados os achados dos demais estudos epidemiológicos. Os meios mais comuns são os periódicos não só da área de epidemiologia, mas também aqueles relacionados com o objeto do estudo, em geral uma doença, e seus correlatos. Assim, se o tema de investigação é a toxoplasmose em escolares, é natural que a publicação dos resultados se faça em revistas especializadas em zoonoses ou infecção, e também em pediatria.

Os resultados também podem ser publicados em teses e livros, de acordo com o planejamento inicial

e com os recursos dispostos pelos patrocinadores do estudo. Entretanto, é fundamental que os resultados de um estudo seccional sejam divulgados da forma mais ampla possível para a população de estudo. Afinal, se uma das características essenciais de um estudo seccional é a existência de uma população de referência, perfeitamente demarcada, é obrigatório que seus achados e conclusões sejam colocados à disposição de seus sujeitos. Ainda mais que, como quase sempre ocorre, as conclusões do estudo podem contribuir para o conhecimento da distribuição dos eventos mórbidos e seus fatores associados.

A divulgação dos resultados e conclusões relativos a agrupamentos de indivíduos pode ser feita pelos meios de comunicação de massa para o conjunto da população de referência e para os grupos específicos, acompanhada de recomendações a respeito da prevenção e resolução dos problemas detectados. Além disto, os resultados individuais de exames complementares, por exemplo, devem ser comunicados a cada um dos examinados pelos mesmos motivos.

BIBLIOGRAFIA

1. Almeida LM, Coutinho ESF. Consumo de Psicofármacos em uma Região Administrativa do Rio de Janeiro: A Ilha do Governador. Rio de Janeiro, Cadernos de Saúde Pública, 1994; 10(1):5-16.
2. Babbie ER. Métodos de Pesquisa de Survey. Belo Horizonte, Editora UFMG, 1999.
3. Barros AJD, Hirakata VN: Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Medical Research Methodology 2003; 3:21, (disponível em: www.biomedcentral.com/1471-2288/3/21).
4. Barros FC, Victora CG. Epidemiologia da Saúde Infantil. São Paulo, Hucitec-Unicef, 1991.
5. Bolfarine H, Bussab WO. Elementos de Amostragem. São Paulo, Editora Edgard Blücher, 2005.
6. Brogan D, Flagg EW, Deming M, Waldman R. Increasing the accuracy of the expanded programme on immunization's cluster survey design. Annals of Epidemiology 1994; 4:302-11.
7. Cochran WG. Sampling Techniques. New York, John Wiley & Sons, 1977.
8. Cochrane AL, Holland WW. Validation of screening procedures. British Medical Bulletin 1971; 27:3-8.
9. Dean AG, Arner TG, Sunki GG, Friedman R, Lantinga M, Sangam S *et al.* Epi Info™, a database and statistics program for public health professionals. Centers for Disease Control and Prevention, Atlanta, Georgia, USA, 2002.
10. Fleiss JL. Statistical methods for rates and proportions. 2 ed., New York, Wiley, 1981.
11. Frerichs RR. Simple analytic procedures for rapid microcomputer-assisted cluster surveys in developing countries. Public Health Reports 1989; 104(1):24-35.

12. Habicht JP. Estandarización de métodos epidemiológicos cuantitativos sobre el terreno. Boletín de la Oficina Sanitaria Panamericana 1974; 76:375-84.
13. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universality approach. Qual Life Res, 1998; 7: 323-35.
14. IBGE (Fundação Instituto Brasileiro de Geografia e Estatística). Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2000/metodologia/> (acessado em: março de 2006).
15. IEA European Questionnaire Group: Epidemiology Deserves Better Questionnaires. Disponível em: <http://www.dundee.ac.uk/iea/EuroQuests.htm> (acessado em: março de 2006).
16. Jewell NP: Statistics for Epidemiology. Chapman & Hall/CRC, 2004.
17. Kahn HA, Sempos CT. Statistical methods in epidemiology. New York, Oxford University Press, 1989.
18. Kalton G. Introduction to survey sampling. California, Sage University Paper 35, Sage Publications, 1985.
19. Kelsey JL, Whittemore AS, Evans AS, Thompson WD. Methods in observational epidemiology. 2 ed., New York, Oxford University Press, 1996.
20. Kish L. Statistical design for research. New York, John Wiley & Sons, 1987.
21. Kish L. Survey sampling. New York, John Wiley & Sons, 1965.
22. Klein CH, Costa EA. Os erros de classificação e os resultados de estudos epidemiológicos. Rio de Janeiro, Cadernos de Saúde Pública 1987; 3(3):236-49.
23. Klein CH, Silva NAS, Nogueira AR, Bloch KV, Campos LHS. Hipertensão arterial na Ilha do Governador, Rio de Janeiro, Brasil. I. Metodologia. Rio de Janeiro, Cadernos de Saúde Pública 1995; 11(2): 187-201.
24. Klein CH, Silva NAS, Nogueira AR, Bloch KV, Campos LHS. Hipertensão arterial na Ilha do Governador, Rio de Janeiro, Brasil. II. Prevalência. Rio de Janeiro, Cadernos de Saúde Pública 1995; 11(3): 389-94.
25. Kleinbaum D, Kupper L, Morgenstern H. Epidemiologic research. California, Lifetime Learning Publications, 1982.
26. Larousse. Grande Encyclopédia Larousse Cultural. São Paulo, Editora Nova Cultural, 1998.
27. Lemeshow S, Robinson D. Surveys to measure programme coverage and impact: a review of the methodology used by the expanded programme on immunization. World Health Statistics Quarterly 1985; 38:66-75.
28. Levy PS, Lemeshow S. Sampling for health professionals. California, Lifetime Learning Publications, 1980.
29. McDowell I, Newell C. Measuring Health. A guide to rating scales and questionnaires. Oxford, Oxford University Press, 1996.
30. Ministério da Saúde. Diretrizes e normas regulamentadoras de pesquisa envolvendo seres humanos – Resolução 196/96 do Conselho Nacional de Saúde. Rio de Janeiro, MS do Brasil – Fiocruz, 1998.
31. Monteiro CA, de Moura EC, Jaime PC et al. Surveillance of risk factors for chronic diseases through telephone interviews. Rev Saude Publica 2005; 39:47-57.
32. Moser CA, Kalton G. Survey Methods in social investigation. Hants, Gower Publishing Co. Ltd., 1986.
33. Noether GE. Introdução à estatística: uma abordagem não-paramétrica. Rio de Janeiro, Editora Guanabara Dois, 1983.
34. Reinchenheim ME, Moraes CL. Buscando a qualidade das Informações em Pesquisas Epidemiológicas. In: Minayo MCS (Org) Caminhos do Pensamento: Epistemologia e Métodos. Rio de Janeiro: Editora Fiocruz, 2002.
35. Rothman KJ, Greenland S. Modern epidemiology. 2 ed., Philadelphia, Lippincott-Raven Publishers, 1998.
36. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. American Journal of Epidemiology 1980; 112(4):467-70.
37. Stata Corporation: Stata, Version 7. Texas, Stata Press, Stata Corporation, 2001.
38. Szklo M, Nieto FJ. Epidemiology – Beyond the Basics. Maryland, Aspen Publishers, 2000.

Estudos Caso-controle

Laura C. Rodrigues
Guilherme L. Werneck

PRINCÍPIOS BÁSICOS

Um estudo caso-controle é um tipo de estudo observacional que se inicia com a seleção de um grupo de pessoas portadoras de uma doença ou condição específica (*casos*), e um outro grupo de pessoas que não sofrem dessa doença ou condição (*controles*). O propósito desse estudo é identificar características (*exposições, ou fatores de risco*) que ocorrem em maior (*ou menor*) freqüência entre casos do que entre controles. A proporção de expostos a um fator de risco é medida nos dois grupos e comparada. Se a proporção de expostos ao fator é maior entre casos do que entre controles, então é possível que esta exposição esteja relacionada a um aumento do risco para a doença em questão. Por outro lado, se esta proporção é menor entre casos, então a exposição sob estudo pode ser considerada um *fator protetor*. A Figura 11.1 mostra a estrutura básica do desenho de um estudo caso-controle:

Com freqüência é útil, ao planejar um estudo caso-controle, tentar identificar a população que produziu os casos e tratá-la como uma coorte (real ou imaginária) a qual o estudo caso-controle corresponde. Em um estudo de coorte, existem pessoas na população de estudo que são expostas, e pessoas que não são expostas aos fatores que se suspeita estarem relacionados à ocorrência de doença. Nos estudos caso-controle, controles são selecionados para estimar a freqüência da exposição nesta população. Neste sentido, nos estudos caso-controle,

os controles podem ser vistos como uma amostra da população que produziu os casos.

Em um estudo de coorte conduzido em uma população fechada (Tabela 11.1), o número de casos expostos (a) e não-expostos (c), e o número de pessoas expostas (P_1) e não-expostas (P_0), são conhecidos, então o risco em expostos (a/P_1) e o risco em não-expostos (c/P_0) também são conhecidos, e a razão de riscos (RR) pode ser diretamente calculada da seguinte maneira: $(a/P_1)/(c/P_0)$. Uma outra medida que pode ser extraída de coortes é a razão de chances, ou *odds ratio* (OR). Isto é, a razão de chances (casos/não-casos) na população exposta (a/b), dividida pela razão de chances (casos/não-casos) na população não-exposta (c/d), ou ad/bc . Observe que $P_1/P_0 \approx (b)/(d)$, se tanto (a) e (c) (casos expostos e não-expostos, respectivamente) ocorrem raramente na base populacional, afetando pouco a magnitude de P_1 e P_0 . Em geral, se a incidência da doença na base populacional é menor do que 10%, esta aproximação tende a funcionar, e a razão de chances (ad/bc) é semelhante à razão de riscos $(a/P_1)/(c/P_0)$.

Agora, vamos imaginar o caso-controle correspondente a esta coorte. Na opção mais simples, nós selecionamos, para constituir o grupo de casos, todos os casos que surgiram de dentro da coorte, e como controles uma amostra da população que não desenvolveu a doença de interesse durante o período de duração da coorte. Neste estudo caso-controle, as únicas medidas que nós temos disponíveis são as proporções de pessoas expostas entre

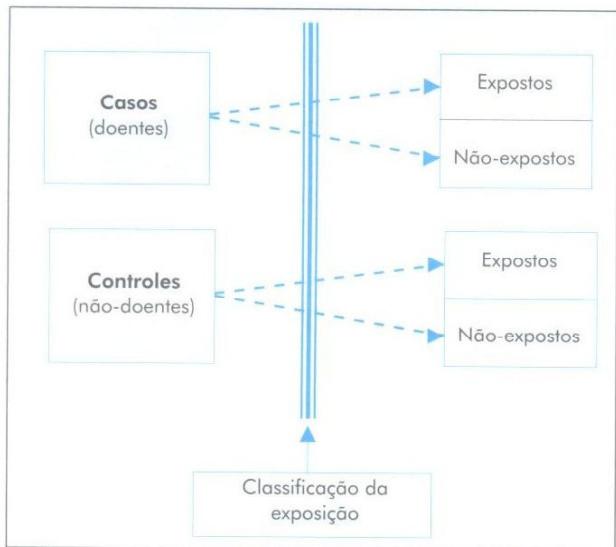


Fig. 11.1 — Estrutura básica de um estudo caso-controle.

Tabela 11.2
Apresentação dos Dados de um Estudo Caso-controle no qual a Exposição é Dicotômica

	Casos	Controles	
Expostos	a	$f.P_1 = b$	$RR = \frac{a}{c}$
Não-expostos	c	$f.P_0 = d$	
Total	$a + c$	$b + d$	

$= P_1/P_0 = b/d$, ou seja, a razão entre controles expostos e não-expostos (b/d) fornece uma boa estimativa da razão entre população exposta e não-exposta (P_1/P_0).

Tradicionalmente, em estudos caso-controle, os casos são comparados aos controles com respeito à freqüência de exposição através de uma medida de associação denominada *razão de chances de exposição*, que é a medida de quanto a exposição é mais freqüente em casos do que em controles. A razão de chances de exposição é a chance de exposição entre casos (a/c) dividida pela chance de exposição entre controles (b/d). Isto é numericamente igual à *razão de chances de doença* em expostos e não-expostos, que, como vimos, pode ser calculada em estudos de coorte (e também em estudos caso-controle!), e que numericamente se aproxima da razão de riscos em doenças raras. Vejamos:

	Casos	Não-casos	Total	
Expostos	a	b	$a + b = P_1$	$RR = \frac{a}{P_1}$
Não-expostos	c	d	$c + d = P_0$	$\frac{c}{P_0}$

os casos e entre os controles. Não temos informação sobre o número de pessoas expostas (P_1) e não-expostas (P_0) na população que produziu os casos. O número de casos e controles, num estudo caso-controle, é definido arbitrariamente pelo investigador e por definição não reflete a relação real entre número de casos e total de pessoas na população. Por causa deste aspecto do desenho de estudos caso-controle, não podemos calcular os riscos de doença em expostos e não-expostos, nem podemos calcular a razão de riscos diretamente. Por causa disto, a apresentação de dados tabulares nos estudos caso-controle não inclui uma coluna de total, somando casos e controles: a soma de casos e controles não tem interpretação, e uma decisão de selecionar por exemplo dois ou três controles para cada caso, em vez de apenas um controle por caso, muda os valores nesta coluna. Os controles são amostras da população que produziu os casos. A fração amostral (f) é a proporção da população que é selecionada como controle no estudo. Quando não há viés de seleção, a fração amostral é a mesma para controles expostos e não-expostos, de modo que $f.P_1/f.P_0$

$$\text{Caso-controle} = \frac{a/c}{b/d} = \frac{a/b}{c/d}$$

$$\text{Coorte} = \frac{a/b}{c/d} = \frac{a/P_1}{c/P_0}$$

(se a doença é rara)

Pode-se perceber que a razão de chance de doença e a razão de chance de exposição são numericamente similares à razão de riscos quando a doença é rara. A razão de chances (*odds ratio*) tende sempre a exagerar o efeito que seria estimado através do risco relativo, porém esta distorção é insignificante quando a premissa de raridade de doença é cumprida.

ESTUDOS CASO-CONTROLE E RARIDADE DA DOENÇA

Ainda que tenha perdurado durante anos como uma regra geral, o princípio clássico de que estudos caso-

controle provêem estimativas válidas do risco relativo apenas quando a doença sob estudo é rara, refere-se a uma situação bem específica. A aplicação deste princípio se dá em um tipo especial de caso-controle no qual casos e controles são selecionados *após* o término do período de risco. Como exemplo toma-se um estudo caso-controle sobre a associação entre tabagismo materno e ocorrência de asma entre crianças de seis a 12 anos de idade realizado no ano de 1990 em uma determinada cidade. Se o grupo de casos é formado pela totalidade de casos de asma ocorridos durante o ano de 1990 e o grupo-controle por crianças sem asma selecionados após o término do período de estudo (p. ex., em janeiro de 1991), então a premissa de raridade de doença é necessária para que a razão de chances seja um bom estimador da razão de riscos.

Entretanto, uma concepção mais geral de estudos caso-controle, mostra que a exigência da raridade de doença não é necessária para que a razão de chances seja um bom estimador do risco relativo. Este “novo olhar” (inicialmente explicitado na década de 1970 por Olli Miettinen e discutido por Laura Rodrigues em 1992) considera que para todo estudo caso-controle existe uma coorte subjacente (imaginária ou concreta) composta pela população de indivíduos de onde os casos se originaram (“base populacional”). Isto será discutido em maior detalhe na seção final deste capítulo.

DEFINIÇÃO E SELEÇÃO DOS CASOS

No processo de definição e seleção de casos para um estudo caso-controle é importante ter critérios claros para definir a doença, a sua gravidade, e se serão incluídos casos incidentes (casos novos) ou casos prevalentes. O estudo é mais fácil de conduzir e analisar quando a fonte inclui todos os casos em uma população claramente definida, de modo que a coorte correspondente seja de fácil identificação. Por exemplo, todos os casos de uma doença que se apresentam a uma rede de centros de saúde, entre membros de um sindicato, funcionários públicos, ou mesmo todos os casos em hospitais de uma cidade. Entretanto, isto não é sempre possível.

DEFINIÇÃO DE CASO (DEFINIÇÃO DO DESFECHO DE INTERESSE)

Critérios acurados para a definição de caso são essenciais em estudos caso-controle. Em geral, os critérios podem ser baseados em achados laboratoriais, p. ex., tipo histológico de câncer, ou podem ser definidos clinicamente, p. ex., um ou mais episódios de depressão

nos últimos cinco anos em mulheres entre 20-65 anos de idade. Neste caso, uma definição clara e padronizada de “depressão clínica” deve ser utilizada.

Gravidade da Doença

O estudo caso-controle vai medir a associação entre o fator de risco e a doença conforme ela for enunciada na definição de caso. Fatores de risco podem variar para os diferentes graus de gravidade de uma doença. Neste sentido, torna-se importante considerar se o evento de interesse é definido em termos de, por exemplo, morbidade, mortalidade, infecção, seqüela etc. Por exemplo, se o interesse é na investigação do papel protetor da vacina anti-sarampo contra casos graves, é melhor fazer um estudo de casos de sarampo hospitalizados do que um estudo com casos de sarampo que ocorrem na comunidade.

Casos Incidentes Versus Casos Prevalentes

*Ver no
livro*

Uma questão importante a ser considerada é se serão incluídos no estudo casos *prevalentes* ou casos *incidentes*. Casos incidentes são casos novos que surgem dentro de um período fixo de tempo, enquanto casos prevalentes são os indivíduos com a doença em questão em um determinado momento no tempo (ou período curto de tempo). Casos prevalentes incluem pacientes que podem estar doentes em longo tempo, e usualmente excluem casos que já morreram, ou se recuperaram, e podem ser diferentes em termos do seu perfil de exposição, quando comparados com casos incidentes. Três pontos importantes devem ser considerados: casos prevalentes ainda estão vivos, ainda estão doentes (e, portanto, o grupo inclui uma proporção maior de casos com doença de longa duração), e podem ter modificado seus hábitos (ou “exposições”) *por causa* da doença. Quando casos prevalentes são utilizados, os primeiros dois pontos levam à identificação de fatores de risco para a doença e também de fatores, que, embora não relacionados à etiologia, têm associação com o tempo de duração da doença. O último ponto pode levar a erro na avaliação da informação sobre o verdadeiro nível de exposição ao qual o caso estava submetido antes de desenvolver a doença.

Fontes de Seleção de Casos

Quais casos serão recrutados para um determinado estudo é um aspecto que precisa ser cuidadosamente considerado. As fontes dos casos podem ser unidades

de atenção à saúde (hospitais, centros de saúde, ambulatórios), sistemas de informação em saúde (notificações de doenças infecciosas, registro de câncer), ou mesmo a busca dos casos que ocorrem na comunidade, podendo eles ter ou não procurado um serviço de saúde. O grupo de casos pode ser composto por pacientes que procuram um determinado hospital e cumprem os critérios para definição de caso, p. ex., todos os recém-nascidos de baixo peso da maternidade Carmela Dutra ocorridos entre 1985-1992, ou todos os casos de sarampo diagnosticados em centros de saúde na rede de centros de saúde na cidade de Santo André. Alternativamente, a fonte de casos pode ser a própria comunidade, sendo retirados de uma população definida durante um período fixo de tempo, p. ex., todos os casos de diarréia identificados pelo estudo em uma comunidade durante um ano. Comunidade neste caso não precisa ser uma área geográfica: pode ser a rede escolar, uma fábrica etc. É comum pensar que se os casos forem selecionados em unidades de saúde, os controles também devem ser selecionados em unidades de saúde. Isto é uma falácia. Independentemente da fonte de casos, a melhor fonte de controles é aquela que permitir selecionar um grupo-controle que corresponda a uma amostra representativa da população que produziu os casos. Em geral, em estudos de base hospitalar a identificação da coorte correspondente é um ponto crítico, qual seja, é difícil delimitar a base populacional que deu origem a estes casos e, por conseguinte, proceder a uma seleção apropriada de controles. Para identificar esta base populacional deve-se considerar como os casos escolhidos se relacionam a todos os casos que preenchem os critérios utilizados para a definição de caso, mas por vários motivos não foram incluídos no estudo devido a recusas ou a diferenças no acesso a serviços de saúde, na sobrevivência de pacientes e nos sistemas de referências para hospitais especializados.

Toda a vez que a definição de caso depende de um procedimento diagnóstico que não é realizado freqüentemente, a população que produziu os casos é a população de pessoas que teria utilizado este procedimento se tivesse apresentado os sintomas. Por exemplo, em estudos para investigar a relação entre dieta e colelitíase, se a definição de caso depende da realização de exames invasivos como a colangiografia venosa, é possível que somente um subgrupo dos casos com sintomatologia mais leve seja diagnosticado, apenas porque estes indivíduos têm maior acesso à atenção médica. Se o acesso a este tipo de serviços de saúde depende do nível socioeconômico da população (o que, por sua vez, afeta os hábitos alimentares das pessoas), os controles apropriados devem representar as pessoas do mesmo nível

socioeconômico, e com mesmo acesso a este nível de atenção médica dos casos.

DEFINIÇÃO E SELEÇÃO DE CONTROLES

DEFINIÇÃO DE CONTROLES

Controles devem preencher os mesmos critérios utilizados para a definição de caso, exceção feita aos critérios relacionados à definição da doença a ser estudada. Por exemplo, se os casos são mulheres de 14 a 44 anos de idade com artrite reumatóide, então os controles devem ser selecionados dentre mulheres com a mesma idade sem a doença em questão. O número de casos e controles não precisa ser igual, mas aumentar o número de controles a mais do que quatro controles para um caso não eleva substancialmente o poder do estudo. Se o número de casos não for limitado, é melhor aumentar o número de casos mantendo a relação de um caso para um controle.

Seleção de Controles e Viés de Seleção

O grupo de controles deve ser composto por uma amostra representativa da população que deu origem aos casos. Quando os controles não são representativos da base populacional que produziu os casos, o *viés de seleção* é introduzido no estudo.

Uma forma de avaliar a possibilidade de viés de seleção é definir a coorte (real ou imaginária) que subjaz ao estudo caso-controle e considerar se os controles selecionados são uma amostra representativa desta coorte. Outra maneira é checar se estes controles seriam selecionados como casos neste estudo se eles viessem a desenvolver a doença sob estudo.

É fundamental garantir que o processo de seleção seja independente do *status* de exposição dos controles. Isto é, controles não devem ser selecionados com maior ou menor freqüência em função do nível de exposição que apresentem, mas devem simplesmente prover informação acerca da distribuição da exposição na base populacional que deu origem aos casos.

Se os casos são selecionados como uma amostra de todos os casos (ou mesmo a totalidade dos casos) incidentes numa população definida durante um determinado período de tempo, então os controles devem ser selecionados para representar esta população. Esta população pode, mas não precisa, ser geograficamente definida. Fontes de populações definidas incluem muitos tipos de listagens: crianças registradas em centros de saúde ou escolas, adultos registrados em clubes,

associações, sindicatos; listas de endereços, listas de habitações do censo, até mesmo listas telefônicas. Se os casos fazem parte desta população definida (e você pode resolver excluir do seu estudo os casos que não fazem parte desta população), os controles devem ser uma amostra representativa desta população. Por exemplo, se a população de estudo for definida como a população com telefone em casa, os casos sem telefone são excluídos, e os controles poderão ser selecionados através de procedimentos aleatórios de amostragem via ramal telefônico. Como, em geral, o prefixo telefônico corresponde a regiões específicas, então é possível identificar os prefixos correspondentes à área geográfica da população que se quer amostrar, e aleatoriamente selecionar os quatro números restantes. Assim, sem ter a necessidade de ir ao local, pode-se travar um primeiro contato com os potenciais controles, explicar os objetivos da investigação, garantir a participação e agendar local e hora para a entrevista. Este procedimento é útil em regiões onde a cobertura telefônica é alta.

Se os casos são recrutados a partir de uma lista de admissões hospitalares, mas este hospital cobre todos os casos da doença em questão (no nível de gravidade especificado na definição de casos do estudo) que ocorrem em uma população definida, então, de novo, os controles devem ser amostrados a partir desta população. Em ambas as situações fica razoavelmente claro qual a base populacional de onde os casos surgiram, e, portanto, de onde os controles devem ser selecionados. Idealmente, uma amostra probabilística (p. ex., amostra aleatória simples) desta população poderia ser utilizada para selecionar controles, porém nem sempre isto é factível devido à ausência de listas completas dos moradores da região, e também por questões de custo. Alternativas incluem amostras em múltiplos estágios, baseadas em setores censitários.

Uma opção alternativa é selecionar controles que moram na mesma vizinhança dos casos, definindo a população base como os moradores da vizinhança dos casos. Nesta situação, para cada caso seleciona-se um (ou mais) controle(s) que mora(m) em residências localizadas a uma certa distância e direção da moradia do doente. Este procedimento diminui os custos do processo de seleção, mas pode ser ineficiente, pois invariavelmente existirão habitações vazias (pessoas trabalhando ou na escola) e/ou vizinhos não dispostos a colaborar, sobretudo em decorrência do aumento da violência nas áreas urbanas. Uma outra opção alternativa é solicitar ao caso o nome de uma pessoa (melhor amigo, parente, vizinho etc.) para ser entrevistada como controle. Provavelmente o nível de colaboração será maior,

mas controles obtidos desta forma tendem a ser mais similares aos casos em termos demográficos e sociais, o que poderia levar a discrepâncias entre a distribuição da exposição que se pretende estimar (a da população que originou os casos) e aquela que efetivamente este tipo de controle oferece.

Existem situações em que não se consegue definir claramente qual a base populacional que deu origem aos casos, em especial quando estes são selecionados em um determinado hospital, tornando a escolha dos controles um procedimento pouco óbvio. As diretrizes são, como sempre, tentar imaginar a população-base — qual é a população de onde estes casos estão se originando? Em geral, isto depende da natureza da doença ou condição sob estudo. Se existe algum mecanismo de seleção para os indivíduos que são atendidos neste hospital ou clínica — p. ex.: em áreas rurais, casos hospitalizados são indivíduos que desenvolveram doenças dentre uma população que consiste em pessoas que têm acesso à assistência médica —, então os controles devem ser selecionados dentre pessoas que dariam entrada neste hospital caso viessem a desenvolver qualquer doença de gravidade semelhante àquela definida para casos.

Esta definição pode incluir também pacientes hospitalizados no passado ou no presente. A escolha de controles “hospitalares” adequados é um desafio e deve-se tomar cuidado redobrado para evitar o viés de seleção. Pessoas hospitalizadas são pessoas doentes e têm maior tendência a terem sido expostas com maior freqüência a uma série de fatores adversos, quando comparadas à população geral.

É surpreendentemente simples escolher de forma inadvertida um grupo-controle que tenha uma prevalência de exposição mais alta, ou mais baixa, quando comparado com a população geral. Por exemplo, em um estudo para investigar o papel do álcool no desenvolvimento do câncer de mama, o uso de controles oriundos do serviço de emergência do mesmo hospital em que os casos foram selecionados pode induzir viés de seleção, porque este grupo de pacientes sabidamente consome mais álcool do que a população em geral. Então, a seleção de controles “hospitalares” deve ser precedida de uma cuidadosa avaliação de qual(is) doença(s) são ou não sabidamente relacionadas com a exposição que o estudo deseja investigar. Por exemplo, ao estudar o papel do tabagismo no câncer de pulmão, não seria indicado selecionar como controles pacientes portadores de doença pulmonar obstrutiva crônica, pois estes sabidamente têm um excesso de exposição ao tabaco quando comparados com a população que deu origem

aos casos. Se insistíssemos em selecionar controles deste tipo, é possível que o estudo fosse incapaz de detectar a associação entre fumo e câncer de pulmão, um exemplo clássico de viés de seleção. Em geral, é preferível incluir no grupo-controle pacientes com uma grande variedade de doenças, as quais se imagina não serem relacionadas com a exposição em questão.

Uma alternativa comum é a utilização de dois ou mais tipos de grupo-controle, mas esta estratégia também pode provocar problemas: os dois grupos-controle podem gerar resultados diferentes, levando a situações nem sempre de fácil interpretação. Por exemplo, em um estudo sobre o efeito potencial da radiação no desenvolvimento de leucemia em crianças, os casos foram selecionados dentre aqueles indivíduos com leucemia linfóide aguda que deram entrada no hospital X e dois grupos-controle foram selecionados dentre pacientes que também deram entrada neste hospital: 1. controles com câncer de cérebro e; 2. controles com traumatismos diversos. Os resultados do estudo foram discrepantes, tendo sido identificada associação entre radiação e leucemia apenas quando os casos foram comparados ao segundo grupo-controle. Qual a explicação para os resultados? Será que a radiação também é um fator de risco para câncer de cérebro? Será que as mães das crianças com câncer cerebral se lembram mais das exposições pré-natais à radiação do que as crianças com traumatismos? Será que o grupo-controle com traumatismo estaria por alguma razão menos exposto à radiação do que a população-base que originou os casos de leucemia?

Mais rara, porém possível, é a introdução de viés na seleção de casos. Isto acontece quando casos expostos têm maior chance de serem selecionados para o estudo do que casos não-expostos (ou o oposto). Uma área onde isto foi descrito é no estudo de reações adversas a vacinas. Um outro exemplo onde esta possibilidade foi discutida é a controvérsia relativa à interpretação dos resultados de vários estudos caso-controle conduzidos no início dos anos 1970 para investigar a associação entre estrogênio artificial e o risco de câncer de endométrio. Alguns desses estudos encontraram uma forte associação ($OR = 9$) para câncer de endométrio entre mulheres que haviam tomado estrogênio. Entretanto, Horowitz e Feinstein argumentaram que um sério viés de seleção poderia explicar estes achados. Estes autores ponderaram o seguinte:

- estrogênio artificial causa sangramento uterino, a despeito da presença ou não de câncer de endométrio;
- este sintoma (sangramento uterino) conduziria a mulher a um exame ginecológico;
- uma investigação ginecológica revelaria a presença de câncer de endométrio que de outra forma poderia passar despercebido;
- consequentemente, a taxa de detecção de câncer de endométrio seria maior entre mulheres que tivessem tomado estrogênio do que entre mulheres que não tivessem tomado. Ou seja, o processo de inclusão de casos no estudo seria função da exposição, gerando, portanto, um viés de seleção que resultaria numa superestimação do risco relativo.

Este argumento foi questionado por diversos autores. O principal contra-argumento era de que, ainda que a administração de estrogênio tornasse mais precoce o diagnóstico, seria esperado que os cânceres de endométrio (não *in situ*) fossem, quase invariavelmente, progredir a ponto de causar sintomas que levariam ao diagnóstico.

Como estamos vendo, a escolha de um grupo-controle adequado é a parte mais difícil na elaboração de um estudo caso-controle. O principal lema a considerar é: controles existem para representar a distribuição da exposição na população de onde os casos se originaram.

AFERIÇÃO DA EXPOSIÇÃO E VIÉS DE INFORMAÇÃO

Dados sobre exposição podem ser coletados de várias maneiras: por meio de entrevistas pessoais, por telefone ou através de questionários enviados pelo correio; pelo exame de registros médicos ou ocupacionais; ou a partir de amostras de material biológico (p. ex.: sangue ou outros tecidos). A questão fundamental é garantir que a informação coletada seja acurada e não influenciada pelo fato de o indivíduo ser um caso ou um controle. Pelo simples fato de que em estudos caso-controle os indivíduos são selecionados com base no *status* de doença, estes estudos são particularmente vulneráveis ao viés de informação, isto é, o fato de o indivíduo ter ou não a doença sob estudo pode influenciar a captação correta da informação sobre se a pessoa foi ou não exposta ao fator de interesse. Isto é particularmente verdadeiro quando a informação sobre exposição depende da memória dos participantes do estudo. É preciso considerar também que casos e controles podem ser mais ou menos motivados para participar do estudo e podem diferir na qualidade da informação que provêem, em especial no que diz respeito a questões “delicadas” (p. ex.: acerca de práticas sexuais, violência familiar, renda, aborto prévio etc.). Em estudos caso-controle é sempre melhor,

quando possível, obter informação a partir da revisão de registros criados antes da doença ter sido diagnosticada por pesquisadores “cegos” ao *status* de caso ou controle da pessoa cujo registro está sendo revisto.

Viés do Respondente

Viés do respondente refere-se à situação em que a informação sobre exposição fornecida pelos participantes do estudo difere em função de ser ele um caso ou um controle. Um exemplo é o *viés de memória*: ser doente pode influenciar as respostas dadas a certas questões. Um exemplo clássico é o de mães de crianças nascidas com malformação congênita. Estas mães tendem a se lembrar mais acerca do uso de medicações durante a gravidez do que as mães cujos filhos nasceram sem malformação. O viés do respondente pode ser minimizado na medida em que os participantes do estudo desconhecem as hipóteses sob investigação. Viés de memória é raro em exposições que não dependem de julgamento subjetivo (“quantos filhos vivos você tem”) e é evitado por obtenção cega da informação de registros anteriores ao diagnóstico. Sempre que possível, deve-se garantir que casos e controles tenham o mesmo incentivo para recordar eventos passados, e sejam entrevistados em locais semelhantes, usando o mesmo questionário.

Viés do Observador

Viés do observador refere-se à situação na qual o processo de obtenção da informação pelo investigador difere se o informante é um caso ou um controle. É plausível que, na medida em que o entrevistador conheça o *status* de caso ou controle e a hipótese sob estudo, ele tenda, em situações limítrofes (p. ex.: participante relata que fuma ocasionalmente, apenas em situações sociais), a classificar os casos como expostos (*i. e.*, são tabagistas) e os controles como não-expostos (*i. e.*, não são tabagistas) apenas porque ele “crê” na hipótese do estudo de que tabagismo é fator de risco para câncer de bexiga. Uma outra maneira pela qual o investigador pode introduzir viés é entrevistando diferentemente os casos e controles. Por exemplo, o investigador pode aceitar uma resposta relatando não-exposição do controle, mas insistir com os casos (“tem certeza que você nunca fumou? Nunca mesmo, nem muito tempo atrás...”). Idealmente, o entrevistador deve desconhecer as hipóteses sob estudo e também desconhecer quem é caso e quem é controle. Em situações reais é muito difícil alcançar uma situação como esta, e os entrevistadores devem ser treinados extensivamente para realizar uma coleta de

dados não-enviesada. A informação deve ser coletada de uma maneira objetiva e o mesmo tipo de formulário e questionário, tipo e local de entrevista, utilizados para casos e para controles.

PAREAMENTO EM ESTUDOS CASO-CONTROLE

Pareamento se refere ao procedimento pelo qual, para cada caso selecionado, são recrutados um ou mais controles idênticos com relação a certas características outras que não o fator sob investigação. Embora controles existam para representar a população que produziu os casos, pareamento é uma maneira de garantir que o grupo de casos e controles venha de uma população semelhante também em termos de certas variáveis. Variáveis demográficas, como idade e sexo, são comumente utilizadas no pareamento, mas pode-se parear por outras variáveis como local de residência, nível socioeconômico, ou paridade da mãe. A decisão sobre parear ou não, e a escolha das variáveis a serem utilizadas no pareamento, devem ser feitas considerando as características de cada estudo. As variáveis a serem utilizadas no pareamento são aquelas suspeitas de serem variáveis de confusão (também denominadas variáveis de confundimento): o pareamento tem como objetivo torná-las equitativamente distribuídas entre casos e controles, desfazendo a associação entre a variável de confusão e a doença (nem todas as variáveis de confusão devem ser pareadas – isto será discutido mais adiante). Em um estudo sobre a relação entre consumo de álcool na gravidez e baixo peso ao nascer, por exemplo, o tabagismo da mãe foi considerado um importante fator de confusão. Neste contexto, pareamento levaria a que para cada caso (filho de baixo peso) cuja mãe tivesse fumado na gravidez, um controle seria selecionado cuja mãe também tivesse fumado (ou, mais precisamente, fumado uma quantidade semelhante) durante a gravidez; e para cada caso cuja mãe era não-fumante, seria recrutado um controle cuja mãe também não tivesse fumado durante a gravidez.

Entretanto, ao contrário do que se observa em estudos de coorte, o pareamento em estudos caso-controle não leva diretamente ao controle do confundimento. Isto se dá porque, em estudos de coorte, o pareamento opera apenas na relação exposição-confundimento, já que a doença ainda está por ocorrer. Em outras palavras, num estudo de coorte o pareamento faz com que a coorte exposta e a coorte não-exposta sejam comparáveis em relação à freqüência do fator de confusão; e a diferença na ocorrência da doença é atribuível à presença da exposição. Em estudos caso-controle, o pareamento torna

casos e controles semelhantes em relação à freqüência da variável de confusão. Mas como casos e controles podem ter uma freqüência diferente de exposição, o pareamento leva a uma distribuição alterada da freqüência de exposição entre controles. Em suma, em um estudo caso-controle, o pareamento opera na relação doença-confundimento-exposição, e, portanto, requer que a comparação de freqüência de exposição se dê dentro de cada par de caso e controle. Se isto não for feito, e na medida em que a variável de pareamento é de fato uma variável de confusão, a associação sob estudo é alterada em direção ao valor nulo (não-associação), já que torna semelhante a freqüência da exposição em casos e controles. Este viés em direção ao valor nulo é imposto pelo pareamento e só pode ser corrigido através de uma análise pareada. Portanto, é fundamental notar que, toda vez que o estudo caso-controle for pareado, a análise também tem que ser pareada, comparando a freqüência de exposição dentro de cada par.

Existem outras maneiras de se controlar para fatores de confusão em estudos caso-controle (p. ex.: restrição, estratificação e modelos de regressão), mas o pareamento para variáveis que são confundidoras pode aumentar muito o poder do estudo, em especial quando a variável de confusão está fortemente associada à doença. Portanto, o objetivo principal do pareamento em estudos caso-controle é evitar a situação de confusão e aumentar a eficiência do estudo. Por exemplo, num estudo de câncer, devido à força de uma associação entre câncer e idade, se não houver pareamento por idade, a maioria dos casos tenderia a ser de idosos e a maioria dos controles de jovens (refletindo a estrutura etária da população brasileira!).

Pareamento para variáveis que não são confundidoras diminui o poder do estudo, em especial quando esta variável é somente associada à exposição ou é uma variável intermediária (ou interveniente) quando se considera a relação entre a exposição e doença, qual seja, uma variável que é determinada pela exposição e, por sua vez, causa a doença.

Finalmente, pareamento pode ser útil para controlar o efeito de variáveis de confusão que são difíceis de especificar e quantificar. Ao utilizar no pareamento variáveis como vizinhança, ser amigo/nomeado pelo caso, ser sócio do mesmo clube, membro da família etc., casos e controles tenderiam a ser mais semelhantes no que diz respeito a fatores tais como nível socioeconômico, exposições ambientais, facilidades de acesso a serviços de saúde, hábitos de vida, fatores genéticos etc. A desvantagem de pareamento por confundidoras não

bem definidas é justamente criar a situação em que não sabemos exatamente o que está sendo controlado.

PAREAMENTO POR FREQUÊNCIA

Uma alternativa ao pareamento individual é o pareamento por freqüência. Nesta abordagem, uma vez escolhida a variável a ser pareada, o processo de recrutamento dos participantes é conduzido monitorando-se o número de casos e controles em cada estrato da variável de pareamento, garantindo que haja a mesma proporção de controles e casos nos estratos. Por exemplo, em um estudo em que se decida parear por sexo, ao fim de cada semana os pesquisadores determinam qual é a proporção de homens entre os casos e entre os controles. Se esta proporção não é a mesma, na semana seguinte recruta-se mais controles do sexo sub-representado. O pareamento por freqüência requer análise estratificada, e não-pareada (Quadro 11.1).

Quadro 11.1

Vantagens e Desvantagens do Pareamento em Estudos Caso-controle

Vantagens do Pareamento

- Quando a variável a ser utilizada no pareamento é um forte fator de confusão, o pareamento aumenta a eficiência do estudo e, portanto, diminui o tamanho da amostra necessária. Particularmente útil em estudos pequenos sobre doenças raras em que existem muitas variáveis de confusão.
- Permite o controle de variáveis que são difíceis de medir (como o “estilo de vida”, perfil genético etc.). Isto é possível quando se usa para o pareamento variáveis como vizinhança, ser amigo ou membro da família.
- O conceito de pareamento, uma vez entendido, é simples e faz sentido, é convincente e mais fácil de ser explicado a um leigo do que, por exemplo, estratégias baseadas em modelos de regressão logística.
- Uma vez pareado (e feita a análise pareada) pode-se “esquecer” esta variável de confusão (mas não completamente, pois há sempre a possibilidade de confundimento residual).

Desvantagens do Pareamento

- Pode ser difícil, caro e demorado encontrar controles com as características necessárias para serem pareados a cada caso. Isto piora à medida que aumenta o número de variáveis a serem usadas no pareamento. Por exemplo, parear por idade, classe social, número de filhos e hábito de fumar simultaneamente pode ser extremamente difícil. É recomendado não usar mais de três variáveis no pareamento.
- Uma vez que uma variável é usada no pareamento, não é possível investigar a associação dela com a doença estudada.
- Para possibilitar a seleção de um controle pareado, normalmente é necessário entrevisitar o caso antes do controle para classificá-lo quanto à variável de confusão, aumentando o risco de viés. E, dependendo da variável de pareamento, pode ser preciso entrevistar os controles potenciais (às vezes muitos), para obter informação sobre as variáveis usadas no pareamento e, finalmente, selecionar os controles a serem incluídos no estudo.
- A apresentação dos dados é mais difícil e a interpretação, menos evidente.
- Para controlar por outras variáveis de confusão é necessário o uso de regressão logística condicional, já que o controle para estas variáveis adicionais não pode ser feito via estratificação na análise pareada.

Superpareamento (overmatching) ou pareamento excessivo pode ocorrer quando a variável utilizada para pareamento fizer parte do processo causal, localizando-se entre a exposição e a doença (variável intermediária ou interveniente) ou, simplesmente, quando a variável confundidora estiver associada muito fortemente à exposição (de forma que o pareamento se aproxima de um pareamento pela exposição). O superpareamento tem o efeito de diminuir muito o poder do estudo ou mesmo, em casos extremos, enviesar o resultado em direção ao valor nulo, mesmo quando a análise é pareada.

ANÁLISE

A principal medida de associação estimada em estudos caso-controle é a razão de chances (*odds ratio* – OR). A razão de chances é uma medida da força da associação entre a exposição e a doença sob estudo. Uma

razão de *chances* (OR) de três, por exemplo, significa que a exposição ao fator sob estudo aumenta em três vezes a *chance* de adoecer. Por outro lado, uma razão de chances de 0,5 significa que a chance de doença cai pela metade quando pessoas estão expostas ao fator sob estudo, neste caso denominado fator de proteção. Uma razão de chances igual a 1 significa que não existe associação entre a exposição e a doença. Ainda que inadequada, é comum a utilização do termo *risco relativo* para denominar genericamente as medidas de associação do tipo razão mais utilizadas em epidemiologia (*i.e.*, a razão de riscos, a razão de taxas e a razão de chances). Neste sentido é que se diz que a razão de chances (OR) *estima* o risco relativo, quando na verdade ela pode ser um bom estimador para a razão de taxas e razão de riscos, para apenas uma delas, ou ainda, para nenhuma das duas medidas.

É importante salientar que, quando uma associação é encontrada, o investigador deve sempre considerar a possibilidade de este resultado ter ocorrido apenas devido ao acaso, viés de seleção, viés de informação ou controle incompleto das variáveis de confusão (confundimento residual).

Também existe a possibilidade de que a exposição seja *resultado* da doença (causalidade reversa). Por exemplo, mesmo se uma associação entre dieta e câncer de estômago é encontrada em um estudo caso-controle, existe a possibilidade de que dieta não seja uma causa mas uma *consequência* do câncer, ou seja, após o diagnóstico do câncer a pessoa modificou seus hábitos alimentares devido à doença.

A razão de chances é calculada de maneiras diferentes em estudos pareados e não-pareados, estratificados e usando regressão logística.

Análise de Estudos de Caso-controle Não-pareados

Em estudos não-pareados, o número de casos e controles expostos ou não ao fator sob investigação pode ser apresentado em uma tabela 2 × 2, como na Tabela 11.3.

Tabela 11.3
Apresentação dos Dados de um Estudo Caso-controle
Não-pareado no qual a Exposição é Dicotômica

	Casos	Controles	Total	
Expostos	a	b	a + b	$\frac{a}{?}$
Não-expostos	c	d	c + d	$\frac{c}{?}$
Total	a + c	b + d	n	$\frac{?}{?}$

Como visto previamente, a *razão de chances* (*odds ratio*) é calculada através da *chance de exposição* (*odds of exposure*) entre casos (a/c), dividida pela *chance de exposição* (*odds of exposure*) entre controles (b/d):

$$OR = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$

Quando o estudo é apropriadamente desenhado, esta medida estima o *risco relativo*.

Testes de Significância Estatística e Intervalos de Confiança para a Razão de Chances

ESTUDOS COM EXPOSIÇÃO DICOTÔMICA E SEM ESTRATIFICAÇÃO

PARA TESTAR A HIPÓTESE NULA $H_0: OR = 1$ (ISILO É, A HIPÓTESE DE QUE NÃO EXISTE ASSOCIAÇÃO ENTRE A DOENÇA E O FATOR ESTUDADO) PODE-SE UTILIZAR O TESTE DE χ^2 (qui-quadrado) com um grau de liberdade para tabelas 2×2 (ver detalhes sobre o qui-quadrado no capítulo Probabilidade e Distribuição de Probabilidades), da seguinte maneira:

$$\chi^2 = \frac{(ad - bc)^2 \times n}{(a+b)(c+d)(a+c)(b+d)}$$

Este teste funciona bem nas situações em que a amostra é grande (em especial quando o número esperado de ocorrências em cada uma das caselas da tabela é maior que cinco). Quando esta premissa não é verdadeira, então testes exatos ou corrigidos (p. ex., o teste com correção de continuidade de Yates) devem ser utilizados.

Um intervalo de 95% de confiança aproximado para a razão de chances (OR) pode ser calculado utilizando-se o método de Wolf, como a seguir:

$$In(OR) \pm 1,96 \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

$$Limite\ inferior: (OR) \times e^{-1,96 \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

$$Limite\ superior: (OR) \times e^{+1,96 \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

ESTUDOS COM MAIS DE UM NÍVEL DE EXPOSIÇÃO E SEM ESTRATIFICAÇÃO

QUANDO HÁ MAIS DE UM NÍVEL DE EXPOSIÇÃO (P. EX., NÃO-FUM ANTES, FUM ANTES E EX-FUM ANTES; NÃO-VACINADOS, VACINADOS COM UMA DOSE, VACINADOS COM DUAS DOSES ETC.), CADA NÍVEL DE EXPOSIÇÃO É COM PARADO COM O NÍVEL NÃO EXPOSTO.

UMA TABELA COM DOIS NÍVEIS DE EXPOSIÇÃO SERIA APRESENTADA CONFORME DESCrito NA TABELA 11.4.

Tabela 11.4 Apresentação dos Dados de um Estudo Caso-controle Não-pareado no qual a Exposição tem mais de Duas Categorias			
	Casos	Controles	Total
Expostos (nível mais alto de exposição – nível 1)	a	b	a + b
Expostos (nível mais baixo de exposição – nível 2)	c	d	c + d
Não-expostos	e	f	e + f
Total	a + c + e	b + d + f	n

Para o nível de exposição 1, assumir $n_1 = a + b + e + f$:

$$OR = \frac{a/e}{b/f} = \frac{a \times f}{b \times e}$$

$$\chi^2 = \frac{(af - be)^2 \times n_1}{(a+b) \times (e+f) \times (a+e) \times (b+f)}$$

Para o nível de exposição 2, assumir $n_2 = c + d + e + f$:

$$OR = \frac{c/e}{d/f} = \frac{c \times f}{d \times e}$$

$$\chi^2 = \frac{(cf - be)^2 \times n_2}{(c+d) \times (e+f) \times (c+e) \times (d+f)}$$

Estudos com Estratificação para Variáveis de Confusão

Estudos caso-controle pareados individualmente não podem ser estratificados. Estudos pareados por

freqüência devem ser analisados usando estratificação pelas variáveis usadas no pareamento de freqüência. Estudos não-pareados e pareados por freqüência podem usar estratificação para o controle do confundimento. Na estratificação, casos e controles são separados em estratos de acordo com níveis de variáveis de confusão e a análise (cálculo de razão de chances, χ^2 e intervalos de confiança) é realizada separadamente para cada estrato. Se a razão de chances for semelhante nos estratos, uma “medida resumo” única pode ser produzida, que é uma média ponderada da razão de chances nos estratos, e então um intervalo de confiança é calculado para esta medida sumária. Esta medida resumo é dita a razão de chances “controlada” para a variável de confusão (Mantel Haenszel *odds ratio*).

Na situação mais simples, em que a variável de exposição é dicotômica e existe apenas uma variável de confusão também dicotômica, a análise estratificada se resume a duas tabelas. A Tabela 11.5 a seguir apresenta um exemplo de estratificação em um estudo caso-controle para investigar o papel do uso de contraceptivos na ocorrência de infarto do miocárdio. O *status* de exposição, neste caso, é definido segundo o uso ou não de contraceptivo oral, casos são pessoas com diagnóstico de primeiro infarto de miocárdio, controles são pessoas sem história de infarto e oriundas da mesma base populacional que deu origem aos casos, e a idade foi considerada um importante fator de confusão.

Tabela 11.5
Apresentação dos Dados de um Estudo Caso-controle no qual a Exposição é Dicotômica e Realiza-se a Estratificação por uma Variável de Confusão Também Dicotômica

	Estrato 1 ($i = 1$) Idade = 30-39 anos			Estrato 2 ($i = 2$) Idade = 40-49 anos		
	Casos	Controles	Total	Casos	Controles	Total
Expostos	13	59	72	12	14	26
Não-expostos	45	720	765	158	663	821
Total	58	779	837	170	677	847

Com base nesta tabela pode-se calcular duas razões de chances (OR_1 e OR_2), uma para cada estrato de idade:

$$OR_1 = \frac{13 \times 720}{59 \times 45} = 3,53$$

$$OR_2 = \frac{12 \times 663}{14 \times 158} = 3,60$$

Na medida em que as duas razões de chance (OR_1 e OR_2) são similares, faz sentido estimar uma medida de associação que resuma os achados das duas tabelas. Esta medida é denominada *odds ratio* de Mantel Haenszel (OR_{MH}). Vale ressaltar que existem testes estatísticos que podem ser aplicados para a avaliação da hipótese de que estas medidas estratificadas sejam ou não similares, como por exemplo o teste de Woolf para avaliar a homogeneidade das OR através de diferentes estratos. Se OR_1 e OR_2 são diferentes, então estamos frente a um fenômeno denominado *modificação de efeito* (muitas vezes também denominado de *interação*) e não faz sentido calcular a OR_{MH} .

A OR_{MH} nada mais é do que uma média ponderada da razão de chances nos estratos. No exemplo anterior a OR_{MH} pode ser calculada conforme descrito a seguir:

Pois bem, neste exemplo, a OR_{MH} é uma medida resumo da associação entre uso de contraceptivo oral e infarto do miocárdio, controlando para o potencial efeito de confundimento da idade. Mas idade é mesmo uma variável de confundimento nesta situação? A maneira operacional mais simples de avaliar esta questão é comparando a OR_{MH} com a OR não ajustada (OR_{na}), ou seja, aquele que obteríamos em uma tabela 2×2 simples, apenas com a exposição e o desfecho, sem levar em consideração a possível variável de confusão.

No exemplo anterior, $OR_{na} = [(13 + 12) \times (720 + 663)] / [(59 + 14) \times (45 + 158)] = 2,33$, valor bem menor do que a OR_{MH} . Pela definição operacional de confundimento, toda a vez que $OR_{na} \neq OR_{MH}$ há confundimento e a medida de associação que deve ser utilizada é aquela ajustada para o efeito do confundimento, ou seja, a OR_{MH} . Como mencionado anteriormente, é importante notar que nem toda diferença entre OR_{na} e OR_{MH} indica situação de confusão, já que pequenas diferenças podem ocorrer apenas em função de variação aleatória. É também comum utilizar critérios que incorporam a magnitude da diferença, por exemplo, somente uma razão entre OR_{MH} e OR_{na} superior a 10% seria considerada indicativa de confundimento.

Dito isto, torna-se necessário testar a hipótese nula de que $OR_{MH} = 1$ através de testes de significância e intervalos de confiança. O cálculo manual de intervalos de confiança para a OR_{MH} é bastante complicado e não será apresentado neste capítulo. O teste de χ^2_{MH} (qui-quadrado de Mantel-Haenszel) pode ser calculado para o exemplo anterior, conforme descrito a seguir:

$$OR_{MH} = \frac{\sum [(a_1 \times d_1) / n_1]}{\sum [(b_1 \times c_1) / n_1]} = \frac{[(13 \times 720) / 837] + [(12 \times 663) / 847]}{[(59 \times 45) / 837] + [(14 \times 158) / 847]} = 3,56$$

$$\chi^2_{MH} = \frac{\{\sum [(a_i \times d_i) / n_i]\}^2}{\sum V_i} = \frac{\{[13 - (72 \times 58 / 837)] + [12 - (26 \times 170 / 847)]\}^2}{[(72 \times 765 \times 58 \times 779) / (837)^2 (836)] + [(26 \times 821 \times 170 \times 677) / (847)^2 (846)]} = 26,37$$

Onde:

$$O_i = a_i$$

$$E_i = [(a_i + b_i) \times (a_i + c_i)] / n_i$$

$$Vi = [(a_i + b_i) \times (a_i + c_i) \times (b_i + d_i) \times (c_i + d_i)] / [(n)^2 \times (n_i - 1)]$$

i = índice para cada estrato, neste exemplo i = 1, 2

Análise de Estudos Caso-controle Pareados Individualmente

Se o estudo é pareado, torna-se necessário realizar uma análise pareada de cunho um pouco mais complexo, tendo em vista o número de pares concordantes e discordantes no que diz respeito à exposição. Por exemplo, um par em que o caso é exposto mas o controle é não-exposto (ou vice-versa) é denominado “par discordante”. Um “par concordante” é aquele em que tanto o caso como o controle têm o mesmo nível de exposição. A Tabela 11.6 apresenta como são construídas tabelas 2 × 2 para apresentação destes dados:

Tabela 11.6 Apresentação dos Dados de um Estudo Caso-controle Pareado Individualmente			
	Controle Exposto	Controle Não-exposto	
Caso exposto	a	b	
Caso não-exposto	c	d	

A soma a + b + c + d é o número total de pares formados por casos e controles. Em estudos pareados, os únicos pares cuja informação é relevante para a estimativa da OR são os chamados pares discordantes em relação à presença da exposição (na Tabela 11.6, b e c). Quanto maior for o número de pares discordantes em que o caso é exposto e o controle não, comparado ao número de pares discordantes em que o controle é exposto mas o caso não, maior será o OR para a exposição investigada. Esta relação é sintetizada pela fórmula:

$$OR_p = \frac{b}{c}$$

Para testar a hipótese nula $H_0: OR_p = 1$ (isto é, a hipótese de que não existe associação entre a doença e o fator estudado) pode-se utilizar o teste de $\chi^2_{McNemar}$ (teste de McNemar), da seguinte maneira:

$$\chi^2_{McNemar} = \frac{(|b - c| - 1)^2}{(b + c)}$$

Um intervalo de 95% de confiança aproximado pode ser calculado utilizando-se o método a seguir:

$$In(OR_p) \pm 1,96 \sqrt{\frac{1}{b} + \frac{1}{c}}$$

$$Limite\ inferior: OR_p \times e^{-196 \sqrt{\frac{1}{b} + \frac{1}{c}}}$$

$$Limite\ superior: OR_p \times e^{+196 \sqrt{\frac{1}{b} + \frac{1}{c}}}$$

Ajuste para Múltiplas Variáveis de Confusão com Regressão Logística

A regressão logística é um poderoso instrumento estatístico para estimar a magnitude da associação (OR) entre um fator de risco e uma doença (ou evento) de interesse, após ajustar simultaneamente para o efeito de variáveis de confusão. Em geral, o uso da regressão logística está indicado quando o desfecho de interesse é dicotômico (p. ex.: doente vs. não-doente), multicategórico (p. ex.: doente, infectado sem sintomas ou sinais clínicos, não-infectado), ou ordinal (p. ex.: não-doente, doente oligossintomáticos, doentes graves e óbitos).

Este modelo é uma variante de uma equação de regressão linear múltipla em que o logaritmo neperiano (\ln) da *odds* (chance) de adoecer (variável dependente) é expresso como uma função de uma série de variáveis independentes (fatores de risco). Os coeficientes obtidos através de regressão logística indicam o efeito de um fator específico sobre o logaritmo da *odds* de adoecer quando todas as outras variáveis são mantidas constantes. A razão de *odds* (razão de chances ou *odds ratio*) para o fator em questão corresponde ao antilogaritmo desse coeficiente e reflete a magnitude da associação investigada após controlar simultaneamente por um número variado de fatores de confusão.

Regressão logística pode ser usada em estudos pareados (regressão logística condicional) e não-pareados (regressão logística não-condicional). As desvantagens da regressão logística são a necessidade de um computador com programas estatísticos, e o fato de que, como em todo modelo, assume-se que os dados têm certas características, por exemplo, que o risco associado com cada variável é multiplicativo e constante em todos os estratos das variáveis que estão sendo controladas e, ainda, que o logaritmo da chance de adoecer aumenta linearmente com a intensidade da exposição à variável. É possível, no entanto, modificar-se o modelo para acomodar mudanças nestas premissas, embora com freqüência os modelos sejam usados sem essas adaptações.

Uma outra desvantagem do uso de regressão logística é a não-visualização dos dados, o que limita a habilidade de julgar a validade das conclusões apresentadas no estudo. O uso de regressão logística deve ser precedido, ou usado em paralelo, de métodos pertinentes mais simples que permitam explorar e familiarizar o investigador e o leitor com os dados existentes.

FRAÇÕES ATRIBUÍVEIS EM ESTUDOS CASO-CONTROLE

Em estudos caso-controle não é possível a estimativa direta das taxas de incidência para as diferentes categorias de exposição ou para a população como um todo, a não ser que se tenham informações sobre a fração de amostragem para casos e controles (em outras palavras, sobre o número de pessoas na população que produziu os casos, e a proporção de todos os casos da população que foi incluída no estudo). Do mesmo modo, também não é possível calcular os riscos (taxas) atribuíveis entre expostos ou para a população total (ver Capítulo 9 – Medidas de Associação e Medidas de Impacto). Entretanto, é possível calcular as frações atribuíveis nos expostos e na população.

A fração atribuível nos expostos (FAE) é calculada da seguinte maneira:

$$FAE = \frac{OR - 1}{R}$$

A fração atribuível na população (FAP) pode ser calculada de duas formas:

$$FAP = \frac{\pi \times (OR - 1)}{OR}$$

$$FAP = \frac{P \times (OR - 1)}{P \times (OR - 1) + 1}$$

onde π é a proporção de expostos entre os casos e P é a proporção de expostos na base populacional, ou seja, entre os controles (desde que os controles sejam representativos da população em termos de idade, sexo etc.).

O NOVO OLHAR, A PREMISSE DA RARIDADE DA DOENÇA E A SELEÇÃO DE CONTROLES

Retomamos aqui a discussão iniciada anteriormente sobre o fato de que a seleção apropriada de controles permite a estimativa do risco relativo (tanto da razão de taxas quanto da razão de incidência) mesmo em doenças comuns.

Este “novo olhar” considera que para todo estudo caso-controle existe uma coorte subjacente (imaginária ou concreta) composta pela população de indivíduos de onde os casos se originaram (“base populacional”).

Estudos caso-controle podem ser conduzidos nas situações em que a coorte subjacente é do tipo fechada. Em estudos de coorte com populações fechadas, a medida de associação de maior interesse é a razão de riscos (risco relativo), ainda que a razão de taxas possa ser calculada. Nesta situação também se pode aplicar a idéia de que o estudo caso-controle nada mais é do que uma estratégia eficiente de amostragem da base populacional.

Vamos imaginar primeiro uma coorte fechada, em que a população base é definida em um momento e seguida ao longo do tempo para detecção da ocorrência de novos casos, e o cálculo das incidências acumuladas (riscos) de acordo com níveis de exposição leva diretamente à estimativa de riscos relativos. Imaginemos que, após o término do seguimento, alguns investigadores desejam avaliar a associação entre o nível de prostaglan-

dina sérica medida no início do estudo e a ocorrência de infarto do miocárdio ao final de dez anos de seguimento. Imaginemos, também, que este exame laboratorial não foi realizado na época do recrutamento da coorte, porém existem amostras de sangue estocadas. Uma solução seria pegar as amostras sanguíneas de todos os membros da coorte e realizar o exame requerido para o estudo. No entanto, este pode ser um exame muito caro e inviável de ser realizado em todos os indivíduos. Uma alternativa é realizar um estudo caso-controle que se origina dentro da coorte. Todos os casos que ocorreram ao longo dos dez anos são selecionados e, após a realização dos exames para mensuração dos níveis de prostaglandina sérica nas amostras de sangue estocadas, irão prover informação sobre a chance de exposição entre casos (a/c). A seguir, amostra-se também uma fração dos indivíduos que estavam presentes no início da coorte (controles). Estes irão prover informação sobre a chance de exposição (b/d) na base populacional que deu origem aos casos. A razão de chances de exposição nesta situação é um estimador da razão de riscos, sem que se necessite da premissa de raridade da doença. Vale notar que entre controles assim selecionados é possível que sejam detectados indivíduos que viriam a se tornar casos, mas isto não é problema, pois também em estudos da coorte o denominador do risco inclui pessoas que eram livres de doença no início do estudo e vieram a se tornar doentes posteriormente. Este subtipo de estudo caso-controle é também denominado “caso-base”. Nesta situação controles são selecionados para representar a população sob risco no início do período de interesse.

Controles também poderiam ser selecionados dentre os membros da coorte (correspondente ao estudo caso-controle) que estão livres de doença no momento em que cada caso ocorre. Nesta situação os controles são selecionados para representar o conjunto das pessoas-tempo na coorte subjacente ao estudo caso-controle. Os controles são pareados aos casos no que diz respeito ao tempo de seguimento: cada controle representa a população em risco no momento em que o caso ao qual ele está pareado foi diagnosticado. Neste subtipo de estudo caso-controle, também chamado “caso-coorte”, a *odds ratio* é um estimador da razão de taxas de incidência sem qualquer premissa acerca da raridade de doença.

Vamos agora considerar uma população dinâmica, em estado de equilíbrio (*steady state*), composta por mulheres acima de 50 anos e residentes de uma cidade em particular. Durante um período de cinco anos a composição da população se modifica com a entrada de novas mulheres em decorrência de imigração ou por completarem 50 anos. Também se modifica pela saída

de mulheres em decorrência de morte ou emigração. No cômputo geral o tamanho total da população não muda. Ao longo destes cinco anos pode ser identificado o total de casos de fraturas de colo de fêmur ocorridas em mulheres acima de 50 anos que usavam ou não terapia de reposição hormonal (fator de exposição). Também se pode calcular o total de pessoas-tempo para as mulheres expostas (P_1) e não-expostas (P_0) à terapia hormonal. A medida de efeito de interesse neste caso é a razão de taxas. Vejamos como uma simples reexpressão algébrica da razão de taxas de incidência (RTI), em geral a medida de associação de interesse em estudos de coorte com população dinâmica, pode nos auxiliar a compreender esta visão:

$$RTI = \frac{\frac{a}{P_1}}{\frac{c}{P_0}} = \frac{\frac{a}{c}}{\frac{P_1}{P_0}}$$

A razão (a/c) é denominada de chance de exposição (*exposure odds*) entre os casos e é observada diretamente tanto em estudos de coorte quanto em estudos caso-controle. Em estudos caso-controle a razão de taxas não pode ser calculada diretamente porque a chance de exposição na base populacional (P_1/P_0) não é observada. No entanto, P_1/P_0 pode ser estimada através da razão entre indivíduos expostos e não-expostos sob risco de adoecer que existem na base populacional de onde os casos se originaram. Amostrando-se indivíduos durante o período em que os casos foram identificados, está-se levando em consideração não apenas seus números relativos, mas também a duração do tempo que eles passaram no estado de exposto ou não-exposto (isto é, pessoa-tempo na base populacional). Então, em um estudo caso-controle, os controles passam a ser vistos como um grupo de pessoas em que o *status* de exposição provê informação acerca da distribuição da exposição na base populacional que deu origem aos casos. Neste caso, a OR é um estimador da razão de taxas de incidência sem levar em conta a premissa de raridade de doença. Objetivamente, em todas estas situações e sem levar em conta a raridade da doença, a razão entre controles expostos e controles não-expostos (b/d), também denominada de chance de exposição entre controles, é uma estimativa de (P_1/P_0) , sejam P_1 e P_0 indivíduos originários de uma coorte fechada, ou pessoa-tempo despendido na base populacional (coorte aberta).

Vantagens e Limitações de Estudos Caso-controle

Vantagens

- Relativamente barato (apenas uma amostra da coorte subjacente, os controles, é necessária para se obter informação sobre a distribuição da exposição na base populacional).
- Permite a investigação simultânea de uma maior variedade de possíveis fatores de risco.
- Útil para o estudo de doenças raras (de novo porque apenas uma amostra da coorte subjacente, os controles, é necessária para se obter informação sobre a distribuição da exposição na base populacional).
- Porque o tamanho da amostra é geralmente menor do que em estudos de coorte, pode-se empregar exames e/ou testes caros e/ou laboriosos para determinar a exposição.
- Relativamente rápido (não há a necessidade de se esperar pelo desenvolvimento da doença, na medida em que os indivíduos são selecionados com base no *status* de doença).
- É relativamente mais rápido, é mais fácil controlar a consistência das técnicas de mensuração.

Limitações

- Mais suscetível a vieses de seleção (porque necessita selecionar controles que sejam representativos da população que deu origem aos casos).
- Mais suscetível a vieses de informação (na medida em que o *status* de exposição é determinado após o diagnóstico da doença).
- Dificuldades para assegurar a correta seqüência de eventos.
- Inadequado para a investigação de exposições raras, a não ser que o risco atribuído à exposição na população de estudo seja muito alto.
- Não é possível obter estimativas de incidência da doença (a não ser que o tamanho da população que produz os casos seja conhecido).

BIBLIOGRAFIA

1. Berkson J. Limitation of the application of the fourfold analysis to hospital data. *Biometrics Bulletin*, 2: 47-53, 1946.
2. Blin P et al. Evaluation of the protective effect of BCG vaccination. *Tuberc.*, 67: 283-88, 1986.
3. Breslow N, Day N. Statistical Methods in Cancer Research, Vol.1. The analysis of case-control studies. International Agency for Research on Cancer, Lyon, 1980.
4. Brinton et al. Oral contraceptives use and risk of invasive cervical cancer. *Int. J. Epidemiol.*, 90(1): 4-11, 1990.
5. Casagrande J et al. Exogenous oestrogens and breast cancer in women with natural menopause. *J. Natl Cancer Inst.*, 56: 839-841, 1976.
6. Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *Brit. Med. J.*, 13(2):1271-86, 1952.
7. Filho VW et al. Effectiveness of BCG vaccination against tuberculous meningitis. *WHO Bull. OMS*, 68: 69-74, 1990.
8. Gardner MJ. Considerations in the choice of expected numbers for appropriate comparisons in occupational cohort studies. *Med. Lav.*, 77: 23-47, 1986.
9. Gordis L. Epidemiology. W. B. Saunders, Philadelphia, 1996.
10. Hennekens CH, Buring JE. Epidemiology in Medicine. Little Brown & Co., 1987, p:132-52.
11. Horowitz RI, Feinstein AR. alternative analytic methods for case control studies of oestrogens and endometrial cancer. *N. Engl. J. MEd.*, 289: 1089-1094, 1978.
12. Kelsey JL, Thompson WD, Evans AS. Methods in Observational Epidemiology. Oxford University Press, p:148-86, 1986.
13. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research. Van Nostrand Reinhold, New York, 1982.
14. Lilienfeld AM, Lilienfeld DE. Foundations of Epidemiology (2nd Edition). Oxford Univ Press, 1980.
15. MacMahon B, Trichopoulos D. Epidemiology: principles and methods. 2nd ed. Little, Brown & Co., Boston, 1996.
16. Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol.* 103:226-235, 1976.
17. Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. *Int J Epidemiol.* 19:205-13, 1990.
18. Rodrigues LC, Smith PG. Case control approach to vaccine evaluation: efficacy and adverse reactions. *Epidemiological Rev.* 21:56-72, 1999.
19. Rosner B. Fundamental of Biostatistics. 4th ed. Duxbury, Belmont, 1995.
20. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Lippincott-Raven, Philadelphia, 1998.
21. Sackett DL. Bias in analytical research. *J. Chron. Dis.*, 32: 51-63, 1979.

22. Schlesselman JJ. Case-control studies - design, conduct, analysis. Oxford University Press, 1982.
23. Shapiro S, Slone D, Rosenberg L et al. Oral-contraceptives use in relation to myocardial infarction. Lancet. 1:743-747, 1979.
24. Szklo M, Javier Nieto F. Epidemiology: beyond the basics. Aspen, Gaithersburg, 2000.
25. Victora et al. Infant feedings and deaths due to diarrhoea. Am J. of Epidemiol, 129:1032-41, 1989.
26. Wacholder S, Boivin JF. External comparisons with the case cohort design. Am J Epidemiol. 73:1-11, 1986.
27. Walker AM. Observation and Inference. Epidemiology Resources Inc., Newton Lower Falls, 1991.

Estudos de Coorte

Cláudia Medina Coeli
Eduardo Faerstein

INTRODUÇÃO

Estudos de coorte são estudos observacionais em que a situação dos participantes quanto à exposição de interesse determina sua seleção para o estudo, ou sua classificação após a inclusão no estudo. Esses indivíduos são monitorados ao longo do tempo para avaliar a incidência de doença ou de outro desfecho de interesse¹.

Assim como outros tipos de pesquisas epidemiológicas, estudos de coorte podem ser conduzidos com diversas finalidades, incluindo a avaliação da etiologia de doenças (ex.: associação entre fumo e câncer de pulmão) ou de sua história natural (ex.: evolução de pacientes HIV-positivos) e o estudo do impacto de fatores prognósticos (ex.: marcadores tumorais e evolução de câncer), intervenções diagnósticas (ex.: impacto da realização de colpocitologia sobre a mortalidade por câncer de colo uterino) e terapêuticas (ex.: impacto do tipo de tratamento cirúrgico de fraturas de colo do fêmur em idosos sobre a mortalidade).

Diferentes tipos de “exposição” podem ser estudadas, incluindo exposições ambientais propriamente ditas, de natureza química ou física (ex.: radiação, exposições ocupacionais), comportamentos relacionados à saúde (ex.: tabagismo, dieta, atividade física), características biológicas (ex.: pressão arterial, colesterol sérico), e fatores socioeconômicos (ex.: escolaridade, renda), entre outros². Em análises específicas de dados prospectivos, cada um desses fatores pode ser explorado como a exposição principal de interesse, ou como

fatores que podem confundir ou modificar a associação entre a exposição principal e o desfecho.

Da mesma forma, vários tipos de desfechos podem ser investigados, tais como a mortalidade (geral ou por uma determinada causa), a incidência de doenças (ex.: câncer de pulmão) e as mudanças ao longo do tempo nos níveis de um marcador biológico de progressão de doença (ex.: decréscimo na contagem de células CD4+ em pacientes HIV-positivos) ou de aspectos comportamentais (ex.: níveis de atividade física)³.

Um determinado fator pode ser tratado ora como exposição, ora como desfecho. Por exemplo: em um estudo hipotético que procure avaliar a associação entre hábitos de dieta e de atividade física com a ocorrência de obesidade, esta última seria tratada como desfecho. Em outro estudo, que fosse desenvolvido para avaliar a associação entre os diferentes níveis de obesidade e os padrões de mortalidade, a obesidade seria tratada como a exposição de interesse.

Em estudos de coorte, o pesquisador define datas de início e de final do seguimento dos participantes; durante este intervalo são colhidas informações sobre a exposição de interesse, sobre covariáveis (variáveis confundidoras e modificadoras) e sobre o(s) desfecho(s). É importante ressaltar que o momento da coleta, pelo investigador, das informações sobre esses eventos e características, nem sempre coincide com o momento da sua ocorrência de fato. Os estudos de coorte são classificados em *concorrentes* (também conhecidos

como prospectivos) e *não-concorrentes* (também denominados retrospectivos ou históricos) em função da relação temporal entre o início do estudo (coleta das informações) e a ocorrência do desfecho. Nos estudos *não-concorrentes*, tanto a exposição como o desfecho já ocorreram antes do início do estudo. Nos estudos *concorrentes*, a exposição pode (ou não) já ter ocorrido antes do início do estudo, mas o desfecho ainda não ocorreu; desta maneira, o pesquisador deverá acompanhar a população do estudo por um tempo suficientemente longo para permitir que os casos do desfecho de interesse ocorram⁴.

Na parte superior da Figura 12.1 apresentamos o diagrama relativo à investigação da associação entre baixo peso ao nascer (peso inferior a 2.500 g) e a mortalidade neonatal precoce (mortalidade entre nascidos vivos durante a primeira semana de vida) avaliada em um estudo de coorte *concorrente* conduzido no Brasil por Victora, Barros & Vaughan⁵. Neste estudo todos os nascimentos hospitalares ocorridos na cidade de Pelotas, Rio Grande do Sul, no ano de 1982, foram investigados por uma equipe de pesquisadores que visitavam diariamente as quatro maternidades da cidade. As mães eram entrevistadas e pesadas no hospital, e os recém-nascidos eram pesados logo após o parto. Os cartórios de registro da cidade eram visitados regularmente para a identificação dos óbitos. Trata-se de um estudo *concorrente*, pois o desfecho (mortalidade neonatal precoce) ainda não havia ocorrido antes do início do estudo. A informação sobre o baixo peso ao nascer também foi obtida concorrentemente, no entanto, a história de aborto prévio, outra exposição avaliada no estudo, foi colhida de forma retrospectiva a partir da informação fornecida pela mãe durante a entrevista.

Em contraste, o trabalho realizado por Handler, Kistin e Davis⁶ é um exemplo de estudo de coorte *não-concorrente*. Estes autores procuraram investigar a associação entre o consumo de cocaína materno e a ocorrência de diferentes desfechos perinatais. Na parte inferior da Figura 12.1 é apresentado o diagrama deste estudo relativo ao desfecho prematuridade (nascimento com menos de 37 semanas de gestação). Os autores utilizaram como fonte de dados o banco de dados de registros perinatais da Universidade de Illinois (*University of Illinois Hospital Perinatal Registry*). Todos os registros de nascimentos ocorridos durante o ano de 1988 foram revisados para a obtenção das informações sobre o uso de cocaína pela mãe, de outras covariáveis de interesse (ex.: tabagismo) e dos desfechos perinatais. A mulher era considerada exposta se o uso de cocaína tivesse sido indicado em algum momento nos registros do pré-natal ou do trabalho de par-

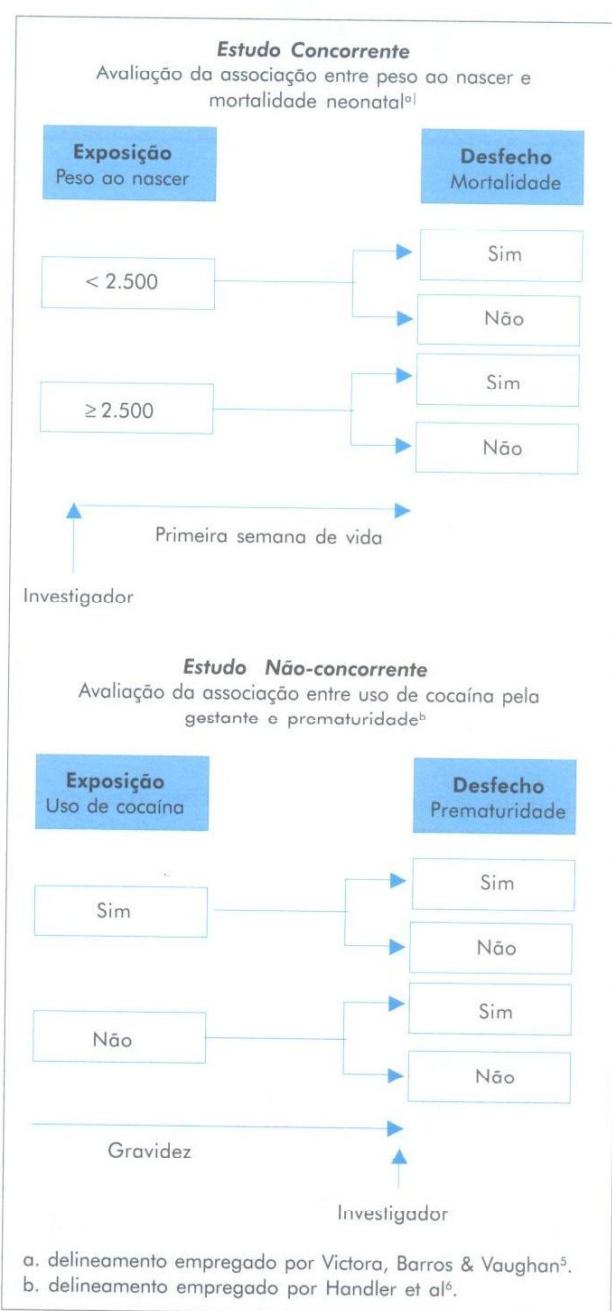


Fig. 12.1 — Estudos de coorte concorrente e não-concorrente.

to. A informação registrada no banco de dados sobre o uso de cocaína foi em alguns casos baseada na informação fornecida pela gestante, enquanto em outros foi baseada na realização de *screenings* toxicológicos na mãe ou na criança. Trata-se de um estudo de coorte *não-concorrente*, pois todas as informações (exposição, covariáveis e desfecho) foram colhidas após o desfecho já ter ocorrido. As vantagens e limitações da utilização de cada um destes delineamentos serão discutidas ao longo deste capítulo.

ASPECTOS DOS DESENHOS DE ESTUDOS DE COORTE

Tipos de População de Estudo

Os estudos de coorte podem investigar a experiência de populações fixas ou dinâmicas. Denomina-se população fixa o conjunto de pessoas que apresentam um *evento* comum (restrito no tempo e no espaço) que caracteriza a sua admissão na coorte (o nascimento, por exemplo)⁷. Portanto, os grupos de expostos e não expostos a esse evento são definidos no momento do início do estudo, não havendo ao longo do seguimento entrada de novos participantes, ou mudanças na sua situação de exposição. Espera-se que todos os participantes sejam seguidos por um tempo fixo definido pelo pesquisador (a primeira semana de vida para avaliar a mortalidade neonatal precoce, por exemplo), embora alguns participantes possam ter um tempo menor de seguimento em função de desenvolverem o desfecho de interesse antes do final do período de seguimento previamente definido, abandonarem o estudo ou morrerem (devido à doença em estudo ou a outra doença).

Entretanto, nem sempre a exposição de interesse em um estudo de coorte é uma característica permanente, como o peso ao nascer. Com freqüência trata-se de um fator modificável, como por exemplo o hábito de fumar. Durante o seguimento, alguns indivíduos podem iniciar ou abandonar o hábito de fumar, ou modificar o consumo diário de cigarros. Na Figura 12.2 é apresentado o diagrama de evolução da exposição e de covariáveis em um participante de uma coorte hipotética acompanhada para avaliar a associação entre consumo de cigarros e incidência de doença coronariana. Percebe-se que tanto a exposição de interesse (consumo de cigarros) como a idade (possível fator confundidor) sofrem modificações ao longo do tempo, enquanto o sexo (possível fator modificador) permanece constante.

Essa natureza variável de alguns fatores de exposição faz com que o estudo de uma população dinâmica seja a maneira mais apropriada de investigar o papel desses fatores em relação a desfechos de interesse. Denomina-se população *dinâmica* ao conjunto de pessoas que apresenta um estado que define a sua participação na população durante o tempo em que apresente esta característica⁷. Uma população dinâmica pode constantemente incorporar ou perder membros durante seu seguimento. Por exemplo, a população de uma cidade é sempre dinâmica, pois está constantemente incorporando novos membros (por nascimento e imigração) e perdendo outros (por emigração e morte) — um indivíduo fará parte da população da cidade enquanto nela residir.

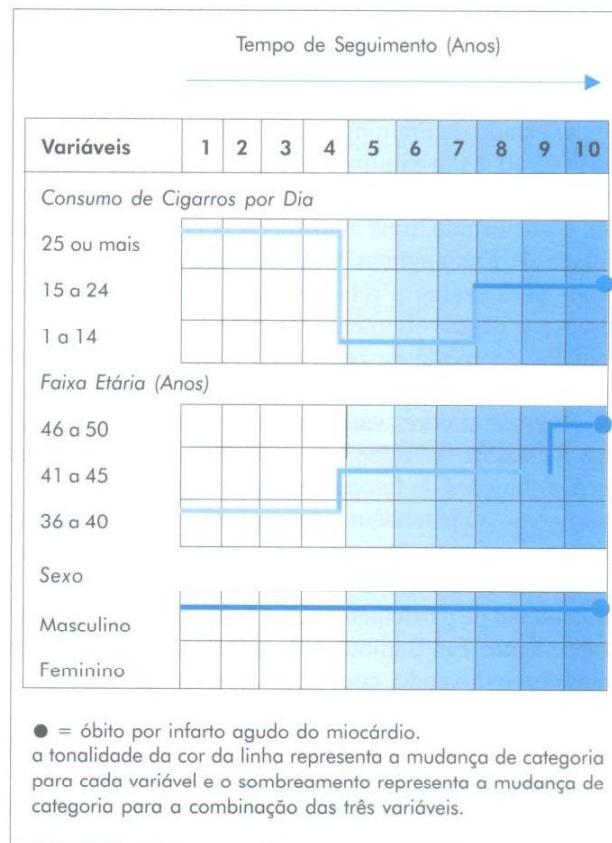


Fig. 12.2 — Evolução da situação da exposição e de covariáveis de um participante que tinha 37 anos no início de um estudo de coorte realizado para avaliar a associação entre fumo e doença coronariana.

Ao se tratar os grupos de exposição como populações dinâmicas, permite-se tanto a adição de novos indivíduos após o início do estudo, como mudanças na situação de exposição ao longo do seguimento (ex. de fumante para não-fumante). Assim, ao contrário do que ocorre na população *fixa*, na população *dinâmica* a exposição não se refere a um evento, mas sim a um estado que pode mudar ao longo do tempo.

Seleção da População de Estudo

Estudos de coorte podem ser realizados através do seguimento da amostra de uma população geral, i.e., de uma área bem definida geográfica ou administrativamente (ex.: a população de uma cidade), ou a partir do estudo de grupos especiais. Estes últimos são selecionados por apresentarem características que facilitem o recrutamento de participantes, a coleta das informações e o seguimento (grupos populacionais restritos, como os trabalhadores de uma indústria, por exemplo), ou em função de estarem especialmente expostos a um fator

que é raro na população geral (grupos de exposição especiais – pacientes expostos à radiação terapêutica, por exemplo).

O estudo de Framingham⁸, iniciado ao final da década de 40, é um exemplo de estudo de coorte baseado no acompanhamento de uma amostra da população geral. Neste estudo, uma amostra da população adulta da cidade de Framingham (Massachusetts, EUA) foi seguida para avaliar a relação entre diversos fatores de risco e o desenvolvimento de doença cardiovascular. Estudos de coorte de população geral apresentam diversas vantagens, como a possibilidade, em princípio, de observar maiores variações nos níveis de exposição aos fatores de interesse, e a possibilidade de conhecer as distribuições de fatores de risco na população geral, bem como suas tendências temporais. Entretanto, tais estudos em geral são caros e de difícil operacionalização. Pesquisadores buscam situações propícias ao recrutamento de participantes com boas oportunidades de seguimento. Em Bambuí, Minas Gerais, pesquisadores vêm desenvolvendo, com poucas perdas de seguimento, um estudo longitudinal entre cerca de 1.700 idosos. Os autores mencionam a importância, nesse caso, da escolha de uma comunidade pequena, demograficamente estável e com tradição de participação em outras pesquisas (Doença de Chagas)⁹.

Uma alternativa que facilita o desenvolvimento de estudos de coorte é o seguimento de grupos populacionais restritos, tais como usuários de seguros de saúde, empregados de empresas e profissionais de saúde. Nestes grupos, a existência de cadastros com informações atualizadas sobre endereços pode facilitar tanto o recrutamento quanto o seguimento dos participantes. Estes grupos apresentam características que facilitam a coleta das informações. Por exemplo, usuários de seguros de saúde costumam buscar atendimento em serviços de saúde específicos contratados pelos planos, o que facilita a revisão de registros de atendimento. Com relação à coleta das informações, os profissionais de saúde constituem um grupo à parte em função da maior capacidade para fornecerem informações acuradas sobre as exposições e os desfechos de interesse. Por exemplo, em 1976 foi iniciado um estudo de coorte em que enfermeiras americanas com idade entre 30 e 55 anos foram acompanhadas para avaliar fatores de risco para câncer e doença cardiovascular (*The Nurse's Health Study*)¹⁰, sendo as informações relativas às exposições e desfechos colhidas a partir do envio de questionários por correio a cada dois anos.

O Estudo Pró-Saúde¹¹, iniciado em 1999, pretende investigar o papel de ampla gama de determinantes sociais e fontes de desigualdade (p. ex., condições socio-

econômicas atuais e na infância, gênero, cor/etnia), eventos de vida e padrões de apoio social, sobre comportamentos de saúde e morbidade física e mental, em uma coorte de cerca de quatro mil funcionários de uma universidade pública no Rio de Janeiro. No caso, considerou-se que essa população restrita deverá possibilitar a realização de um estudo com boas taxas de participação e seguimento, devido à sua relativa estabilidade e natureza “cativa”. Os patamares de escolaridade (maioria com segundo grau ou mais) deverão permitir a coleta de informação de qualidade adequada, com a utilização de uma variedade de métodos de coleta de dados, com eficiência de gastos financeiros (ex.: questionário autopreenchível, com apoio de aplicadores).

Tanto os estudos de coorte que acompanham amostras da população geral, como os baseados no seguimento de grupos populacionais restritos, são úteis para a avaliação de fatores de exposição freqüentes, tais como fumo, atividade física e características da dieta. O processo de formação dos grupos de comparação envolve em primeiro lugar a definição da população que será acompanhada (ex.: amostra de adultos residentes em Framigham com idade variando entre 30 e 59 anos) e posterior classificação segundo os diferentes níveis da variável de exposição de interesse (colesterol plasmático em mg/dL < 200; 200 a 219; 220 a 239; 240 a 259; 260 ou +). Esta estratégia permite que a formação dos grupos de comparação (expostos e não-expostos) se faça no interior da coorte, garantindo, portanto, que esses grupos pertençam à mesma base populacional (Fig. 12.3). Estes estudos permitem que várias hipóteses possam ser avaliadas simultaneamente, isto é, tanto pode ser avaliada a associação de um determinado fator de risco com múltiplos desfechos, como também podem ser exploradas as associações de diferentes fatores de risco com diversos desfechos.

Uma tendência mais recente é o desenvolvimento de estudos multicêntricos. A seleção de participantes em centros localizados em áreas geográficas distintas contribui para uma maior generalização dos resultados alcançados¹². Em 2005 foi formado o consórcio multidisciplinar ELSA, reunindo cerca de 40 pesquisadores brasileiros de seis estados (RJ, BA, ES, MG, RS, SP), com a meta de implementar o Estudo Longitudinal de Saúde do Adulto (ELSA): um estudo de coorte multicêntrico com cerca de 15 mil funcionários de instituições públicas de ensino superior e pesquisa, para investigar o desenvolvimento de doenças crônicas, principalmente as doenças cardiovasculares e o diabetes. A natureza do projeto é multidisciplinar, envolvendo os campos da epidemiologia, clínica médica (cardiologia, endo-

crinologia), oftalmologia, psiquiatria, fisiologia, biologia molecular, imunologia, nutrição, bioestatística e antropologia entre outros. (http://www.famed.ufrgs.br/ppg/PPG_Epi/elsa – acessado em: 06/04/2006).

Quando há interesse em investigar o papel de exposições raras na população geral (radiação ionizante, por exemplo), é mais eficiente o seguimento de grupos populacionais altamente expostos ao suposto fator de risco, sejam vítimas da bomba atômica expostas à radiação ionizante, residentes próximos a áreas de despejos tóxicos, ou trabalhadores sujeitos a exposições ocupacionais (trabalhadores de fábricas de asbesto, por exemplo). Nesses casos, o acompanhamento de coortes de população geral (ou mesmo de grupos populacionais restritos) é ineficiente, pois tende a haver grande desequilíbrio nas proporções de indivíduos expostos e não-expostos nessas populações. Em consequência, seria necessário estudar amostras muito grandes para garantir um número razoável de indivíduos expostos.

Tipos de População de Estudo

- População fechada
- População dinâmica

Seleção da População de Estudo

- População geral
- Grupos populacionais restritos
- Grupos especiais de exposição

Quando se utilizam grupos especiais de exposição, o processo de formação dos grupos de comparação (expostos e não-expostos) se faz de forma diferente daquela empregada quando se acompanham coortes de população geral ou de grupos populacionais restritos. Em vez de se identificar a coorte e internamente classificar seu membros em expostos e não-expostos, a seleção dos participantes é feita segundo a situação de exposição. Em outras palavras, em primeiro lugar é identificada a coorte altamente exposta ao fator de interesse e após se escolhe uma coorte não-exposta (porém similar à coorte exposta em relação a outras características) para servir de grupo-controle (por exemplo, em estudos ocupacionais escolhe-se como controle profissionais da mesma indústria que trabalham em atividades não-ligadas à exposição) (Fig. 12.3).

Outra estratégia que pode ser utilizada é comparar a incidência do desfecho de interesse (óbito, por exem-

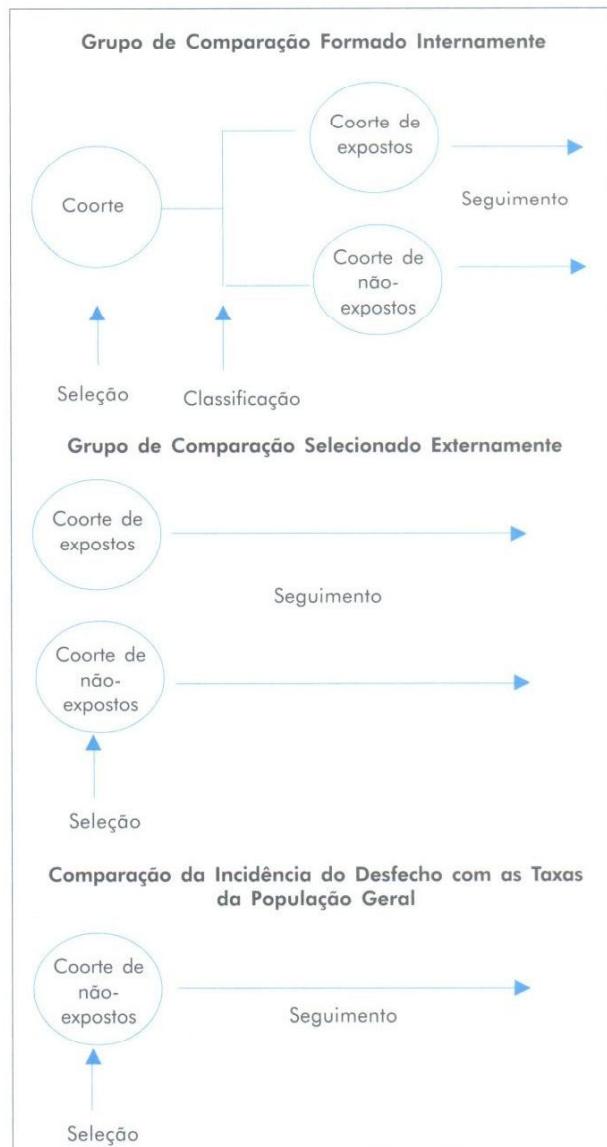


Fig. 12.3 — Seleção do grupo de comparação (não-expostos) em estudos de coorte.

pleno) na coorte de exposição especial com a incidência observada na população geral no período em que a coorte está sendo acompanhada. Dessa forma, avalia-se se a ocorrência do desfecho observada na coorte exposta difere daquela que seria esperada caso a exposição estudada não fosse um fator de risco* (Figura 12.3). Johansen & Oslen¹³ avaliaram a associação entre exposição a campos magnéticos e asbesto e a ocorrência

* Esta estratégia implica no cálculo de taxas de incidência padronizadas pelo método indireto, tema que é abordado no Capítulo 3 – Indicadores de Saúde.

de câncer, comparando a ocorrência deste desfecho em uma coorte de trabalhadores de empresas de eletricidade da Dinamarca com aquela observada na população total da Dinamarca. Este estudo foi possível de ser realizado em função de a Dinamarca contar com um sistema de registro de câncer de base populacional (*The Danish Cancer Registry*). Esta é uma das limitações deste tipo de estratégia de avaliação, isto é, ela é restrita ao estudo da mortalidade e da incidência de doenças que contem com registros de base populacional. As limitações do uso de grupos de exposição especial serão discutidas na seção sobre validade, deste capítulo.

Fontes de Informação

Diferentes fontes de informação podem ser utilizadas em estudos de coorte; sua escolha depende de fatores variados, tais como a disponibilidade de dados de registros (registros médicos ou industriais, por exemplo), a natureza da exposição e do desfecho a serem investigados, e os recursos disponíveis para a realização do estudo.

As fontes em geral empregadas englobam dados de registros, questionários de autopreenchimento, entrevistas por telefone, entrevistas pessoais, exame físico, testes médicos (electrocardiograma, dosagem de glicemia, contagem de células CD4+, por exemplo) e medidas ambientais (ex.: amostras de ar e água). Em sua maioria, essas fontes também podem ser empregadas para a coleta de informações sobre covariáveis e desfechos.

O pesquisador deve sempre procurar escolher fontes onde as informações registradas apresentem a melhor qualidade possível, e que possam, em princípio, ser obtidas para todos os participantes do estudo ao longo de todo o período de seguimento. Muitas vezes são utilizadas mais de uma fonte de dados em um único estudo, em função de estarem sendo estudadas exposições e desfechos de naturezas diversas. No estudo de Framingham⁹ as informações sobre os fatores de risco foram colhidas através de entrevista (ex.: hábito de fumar), exame físico (ex.: pressão arterial) e exames bioquímicos (ex.: dosagem de colesterol). Outro motivo para a utilização de mais de uma fonte de dados é melhorar a qualidade da informação. No Nurses¹⁰ Health Study, sempre que uma participante relata um episódio de acidente vascular cerebral nos questionários de seguimento (autopreenchimento), os pesquisadores solicitam sua autorização para revisão dos registros médicos correspondentes.

Os dados de registros podem ser empregados em estudos de coorte *concorrentes* e *não-concorrentes* – nestes últimos, representam a única fonte de informa-

ção possível. No estudo realizado por Johansen & Oslén¹⁴, as informações demográficas dos trabalhadores e aquelas necessárias para a classificação dos níveis de exposição ao asbesto e ao campo magnético (data de admissão e demissão do emprego, função, área física de atuação dentro da empresa) foram obtidas a partir dos registros das empresas de eletricidade da Dinamarca, do banco de dados do fundo de pensão obrigatório e da folha de pagamento dos funcionários públicos. Dados sobre a incidência de câncer foram obtidos do registro de câncer de base populacional daquele país.

Os dados de registros representam a fonte de menor custo, especialmente quando os sistemas de registro se encontram informatizados. Quando a disponibilidade de dados em bancos informatizados se refere à exposição e ao desfecho, é possível eliminar simultaneamente os problemas da recusa em participar do estudo e de abandono do seguimento, ficando a perda de participantes restrita aos casos em que por algum problema não seja possível localizar algum registro.

Dados de registros médicos podem ser melhores que entrevistas ou questionários de autopreenchimento quando se necessita colher informações detalhadas sobre procedimentos médicos. Outra vantagem é o fato de que muitas informações são registradas na mesma época da sua ocorrência, o que reduz a possibilidade de viés de memória associado a entrevistas conduzidas posteriormente. Por exemplo, em um estudo de coorte não-concorrente desenvolvido com o objetivo de avaliar fatores de risco para a mortalidade um ano após cirurgia para correção de fratura de colo do fêmur poderiam ser obtidas informações detalhadas sobre o tipo de procedimento cirúrgico realizado e o risco cirúrgico do paciente a partir da revisão dos registros da hospitalização.

Estudos de coorte baseados na utilização de registros preexistentes (médicos ou de outra natureza) podem estar limitados em relação a vários aspectos. A qualidade dos dados pode ser inadequada, uma vez que a informação não é colhida de forma padronizada, para fins científicos. Além disso, dados sobre covariáveis de interesse (potencialmente confundidoras ou modificadoras de efeito) em geral não estão disponíveis, limitando a interpretação dos resultados.

Investigações que incluem contatos com os participantes, através da utilização de diversas fontes possíveis, trazem como vantagens a possibilidade de colher informações de melhor qualidade sobre uma maior variedade de fatores de interesse. Essas fontes, entretanto, não são imunes a problemas na qualidade. Questionários enviados por correio, por exemplo, costu-

mam apresentar uma elevada proporção de não-resposta e de não-preenchimento de dados. Em entrevistas pessoais, o comportamento dos entrevistadores pode influenciar as respostas dos participantes; as informações sobre exposições ocorridas anteriormente à entrevista podem não ser relatadas de forma correta e as exposições que se referem a comportamentos estigmatizáveis (ex.: consumo de álcool) podem ser informadas de forma incorreta.

Em geral, estudos baseados em contatos com os participantes geram custos mais elevados do que aqueles baseados em dados já registrados. O custo é menor quando o contato é realizado através de questionários de autopreenchimento enviados pelo correio, e cresce de forma progressiva com a utilização de entrevistas por telefone, entrevistas pessoais, exame físico e exames médicos. Alguns estudos podem se tornar especialmente complexos e caros, quando seus participantes são submetidos a exames mais dispendiosos. Por exemplo, na coorte do *Multi-Ethnic Study of Atherosclerosis* (MESA), que investiga a progressão da doença cardiovascular subclínica, todos os 6.500 participantes realizaram, durante a coleta de dados de base do estudo, tomografia computadorizada para medida do cálcio nas artérias coronarianas, ressonância magnética cardíaca para medida da massa e função ventricular, e ultrasonografia das artérias carótidas, entre outros exames. O custo também tende a aumentar quando múltiplos contatos são necessários ao longo do seguimento para a atualização das informações sobre exposição, covariáveis e desfecho. Finalmente, os indivíduos selecionados podem se negar a participar do estudo ou podem abandoná-lo ao longo do seguimento. Algumas medidas que podem ser implementadas visando reduzir estes problemas serão discutidas na seção sobre validade.

Fontes de Informação em Estudos de Coorte

- Dados de registro
- Questionários de autopreenchimento
- Entrevistas por telefone
- Entrevistas pessoais
- Exame físico
- Testes médicos
- Medidas ambientais

Análise de Dados

Em estudos de coorte a população a ser analisada é definida segundo a situação de exposição, sendo seguida para a observação da ocorrência de casos novos de

doença (ou de outro desfecho de interesse). Estas características permitem o cálculo direto das medidas de incidência. A maioria dos estudos de coorte tem como objetivo primário a comparação da incidência do desfecho entre indivíduos expostos (I_E) e não-expostos ($I_{\bar{E}}$), o que é feito a partir do cálculo de medidas de associação baseadas em diferença de incidências ($I_E - I_{\bar{E}}$) ou em razão de incidências ($I_E/I_{\bar{E}}$).

Dependendo do tipo de população de estudo (fixa ou dinâmica), e da duração do período de seguimento, define-se o tipo de medida de incidência (incidência acumulada ou taxa) e de associação (diferença ou razão de incidências acumuladas; diferença ou razão de taxas) mais indicadas para serem avaliadas em um determinado estudo.

Nesta seção apresentamos de forma sumária o uso destas medidas em situações com freqüência observadas nos estudos de coorte. Uma discussão mais aprofundada sobre as medidas de freqüência e de associação podem ser encontradas, respectivamente, no Capítulo 2 – Medidas de Freqüência de Doença e Capítulo 9 – Medidas de Associação e Medidas de Impacto. Adicionalmente, aspectos da análise de estudos de coorte, incluindo o cálculo de intervalos de confiança das medidas de freqüência de doenças e de associação, testes estatísticos e controle de confundimento, são apresentados no Capítulo 24 –Associação Estatística em Epidemiologia: Análise Bivariada.

A incidência acumulada é calculada em cada grupo de comparação (em expostos e não-expostos, ou nas diversas categorias de exposição) dividindo-se o número de indivíduos que desenvolvem o desfecho durante o período de observação pelo total de participantes no início do período de observação. Esta medida de incidência, assim como as medidas de associação correspondentes (diferença e razão entre incidências acumuladas), podem ser empregadas em estudos em que a exposição ocorra de modo bastante delimitado no tempo e o período de seguimento seja curto, portanto com reduzidas perdas de participantes. Por exemplo, a incidência acumulada poderia ser empregada em um estudo de coorte para identificar a fonte de um surto de botulismo. O botulismo é uma doença neuroparalizante causada por toxinas produzidas pelo *Clostridium botulinum*, podendo ser transmitida pela ingestão de alimentos contaminados; como os sintomas aparecem geralmente cerca de 12 a 36 horas após a exposição, um curto período de observação torna-se suficiente para observar a ocorrência dos casos.

A incidência acumulada não deve ser utilizada em situações que envolvam exposições pontuais, mas que demandem longos períodos de seguimento (ex.: avaliação de

incidência de câncer após um acidente nuclear). O prolongado período de seguimento aumenta a chance de ocorrem perdas de participantes, que mesmo sendo aleatórias (ver discussão sobre validade), levariam a estimativas enviesadas da incidência acumulada, pois os casos perdidos de observação não seriam contabilizados no numerador, mas continuariam sendo contabilizados no denominador.

Nestas situações, é mais adequado empregar as mesmas técnicas de análise utilizadas para a avaliação de populações dinâmicas, sendo indicado o cálculo das taxas de incidência nas coortes de expostos e não-expostos e das medidas de associação correspondentes (diferença e razão de taxas). A taxa é calculada dividindo-se o número de indivíduos que desenvolvem o desfecho durante o período de observação pelo total de pessoas-tempo em observação em cada grupo de exposição. Este último, por sua vez, é calculado a partir da soma do tempo em que cada indivíduo da coorte permanece em observação no respectivo grupo de exposição sem apresentar o desfecho.

A utilização de pessoas-tempo em observação no denominador do cálculo das taxas permite que as perdas de participantes, desde que aleatórias (ver seção sobre validade), não enviem o cálculo das taxas e das medidas de associação correspondentes, pois os indivíduos só são contabilizados no denominador durante o período que permanecem sob observação. De modo similar, acomodam-se mudanças na situação de exposição a partir da contabilização do total de pessoas-tempo e do cálculo da taxa da incidência dentro de cada estrato.

Por exemplo, o indivíduo representado na Figura 12.2, que no início do período de observação tinha 37 anos, contribuiu com quatro anos para o total de pessoas-tempo do estrato de “sexo masculino, consumo de 25 ou mais cigarros por dia e faixa etária de 36 a 40 anos” e com um ano para o estrato “sexo masculino, consumo de 15 a 24 cigarros por dia e faixa etária de 46 a 50 anos”. Neste último estrato ele desenvolveu um episódio de infarto agudo do miocárdio, sendo contabilizado também no numerador para o cálculo da taxa de incidência no estrato.

Técnicas de análise de sobrevida¹⁶ (ver Capítulo 26 – Análise de Sobrevida) também podem ser empregadas para a análise de estudos de coorte, onde perdas de participantes e mudanças na situação de exposição e de covariáveis ocorram ao longo do seguimento. Mais recentemente, métodos de análise longitudinal vêm sendo utilizados em estudos de coorte que envolvam medidas repetidas (mais de duas) de exposição e desfecho. Estas técnicas são úteis, por exemplo, para a avaliação de marcadores de progressão de doença observados repetidamente em membros da coorte (ex.: decréscimo na contagem de células CD4+ em pacientes HIV-positivos)³.

Por fim, a razão de incidência (ou de mortalidade) padronizada é utilizada quando se deseja comparar a incidência do desfecho de interesse (óbito, por exemplo) na coorte com a incidência observada na população geral no período em que a coorte está sendo acompanhada. Trata-se de uma razão entre os casos (ou óbitos) observados na coorte e aqueles que seriam esperados caso a coorte apresentasse as mesmas taxas de incidência (ou de mortalidade) estrato-específicas (por exemplo, por sexo e faixa etária) da população geral (para detalhes do cálculo, ver Capítulo 3 – Indicadores de Saúde).

Na Tabela 12.1 são apresentados os resultados de um estudo que avaliou a mortalidade 30 dias após cirurgia para correção de fratura de colo de fêmur em uma coorte de 606 idosos¹⁷. Esse estudo segue uma população fixa por um curto período de tempo, não sendo observadas perdas de seguimento. No exemplo, buscou-se comparar a mortalidade segundo o sexo (comparação interna à coorte) e a mortalidade na coorte como um todo (homens + mulheres) em relação à mortalidade da população geral. Para a primeira comparação pode ser empregada tanto a razão de taxas como a razão de incidências acumuladas, já que o tempo de seguimento é curto e não ocorreram perdas de seguimento. Nesse exemplo, embora as duas medidas tenham o mesmo valor numérico (1,5), a interpretação delas é diferente. Uma razão de incidências acumuladas de 1,5 indica que os homens apresentaram em relação às mulheres uma probabilidade 50% maior de morrer durante os primeiros 30 dias após a cirurgia, enquanto a razão de taxas indica uma força de mortalidade 50% maior nos homens, quando comparados às mulheres. Já a razão de mortalidade padronizada indica que para cada 11 óbitos observados na coorte seria esperado apenas um óbito, caso a coorte apresentasse as mesmas taxas de mortalidade específicas por sexo e faixa etária que a população geral.

Validade

No planejamento e análise de estudos de coorte, devemos levar em consideração que a observação (ou não) de uma associação entre a exposição e o desfecho de interesse pode ser parcial ou exclusivamente resultante da ocorrência de confundimento e/ou de vieses (seleção e informação).

Confundimento

O confundimento surge em função de diferenças inerentes nas probabilidades de adoecer entre a população de expostos e não-expostos, isto é, as probabilidades de adoecer dos grupos de expostos e não-expostos

seriam diferentes mesmo na ausência da exposição (falta de comparabilidade entre os grupos). Ele é gerado por fatores (variáveis confundidoras) que são simultaneamente preditores da ocorrência da doença entre os não expostos, apresentam-se associados com a exposição na base populacional de onde se derivam os casos e não representam passos intermediários na cadeia de eventos que liga a exposição ao desfecho. Por exemplo, em um estudo buscando avaliar a associação entre o nível de atividade física e a mortalidade, a idade pode confundir a relação entre estas duas variáveis, pois indivíduos mais velhos costumam ser mais inativos e o envelhecimento aumenta a probabilidade da ocorrência de óbito.

O tema do confundimento está presente em uma das mais importantes controvérsias médicas da virada do século: o efeito da terapia de reposição hormonal (TRH) sobre o risco cardiovascular entre mulheres após a menopausa. No inicio da década de 1990, uma meta-análise de estudos de coorte estimou que usuárias de TRH tinham risco de doença coronariana reduzido à metade, quando comparadas a não-usuárias ($RR = 0,50$; IC 95% = 0,43-0,56). Entretanto, estudos experimentais não corroboraram essas evidências observacionais¹⁸. Alguns autores salientam que, possivelmente, essa discrepância deveu-se à existência de confundimento residual por posição socioeconômica (PSE) nas estimativas obtidas nos estudos observacionais¹⁵. Esse argumento considera o fato de que a PSE está fortemente associada tanto ao uso de TRH como ao risco de doença coronariana, e que o ajuste habitual das estimativas, nos estudos epidemiológicos, por um (ou pouco mais do que isso) marcador de PSE (ex.: renda, ocupação) na vida adulta não captaria a complexidade do papel da PSE ao longo da vida, ex.: efeitos da PSE na infância e mesmo efeitos intergeracionais associados tanto à exposição como ao desfecho de interesse.

Assim como em outros estudos observacionais, nos estudos de coorte o confundimento só pode ser prevenido na fase de desenho do estudo através da realização da restrição ou do pareamento. Na restrição adotam-se critérios de inclusão e exclusão de participantes, fazendo com que as coortes de expostos e não-expostos se tornem homogêneas em relação a um possível fator confundidor (ex.: exclusão de idosos do estudo). No pareamento, selecionam-se indivíduos não-expostos que sejam idênticos aos expostos em relação a um ou mais fatores confundidores (ex.: idade). O fato de estudos de coorte envolverem a seleção de um número grande de participantes faz com que o pareamento se torne uma estratégia difícil de operacionalizar, sendo, portanto, pouco utilizada. Em estudos de coorte em geral se opta pelo controle do confundimento durante a análise dos dados.

A prevenção do confundimento através do emprego da restrição e do pareamento, ou seu controle durante a fase da análise, só são possíveis através do conhecimento prévio dos possíveis fatores confundidores e da utilização de fontes de dados que permitam a coleta das informações sobre estas variáveis. Estudos de coorte *não-concorrentes* muitas vezes impedem o controle adequado do confundimento em função da ausência das informações necessárias nas fontes de dados de registro. Da mesma forma, em estudos concorrentes que se baseiam em fontes de dados de registro pode ser necessário utilizar fontes adicionais objetivando a coleta de informações sobre variáveis confundidoras, embora neste caso algumas estratégias possam ser empregadas visando à redução dos custos do estudo (ver seção sobre custos). Por fim, o seguimento apenas da coorte exposta comparando a incidência de ocorrência de doenças neste grupo com as taxas de incidência da população geral pode ser limitado pelo fato de as informações sobre possíveis variáveis confundidoras não se encontrarem disponíveis em nível populacional.

Viés de Seleção

Quando ocorre um viés de seleção a associação entre a exposição e o desfecho observada na população de estudo é diferente daquela que seria verificada caso fossem estudados todos os indivíduos da base populacional, *i. e.*, todos os indivíduos que seriam elegíveis como participantes do estudo. Ao contrário do confundimento, onde a associação espúria (não-válida) é observada na base populacional, a associação espúria gerada por um viés de seleção é verificada na população do estudo, mas não na base populacional²⁰. O viés de seleção surge como consequência das estratégias adotadas para a seleção de participantes e de fatores que influenciam a participação no estudo.

Fatores que Influenciam na Validade de Resultados em Estudos de Coorte

- Confundimento
- Viés de Seleção
 - Efeito do trabalhador saudável
 - Auto-seleção
 - Perda seletiva de seguimento
 - Não-resposta
- Viés de Informação
 - Não-diferencial
 - Diferencial

Estudos de coorte baseados em amostras de população geral ou de grupos populacionais restritos são menos sujeitos à ocorrência de vieses de seleção, pois nestes casos a base populacional é identificada primariamente, sendo selecionada uma amostra desta base cujos membros são, então, classificados internamente em expostos e não-expostos. Quando se utilizam grupos de exposição especiais, em primeiro lugar é identificada a coorte altamente exposta ao fator de interesse e posteriormente se escolhe uma coorte não-exposta para servir de grupo controle ou se compara a ocorrência do desfecho neste grupo com as taxas de incidência da população geral. Sendo assim, sempre existe a possibilidade de os grupos de expostos e não-expostos não serem oriundos da mesma base populacional, não sendo, portanto, comparáveis. Ou seja, se estes grupos diferirem em relação à distribuição de fatores que são preditores para o desfecho estudado, uma associação espúria entre exposição e desfecho será observada na população de estudo.

Por exemplo, supondo que se deseja avaliar a existência de uma possível associação entre trabalhar como pedreiro e mortalidade a partir da comparação das taxas de mortalidade de pedreiros de uma empresa com as taxas da população onde a empresa se localiza. Mesmo admitindo que existissem fortes evidências que apontassem para uma possível associação entre esta atividade profissional e a mortalidade, é bem provável que o estudo anteriormente descrito mostrasse uma menor mortalidade entre os pedreiros. Os pedreiros provavelmente terão sido submetidos a um rigoroso exame médico antes de serem admitidos na empresa, só sendo contratados aqueles saudáveis e aptos para a realização da função. Por outro lado, aqueles trabalhadores que desenvolveram algum tipo de doença que comprometesse a realização das atividades foram provavelmente aposentados ou remanejados para a realização de atividades mais leves. Como consequência destes fatores de seleção, os pedreiros em atividade seriam em média mais saudáveis do que a população geral, que é formada por uma mistura de indivíduos saudáveis e doentes.

Este fenômeno, denominado de efeito do trabalhador saudável (*healthy-worker effect*), é em geral observado em estudos onde é comparada a mortalidade de trabalhadores ativos em relação à mortalidade na população geral. Uma alternativa seria classificar a coorte de pedreiros internamente em diferentes níveis de exposição (p. ex., segundo horas gastas por dia na realização de determinadas tarefas) e comparar a mortalidade entre os pedreiros classificados em categorias de “alta” ou “baixa” exposição.

A auto-seleção, a perda seletiva de seguimento e a não-resposta são outras situações que podem provocar

a ocorrência de viés de seleção em estudos de coorte, mesmo quando os grupos de comparação (expostos e não-expostos) são formados internamente.

A auto-seleção ocorre quando os indivíduos escolhem voluntariamente participar do estudo e os fatores que determinaram a participação voluntária estão associados ao desenvolvimento do desfecho que está sendo investigado. Por exemplo, pessoas que se voluntariam a participar de programas de rastreamento de doenças podem fazer essa escolha justamente por serem mais cuidadosas com sua saúde e, portanto, podem apresentar um menor risco de mortalidade do que aqueles que optam por não participar, independentemente de um possível benefício do programa de rastreamento (em alguns casos a situação inversa pode ser verdade, i.e., pessoas com maior risco optam pela participação do programa de rastreamento). Em função desse possível viés de seleção, estudos randomizados seriam mais indicados para a avaliação de programas de rastreamento²¹.

Durante o desenvolvimento de estudos de coorte é esperado que alguns participantes deixem o estudo. Os motivos para a perda de participantes ao longo do seguimento incluem a mudança de endereço, o abandono e a morte por outras causas que não a doença sob investigação. Estas perdas raramente são aleatórias, i. e., são geralmente associadas à exposição, ao desfecho, ou a ambos. Por exemplo, fumantes costumam abandonar mais o estudo do que não-fumantes. Alguns pacientes podem abandonar o estudo justamente por terem desenvolvido a doença que está sendo avaliada e optarem por migrar para outra localidade em busca de tratamento.

Quando a perda é seletiva somente em relação à exposição, nenhum enviesamento é observado nas taxas de incidência* do desfecho nas coortes de expostos e não-expostos. Quando a perda é seletiva somente em relação ao desfecho, i. e., a probabilidade de desenvolver o desfecho no grupo de indivíduos perdidos de observação é diferente da observada para o grupo de indivíduos que permanecem no estudo, observa-se um enviesamento das taxas de incidência no grupo de expostos e de não-expostos. Entretanto, como a perda é equivalente nos dois grupos, as medidas de associação baseadas na razão de incidências em geral se apresentam apenas um pouco enviesadas. Quando a perda é simultaneamente seletiva em relação à exposição e ao desfecho, ocorrerá o enviesamento tanto das taxas de

* As estimativas de incidência acumulada serão enviesadas mesmo no caso da perda não ser associada ao desfecho. Nestes casos devem ser calculadas taxas, ou utilizadas técnicas de análise de sobrevivência.

incidência nos grupos de expostos e não-expostos, como da razão entre elas.

Nem todos os indivíduos que são selecionados para um estudo aceitam participar do mesmo. Os efeitos da não-resposta sobre a validade dos resultados são muito semelhantes aos observados com a perda seletiva de participantes ao longo do seguimento. Ou seja, se a não-resposta for associada somente à exposição, perde-se a capacidade de estimar a distribuição do fator de exposição de interesse na base populacional de onde a amostra foi selecionada, mas as estimativas das taxas incidência do desfecho e da razão entre taxas não serão enviesadas. Se a não-resposta for associada somente ao desfecho, as estimativas das taxas serão enviesadas, mas a razão entre elas será apenas ligeiramente enviesada. Por fim, se a não-resposta for seletiva simultaneamente em relação à exposição e ao desfecho, tanto as estimativas das taxas como a razão entre elas serão enviesadas.

Dificilmente é possível identificar de modo conclusivo se as perdas observadas em um determinado estudo ocorreram de forma aleatória ou seletiva. Indivíduos que se recusam a participar podem ser comparados aos que aceitam em termos de algumas informações existentes em dados de registro (ex.: idade, sexo). Da mesma forma, indivíduos que abandonam o estudo podem ser comparados aos que permanecem em relação aos dados colhidos até o momento da perda. Adicionalmente, quando dados de registros sobre desfechos se encontram disponíveis (ex.: sistema de informações de óbitos) é possível avaliar se a probabilidade de desenvolvimento do desfecho difere entre os membros que aceitam e permanecem no estudo, e aqueles que não aceitam ou abandonam o estudo. Estas análises são, entretanto, limitadas, pois as informações disponíveis para estas comparações são geralmente restritas.

É importante, portanto, que uma série de medidas sejam adotadas visando reduzir a não resposta e maximizar a adesão ao estudo^{6,22}. Deve ser mencionada a necessidade de, entre outras iniciativas, obter-se apoio ativo de representantes das populações que se deseja estudar (ex.: representantes de comunidades ou de grupos profissionais); explicitar a relevância e os objetivos do estudo, e as atividades previstas; dispor de telefone de amigos e parentes visando auxiliar a localização de participantes que tenham sido perdidos de observação; criar um nome e logotipo para o estudo e enviar periodicamente informes e boletins; desenvolver um bom relacionamento entre membros da equipe e participantes; enviar pequenas lembranças (ex.: cartetas, cartões de aniversário).

Viés de Informação

Este viés surge em função de erros sistemáticos que ocorrem no momento da coleta das informações sobre as variáveis de interesse (exposição, covariáveis e desfecho).

Nos estudos de coorte *concorrentes*, a informação sobre a exposição é colhida antes da ocorrência do desfecho. Em consequência, erros na coleta da informação sobre a exposição são semelhantes entre os indivíduos que posteriormente irão ou não desenvolver o desfecho. Por outro lado, o conhecimento da situação de exposição pode influenciar a coleta das informações sobre o desfecho. Por exemplo, o conhecimento de que um participante faz parte do grupo altamente exposto ao fumo pode influenciar o clínico a diagnosticar a presença de doença coronariana.

Nos estudos *não-concorrentes*, a informação sobre o desfecho é registrada antes do início do estudo, em geral, portanto, sem o conhecimento da situação de exposição do participante. Os erros na coleta da informação sobre o desfecho tendem a ocorrer de forma equivalente no grupo de expostos e de não-expostos. Com relação à exposição, em algumas situações (ex.: avaliações ocupacionais) o dado que se encontra registrado é empregado para classificar os indivíduos em diferentes categorias de exposição. Embora sempre se busque empregar regras claras para a classificação, este processo geralmente está sujeito a um certo grau de subjetividade (por exemplo, na hora de classificar indivíduos com valores limites entre duas classes). Quando os responsáveis pela classificação conhecem a situação de desfecho de cada participante, este fato pode levar à classificação dos participantes que desenvolveram o desfecho nas categorias de exposição mais intensa.

As consequências dos erros na medida do desfecho sobre as estimativas das medidas de associação serão diferentes na dependência de eles ocorrerem de forma igual no grupo de expostos e não-expostos (erro de classificação *não-diferencial*), ou de serem maiores em um grupo do que no outro (erro de classificação *diferencial*). Da mesma forma, o impacto dos erros na medida de exposição sobre as estimativas das medidas de associação também dependerá de eles ocorrerem de forma equivalente nos participantes que desenvolveram ou não o desfecho (erro de classificação *não-diferencial*), ou de serem maiores em um grupo do que no outro (erro de classificação *diferencial*).

Os erros de classificação *não-diferenciais* nas situações de exposições com apenas duas categorias (ex.: atividade física regular, sim ou não) e nos desfechos,

em geral levam a uma atenuação da magnitude da associação, ou seja, na direção da ausência de associação (ou “hipótese nula”), enquanto os erros *diferenciais* podem tanto acentuar como atenuar a magnitude da associação. No caso dos desfechos é interessante observar que as razões de incidência (embora não as medidas de diferença) podem não sofrer nenhum enviesamento se o processo de classificação do desfecho apresentar especificidade de 100%, mesmo na presença de sensibilidade inferior a 100%. Por fim, erros na medida de variáveis confundidoras podem limitar o controle do confundimento durante a análise dos dados¹.

Vários tipos de problemas que podem levar a erros na classificação da situação de exposição, das covariáveis e do desfecho foram comentados anteriormente na seção de fontes de informação. Algumas estratégias podem ser implementadas visando à prevenção destes erros e, consequentemente, da ocorrência de viés de informação, a saber: avaliar a confiabilidade e a validade dos questionários e dos testes médicos a serem empregados; treinar os entrevistadores e técnicos responsáveis pela realização dos testes; garantir que os responsáveis pela coleta das informações do desfecho desconheçam a situação de exposição dos participantes; efetuar a revisão e o controle da entrada de dados, e realizar a análise periódica dos dados, buscando identificar a ocorrência de inconsistências nos dados^{22,23}.

Custos

Estudos de coorte em geral implicam custos elevados, sobretudo quando se trata de estudos de natureza etiológica, ou seja, destinados a avaliar associações entre possíveis fatores de exposição e a incidência de doenças. Nesses estudos, com freqüência é necessário acompanhar indivíduos em maior número e/ou durante mais tempo, pois a ocorrência de doenças em uma população previamente saudável é um fenômeno bem mais raro do que a ocorrência de uma complicação secundária em pacientes previamente doentes, como no caso do estudo (de natureza prognóstica) da mortalidade após um episódio de infarto agudo do miocárdio.

Várias estratégias podem ser empregadas para redução de custos em estudos de coorte. Entre elas estão a utilização do delineamento *não-concorrente*, o seguimento de grupos populacionais especiais, o seguimento apenas da coorte exposta comparando a incidência de ocorrência de doenças neste grupo com as taxas de incidência da população geral e a seleção de fontes de informação de menor custo (p. ex., dados de registro e questionários de autopercepção). Nas últimas dé-

cadas, dois novos tipos de estratégias vêm sendo empregadas visando à redução dos custos em estudos de coorte: a utilização de bancos de dados informatizados em conjunto com técnicas de relacionamento (*linkage*) de registros e a aplicação de estratégias analíticas do tipo caso-controle no contexto de estudos de coorte.

Grandes bases informatizadas com dados sobre mortalidade, atendimentos em serviços de saúde e registro de base populacional de doenças (ex.: câncer) passaram a ser empregadas para o monitoramento da ocorrência de desfechos em estudos de coorte, reduzindo ou mesmo eliminando a necessidade de contatos repetidos com os membros da coorte para o monitoramento da evolução dos mesmos. O relacionamento dos registros dos participantes da coorte com os registros nestes bancos de dados é realizado através de um identificador comum (número da segurança social, por exemplo). Mesmo quando este identificador não existe, o relacionamento pode ser realizado empregando-se o método de relacionamento probabilístico de registros (*probabilistic record linkage*), que se baseia na utilização de uma combinação de atributos menos específicos (nome, data de nascimento e endereço, por exemplo), juntamente com o emprego de algoritmos destinados à comparação aproximada de nomes, que levam em consideração possíveis erros fonéticos e de digitação (ex.: Manoel e Manuel seriam reconhecidos como iguais), e de técnicas estatísticas¹³.

Outra estratégia que permite a redução de custos é a identificação e o seguimento de uma coorte para a avaliação da ocorrência do desfecho de interesse. Entretanto, em vez de se colher as informações sobre a exposição, variáveis confundidoras e modificadoras de efeito para todos os membros da coorte, isto é feito apenas em uma parte dos mesmos, empregando-se para tal um delineamento do tipo caso-controle. Duas estratégias amostrais são mais freqüentemente utilizadas, recebendo o estudo caso-controle uma denominação especial na dependência da estratégia empregada*. Denomina-se estudo caso-controle aninhado (*nested case-control study*) àquele em que a coorte é seguida no tempo e a cada momento em que um caso ocorre, um ou mais controles são selecionados entre os membros originais da coorte que se encontram sob risco (*i. e.*, não foram perdidos de observação e se encontram sem a doença) no momento em que o caso ocorre. No segundo caso,

*Na literatura são empregadas diferentes terminologias para identificar os estudos caso-controle que fazem uso das estratégias amostrais acima comentadas. Neste texto optamos por empregar a terminologia adotada por Langholz & Thomas²⁴.

denominado de estudo caso-coorte, uma amostra aleatória de indivíduos da coorte é selecionada no início do período de observação, e esta é estudada juntamente com os casos que venham a ocorrer durante o período de observação (*i. e.*, indivíduos da coorte não originalmente selecionados na amostra, mas que evoluem como casos, também são estudados).

Ross e cols.²⁵ conduziram estudo de coorte entre 18.244 homens em Xangai, China, com o objetivo de analisar, entre outras hipóteses, o papel de determinadas aflatoxinas na etiologia do câncer hepático. Estas análises constituíram um exemplo de estudo caso-controle aninhado na coorte. Durante o seguimento dos participantes, sempre que algum caso novo de câncer era diagnosticado (no total, 22), cinco a dez controles eram selecionados ao acaso entre os membros da coorte livres de câncer naquele momento, e apenas nestes casos e controles eram realizadas as dosagens detalhadas (e caras) dos marcadores biológicos de interesse.

Um exemplo de estudo caso-coorte foi o desenvolvido por Duncan e cols.⁶, com o objetivo de avaliar a associação entre marcadores de inflamação sistêmica e incidência de diabetes tipo 2. O estudo captou a experiência de 9 anos de seguimento de 10.275 participantes da coorte do estudo ARIC (*Atherosclerosis Risk in Communities Study*). As amostras de plasma armazenadas de 571 participantes que desenvolveram diabetes ao longo do período de acompanhamento e de 572 participantes que não desenvolveram esse desfecho, foram analisadas com relação à presença de diferentes marcadores de inflamação. O estudo evidenciou que a presença de inflamação sistêmica era um bom preditor para o desenvolvimento de diabetes tipo 2.

Nesta seção foram apresentadas algumas estratégias que permitem a redução dos custos em estudos de coorte. A adoção das mesmas dependerá da particular hipótese que se deseja estudar e mais uma vez é necessário avaliar o impacto destas opções sobre a validade do estudo. Por exemplo, se uma determinada fonte de dados de registro sabidamente apresenta uma baixa qualidade nas informações registradas, não se justifica empregá-la, ainda que a utilização da mesma implique em uma redução considerável nos custos com a coleta de dados.

CONCLUSÃO

Neste capítulo buscamos apresentar alguns aspectos do desenho, condução e interpretação de resultados de estudos de coorte. A opção pelo desenho de estudo de coorte para a avaliação de uma determinada hipótese a

ser investigada deverá levar em consideração as vantagens e limitações deste delineamento.

Em resumo, estudos de coorte são menos sujeitos à ambigüidade na determinação da seqüência temporal entre variáveis exploradas como fatores de exposição e desfechos. Como os indivíduos são selecionados ou classificados segundo a situação de exposição para posterior avaliação da ocorrência do desfecho, em geral não existem dúvidas quanto ao fato de a exposição realmente preceder a ocorrência do desfecho. Permite o cálculo direto das medidas de incidência nas coortes de expostos e não-expostos, e a avaliação de exposições raras. Vários desfechos podem ser investigados simultaneamente, sendo também possível testar diversas hipóteses envolvendo diferentes exposições e desfechos. São menos sujeitos a viés de seleção que os estudos de caso-controle, sobretudo quando os grupos de expostos e não-expostos são classificados internamente à coorte. A situação de doença não influencia a medida da exposição, especialmente em estudos prospectivos.

Estratégias para Redução de Custos em Estudos de Coorte

- Delineamento *não-concorrente*
- Seguimento de grupos populacionais especiais
- Seguimento apenas da coorte exposta
- Seleção de fontes de informação de menor custo
- *Linkage* de banco de dados
- Estudos caso-controle aninhado e caso-coorte

Estudos de coorte são ineficientes para a avaliação de desfechos raros ou que apresentem um período de indução longo. São geralmente caros e difíceis de operacionalizar, embora algumas estratégias possam ser adotadas visando à redução de custos. A perda de participantes ao longo do seguimento pode comprometer a validade dos resultados. Por fim, como não é realizada a alocação aleatória de intervenções (terapêuticas ou preventivas), são menos indicados que os estudos experimentais para a avaliação dessas intervenções, pois fatores de seleção e confundidores não controlados podem comprometer a validade do estudo.

BIBLIOGRAFIA

1. Rothman KJ, Greenland S. Types of epidemiologic studies. In: Rothman KJ, Greenland S. Modern Epidemiology 2 ed. Philadelphia: Lippincott-Raven Publishers, 1998.

2. White E, Hunt JR, Casso D. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiol Rev*, 1998; 20:43-56.
3. Tager IB. Outcomes in cohort studies. *Epidemiol Rev*, 1998; 20:15-27.
4. MacMahon B, Trichopoulos D. Cohort studies. In: MacMahon B, Trichopoulos D. *Epidemiology principles and methods* 2nd ed. Boston: Little, brown and Company, 1996; 165-227.
5. Victora CG, Barros FC, Vaughan JP. Mortalidade perinatal. In: Victora CG, Barros FC, Vaughan JP. *Epidemiologia da desigualdade* 2 ed. São Paulo: Hucitec, 1989; 59-71.
6. Handler A, Kistin N, Davis F. Cocaine use during pregnancy: perinatal outcomes. *Am J Epidemiol*, 133:818-25, 1991.
7. Miettinen OS. Design of the study base. In: Miettinen OS. *Theoretical epidemiology* 1st ed. New York: John Wiley & Sons, 1985; 46-67.
8. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies. *Annals New York Academy of Sciences*, 1963; 107:539-556.
9. Costa MFFL, Uchoa E, Guerra H et al. The Bambu health and ageing study (BHAS): methodological approach and preliminary results of a population-based cohort study of the elderly in Brazil. *Rev. Saúde Pública*, 2000; 34: 126-135.
10. Hu FB, Stampfer MJ, Colditz GA, Ascherio A, Rexrode KM, Willett WC, Manson JE. Physical activity and risk of stroke in women. *JAMA*, 2000; 283:2961-67.
11. Faerstein E, Chor D, Lopes CS, Werneck GL. Estudo Pró-Saúde: características gerais e aspectos metodológicos. *Rev. Bras Epidemiol* 2005; 8:454-466.
12. Samet JM & Muñoz A. Evolution of the Cohort Study. *Epidemiol Rev* 1998; 20: 1-14.
13. Camargo Jr. KR, Coeli CM. Reclink: aplicativo para o relacionamento de bases de dados, implementando o método probabilistic record linkage. *Cad Saúde Pública*, 2000; 16:439-447.
14. Johansen C, Oslen JH. Risk of cancer among Danish Utility workers – a nationwide cohort study. *Am J Epidemiol*, 1998; 147:548-55.
15. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez-Roux AV, Folsom AR. Multi-Ethnic Study of Atherosclerosis: Objectives and Design. *Am J Epidemiology* 2002; 156:871-81.
16. Carvalho MS, Andreozzi VL, Codeço CT. *Análise de sobrevida: teoria e aplicações em saúde*. Rio de Janeiro: Editora Fiocruz, 2005.
17. Vidal EI, Coeli CM, Pinheiro RS, Camargo KR Jr. Mortality within 1 year after hip fracture surgical repair in the elderly according to postoperative period: a probabilistic record linkage study in Brazil. *Osteoporos Int*. 2006 (in press Jul 27).
18. Pettitt D. Commentary: Hormone replacement therapy and coronary heart disease: four lessons. *International Journal of Epidemiology* 2004; 33:461-463.
19. Lawlor DA, Davey-Smith G, Ebrahim S. Commentary: The hormone replacement – coronary heart disease conundrum: is this the death of observational epidemiology? *International Journal of Epidemiology* 2004; 33:464-467.
20. Koopman JS. Stratification of Exposure-Disease relationships upon a third variable and the assessment of joint effects. March 19 Lecture: Part 1. Dr. Koopman [acessado em: 15 de junho de 2000]. Available from URL: <http://www.sph.umich.edu/group/epid/pid655/1997/lecmr19a.html>.
21. Gordis L. *Epidemiology* 2 ed. Philadelphia: W.B. Saunders Company 200; 262.
22. Willet WC, Colditz GA. Approaches for conducting large cohort studies. *Epidemiol Rev*, 1998; 20:57-70. (Footnotes)
23. Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. *Epidemiol Rev*, 1998; 20:71-80.
24. Langholz B, Thomas DC. Nested case-control and case-cohort methods of sampling from a cohort: a critical comparison. *Am J Epidemiol*, 1990; 131:169-76.
25. Ross RK, Yuan JM, Yu MC et al. Urinary aflatoxin biomarkers and risk of hepatocellular carcinoma. *Lancet*, 1992; 339:943-46.
26. Duncan BB, Schmidt MI, Pankow JS, Ballantyne CM, Couper D, Vigo A, Hoogeveen R, Folsom AR, Heiss G; Atherosclerosis Risk in Communities Study. Low-grade systemic inflammation and the development of type 2 diabetes: the atherosclerosis risk in communities study. *Diabetes*. 2003; 52:1799-805.
27. Hunt JR, White E. Retaining and tracking cohort study members. *Epidemiol Rev*, 1998; 20:57-70.

Estudos de Intervenção

Claudia Caminha Escosteguy

INTRODUÇÃO

O termo “estudo de intervenção” envolve, de uma forma geral, aqueles estudos em que o pesquisador manipula o fator de exposição (a intervenção), ou seja, provoca uma modificação intencional em algum aspecto do estado de saúde, através da introdução de um esquema profilático ou terapêutico. Tais estudos podem ter ou não grupo-controle e referir-se a indivíduos ou a comunidades.

Um exemplo de estudo de intervenção não-controlado é o chamado estudo “antes e após” (*before and after*), nos quais todos os pacientes recebem o mesmo tratamento e sua condição é verificada antes do início e em vários momentos após o tratamento¹². Em alguns casos em que o curso clínico da doença pode ser predition com bastante confiança, tais estudos podem ser válidos, como os exemplos históricos da melhora observada na evolução da endocardite bacteriana subaguda após a introdução da penicilina e melhora do hipotireoidismo a partir do uso do hormônio tireoidiano. Entretanto, a maioria das decisões terapêuticas não envolve este tipo de situação. Outro exemplo de estudo “antes e após” seria um estudo do efeito de um tratamento sobre o nível de pressão arterial em um mesmo grupo de pacientes hipertensos, nos quais seria aferida a pressão arterial antes do início e em vários momentos após a introdução do tratamento. Contudo, as aferições posteriores da pressão arterial poderiam melhorar por razões puramente estatísticas, de regressão dos valores à média, sem necessariamente representar um efeito verdadeiro da intervenção naquele indivíduo. Ou seja, em um estudo não-controlado pode ser difícil implicar ao tratamento os efeitos observados.

Alguns autores usam “estudo de intervenção” como sinônimo de “estudo experimental”; outros consideram o termo “experimental” como mais bem aplicado àqueles estudos controlados e com alocação aleatória do fator de intervenção, ou seja, os ensaios clínicos controlados randomizados.

Segundo Pocock²⁵, o termo “ensaio clínico”, de uma forma genérica, pode ser aplicado a qualquer experimento planejado que envolva pacientes e que seja desenhado para elucidar o tratamento mais apropriado para futuros pacientes com uma dada condição médica. Alguns autores usam, ainda, o termo “ensaio clínico não-controlado” para descrever um estudo em que todos os participantes recebem o fator de intervenção. Na verdade, este seria apenas um estudo descritivo dos efeitos de uma intervenção num grupo, sendo que a maioria dos autores não considera tal estudo como um ensaio clínico. Outros autores guardam o termo ensaio clínico apenas para os ensaios controlados randomizados, não aceitando o uso do termo para ensaios controlados, porém sem randomização²⁵.

Na taxonomia dos desenhos epidemiológicos, experimento refere-se a um estudo no qual o pesquisador intencionalmente altera um ou mais fatores, sob condições controladas, com o objetivo de investigar os efeitos dessa alteração. Nesses estudos, o investigador faz uso de estratégias que visam o controle de outras variáveis que não o fator de interesse sobre os resultados. Adiante veremos que a melhor dessas estratégias é a randomização ou alocação aleatória, na qual os participantes do estudo têm a mesma chance de pertencerem ao grupo de intervenção ou ao grupo-controle.

Quando o fator de interesse (intervenção) é manipulado pelo investigador, mas a alocação dos indivíduos nos grupos a serem comparados não se dá de modo aleatório (ao acaso), esses estudos são referidos como quase-experimentos²⁹. Um exemplo histórico de um estudo quase-experimental foi o estudo de James Lind sobre intervenções no escorbuto, em 1753, no qual o pesquisador distribuiu algumas intervenções entre 12 pacientes marinheiros com escorbuto, concluindo pelo efeito benéfico do uso de laranjas e limões²⁵. Menos citado, mas bem anterior a esse exemplo, é a descrição bíblica encontrada no primeiro capítulo do Livro de Daniel, sobre a comparação dos efeitos benéficos de uma dieta vegetariana com os de uma dieta real babilônica durante 10 dias, que constitui o provável primeiro registro conhecido de um estudo de intervenção controlado, realizado em torno do ano 600 a.C¹⁹.

Os estudos experimentais em epidemiologia podem ser divididos em profiláticos ou terapêuticos, se o objetivo é investigar uma intervenção preventiva (ex.: vacina) ou terapêutica (ex.: hipoglicemizante). Outra classificação é quanto à unidade de análise, indivíduos, em ensaios clínicos (*clinical trials*), ou comunidades, ensaios de comunidade (*community trials*). No ensaio de comunidade a unidade de alocação para receber a intervenção é uma comunidade inteira (por exemplo, uma cidade). Dois exemplos seriam os ensaios de comunidade para avaliar os efeitos da fluorização da água sobre a incidência de cáries nos Estados Unidos e Canadá³, e o de prevenção de doença coronariana realizado na Karelia do Norte, Finlândia²⁶. No estudo finlandês, o setor saúde, a indústria alimentícia e os meios de comunicação de massa desencadearam iniciativas de promoção de hábitos saudáveis na comunidade (alimentação saudável, atividade física e redução de tabagismo).

ENSAIOS CLÍNICOS CONTROLADOS RANDOMIZADOS: HISTÓRICO E CONCEITO

O crédito ao primeiro ensaio clínico randomizado moderno é geralmente dado a Sir Austin Bradford Hill, publicado no final da década de 1940, no qual o estatístico alocou aleatoriamente pacientes com tuberculose pulmonar em dois grupos: um para ser tratado com estreptomicina e outro sem estreptomicina, podendo então avaliar a eficácia da droga.

O ensaio clínico randomizado é um estudo prospectivo em seres humanos, que compara o efeito e o valor de uma intervenção (profilática ou terapêutica) com controles. O investigador distribui o fator de interven-

ção a ser analisado de forma aleatória através da técnica da *randomização*; dessa forma, os grupos experimental e controle são formados por um processo aleatório de decisão. A intervenção pode referir-se a fármacos, técnicas ou procedimentos¹². Portanto, trata-se de um estudo experimental, de desenho prospectivo, controlado e randomizado, considerado a melhor fonte de determinação da eficácia de uma intervenção.

O termo eficácia refere-se ao resultado de uma intervenção realizada sob condições ideais, bem controladas, como nos ensaios clínicos randomizados ou em “centros de excelência”. Efetividade refere-se ao resultado de uma intervenção aplicada sob as condições habituais da prática médica, que incluem as imperfeições de implementação que caracterizam o mundo cotidiano¹⁴.

Os ensaios clínicos controlados randomizados constituem um dos principais avanços científicos entre os métodos de pesquisa durante o século XX, tornando-se uma ferramenta fundamental para a avaliação de intervenções para a saúde, sejam elas terapêuticas ou profiláticas. O seu crescimento tem sido exponencial; Olkin estimou a publicação de aproximadamente 9.000 ensaios randomizado anualmente nas revistas médicas e em 1998 a Cochrane Library já registrava mais de 150.000¹⁹. Tais estudos constituem a base metodológica para o que atualmente é denominado de “medicina baseada em evidências”^{10,31}.

Como já se assinalou, o principal objetivo dos estudos randomizados é fazer com que variáveis que podem influenciar no desfecho em estudo sejam distribuídas de forma homogênea, tornando os grupos comparáveis. Por exemplo, ao se investigar a eficácia de uma nova droga anti-hipertensiva, não é interessante que em um dos grupos predominem casos leves de hipertensão, e no outro, casos mais graves. Nesse cenário, ao se identificar um padrão de resposta diferente nos dois grupos, seria difícil concluir se este se deve ao uso de um novo tratamento ou às características clínicas dos grupos em comparação.

Existe consenso de que a melhor maneira de reduzir a influência dos fatores estranhos é através da randomização, ou alocação ao acaso, dos participantes nos grupos de intervenção e de controle. Esta estratégia garante que os participantes tenham a mesma chance de serem incluídos em qualquer um dos grupos.

Segundo Feinstein¹⁰, a idéia da distribuição de um tratamento através da randomização foi proposta por Fisher, em 1923, aplicada à pesquisa agrícola. Em 1926, essa idéia foi pela primeira vez usada clinicamente por Amerson e cols., para testar o valor de um composto

de ouro no tratamento da tuberculose; este estudo foi também o primeiro relatado sob a condição mascarada ou “cega” (*blinded*), isto é, os pacientes desconheciam qual o tratamento recebido, sendo que os controles receberam uma injeção de água destilada. Entretanto, o termo placebo foi usado pela primeira vez no estudo de Diehl, em 1938, sobre vacinação contra resfriado¹⁰. O mascaramento tem por intuito reduzir o grau de subjetividade por parte do paciente e do investigador, por ocasião da avaliação do desfecho.

Os métodos de randomização e de mascaramento serão discutidos adiante.

FASES DA EXPERIMENTAÇÃO OU DOS ENSAIOS CLÍNICOS

Os ensaios clínicos com drogas são freqüentemente classificados em quatro fases principais da experimentação^{19,25}, que em geral são conduzidas durante a avaliação de uma nova droga.

- *Fase I:* são ensaios de farmacologia clínica e toxicidade no homem, primariamente relacionados à segurança e não à eficácia, e em geral realizados em voluntários. O principal objetivo é determinar uma dose aceitável da droga que está sendo testada, isto é, que possa ser dada sem causar efeitos colaterais sérios. Tal informação com freqüência é obtida a partir de experimentos de doses escalonadas, nos quais voluntários são submetidos a doses crescentes da droga de acordo com um cronograma predeterminado. A fase I também envolve estudos do metabolismo e da biodisponibilidade da droga. Após estudos em voluntários normais, os ensaios iniciais em pacientes também constituirão parte da fase I. Pacientes com doenças sem cura conhecida, como Aids e alguns tipos de câncer, com freqüência participam de estudos de fase I. Tipicamente, estudos de fase I podem requerer entre 20 a 80 indivíduos e pacientes. Em geral não são randomizados e nem mesmo controlados; costumam ser apenas séries de casos, nas quais os participantes são cuidadosamente monitorados pelos investigadores.
- *Fase II:* são ensaios iniciais de investigação clínica do efeito do tratamento, constituindo investigações ainda em pequena escala da eficácia e segurança da droga, com monitoramento cuidadoso de cada paciente; são estudos-piloto de eficácia. No caso das vacinas, o foco é sobre a imunogenicidade. Algumas vezes, os ensaios de fase II são conduzidos como

um processo de seleção de quais são aquelas drogas que têm realmente um efeito clínico potencial, entre as inúmeras drogas investigadas que são inativas ou excessivamente tóxicas, de tal modo que as drogas selecionadas possam passar à fase III. Raramente a fase II vai além de 100 a 200 pacientes por droga. Com freqüência são estudos não-randomizados.

- *Fase III:* avaliação em larga escala do tratamento. Após a droga ter sido demonstrada como razoavelmente eficaz, é essencial compará-la em larga escala com o(s) tratamento(s) padrão(ões) disponível(eis) para a mesma condição médica, em um ensaio clínico controlado envolvendo um número suficientemente grande de pacientes. Para alguns autores, o termo *ensaio clínico* seria sinônimo desses ensaios de fase III, que vêm a constituir a forma mais rigorosa de investigação clínica de um novo tratamento. A maioria dos ensaios de fase III é randomizada.
- *Fase IV:* fase de vigilância pós-comercialização (*postmarketing surveillance*). Esta fase é conduzida após a droga ter sido aprovada para a distribuição ou comercialização, no caso do Brasil, por exemplo, pelo Ministério da Saúde, e no caso dos Estados Unidos da América (EUA) pela *Food and Drug Administration* (FDA). Nesta etapa, ainda existem questões a serem consideradas com relação a monitoramento de efeitos adversos e estudos adicionais, em larga escala, em longo prazo, de morbidade e mortalidade. Estudos de fase IV não são ensaios clínicos randomizados. No Brasil, o Ministério da Saúde, através da Agência Nacional de Vigilância Sanitária², tem investido na formação de uma rede de farmacovigilância, sobretudo através dos hospitais sentinelas, sensibilizando os profissionais para a notificação de suspeita de reação adversa, que deverá ser investigada.

Algumas vezes o termo *ensaio de fase IV* tem sido usado para descrever estratégias de promoção de uma nova droga, dirigidas ao público médico em geral, o que não faz parte e não deve ser confundido com a pesquisa dos ensaios clínicos propriamente dita.

É importante lembrar que, precedendo os ensaios clínicos, deve existir um programa anterior igualmente importante de pesquisa pré-clínica, incluindo a síntese de novas drogas e os estudos em animais referentes ao metabolismo, eficácia e, sobretudo, toxicidade potencial. Esta fase pré-clínica é responsável, na verdade, pela maior parte do gasto estimado com a pesquisa sobre drogas.

Atualmente, a maioria dos ensaios clínicos relaciona-se à avaliação de novas drogas, e é principalmente

sustentada pela indústria farmacêutica. Estima-se que do universo de novas drogas sintetizadas em laboratórios, apenas uma em cada 10.000 atinge a fase de estudos clínicos, e no máximo 20% das que vão a ensaios clínicos são eventualmente comercializadas. A duração média de um programa inteiro de pesquisa relacionado a uma droga é de sete a dez anos, dos quais aproximadamente a metade é gasta em ensaios clínicos, e envolve milhões de dólares^{26,31}. Portanto, o papel da indústria farmacêutica e dos conflitos de interesse não pode ser colocado em segundo plano.

A classificação de ensaios de fases I a III pode servir como um guia para ensaios clínicos não-relacionados à indústria farmacêutica, com adaptações. Por exemplo, em programas de pesquisa em quimioterapia e radioterapia para câncer, os ensaios de fase I são necessariamente em pacientes e não em voluntários normais, devido à natureza altamente tóxica dos tratamentos. Já no desenvolvimento de novos procedimentos cirúrgicos, a fase I pode ser considerada como o desenvolvimento básico das técnicas cirúrgicas; entretanto, são raros os ensaios clínicos de fase III bem delineados em cirurgia.

CARACTERÍSTICAS DOS ENSAIOS CLÍNICOS RANDOMIZADOS

Randomização

A randomização é um processo de decisão que permite que os grupos de tratamento e controle sejam alocados de forma aleatória (ao acaso), garantindo que todos os participantes tenham a mesma chance de serem incluídos em qualquer um dos grupos estudados. O seu propósito é constituir grupos com características muito semelhantes (comparáveis) no início do estudo, exceto pela intervenção a ser avaliada. Dessa forma, é possível atribuir as diferenças observadas entre os grupos às intervenções que estão sendo comparadas. É a melhor técnica para evitar viés de seleção, além de reduzir a possibilidade de viés de confundimento. Conforme aumenta o tamanho da amostra, aumenta a chance de que a randomização faça com que tanto fatores conhecidos quanto desconhecidos que influenciam o desfecho (*outcome*) de interesse sejam distribuídos de forma semelhante entre os grupos tratado e controle, evitando resultados e conclusões espúrias acerca da intervenção^{10,12,8}.

Estratégias de alocação baseadas em ordem de chegada dos participantes, dias da semana e numeração corrida não devem ser usadas, pois facilitam a identificação do grupo em que o indivíduo será incluído. A identificação pode facilitar a manipulação da alocação, que

pode ocorrer não necessariamente por má fé. Um exemplo seria a hipótese de que o pesquisador acredite que o novo tratamento seja superior ao convencional. Se ele souber que um paciente mais grave será alocado no grupo de tratamento convencional, ele pode não incluir esse indivíduo no estudo, aguardando a chegada de outro paciente e interferindo na comparabilidade dos grupos. O uso de uma seqüência aleatória de números, obtida através de tabelas de números aleatórios ou de algoritmos computacionais, dificulta a identificação da seqüência de alocações. Infelizmente, com freqüência as descrições dos métodos de alocação nos estudos descritos como randomizados apresentadas na literatura científica têm sido insuficientes e limitadas. Em 5-10% dos ensaios em que a metodologia descreve detalhes dos métodos de alocação, não são usadas técnicas que gerem seqüências aleatórias.

Há várias técnicas de randomização²⁵:

- *randomização simples*: a mais empregada, na qual os participantes são alocados diretamente nos grupos de estudo e controle, sem etapas intermediárias. Pode ser obtida, de forma muito simples, jogando-se “cara ou coroa” com uma moeda, em que cara representasse o tratamento, e coroa, o controle, ou vice-versa. Entretanto, esta técnica não é prática, sendo mais conveniente, em geral, usar uma tabela de números aleatórios, em que, por exemplo, os números ímpares vão para o grupo da intervenção e os pares para o grupo-controle. Pode-se ainda usar programas de computador para gerar seqüências aleatórias de números;
- *randomização em blocos*: caracteriza-se pela formação de blocos de número fixo de indivíduos, dentro dos quais se tem a mesma quantidade de indivíduos alocados em cada grupo de comparação. Por exemplo, pode-se definir blocos de tamanho seis e usar os números 1,2,3,4,5,6 seqüenciados aleatoriamente em cada bloco, de modo que a cada seis indivíduos randomizados, teremos exatamente três no grupo de intervenção e três no grupo-controle. Esse método tem a vantagem de garantir um número igual de participantes nos grupos de estudo e controle, mesmo que o ensaio seja interrompido antes do final previsto. É útil, também, em estudos com número reduzido de pacientes, pois a randomização simples feita a partir de uma tabela de números aleatórios só garante a homogeneidade entre os grupos quando há um grande número de participantes a serem randomizados;
- *randomização pareada*: inicialmente são formados pares de participantes e a alocação aleatória é feita

- no interior do par, de tal forma que um indivíduo receba o tratamento em estudo, e o outro, o controle;
- *randomização estratificada*: são formados inicialmente estratos segundo características que possam interverir no desfecho em estudo (ex.: idade, sexo, estadiamento de tumor) e a alocação aleatória é feita dentro de cada estrato para aumentar a chance de grupos balanceados quanto a esses fatores;
 - *randomização por minimização*: usa-se, de início, a randomização simples; após alguns indivíduos terem sido alocados, as características dos grupos são analisadas e os cálculos são refeitos à medida que novos participantes são recrutados, os quais passam a ter maior probabilidade de serem alocados em um dos grupos de modo a diminuir as diferenças que forem detectadas, ou para manter o equilíbrio já alcançado. É uma técnica nova, em que a informática permite o uso de muitas variáveis a serem acompanhadas, de forma a obter o mínimo de diferenças entre os grupos.

Tamanho da Amostra

O ensaio deve recrutar um número suficiente de pacientes para obter uma estimativa de razoável precisão da resposta a cada tratamento envolvido. Embora existam considerações de ordem prática e ética quanto ao tamanho da amostra, uma abordagem estatística padrão refere-se às estimativas do poder do estudo. Há cinco questões relevantes com relação ao tamanho da amostra²⁵:

- *qual é o principal objetivo do ensaio*: por exemplo, verificar se a aspirina tem valor em prevenir a morte pós-infarto, o que é diferente de verificar se previne reinfarto ou se previne morte e reinfarto;
- *qual é a principal medida de desfecho*: por exemplo, morte por qualquer causa durante o primeiro mês após o tratamento, o que é diferente de morte por causa cardiovascular;
- *como os dados serão analisados para que seja detectada uma diferença de tratamento*: a forma mais simples é a comparação entre percentagens, por exemplo, de mortes no grupo tratado e no grupo-controle; um teste qui-quadrado será usado e o nível de significância de 5%, por exemplo, será aceito como mostrando evidência de uma diferença de tratamento;
- *que tipos de resultados são antecipados com o tratamento padrão*: por exemplo, espera-se que

10% dos pacientes no grupo-controle morram no primeiro mês após o tratamento;

- *qual a menor diferença de tratamento considerada importante para ser detectada e com que grau de precisão*: é importante lembrar que a demonstração de reduções moderadas (por exemplo, 20-25%) da ocorrência do evento de interesse pode exigir o recrutamento de milhares de pacientes^{35,36}.

No cálculo do tamanho da amostra devem ser levados em consideração ainda o nível de significância desejado para se detectar uma diferença de tratamento e o poder do estudo. Erro α ou erro do tipo I é a probabilidade de detectar uma diferença que na verdade não existe, ou seja, a probabilidade de um resultado falso-positivo; geralmente se estipula o valor dessa probabilidade em $\alpha = 0,05$; Erro β ou erro do tipo II é a probabilidade de não detectar uma diferença quando ela realmente existe. O poder do estudo, ou seja, a probabilidade de que a diferença de uma certa magnitude entre os tratamentos será detectada, caso ela exista realmente, $1 - \beta$, é geralmente estipulado como 0,80 ou 0,90^{7,25} (maiores detalhes no Capítulo 24 – Associação Estatística em Epidemiologia: Análise Bivariada).

Quando o tamanho necessário da amostra é muito grande, o ensaio pode ser realizado em vários centros, constituindo o chamado ensaio multicêntrico, que evidentemente impõe questões particulares de organização e monitoramento. O termo *mega-trial* tem sido usado para descrever ensaios que incluem milhares de participantes, em geral com um desenho simples, altamente pragmático e com coleta de dados limitada. O principal objetivo desses ensaios é atingir um maior poder estatístico e uma maior possibilidade de generalização de resultados. Frequentemente esses mega-ensaços são multicêntricos, muitas vezes envolvendo países diferentes.

No outro extremo, o chamado ensaio “*n-of-1*”, ou ensaio de um indivíduo, é basicamente um desenho *crossover* no qual um participante recebe a intervenção em estudo e o controle, em pares, em várias ocasiões e numa ordem randômica. Estes estudos provêm resultados individuais e não-generalizáveis, mas podem ser úteis em determinadas situações em que não está claro se um tratamento será útil para um paciente em particular. Por exemplo, no caso de um paciente com uma doença rara para a qual não existam ensaios que dêem conta daquela situação particular; ou quando o tratamento foi avaliado por ensaios que incluíram pacientes muito diferentes do caso em questão. Tipicamente, o número de pares de intervenções usados não é pré-especificado, de tal forma que o clínico e o pa-

ciente podem decidir parar quando estão convencidos de que há (ou não) diferenças importantes entre as intervenções. O sucesso desse desenho depende largamente da cooperação do paciente e do seguimento das regras do desenho *crossover*.

Organização, Planejamento e Monitoramento do Ensaio

Em relação à organização e ao planejamento, é fundamental definir com precisão:

- quais são os pacientes elegíveis, através de critérios de inclusão e exclusão bem definidos;
- que tratamento está sendo avaliado;
- quais são os desfechos de interesse a serem analisados;
- como a resposta de cada paciente será verificada.

Quanto ao monitoramento do processo do ensaio, é necessário acompanhar a adesão ao protocolo, os efeitos adversos, o processamento dos dados e as análises interinas da comparação entre os tratamentos. Possíveis violações e desvios do protocolo devem ser cuidadosamente analisados, tais como não-adesão ao tratamento, perda de participantes, avaliação incompleta, cruzamento entre os grupos tratado e controle após a randomização. Este último problema ocorreu, por exemplo, no braço nitrito *versus* controle do ensaio clínico multicêntrico GISSI-3, de intervenções no infarto agudo do miocárdio, em que 57% dos pacientes randomizados para controle receberam nitrito, diminuindo o poder do estudo em detectar uma possível diferença entre os dois grupos¹³.

A análise das co-intervenções, ou seja, de outras intervenções além da estudada, aplicadas aos pacientes randomizados, é importante, pois se distribuídas de forma desigual entre os grupos de estudo, podem introduzir viés.

As análises interinas são importantes para evitar prolongamento desnecessário de um ensaio clínico em que as diferenças entre os grupos já estejam evidentes, garantindo, portanto, que nenhum paciente no ensaio venha a receber um tratamento sobre o qual já existam evidências de que seja inferior. Constituem análises intermediárias dos resultados de um ensaio que ainda está em andamento, no sentido de detectar se já existem evidências definitivas de diferenças entre os grupos sob estudo²⁵. O número de análises intermediárias a serem realizadas em um estudo deve ser pré-definido e os testes de significância estatística devem levar em consideração o efeito de múltiplas comparações¹⁹.

Tipos de Análise

A análise pode ser feita de duas formas principais²⁶:

- entre aqueles que de fato completaram o tratamento em cada um dos grupos;
- segundo a “intenção de tratar” (*intention-to-treat*), na qual são incluídos todos os que foram randomizados para formar os grupos, independentemente de terem ou não completado o tratamento. Esta forma tem sido a preferida, pois garante a manutenção dos grupos aleatórios, e avalia o tratamento em condições reais, com suas imperfeições. Porém, é necessário saber o que ocorreu com os que não completaram o tratamento, assim como se houve cruzamento entre os grupos, e a dimensão desses fatos, pois se for muito grande poderá introduzir viés.

Outro ponto importante refere-se à chamada análise de subgrupo. O resultado fundamental de um ensaio clínico é a descrição do principal resultado de interesse em cada um dos principais grupos de tratamento. Embora seja tentador analisar os resultados em subgrupos específicos de pacientes, há grandes riscos inerentes a esta análise. O primeiro deles é um número inadequado de pacientes, se a referida análise não fazia parte do plano amostral inicial. O segundo é o risco de viés, já que os subgrupos selecionados por características que são levantadas após a alocação do tratamento podem não ser comparáveis, ainda que tenham sido retirados de grupos inicialmente randomizados. Em terceiro, quando um grande número de subgrupos é examinado, há um aumento da chance de que alguns deles venham a mostrar uma diferença estatisticamente significativa espúria. Um exemplo clássico dessa possibilidade de associação espúria surgiu na análise do efeito do signo astrológico no estudo ISIS-2, sugerindo que o uso da aspirina no infarto miocárdio era benéfico para pacientes de todos os signos exceto Libra e Gêmeos, para os quais havia aparentemente dano^{18,36}.

Mascaramento e Uso de Placebo

A técnica do mascaramento, avaliação cega (*blinding*) ou ocultamento objetiva evitar que os participantes do ensaio saibam o tratamento que é administrado; a sua justificativa reside no potencial para viés que ocorre quando todas as pessoas envolvidas no ensaio conhecem qual é o tratamento que o paciente está recebendo. O mascaramento pode referir-se ao paciente, ao grupo de profissionais que aplica o tratamento e ao avaliador²⁵.

O chamado efeito Hawthorne diz respeito à tendência dos indivíduos mudarem seu comportamento porque são alvos de interesse e atenção especial, não importa a natureza específica da intervenção que estão recebendo. Uma forma de controlá-lo é através do mascaramento e uso de placebo¹².

O fato de o paciente saber que está recebendo um novo tratamento pode ter efeito benéfico psicológico, e ao contrário, o fato de o paciente saber que está recebendo um tratamento convencional ou nenhum tratamento pode exercer um efeito desfavorável. É claro que o impacto depende do tipo de doença e da natureza do tratamento, mas tal possibilidade não deve ser subestimada mesmo em doenças não-psiquiátricas.

Em relação ao grupo que aplica o tratamento, decisões quanto à modificação de dose, detalhamento do exame do paciente, a continuação do tratamento do ensaio e a necessidade de outros tratamentos adicionais são geralmente da responsabilidade do médico assistente, que pode influenciar o curso do tratamento de várias maneiras. Essas decisões podem ser influenciadas pelo conhecimento do grupo do ensaio ao qual o paciente pertence. Também o entusiasmo por algum novo tratamento pode ser transferido ao paciente e ocasionar mudança de atitude do mesmo, aumentando a adesão, por exemplo.

Quanto ao investigador que avalia os resultados, se é informado quanto ao tratamento de cada paciente, há um risco potencial, por exemplo, de registrar respostas mais favoráveis para o tratamento que acredita ser superior. O não conhecimento dos grupos do ensaio ajuda a evitar vieses de aferição, que também são minimizados quando se define o evento final de interesse da forma mais objetiva possível. Viés de aferição é particularmente passível de ocorrer quando a avaliação da resposta ao tratamento requer julgamento clínico. Mesmo em eventos aparentemente bem definidos, como infarto agudo do miocárdio, muitas vezes surge a necessidade do julgamento clínico em casos limítrofes, nos quais o conhecimento sobre o grupo ao qual o paciente pertence poderia fazer com que o avaliador direcionasse o diagnóstico final a favor ou contra infarto, por exemplo.

O termo duplo-cego refere-se àqueles ensaios nos quais nem o paciente nem aqueles responsáveis pela sua assistência e avaliação conhecem qual o tratamento que ele está recebendo. Na verdade, neste caso, os três tipos de participantes estão cegos quanto à condição do tratamento; porém, como freqüentemente os mesmos clínicos que trabalham com a terapêutica são os que avaliam o paciente, o termo duplo-cego é adequado. No ensaio triplo-cego, três grupos de participantes estão cegos quanto à condição do tratamento: o

paciente, o investigador que administra a intervenção e aquele que avalia o efeito; ou o paciente, o investigador que administra e ao mesmo tempo avalia o efeito, e o investigador que analisa os dados. No quádruplo-cego haveria quatro categorias de participantes cegos quanto à condição do tratamento. Na verdade, embora seja teoricamente fácil transformar um duplo-cego em um triplo ou quádruplo-cego, raramente isto acontece. Na prática, é mais comum referir-se a duplo-cego^{19,25}.

Um ensaio em que não há tentativa de mascaramento é chamado de aberto (*open* ou *open label*). Para muitas questões clínicas importantes não é possível o mascaramento, como na avaliação dos efeitos de cirurgias, radioterapia ou dieta¹², cabendo o uso de ensaios abertos.

Quando o mascaramento é possível, em geral é alcançado com o uso de um placebo. O placebo constitui uma substância de aparência, forma e administração semelhante ao tratamento que está sendo avaliado, porém sem ter o princípio ativo do mesmo. A principal razão para introduzir controles com placebo é fazer com que as atitudes dos pacientes do ensaio sejam tão parecidas quanto possível nos grupos tratado e controle. O efeito placebo é uma resposta a uma intervenção médica, que apesar de ser definitivamente um resultado da mesma, *não* tem relação com seu mecanismo de ação específico¹². Um princípio básico a ser considerado é que, eticamente, pacientes não podem ser alocados para receber placebo se existe um tratamento padrão alternativo de eficácia estabelecida.

TIPOS DE ENSAIOS CLÍNICOS

Os ensaios clínicos podem ser usados para avaliar diferentes tipos de intervenções em diversas populações de participantes, em distintas circunstâncias e para diferentes propósitos. Uma vez que o investigador tenha garantido a alocação randômica, o estudo pode ser desenhado através de várias estratégias, que levam em conta recursos disponíveis, motivações acadêmicas, políticas, de *marketing* ou clínicas. Ao longo dos anos, múltiplos termos têm sido empregados para descrever diversos tipos de ensaios clínicos, freqüentemente se referindo a aspectos do seu delineamento. A seguir serão descritos alguns dos mais citados, a partir de uma adaptação da proposta de classificação sugerida por Jadad¹⁹, com o objetivo de facilitar o entendimento desse espectro de termos.

Ensaio em Paralelo, Desenho Fatorial e Ensaios do Tipo Crossover

De acordo com o modo como os participantes são expostos à intervenção, os ensaios clínicos podem apre-

sentar um desenho em paralelo, fatorial ou do tipo *crossover*. A maioria tem um desenho em paralelo, no qual cada grupo de participantes é exposto a apenas uma das intervenções estudadas¹⁹.

No desenho fatorial os efeitos de vários fatores são verificados em um único ensaio. Por exemplo, no estudo das drogas A e B, um desenho fatorial avaliará quatro grupos de tratamento: A, B, A + B e controle sem nenhuma delas. Um exemplo é o ensaio clínico multicêntrico ISIS-2, onde foram avaliados os efeitos da aspirina, da estreptoquinase, de ambas e de nenhuma das duas drogas, em pacientes com suspeita de infarto agudo do miocárdio¹⁸.

Algumas vezes pode ser adequado constituir comparações seqüenciais no mesmo paciente, como ocorre nos ensaios do tipo *crossover* ou seqüencial, nos quais cada paciente receberá mais de um tratamento, de forma alternada. No caso de dois tratamentos A e B, metade dos indivíduos é randomizada para receber os tratamentos na ordem A, B, e metade para recebê-los na ordem B, A. Ensaio do tipo seqüencial permitem comparações para cada participante, de tal forma que cada um age como seu próprio controle. Assim, esses ensaios podem produzir resultados estatisticamente adequados com um menor número de participantes do que seria necessário com um desenho de grupos paralelos convencionais. Por outro lado, o ensaio seqüencial tem a desvantagem de que o efeito do primeiro tratamento pode ocorrer no período no qual o segundo está sendo administrado^{19,21,25}. Para este desenho, é necessário que a doença ou condição seja estável e o efeito da intervenção tenha início rápido e duração curta.

A Figura 13.1 exemplifica esquematicamente um ensaio de agentes hipertensivos com um desenho do tipo *crossover*.

Ensaios Pragmáticos (de Efetividade) e Ensaios Explanatórios (de Eficácia)

O termo ensaio pragmático tem sido usado para referir-se àquele desenho com objetivo não apenas de

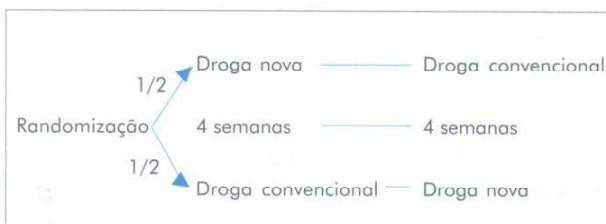


Fig. 13.1 – Desenho do tipo crossover em um ensaio clínico de agente hipertensivo.

determinar se uma intervenção funciona, mas também de descrever todas as consequências de seu uso, sejam boas ou más, sob circunstâncias que se aproximam da prática clínica “real”. Os ensaios pragmáticos tendem a ter critérios menos rígidos, incluindo participantes de características heterogêneas, semelhantes aos vistos pelos clínicos na sua prática diária. Tendem a usar controles que recebem drogas ativas e não placebo, e têm regimes de prescrição mais flexíveis. O desenho é mais simples e menos rígido; pode utilizar desfechos cuja padronização é mais difícil, como medida da qualidade de vida ou da função sexual^{25,28}. Ensaio clínico com o objetivo de avaliar a efetividade de uma intervenção tendem a ser pragmáticos, pois tentam avaliar os efeitos em circunstâncias mais próximas da prática clínica diária. Para assegurar a capacidade de generalização dos resultados de um ensaio pragmático, ele deve incluir, tanto quanto possível, aqueles pacientes para os quais o tratamento será aplicado.

O termo ensaio explanatório tem sido usado por alguns autores para referir-se àquele desenho com objetivo de avaliar se a intervenção funciona ou não, dentro de uma abordagem que se preocupa com princípios da farmacologia ou fisiologia, pretendendo estabelecer como a intervenção funciona, e obedecendo a critérios mais rígidos. O ensaio explanatório tem critérios de inclusão estritos que produzem grupos de estudos altamente homogêneos, tendendo a usar placebo, regimes de prescrição fixos e desfechos precisos e bem padronizados (por exemplo, pressão arterial mensurada em momentos pré-estabelecidos e segundo técnica padronizada). Os ensaios explanatórios são desenhados para avaliar se um tratamento é eficaz sob determinadas condições experimentais ideais.

Na literatura da analgesia, há vários exemplos de abordagem explanatória e pragmática; investigadores na década de 1960 como Beecher e Houde desenvolveram comparações de dose única em pacientes no pós-operatório ou com câncer, supondo que tais pacientes forneceriam um “modelo” para avaliar a eficácia da analgesia em qualquer condição algica e usando grupos-controle variados. Estes estudos explanatórios constituíram a base para as tabelas de potência relativa dos analgésicos modernos. Mais recentemente, grupos de anestesiologistas têm desenvolvido estudos pragmáticos avaliando a utilidade de técnicas como anestesia controlada pelo paciente, uso epidural de opiáceos e anestésicos, e bloqueio de nervos periféricos durante cirurgia, em condições que refletem a prática diária¹⁹.

Ressalta-se que muitas vezes, ao incorporar-se uma tecnologia eficaz segundo um ensaio clínico randomizado, pode ocorrer perda de efetividade e os resultados na prática cotidiana não serem tão satisfatórios em relação a outras escolhas. Um exemplo que tem sido muito discutido em cardiologia diz respeito à escolha da melhor técnica de reperfusão coronária no infarto agudo do miocárdio (IAM). Segundo meta-análise de 24 ensaios clínicos randomizados, a angioplastia primária é mais eficaz do que a trombólise farmacológica na reperfusão coronária no IAM²⁰. Entretanto, quando são avaliados resultados fora do contexto dos ensaios clínicos, muitas vezes a efetividade de ambas se equipara ou até favorece o trombolítico, por diversas razões relacionadas à implementação das intervenções no “mundo real”, ou seja, fora do ambiente controlado dos ensaios. Neste caso, cabem reflexões inclusive de ordem ética na escolha das intervenções, que devem ser eficazes mas também efetivas³². Ao tomar-se a decisão de incorporar uma intervenção ou tecnologia, deve-se levar em consideração não apenas a eficácia, mas também a efetividade e ainda a eficiência. O conceito de eficiência leva em consideração a relação entre custos e consequências, ou custo *versus* benefício.

VIESSES EM POTENCIAL

Embora o ensaio clínico randomizado seja o padrão de referência dos métodos de pesquisa clínica e epidemiológica, também é sujeito a vieses, que devem ser considerados. É importante considerar que são fontes em potencial de viés o processo de seleção dos grupos, a alocação do tratamento, a realização da intervenção da forma proposta, e a aferição dos resultados. A randomização bem realizada controla os dois primeiros. Como já foi visto, problemas que ocorrerem após a randomização podem introduzir diferenças sistemáticas entre os grupos sob observação; alguns exemplos são cruzamentos entre os grupos.

Perdas de seguimento e não-cooperação de participantes podem introduzir viés, sobretudo se diferentemente distribuídas entre os grupos tratado e controle, e devem ser sempre mencionadas.

Viés de aferição (também dito de informação, observação ou medida) resulta de diferenças sistemáticas no modo com que os dados sobre o evento de interesse são obtidos dos vários grupos em estudo. São minimizados quando se usa a técnica duplo-cego com placebo; porém nem sempre é possível, mesmo com essa técnica, ocultar de observadores e observados os grupos a que estes últimos pertencem.

Outro viés relaciona-se com a publicação dos ensaios, e não com a realização propriamente dita dos mesmos; trata-se do viés de publicação, que é a tendência a serem mais publicados os estudos com resultados positivos.

ESTIMATIVA DO TAMANHO DO EFEITO DO TRATAMENTO

A estimativa do tamanho do efeito do tratamento pode ser feita de diversas maneiras⁴. A Tabela 13.1 mostra a distribuição teórica de possíveis efeitos nos grupos tratado e controle:

Tabela 13.1 Tabela de Distribuição de Possíveis Efeitos nos Grupos Tratado e Controle			
Grupo	Evento de Interesse		Risco do Evento
	Presente	Ausente	
Tratado	a	b	RT = a/(a + b)
Controle	c	d	RC = c/(c + d)

RT = risco no grupo tratado.

RC = risco no grupo-controle.

A seguir, são resumidas algumas formas de se medir o tamanho do efeito do tratamento:

A redução absoluta de risco (RAR) é a diferença de risco entre o grupo-controle (RC) e o grupo tratado (RT), em termos absolutos:

$$\text{RAR} = \text{RC} - \text{RT}$$

O risco relativo (RR) é a razão entre o risco no grupo tratado (RT) e o risco no grupo-controle (RC):

$$\text{RR} = \text{RT}/\text{RC}$$

A redução relativa de risco (RRR) ou *relative risk reduction* expressa a redução percentual de eventos no grupo tratado (RT) em relação aos controles (RC). É também conhecida como eficácia, que representa a redução relativa do risco obtida com a intervenção:

$$\text{RRR} = (1 - \text{RT}/\text{RC}) \times 100 = (1 - \text{RR}) \times 100$$

$$\text{RRR} = [(RC - RT)/RC] \times 100$$

O número necessário para tratar (NNT) ou *number needed to treat* é o inverso da redução absoluta de risco. Esta medida expressa o número de pacientes que devem ser tratados a fim de que um evento adverso adicional seja evitado. Por exemplo, se uma droga tem um NNT igual a cinco, em relação ao evento morte, significa que cinco pacientes devem ser com ela tratados para que uma morte adicional seja evitada^{3,22,29}. O NNT tem recebido ênfase crescente na literatura médica como um meio de medir o impacto de uma intervenção. Como a RAR é influenciada pela freqüência do evento que se está avaliando e o NNT é o inverso da RAR, ele também será influenciado pela freqüência do evento.

$$NNT = 1/RAR$$

$$NNT = 1/(RC - RT)$$

Um NNT muito pequeno, próximo a 1, significa que um desfecho favorável ocorre em quase todo paciente que recebe aquele tratamento e em poucos do grupo-controle; portanto, que a intervenção é bastante eficaz. NNTs próximos a 1 quase nunca são encontrados na prática, mas NNTs pequenos ocorreram em alguns ensaios clínicos, como os que compararam antibióticos com placebo na erradicação de infecção por *Helicobacter pylori* em pacientes com úlcera péptica. Em contraste, determinadas intervenções podem ter NNTs tão altos como 20 a 40 e ainda assim serem clinicamente relevantes. Um exemplo é a associação de aspirina e estreptoquinase para reduzir a mortalidade pós-infarto agudo do miocárdio (NNT = 20).

Não é apropriado comparar NNTs aplicados a diferentes condições, principalmente se os desfechos de interesse variam. Por exemplo, um NNT de 30 para prevenir trombose venosa profunda pode ser valorizado de forma diferente de um NNT de 30 para evitar um acidente vascular encefálico ou morte. O conceito expresso pelo NNT é de freqüência e não de utilidade; o valor numérico é uma função da doença, da intervenção e do desfecho. Só é possível comparar diretamente os NNTs de diferentes intervenções quando se referem à mesma doença (e gravidade) e ao mesmo desfecho^{4,23,31}.

Conceito análogo pode ser aplicado à produção de efeitos indesejados: número necessário para produzir dano ou *number needed to harm* (NNH).

Odds ratio (OR) ou razão de chances: é outra medida de associação; trata-se de uma razão entre as chances de ter o desfecho esperado no grupo de intervenção e no grupo-controle, que tem a propriedade matemati-

camente demonstrável de aproximar-se do valor do RR quanto mais raro for o evento de interesse (ver Capítulo 9 – Medidas de Associação e Medidas de Impacto).

$$OR = (a/b) \div (c/d) = a \cdot d / b \cdot c$$

Em seguida, um exemplo do cálculo das medidas do tamanho do efeito da trombólise intravenosa (IV) sobre a mortalidade de cinco semanas no infarto agudo do miocárdio (IAM) é construído com dados estimados a partir de uma meta-análise de ensaios clínicos randomizados multicêntricos, resumidos na Tabela 13.2:

Tabela 13.2 Efeito da Trombólise IV em Pacientes com Supradesnívelamento de ST ou Bloqueio de Ramo e Delta Tempo até Seis Horas do Início dos Sintomas de IAM				
Grupo	Resultado			
	Morte	Não	Total	Letalidade
Tratado	1.352	13.086	14.438	9,4%
Controle	1.773	12.613	14.386	12,3%

A estimativa do efeito da trombólise IV no IAM a partir dos resultados apresentados na Tabela 13.2 pode ser expressa através de:

- *redução absoluta de risco – RAR* = $12,3\% - 9,4\% = 2,9\%$. Ou seja, a diferença na letalidade entre o grupo-controle e o grupo tratado com trombolítico foi de 2,9 unidades percentuais;
- *risco relativo – RR* = $9,4/12,3 = 0,76$. Assim, o grupo tratado com trombolítico apresentou um risco relativo de morte de 0,76 em relação ao controle, tratando-se, portanto, de uma intervenção benéfica;
- *redução relativa de risco – RRR* = $(1 - RR) \times 100 = 24\%$. Isto significa que a trombólise intravenosa possibilitou uma redução na letalidade pós-IAM de 24% em relação ao grupo-controle;
- *number needed to treat – NNT* = $1/RAR = 1 / 0,029 = 34$. A interpretação, neste caso, é de que seria necessário tratar 34 pacientes com trombolítico para se evitar a morte de um deles, ou seja, a cada 34 pacientes tratados, uma morte é prevenida.
- $OR = (1.352 \times 12.613) / (13.086 \times 1.773) = 0,73$. A chance de morrer no grupo tratado é 0,73 em relação ao grupo-controle.

Maiores detalhes podem ser encontrados no Capítulo 9 – Medidas de Associação e Medidas de Impacto.

VALIDADE INTERNA VERSUS VALIDADE EXTERNA

Todas as questões até então discutidas dizem respeito à validade interna do ensaio clínico. A difusão dos ensaios randomizados e seu uso como padrão para demonstração de eficácia terapêutica têm possibilitado que evidência científica de boa qualidade esteja disponível antes que novos agentes terapêuticos sejam introduzidos à clínica. Entretanto, outro aspecto fundamental a ser discutido refere-se à possibilidade ou não de generalização dos resultados do ensaio. A validade externa de um estudo implica na possibilidade de generalizarem-se os resultados da amostra estudada para outras populações que não a população-alvo do estudo. Envolve variações de pacientes, de fatores de gravidade e variações etnoculturais, além de considerações de eficácia x efetividade, custo x benefício, risco, infra-estrutura, e assim por diante. Só se justifica tal consideração após a constatação da validade interna do estudo.

O Quadro 13.1 apresenta um resumo adaptado do *Evidence Based Medicine Working Group*^{15,16,17}, que é útil para a avaliação das principais questões relacionadas à validade interna e externa de um ensaio clínico. Tais questões devem ser consideradas durante o processo de decisão médica de aplicar ou não determinado resultado de um estudo de intervenção a um paciente em questão.

Deve haver bom senso no uso da literatura científica para resolver uma questão terapêutica ou profilática.

Quadro 13.1 Como Avaliar um Estudo de Intervenção

1. Os resultados são válidos? **Avaliar se:**
 - houve randomização
 - o seguimento (*follow up*) foi completo
 - os pacientes foram analisados nos grupos para os quais foram randomizados
 - o estudo foi cego
 - os grupos eram semelhantes no início do ensaio;
 - os grupos foram tratados de forma semelhante, à exceção do tratamento sob investigação
2. Quais são os resultados? **Avaliar:**
 - qual é o tamanho do efeito do tratamento
 - qual a precisão da estimativa do efeito do tratamento
3. Os resultados podem ajudar na assistência ao paciente em questão? **Avaliar se:**
 - os resultados podem ser aplicados ao tipo de paciente em questão
 - os eventos clinicamente importantes foram considerados
 - os benefícios do tratamento irão suplantar os possíveis danos e custos

De início, o problema deve ser claramente definido e uma estratégia adequada de pesquisa bibliográfica deve ser utilizada para obter a melhor evidência disponível. É crescente a realização e a publicação de ensaios clínicos na área médica, representando uma gigantesca quantidade de informação a ser administrada. Para exemplificar, Cheng, autor que estuda acrônimos de ensaios clínicos em cardiologia, discutiu o crescimento exponencial de aproximadamente 200 ensaios em 1992 para mais de 2.300 em 1998, crescimento este que continua ocorrendo. Neste sentido, metodologias de revisão sistemática e meta-análise podem trazer uma contribuição importante para se fazer frente ao volume crescente de dados a serem considerados na decisão clínica sobre que intervenção utilizar (maiores detalhes no Capítulo 16 – Revisão Sistemática e Metanálise).

Uma vez que um artigo relevante tenha sido encontrado, a qualidade da evidência nele contida deve ser avaliada. Se a evidência é pobre, qualquer inferência subsequente será enfraquecida, assim como a decisão clínica gerada a partir dela. Se a qualidade do artigo é adequada, é necessário determinar as margens dentro das quais o efeito do tratamento provavelmente ocorre, e então considerar até que ponto os resultados são generalizáveis para o paciente em questão e se os desfechos que foram medidos são relevantes. Se há dúvidas sobre a generalização ou sobre a importância dos desfechos, a base para uma recomendação terapêutica ficará enfraquecida. Os riscos de efeitos adversos devem ser considerados. Deve ser feito um balanço entre a probabilidade de benefício, os custos associados (incluindo financeiros e relativos à conveniência do tratamento) e os riscos, além das preferências do paciente. Este julgamento guiará a decisão terapêutica. Tornar este processo mais explícito e aumentar a capacidade de identificar e analisar a qualidade da evidência disponível levará a uma melhor tomada de decisão. Isso tem sido feito através de iniciativas como a declaração Consort, que é uma importante ferramenta elaborada para melhorar a qualidade dos relatos de ensaios randomizados. A declaração está disponível em várias línguas e recebeu o endosso de periódicos prominentes. Consiste basicamente de um *checklist* de itens que precisam ser descritos e de um fluxograma para fornecer uma imagem clara do progresso de todos os participantes do ensaio, desde a randomização até o final da participação deles no estudo³³.

LIMITAÇÕES

Embora os ensaios clínicos randomizados sejam considerados a melhor fonte de evidência científica na

pesquisa clínica, não respondem a todas as questões clínicas, havendo várias situações em que não podem ser realizados, ou não são necessários, apropriados ou mesmo suficientes para ajudar a resolver problemas importantes. Este desenho é ideal para responder questões relacionadas a efeitos de intervenções de tamanho pequeno a moderado. Além disso, há muitas situações que poderiam ser estudadas por ensaios clínicos, para as quais eles não estão disponíveis. A realização dos ensaios é ainda limitada por questões financeiras, por taxas de aderência baixas ou por perda de seguimento elevadas. Finalmente, mesmo quando um ensaio clínico randomizado de boa qualidade está disponível, dificilmente ele fornecerá todas as respostas necessárias para a decisão clínica. Esta complexa decisão necessita da evidência científica, da interface com o paciente e da reflexão sobre o contexto do problema, sendo modulada pela experiência clínica. Como bem alerta Bernard Lown²², cardiologista e cientista pioneiro na pesquisa da morte súbita, no uso do desfibrilador cardíaco, da lidocaína nas arritmias ventriculares e das unidades coronarianas: “O paradoxo de minha vida está na ironia de ter minha pesquisa contribuído para aquilo que eu mais deploro... De acordo com alguns médicos, a tecnologia substitui satisfatoriamente a conversa com o paciente”.

QUESTÕES ÉTICAS

O primeiro código de conduta em pesquisas internacionalmente aceito foi o Código de Nuremberg (1947), que estabeleceu que nenhum ser humano poderia ser submetido a qualquer pesquisa sem que tivesse dado seu consentimento explícito para isso. Posteriormente, a necessidade de regulamentar as pesquisas em seres humanos e a pouca influência daquele código levaram à aprovação do que viria a constituir o documento internacional básico para a discussão ética dos ensaios clínicos: a Declaração de Helsinque, 1964, revista em Tóquio em 1975. No Brasil, em 10/10/1996, o Conselho Nacional de Saúde aprovou as Diretrizes e Normas Regulamentadoras de Pesquisa Envolvendo Seres Humanos (Resolução 196/96)⁶, que definiu os princípios básicos para apreciação ética dos protocolos de pesquisa, criou os Comitês de Ética em Pesquisa (CEPs) e a Comissão Nacional de Ética em Pesquisa (Conep). Até então a regulamentação nacional sobre ética em pesquisa estava restrita ao Código de Ética Médica⁶ e de outros conselhos de profissionais de saúde (maiores detalhes no Capítulo 34 – Eticidade da Pesquisa Científica: o caso

da Regulamentação Brasileira da Ética em Pesquisa Envolvendo Seres Humanos).

Mesmo quando a investigação é plenamente justificada, algumas questões merecem reflexão:

- um dos pontos centrais é a privação do grupo-controle de um novo tratamento para o qual haja evidências de ser nitidamente superior ao tratamento habitual. A não-administração aos pacientes de um tratamento eficaz só é eticamente aceitável se há dúvidas quanto à eficácia do tratamento;
- deve ser usado o menor tamanho de amostra informado pelos cálculos que seja adequado para responder à pergunta em investigação. O estudo deve ser interrompido se há evidência definitiva, durante sua condução, do benefício ou da ausência de benefício do tratamento em questão;
- o consentimento informado do paciente deve estar sempre presente;
- comitês de monitoramento do ensaio devem ser constituídos para acompanhar o progresso do estudo, avaliando quebras do protocolo, análises intermediárias e indicações de interrupção precoce do mesmo.

Além das questões até aqui discutidas, é de extrema relevância a percepção da existência do chamado conflito de interesses e sua possível influência sobre a realização de pesquisas e na própria decisão médica. A influência de questões econômicas sobre quais pesquisas devem ser realizadas necessita ser reconhecida. Como já foi dito o gasto em pesquisa de novas drogas é enorme. Estratégias para recuperar esses gastos vêm embutidas, por exemplo, em estratégias de *marketing* e propaganda. No primeiro semestre de 1998, só nos EUA, a indústria farmacêutica gastou 3,1 bilhões de dólares em propaganda de drogas para médicos e consumidores; desse montante, 1,8 bilhão foi gasto com propaganda direta ao médico, “face a face”, por exemplo, através dos propagandistas²⁷. No ano 2000, naquele mesmo país, a indústria farmacêutica gastou 2,5 bilhões de dólares em propaganda direta ao consumidor, sendo que 2,26 bilhões com menção explícita do nome comercial da droga. Alguns estudos mostram que essas estratégias têm impacto sobre a prática médica³⁴. Além disso, a existência do chamado viés de publicação, que representa a maior tendência em publicar resultados positivos (“que deram certo”) do que os negativos, também deve ser reconhecida. O papel da mídia divulgando resultados de estudos muitas vezes sem isenção ou conhecimento técnico também tem sido discutido.

BIBLIOGRAFIA

1. Abreu LM, Escosteguy CC, Amaral W, Monteiro Filho MY. Tratamento Trombolítico do Infarto na Emergência com Tele-consultoria (TIET): resultados de cinco anos. Revista da Socerj 2005; 18(5):418-28.
2. Agência Nacional de Vigilância Sanitária (ANVISA). Ministério da Saúde. <http://www.anvisa.gov.br/notifica/index.htm>; acesso em 15/02/2006.
3. Centers for Disease Control. Achievements in Public Health, 1900-1999: Fluoridation of Drinking Water to Prevent Dental Caries. MMWR (Morbidity and Mortality Report 1999; 48(41):933-940.
4. Chatellier G, Zapletal E, Lemaitre D, Menard J, Degoulet P. The number needed to treat: a clinically useful nomogram in its proper context. British Medical Journal 1996; 312:426-9.
5. Cheng TO. Acronyms of clinical trials in cardiology-1998. American Heart Journal 1999; 137(4):726-65.
6. Conselho Nacional de Saúde - Normas de Pesquisa envolvendo Seres Humanos. Resolução CNS 196/96. Brasília: Ministério da Saúde, 1996.
7. Conselho Regional de Medicina do Estado do Rio de Janeiro - Código de Ética Médica. Legislação dos Conselhos de Medicina, 1988.
8. Coutinho ESF, da Cunha GM. Conceitos básicos de epidemiologia e estatística para a leitura de ensaios clínicos controlados. Revista Brasileira de Psiquiatria 2005; 27(2): 146-51.
9. Evidence-based Medicine Working Group. Evidence-based medicine a new approach to teaching the practice of medicine. JAMA 1999; 268: 2420-25.
10. Feinstein AR. Clinical epidemiology. The architecture of clinical research. Philadelphia: W.B. Saunders Company, 1985.
11. Fibrinolytic Therapy Trialists' (FTT) Collaborative Group - Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and morbidity results from all randomized trials of more than 1000 patients. The Lancet 1994; 343: 311-22.
12. Fletcher RH, Fletcher SW, Wagner EH. Epidemiologia clínica: elementos essenciais. Porto Alegre: Artes Médicas, 1996.
13. ISIS-3 - Effects of lisinopril and transdermal glyceril trinitrate singly and together on 6 week mortality and ventricular function after acute myocardial infarction. The Lancet 1994; 343:1115-221.
14. Goodman C. It's Time to Rethink Health Care Technology Assessment. International Journal of Technology Assessment in Health Care 1992; 8(2): 335-58.
15. Guyatt GH, Sackett D, Cook D, for the Evidence Based Medicine Working Group. How to Use an Article About Therapy or Prevention . Based on the Users' Guides to Evidence-based medicine and reproduced with permission from JAMA. (1993; 270: 2598-2601 and 1994; 271: 59-63). Copyright 1995, American Medical Association. Disponível em: <http://www.ccbe.net/usersguides/therapy.asp>; acessodo em: 11/11/2006
16. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? JAMA 1993; 270: 2598-2601.
17. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? JAMA 1993; 271: 59-63
18. ISIS-2 Collaborative Group - Randomised trial of intravenous streptokinase, oral aspirin, both or neither among 17187 cases of suspected acute myocardial infarction. The Lancet 1988; I:349-60.
19. Jadad A. Randomised controlled trials. A user's guide. 1a ed. London: BMJ Books, 1998.
20. Keeley EC, Boura JA, Grines CL. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomized trials. The Lancet 2003; 361:13-20.
21. Last JM. A dictionary of epidemiology. 4 ed. New York: Oxford University Press, 2001.
22. Lown B. The Lost Art of Healing. Practicing compassion in Medicine. New York: Houghton Mifflin, 1996.
23. McQuay HJ, Andrew Moore R. Using Numerical Results from Systematic Reviews in Clinical Practice. Annals of Internal Medicine 1997; 126:712-720.
24. Olkin I. Statistical and theoretical considerations in meta-analysis. J Clin Epidemiol 1995; 48: 133-46.
25. Pocock SJ. Clinical Trials. A practical approach. Brisbane: John Wiley & Sons, 1991.
26. Pusker P, Nissinen A, Tuomilehto J, Salonen JT, Koskela K, McAlister A et al. The community-based strategy to prevent coronary heart disease. Conclusions from the ten years of the North Karelia Project. Annual Review of Public Health 1985; 6:147-93.
27. Reuters Health. Pharmaceutical industry spent \$3.1 billion on drug promotion in first half of 1998. Professional Medical News 12/11/1998. Disponível em: [Http://www.reutershealth.com/en/index.html](http://www.reutershealth.com/en/index.html); acessodo em: 14/03/2003.
28. Roland M, Torgerson DJ. Understanding controlled trials: What are pragmatic trials? British Medical Journal 1988; 316:285-5.
29. Rothman KJ & Greenland S. Modern Epidemiology. 2 ed. Philadelphia: Lippincott-Raven, 1998.
30. Sackett DL, Haynes RB, Guyatt GH, Tugwell. Clinical epidemiology: a basic science for clinical medicine. 2a ed. Boston: Little Brown, 1991.
31. Sackett DL, Richardson WS, Rosenberg W, Haynes W. Evidence-based Medicine. How to practice and teach EBM. 2a ed. New York: Churchill Livingstone, 2000.
32. Schramm F R, Escosteguy CC. Bioética e avaliação tecnológica em saúde. Cad. Saúde Pública 2000; 16(4):951-61.
33. The CONSORT Group. Consolidated Standards of Reporting Trials (CONSORT). Disponível em <http://www.consort-statement.org/statement/revisedstatement.htm>; acesso em 30/03/2007
34. The National Institute for Health Care Management. Prescription drugs expenditures in 2001. 2002. Disponível em: www.nihcm.org; acessodo em: 11/11/2006.
35. Yusuf S, Wittes J, Friedman L. Overview of results of randomized clinical trials in heart disease. JAMA 1988; 260 (14): 2088-93.
36. Yusuf S, Wittes J, Probstfield J, Tyroler H. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. JAMA 1991; 266(1): 93-8.

Estudos Ecológicos

Roberto de Andrade Medronho

DEFINIÇÃO

São estudos em que a unidade de análise é uma população ou um grupo de pessoas, que geralmente pertence a uma área geográfica definida, como por exemplo um país, um estado, uma cidade, um município ou um setor censitário.

Segundo Greenland¹, esta denominação não seria a mais adequada, pois a palavra “ecológico” pode sugerir que estes estudos se referem especificamente à análise do impacto de fatores ambientais, incluindo as características sociais, quando na verdade são estudos de agregados de indivíduos.

Os estudos ecológicos são freqüentemente realizados combinando-se bases de dados referentes a grandes populações. Em função disto são, em geral, mais baratos e mais rápidos do que estudos envolvendo o indivíduo como unidade de análise.

Os estudos ecológicos procuram avaliar como os contextos social e ambiental podem afetar a saúde de grupos populacionais. Neste caso, as medidas coletadas no nível individual são incapazes de refletir adequadamente os processos que ocorrem no nível coletivo. Por exemplo, o nível de desorganização social de uma comunidade pode contribuir para que uma determinada epidemia ocorra de maneira mais intensa.

OBJETIVOS

Os principais objetivos dos estudos ecológicos são:

- gerar hipóteses etiológicas a respeito da ocorrência de uma determinada doença. Este é o objetivo mais comum dos estudos ecológicos;
- testar hipóteses etiológicas. Este objetivo, muitas vezes, é prejudicado por dificuldades de se controlar confundimento, como será visto mais adiante;
- avaliar a efetividade de intervenções na população, isto é, testar a aplicação de um determinado procedimento para prevenir doença ou promover saúde em grupos populacionais.

NÍVEIS DE ANÁLISE

Em uma *análise de nível individual* (estudo de coorte, caso-controle etc.), o valor de cada variável é assinalado para cada sujeito no estudo. Assim, é possível conhecer, por exemplo, entre os expostos, aqueles que adoeceram (a) e entre os não-expostos, aqueles que não adoeceram (d), conforme observado na Tabela 14.1.

Em uma *análise ecológica*, todas as variáveis são medidas agrupadas. Estas medidas não discriminam, por exemplo, os indivíduos expostos que adoeceram. Assim, dentro de cada grupo não é possível conhecer a distribuição conjunta de cada combinação de variáveis, ou seja, as freqüências de casos-expostos (a) e não-expostos (c) e as freqüências de não casos expostos (b) e não-expostos (d); apenas se conhece a distribuição marginal de cada variável: o total de indivíduos expostos (e_1) e não-expostos (e_0) e o total de casos (m_1) e não-casos (m_0). A Tabela 14.2 exemplifica os resultados de um estudo ecológico².

Tabela 14.1
Estudo de Nível Individual

Fator em Estudo (X)	Doença (Y)		Total
	Caso	Não-caso	
Exposto	a	b	e_1
Não-exposto	c	d	e_0
Total	m_1	m_0	n

Tabela 14.2
Estudo Ecológico

Fator em Estudo (X)	Doença (Y)		Total
	Caso	Não-caso	
Exposto	?	?	e_1
Não-exposto	?	?	e_0
Total	m_1	m_0	n

Em um estudo ecológico, a variável independente (X) é a proporção de indivíduos expostos dentro do grupo (e_1/n) e a variável dependente (Y) é a taxa de incidência da doença (m_1/n). Note que estas variáveis são contínuas, a despeito dos dois fatores serem dicotômicos no nível individual.

TIPOS DE VARIÁVEIS UTILIZADAS

As variáveis utilizadas nos estudos ecológicos podem envolver mensurações efetuadas diretamente no nível individual (idade, sexo, pressão arterial etc.), como também envolver observações de grupos, organizações ou lugares (densidade demográfica, desorganização social, poluição etc.). Estas variáveis podem ser classificadas em três tipos diferentes³:

- *medidas agregadas*: sintetizam as características individuais dentro de cada grupo. São construídas a partir da agregação das mensurações efetuadas no nível individual, como por exemplo, a proporção de fumantes, a taxa de incidência de uma doença e a renda familiar média.
- *medidas ambientais*: são características físicas do lugar onde os membros de cada grupo vivem ou trabalham (por exemplo, nível de poluição do ar, horas de exposição à luz solar). Cada medida ambiental tem uma análoga no nível individual, entretanto o nível de exposição individual pode variar entre os membros de cada grupo.

- *medidas globais*: são atributos de grupos, organizações ou lugares para os quais não existem análogos no nível individual (por exemplo, densidade demográfica, nível de desorganização social, existência de um determinado tipo de sistema de saúde).

TIPOS DE DESENHO

Os estudos ecológicos podem ser classificados em duas dimensões distintas, uma referente ao *método de mensuração da exposição* e a outra referente ao *método de agrupamento*. Em relação à primeira dimensão, os estudos ecológicos podem ser divididos em exploratórios, quando não existe uma exposição específica de interesse ou a exposição de interesse não é incluída no estudo; e em analíticos, se a exposição é mensurada e incluída na análise. Em relação à segunda dimensão, os estudos ecológicos podem utilizar grupos populacionais identificados a partir das diversas localidades geográficas (desenho de múltiplos grupos), através de diferentes períodos de tempo (desenho de séries temporais) ou combinando-se tempo e lugar (desenhos mistos). A seguir são descritos os diferentes tipos de desenho de estudo ecológico.

Desenhos de Múltiplos Grupos

- *Estudo exploratório*: as taxas de doença ou de outros agravos à saúde entre diversas regiões durante o mesmo período de tempo são comparadas entre si. O objetivo é identificar padrões espaciais desses eventos, gerando hipóteses de uma possível etiologia ambiental ou uma etiologia mais específica, relacionada às características da população, como as genéticas. Por exemplo: Medronho e cols.⁵ mapearam a soroprevalência de anticorpos contra o vírus da hepatite A resultante de um inquérito soeroepidemiológico realizado em 21 setores censitários de uma região carente do Município de Duque de Caxias, Rio de Janeiro (Fig. 14.1).

Freqüentemente, o mapeamento das taxas de doença para comparação entre as regiões pode conter dois tipos de problemas:

- regiões com poucos casos mostram grande variação na taxa da doença;
- regiões vizinhas tendem a ter taxas mais semelhantes do que regiões mais distantes. Este fenômeno é denominado de autocorrelação espacial⁶ e ocorre porque os fatores de risco tendem a se aglomerar no espaço.

- *estudo analítico*: avalia a associação entre o nível de exposição médio e a taxa de doença entre diferentes grupos. Este é o estudo ecológico mais comum. Por exemplo, Lefkowitz e Garland⁷, estudando a associação entre a radiação solar média e as taxas de mortalidade por câncer ovarianos específicas por idade nas 100 maiores cidades dos EUA, encontraram uma relação inversamente proporcional, corroborando a hipótese de que a radiação solar poderia ter um efeito protetor para câncer ovarianos (Fig. 14.2).

Desenho de Séries Temporais

- *estudo exploratório*: avalia a evolução das taxas de doença ao longo do tempo em uma determinada população geograficamente definida. Pode ser utilizado também para prever tendências futuras da doença ou avaliar o impacto de uma intervenção

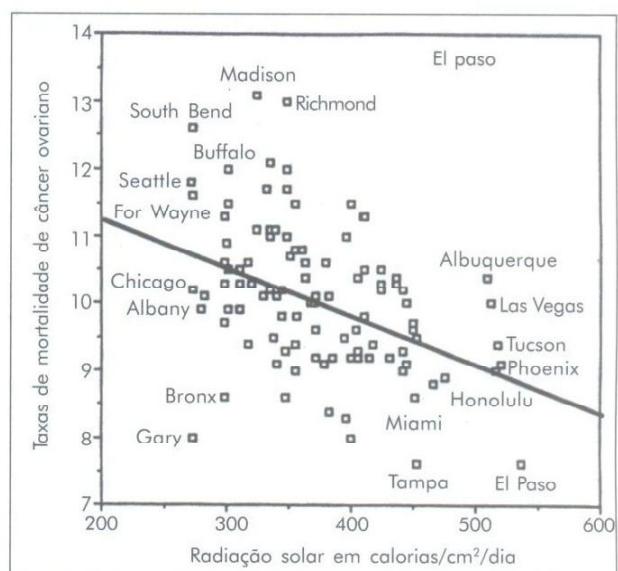


Fig. 14.2 – Taxa de mortalidade por 100.000 mulheres de câncer ovariano ajustada por idade em mulheres com idade maior ou igual a 35 anos de 100 cidades dos EUA, 1979-1988.

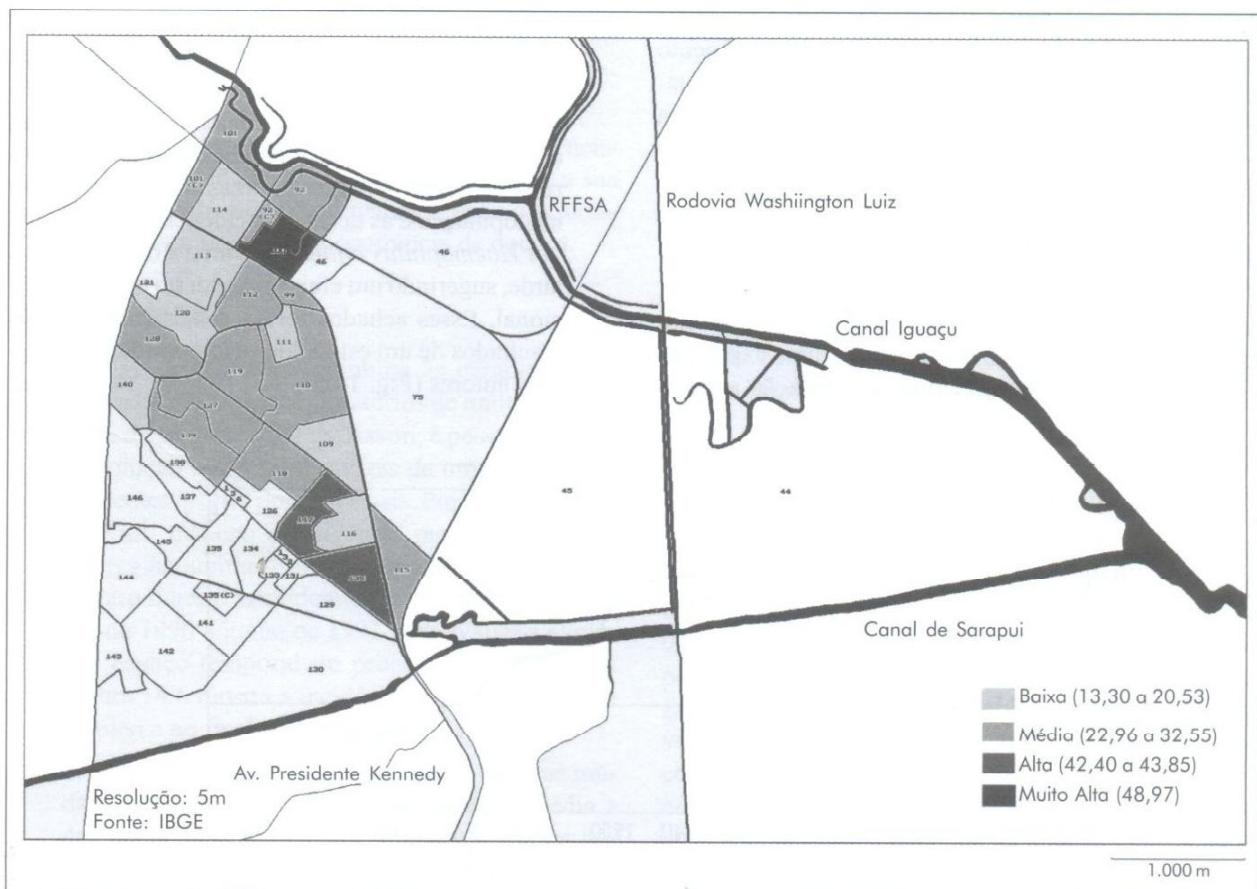


Fig. 14.1 — Taxas de prevalência padronizadas por idade em menores de 10 anos por setores censitários Parque Fluminense, Duque de Caxias, RJ.

populacional. Por exemplo, Antunes e Waldman⁸ analisaram as tendências da mortalidade por tuberculose no município de São Paulo no período de 1900 a 1997. Foram verificadas três tendências distintas: de 1900 a 1945, alta mortalidade com uma tendência estacionária; de 1945 a 1985, uma acentuada redução na mortalidade (7,41% por ano); e de 1985 a 1995, uma retomada do crescimento da mortalidade (4,08% por ano). Os autores ressaltam que a queda acentuada na mortalidade da tuberculose no período de 1945 a 1985 foi produzida por melhorias sociais, pela introdução de recursos terapêuticos eficientes e pela ampliação da cobertura dos serviços de saúde, e que o recrudescimento da tuberculose, de 1985 a 1995, pode estar refletindo a crescente prevalência da co-infecção por *Mycobacterium* e por HIV, além de uma aparente perda de qualidade nos programas de controle da doença (Fig. 14.3).

Um tipo de estudo especial é a análise de coorte de nascimento, que envolve a coleta de dados retrospectivos de uma grande população em um período geralmente maior que 20 anos com o objetivo de estimar os efeitos da idade, período (ano calendário) e ano de nascimento na taxa de doença. Por exemplo, Frost⁹ analisou as taxas de mortalidade por tuberculose em homens para Massachusetts em décadas sucessivas, segundo o ano de nascimento, e demonstrou que as curvas de mortalidade eram muito similares para todos os grupos, com um pico entre as idades de 20 e 30 anos e um declínio uniforme a seguir (Fig. 14.4).

- *Estudo analítico:* avalia a associação entre as mudanças no tempo do nível médio de uma exposição e das taxas de doença em uma população geogra-

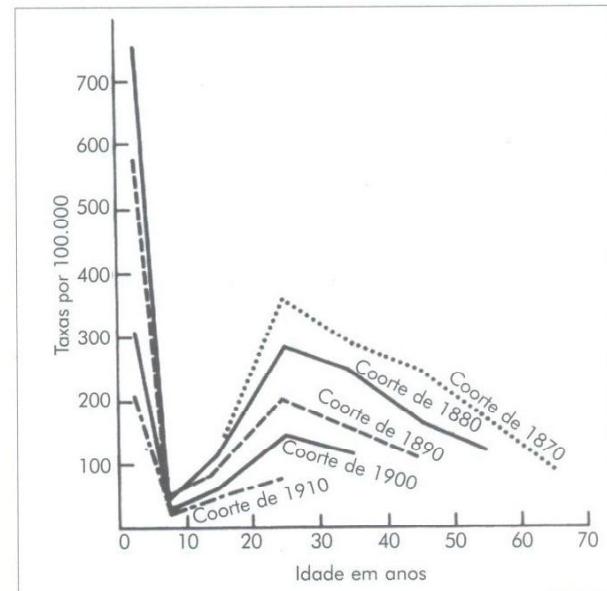


Fig. 14.4 – Taxas de mortalidade por tuberculose – todas as formas – específicas por idade, em homens, para coortes sucessivas de décadas, Massachusetts, EUA.

ficamente definida. Por exemplo, Silfverdal e cols.¹⁰ estudaram a relação entre taxas de incidência de meningite por *Haemophilus influenzae* entre 1956 e 1992 e taxas de amamentação em uma população sueca. Os autores encontraram uma forte correlação negativa entre as taxas de amamentação na população e as taxas de incidência de meningite por *Haemophilus influenzae* cinco a dez anos mais tarde, sugerindo um efeito protetor no nível populacional. Esses achados foram consistentes com os resultados de um estudo de caso-controle dos mesmos autores (Fig. 14.5).

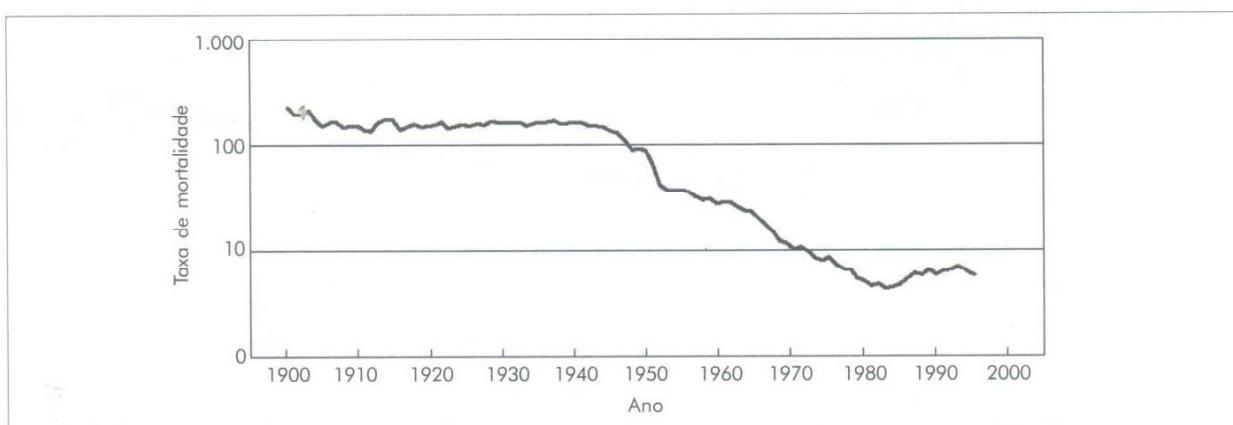


Fig. 14.3 – Séries temporais para a taxa de mortalidade por 100.000 habitantes (escala logarítmica) de tuberculose (todas as formas) padronizada por idade, São Paulo, Brasil, 1990-1997.

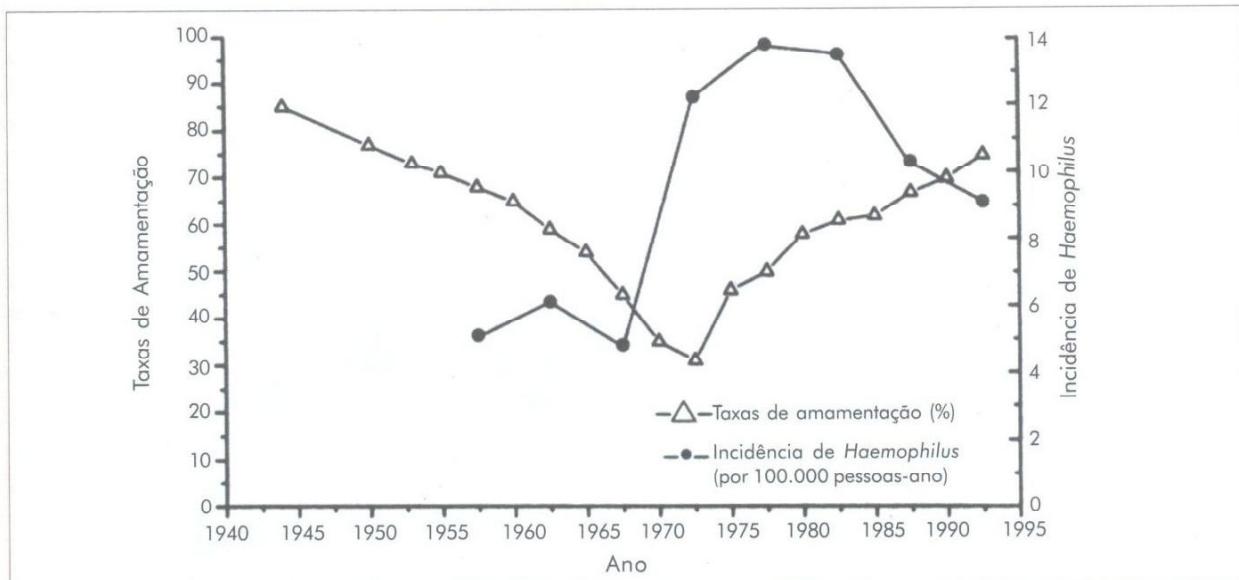


Fig. 14.5 – Séries temporais para taxas de amamentação e incidência de meningite por *Haemophilus influenzae*, Örebro, Suécia.

O processo de inferência causal dos estudos analíticos de séries temporais pode apresentar dois problemas:

- os critérios diagnósticos e de classificação das doenças podem se modificar ao longo do tempo e distorcer a tendência observada na taxa da doença;
- uma doença que possua grande período de latência/indução entre a exposição ao fator de risco e a sua detecção pode dificultar a avaliação entre a associação do fator de risco e a ocorrência da doença.

Desenhos Mistas

- *Estudo exploratório*: combina as características básicas dos estudos exploratórios de múltiplos grupos e de séries temporais. Assim, é possível avaliar a evolução temporal das taxas de uma doença em diferentes grupos populacionais. Por exemplo, Medronho¹¹ mapeou a incidência mensal de dengue por região administrativa do município do Rio de Janeiro durante a epidemia ocorrida entre novembro de 1990 a junho de 1991 para avaliar a evolução espaço-temporal do processo epidêmico. A Figura 14.6 mostra a incidência mensal de dengue no pico e no final da epidemia da doença.
- *Estudo analítico*: avalia a associação entre as mudanças no tempo do nível de exposição média e das taxas de doença entre diferentes grupos populacionais. Este procedimento potencializa a interpretação dos efeitos estimados, porque analisa simultaneamente as mudanças no nível de exposição

médio e nas taxas de doença em função do tempo dentro de grupos e as diferenças entre os grupos. Por exemplo, Crawford e cols.¹² avaliaram a hipótese de que o consumo de água dura (água com alta concentração de cálcio e magnésio) era um fator de proteção para mortalidade por doença cardiovascular (DCV). Os autores compararam as mudanças nas taxas de mortalidade por DCV entre 1948 e 1964 em 83 cidades britânicas, segundo as alterações na composição daqueles íons na água, idade e sexo, encontrando uma associação inversa entre a tendência na composição da água e as taxas de mortalidade por DCV em todos os grupos de idade e sexo, especialmente no sexo masculino.

INFERÊNCIA

O principal objetivo de um estudo epidemiológico é fazer inferências biológicas sobre os efeitos de uma determinada exposição no risco individual de adoecer ou fazer inferências ecológicas sobre os efeitos do nível de exposição médio nas taxas de doença em grupos populacionais^{4,13}. Entretanto, o nível utilizado para realizar a inferência causal nem sempre combina com o nível de análise. Por exemplo, se o propósito de uma análise ecológica é realizar uma inferência biológica sobre o efeito de uma exposição específica no risco de doença no nível individual, com freqüência pode incorrer em diversos problemas relacionados, principalmente a dificuldade de se controlar confundimento. Por exemplo, se o objetivo de um estudo é estimar o efeito físico do uso da

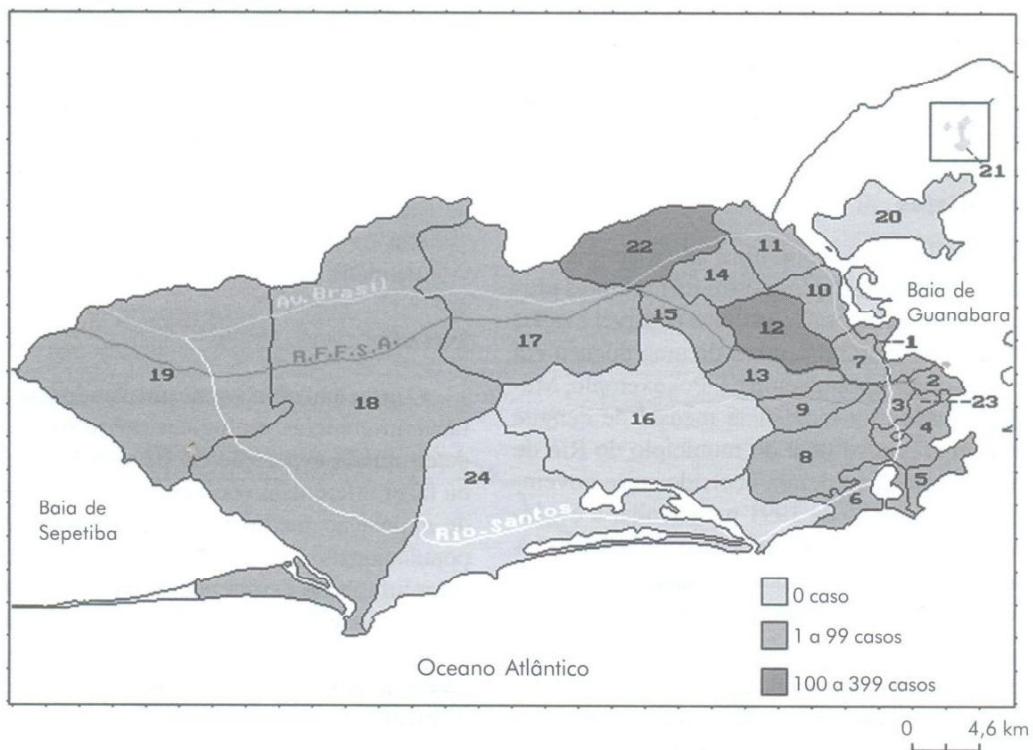
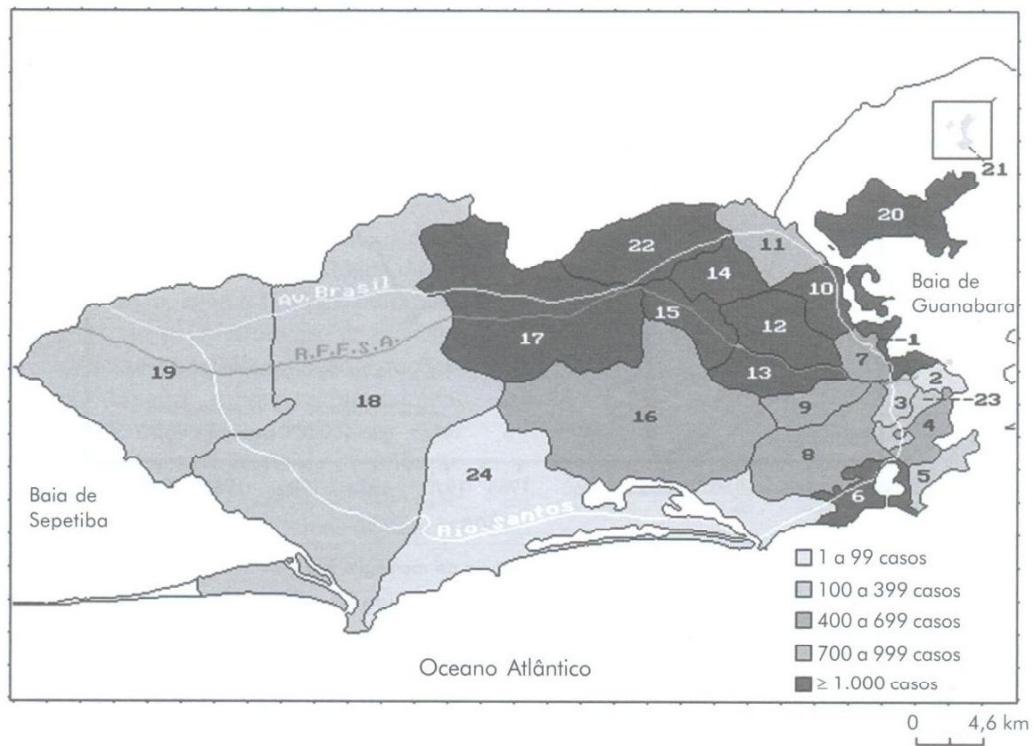


Fig. 14.6 — Incidência de dengue por região administrativa jan/91 e jun/91 — Município do Rio de Janeiro.

camisinha na prevenção da transmissão sexual da Aids, o nível-alvo de inferência causal é o individual. Por outro lado, se o objetivo é estimar o efeito ecológico dos programas de controle da Aids na taxa de incidência da doença nos diferentes Estados, o nível-alvo de inferência causal é o ecológico. Certamente, a magnitude do efeito ecológico depende não somente do efeito protetor do uso da camisinha, como também da efetividade do programa de controle em cada Estado. Além disso, a validade do efeito ecológico estimado vai depender da habilidade em controlar as diferenças entre os Estados na distribuição conjunta de confundidores, incluindo variáveis de nível individual, tais como idade, renda etc.

A maior limitação da análise ecológica para testar hipóteses ecológicas é o potencial para substancial viés na estimativa do efeito. O problema central, conhecido como viés ecológico (também chamado de falácia ecológica), resulta na realização de uma inferência causal inadequada sobre fenômenos individuais na base de observações de grupos, já que uma determinada associação verificada entre variáveis no nível agregado não necessariamente significa que exista essa associação no nível individual. O principal problema neste tipo de análise é a suposição de que os mesmos indivíduos são simultaneamente portadores do problema de saúde e do atributo associado, ou seja, este problema ocorre por não se conhecer a informação individual sobre a distribuição conjunta da exposição e doença. Por exemplo, Durkheim¹⁴, estudando províncias européias predominantemente protestantes no século XIX, observou taxas de suicídio maiores do que em províncias predominantemente católicas. Este achado pode sugerir que os protestantes tenderiam mais ao suicídio. Entretanto, poderiam ser os católicos residentes em províncias predominantemente protestantes os que mais se suicidavam. Dessa forma, a hipótese inicial seria em decorrência de um viés ecológico. Outro exemplo: suponha um estudo que analisou a associação entre renda média de determinadas cidades e ocorrência de acidentes de trânsito com morte no ano de 2007. Neste estudo, encontrou-se uma associação positiva entre a renda média e a proporção de acidentes graves nas cidades estudadas, ou seja, quanto maior a renda média da cidade, maior a proporção de acidentes de trânsito com óbitos. Uma conclusão aparentemente plausível poderia ser que quanto maior a renda do indivíduo, maior o seu risco de morrer de acidente de trânsito. Na verdade, quando se analisaram os dados no nível individual percebeu-se que os indivíduos de mais baixa renda eram os que mais morriam em decorrência dos acidentes de trânsito, mesmo nas cidades com renda média elevada, por estarem mais expostos. Assim, não se pode fazer

inferência para o nível individual diretamente de achados encontrados no nível de grupo.

Uma forma de minimizar o viés ecológico é através da utilização de dados agrupados em unidades de análise geográfica tão menores quanto possível, tornando-as mais homogêneas. Entretanto, deve-se atentar para a possibilidade de ocorrência de migração dentro do grupo e de estimativa de taxas instáveis, em função de se estar trabalhando com pequenos grupos populacionais.

Outro interesse de um estudo ecológico é estimar o *efeito contextual* de uma exposição ecológica no risco individual, como o efeito da poluição ambiental na produção de doença respiratória, por exemplo. Efeitos contextuais são fundamentais em epidemiologia das doenças infecciosas, onde o risco de doença depende da sua ocorrência em outros indivíduos com os quais se tem contato¹⁵.

Os achados de uma análise ecológica devem ser comparados com os achados de outros estudos observacionais, tanto ecológicos como individuais, desenhados para testar a mesma hipótese etiológica. Se o efeito estimado for consistente entre os estudos envolvendo diferentes desenhos e populações, uma interpretação causal pode ser reforçada.

Deve ser ressaltado que em função da análise ecológica utilizar medidas agregadas, isto introduz uma importante fonte de incerteza na inferência ecológica, pois existe uma perda de informação devido à agregação dos dados¹. Por exemplo, ao se comparar duas áreas A e B, sendo que a área A sofreu uma intervenção de um programa de saúde para o controle de uma doença Y e a área B não, constatou-se que a despeito do programa de saúde ter efetividade comprovada no nível individual, não houve diferença na taxa de incidência da doença entre as áreas A e B ao final do estudo. Entretanto, em um estudo agregado não é possível identificar a distribuição de um determinado fator de risco X, que era mais prevalente na área A do que na B. Assim, o efeito benéfico do programa de saúde na área A ficou obscurecido pela maior prevalência do fator de risco X nesta área.

ESTIMATIVA DO EFEITO

Muitos estudos epidemiológicos têm como objetivo estimar o efeito de uma exposição na ocorrência de uma determinada doença em uma população sob risco. Em estudos no nível individual, os efeitos são geralmente estimados pela comparação de taxas de incidência da doença em populações expostas e não-expostas, através de uma razão ou diferença entre as taxas. Em estudos ecológicos em múltiplos grupos, entretanto, como

não se conhece a informação sobre a distribuição conjunta entre exposição e doença dentro dos grupos, não se podem estimar os efeitos dessa forma. Nesse caso, deve-se realizar um procedimento de regressão das taxas de doenças (Y) nos níveis médios de exposição (X). O modelo mais comum para analisar dados ecológicos é o modelo linear, embora existam outros modelos. No modelo linear, o método dos mínimos quadrados ordinários pode ser usado para produzir a seguinte equação de predição: $\hat{Y} = B_0 + B_1 X$, onde B_0 e B_1 são o intercepto estimado e a angulação da reta, respectivamente. Uma estimativa do efeito da exposição no nível individual pode ser derivada dos resultados da regressão. Entretanto, isso somente é possível assumindo a predição de taxas de doença quando todos forem expostos (isto é, $X = 1$) e quando todos forem não expostos ($X = 0$).

A taxa de doença predita ($\hat{Y}_{X=1}$) em um grupo que é inteiramente exposto é $B_0 + B_1(1) = B_0 + B_1$ e a taxa de doença predita ($\hat{Y}_{X=0}$) em um grupo que é inteiramente não-exposto é $B_0 + B_1(0) = B_0$. Assim, a diferença de taxas estimadas é $B_0 + B_1 - B_0 = B_1$ e a razão de taxas estimadas é $B_0 + B_1/B_0 = 1 + B_1/B_0$. Note que o ajuste de um modelo linear poderia levar a uma estimativa negativa da razão de taxas, quando $B_1/B_0 < -1$, não possuindo qualquer significado biológico ou ecológico.

Para fins de exemplo, considere que um determinado estudo foi realizado em diversas cidades de um país e encontrou uma associação entre a proporção de expostos a um suposto fator de risco e a taxa de uma determinada doença. A regressão linear entre essas duas variáveis mostrou a seguinte equação:

$$\hat{Y} = 10 + 20X$$

Conseqüentemente, pode-se calcular a taxa de doença predita para um grupo inteiramente exposto e para um grupo inteiramente não-exposto:

Grupo exposto: $X = 1$

$$\hat{Y}_{X=1} = 10 + 20(1)$$

$$\hat{Y}_{X=1} = 30$$

Grupo não-exposto: $X = 0$

$$\hat{Y}_{X=0} = 10 + 20(0)$$

$$\hat{Y}_{X=0} = 10$$

Assim, a diferença de taxas estimadas e a razão de taxas estimadas serão:

Diferença de taxas: $30 - 10 = 20$

$$\text{Razão de taxas: } \frac{30}{10} = 3$$

Ressalta-se que essas estimativas não necessariamente irão refletir as estimativas no nível individual, pelas razões já expostas anteriormente.

VANTAGENS

As principais vantagens dos estudos ecológicos que contribuem para o uso freqüente como primeiro passo na investigação de uma possível associação entre exposição e doença são:

- os estudos ecológicos são geralmente de baixo custo e de execução rápida, devido às várias fontes de dados secundários já disponíveis, cada uma envolvendo diferentes informações necessárias para análise. Por exemplo, um estudo ecológico para se analisar a associação entre o nível socioeconômico e a incidência de tuberculose entre os diversos municípios do Estado do Rio de Janeiro envolveria dados obtidos do Censo Demográfico do IBGE e do Sinan, respectivamente;
- freqüentemente não é possível medir de maneira acurada exposições para grandes números de indivíduos, em função dos recursos e tempo disponíveis. Assim, uma forma prática é medir a exposição no nível ecológico^{2,16};
- estudos de nível individual não conseguem estimar bem os efeitos de uma exposição, quando ela varia pouco na área de estudo. Como os estudos ecológicos cobrem áreas muito grandes, podem com mais facilidade encontrar uma maior variação na exposição média entre as diferentes regiões;
- um estudo pode ter como objetivo mensurar um efeito ecológico. Por exemplo: a implantação de um novo programa de saúde ou uma nova legislação em saúde na melhoria das condições de saúde de uma população.

LIMITAÇÕES

As principais limitações dos estudos ecológicos são:

- não é possível associar exposição e doença no nível individual;
- dificuldade de controlar os efeitos de potenciais fatores de confundimento;

- dados de estudos ecológicos representam níveis de exposição média em vez de valores individuais reais;
- os dados são provenientes de diferentes fontes, o que pode significar qualidade variável da informação;
- a falta de disponibilidade de informações relevantes é um dos mais sérios problemas na análise ecológica.

MODELOS MULTINÍVEL

Uma população humana não é um mero somatório de indivíduos, pois eles estabelecem complexas relações sociais entre si. Da mesma forma, um indivíduo não é um mero agregado de diferentes células. Assim, muitos estudos no nível individual podem apresentar problemas de inferência se os indivíduos forem descontextualizados de seu meio social. Da mesma forma, um estudo no nível ecológico também pode apresentar problema se for considerado que o somatório (agregado) de indivíduos é capaz de captar as complexas relações que se estabelecem dentro dos grupos e entre os grupos sociais. Deve-se sempre levar em conta que um determinado grupo exerce influência sobre cada indivíduo e que o indivíduo também desempenha influência na população a que pertence.

Um bom exemplo pode ser um processo epidêmico de uma doença infecciosa contagiosa. O risco de adoecer de um indivíduo está relacionado com a incidência da doença no grupo social a que ele pertence. Por outro lado, a presença de um indivíduo infectado em um grupo suscetível vai contribuir para a disseminação da doença no grupo a que ele pertence.

A análise de dados agregados de uma população deve levar em conta que eles representam pontualmente um acúmulo de complexos processos históricos, geográficos e sociais. Diversas técnicas quantitativas vêm sendo desenvolvidas para incorporar não apenas a variabilidade entre os indivíduos, mas também a variabilidade existente entre os diferentes grupos sociais a que eles pertencem. Neste sentido, a utilização do *modelo multinível* permite a integração entre os distintos níveis: o fisiológico, que avalia a resposta dos sistemas orgânicos às exposições; o *individual*, que avalia a resposta dos indivíduos às exposições e o *agregado ou contextual*, que avalia a resposta de grupos de indivíduos às exposições, destacando que esses indivíduos pertencem a localidades e sociedades¹. Na análise multinível a explicação da variabilidade na variável dependente em um nível é função das variáveis definidas em diversos níveis, além das interações dentro e entre os níveis.

Segundo Susser¹⁷, os epidemiologistas têm levado cada vez mais em conta os múltiplos níveis de causalidade. Assim, as causas de um desfecho de saúde podem ser pesquisadas no nível macro, através da distribuição da riqueza entre e dentro de sociedades; no nível individual, através do estudo do comportamento dos indivíduos e no nível micro, na expressão de genes das células. Note que uma determinada doença se expressa em um indivíduo, entretanto, sua prevenção pode ser dirigida a qualquer um desses níveis. É importante utilizar diferentes abordagens para se identificar os processos causais em cada um desses níveis, mas também é preciso examinar como os processos de um nível (por exemplo, desagregação social ou alterações celulares) podem se manifestar em outro nível (por exemplo, doença em um indivíduo).

A noção básica é que os indivíduos são influenciados por seu contexto social e que o processo saúde-doença ocorre neste contexto. Entretanto, a pesquisa dos determinantes de saúde freqüentemente é focada no indivíduo, ou seja, a explicação dos desfechos no nível individual é feita na maioria das vezes utilizando-se variáveis independentes de nível individual e as variáveis de grupo são utilizadas somente como *proxies* quando os dados individuais não estão disponíveis.

O modelo multinível (também chamado de modelo hierárquico) é uma versão do modelo de regressão múltipla para dados que possuem níveis hierárquicos. A noção básica é que os indivíduos que pertencem às áreas sob estudo interagem entre si e são influenciados pelo nível hierarquicamente superior. Por exemplo, o indivíduo sofre influencia do nível municipal que, por sua vez, é influenciado pelo nível estadual. Por outro lado, as características do nível estadual influenciam seus municípios, que por sua vez vão influenciar os indivíduos residentes nestes municípios. Assim, os indivíduos, os municípios e o estado formam um sistema hierárquico em três níveis. A suposição é a de que os indivíduos pertencentes ao Município tendem a ser mais semelhantes entre si do que os indivíduos residentes em outro Município do mesmo Estado. Da mesma forma os Municípios que pertencem a um Estado são mais semelhantes entre si do que aqueles pertencentes a outro Estado. Por exemplo, os indivíduos residentes em um Município do Estado do Rio Grande do Norte tendem a ser mais semelhantes que os residentes em outro Município do mesmo Estado, assim como, os Municípios do Estado do Rio Grande do Norte tendem a ser mais semelhantes entre si do que os Municípios do Estado do Rio Grande do Sul¹⁸.

Note que o uso de um modelo de regressão tradicional pode ser inadequado para a situação descrita anteriormente porque ele pressupõe a independência entre as observações e a homogeneidade da variância (homocedasticidade). Neste caso, o valor de uma determinada variável contextual será o mesmo para todas as observações do primeiro nível pertencentes a uma mesma unidade de observação do segundo nível, pois este nível não é levado em consideração neste tipo de modelo. Assim, ocorrerá uma redução artificial da variabilidade do coeficiente estimado para esta variável “contextual”, podendo produzir uma significância estatística espúria no modelo¹⁹.

Na área da saúde pública, o uso do modelo multinível vem crescendo em função do interesse pelos macrodeterminantes da saúde, da importância de se levar em conta os diferentes níveis de análise em estudos epidemiológicos e do acelerado desenvolvimento de métodos estatísticos e de softwares para a abordagem de dados com estrutura hierárquica.

BIBLIOGRAFIA

1. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol* 2001; 30(6):1343-1350.
2. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health* 1982; 72: 1336-1344.
3. Morgenstern H. Ecologic Studies. In: Rothman KJ, Greenland S. Modern Epidemiology. 2. ed. Philadelphia: Lippincott-Raven Publishers, 1998.
4. Morgenstern H. Ecologic studies in Epidemiology: concepts, principles, and methods. *Annu Rev Public Health* 1995; 16: 61-81.
5. Medronho RA, Gomide M, Guimarães LB, Oliveira LM, Braga, RCC. Uso do geoprocessamento na investigação do papel do meio ambiente na produção de hepatite A. *Cad Saúde Coletiva* 1998; 6(supl 1): 61-78.
6. Morgenstern H, Thomas DC. Principles of study design in environmental epidemiology. *Environ Health Perspect* 1993; 101(suppl 4): 23-38.
7. Lefkowitz ES, Garland CF. Sunlight, vitamin D, and ovarian cancer mortality rates in US women. *Int J Epidemiol* 1994; 23(6): 1133-1136.
8. Antunes JLF, Waldman EA. Tuberculosis in the twentieth century: time-series mortality in São Paulo, Brazil, 1900-97. *Cad Saúde Pública* 1999; 15(3): 463-476.
9. Frost WH. The age selection of mortality from tuberculosis in successive decades. *Am J Hyg* 1939; 30: 91-96.
10. Silfverdal SA, Bodin L, Olcen P. Protective effect of breastfeeding: an ecologic study of *Haemophilus influenzae* meningitis and breastfeeding in a Swedish population. *Int J Epidemiol* 1999; 28(1):152-156.
11. Medronho RA. Geoprocessamento e saúde: uma nova abordagem do espaço no processo saúde-doença. Rio de Janeiro: NECT/FIOCRUZ, 1995.
12. Crawford MD, Gardner MJ, Morris JN. Changes in water hardness and local death-rates. *Lancet* 1971; 2: 327-329.
13. Cressie NAC. Statistics for spatial data. New York: Wiley, 1991.
14. Durkheim E. Suicide: a study in sociology. New York: Free Press, 1951.
15. Koopman JS, Longini IM Jr. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am J Public Health*; 84: 836-842, 1994.
16. Szkló M, Nieto J. Epidemiology: Beyond the Basics, 2 ed. Maryland: Aspen Publishers, 2006.
17. Susser E. Eco-epidemiology: thinking outside the black box. *Epidemiology* 2004; 15:519-20.
18. Hox JJ. 1995. Applied multilevel analysis. Amsterdam: TT-Publikaties.
19. Diez-Roux AV, 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Pub Health* 88(2): 216-222.