

Princípios de Epidemiologia Clínica Aplicada à Cardiologia

Mário Coutinho

Florianópolis, SC

Através dos tempos, a experiência pessoal tem guiado o médico nas suas decisões. Entretanto, com o aumento exponencial, em volume e complexidade, da literatura médica, a figura do médico depositário de todo conhecimento, tornou-se anacrônica. Diante deste quadro, fez-se necessária a definição de critérios de avaliação da evidência científica, que orientassem o clínico a identificar os estudos válidos e aplicáveis aos cuidados dos pacientes. Porém, estamos distantes da situação ideal. Pois, estima-se que, cerca de 50% das práticas médicas correntes, foram incorporadas na prática sem uma avaliação científica rigorosa¹.

A Cardiologia é uma especialidade afortunada, pois dispõe de grande volume de evidência científica de alta qualidade, adquirida nas últimas duas décadas, com a realização dos chamados *mega-trials*. Estes estudos randomizados controlados nos ensinaram algumas lições importantes, como, por exemplo, que uma hipótese bem fundamentada, do ponto de vista fisiopatológico, pode ser negada pelos resultados do ensaio clínico.

O caso dos anti-arrítmicos no pós-infarto², e o das drogas inotrópicas positivas não-digitálicas, na insuficiência cardíaca³ são bastante ilustrativos. Ambos testaram hipóteses baseadas em mecanismos e processos fisiológicos (supressão de arritmias no pós-IAM e aumento da fração de ejeção na ICC) utilizando desfechos clinicamente importantes, tais como mortalidade. Os resultados mostraram um efeito na direção contrária. As drogas testadas aumentaram a morbidade e mortalidade dos pacientes, motivo pelo qual os estudos foram suspensos precocemente.

Com base nesses fracassos terapêuticos, cresceu entre os cardiologistas, a necessidade de conhecer os conceitos que regem a pesquisa clínica, com o objetivo de melhor avaliar a literatura médica e aplicar seus resultados, de maneira adequada, na clínica.

Epidemiologia clínica: uma nova ciência básica para o clínico

A Epidemiologia Clínica surgiu da necessidade de tornar a medicina-arte, mais científica. Vários fatores contri-

buíram para este surgimento: a) a grande variabilidade de condutas clínicas, sem diferenças substanciais nos resultados clinicamente importantes; b) o crescimento exponencial do conhecimento e do volume de publicações na área médica; c) a heterogeneidade qualitativa daquelas publicações; d) os gastos crescentes no setor da saúde⁴.

Criou-se, assim, a necessidade de se definir métodos mais rigorosos de avaliação da evidência científica, e de basear a prática clínica na melhor evidência disponível. Além disso, segundo este novo paradigma, as pesquisas clínicas devem priorizar os desfechos, clinicamente relevantes (mortalidade, morbidade e incapacidade), deixando, em segundo plano, as pesquisas que avaliam desfechos intermediários e fisiológicos.

A Epidemiologia Clínica reúne conceitos da medicina clínica e da epidemiologia tradicional, e tem por objetivo, auxiliar o clínico na solução de questões diagnósticas, terapêuticas e prognósticas, que se apresentam diariamente na prática clínica⁵.

A era científica da arte da Medicina, iniciou-se no pós-guerra, com a realização dos primeiros ensaios randomizados⁶ e, desde então, vem evoluindo. Antes desta nova era, por falta de evidências científicas sólidas, o tratamento das doenças era feito com base em estudos observacionais, não-controlados, abertos e com amostras de tamanho pequeno. Além disso, o saber fisiopatológico exercia uma forte influência sobre as condutas clínicas adotadas na prática.

Através do conhecimento da metodologia dos estudos clínicos publicados e da sua análise crítica, o médico pode decidir sobre a sua validade e aplicabilidade na clínica, sendo estes, alguns dos objetivos desta nova “ciência básica para o clínico”⁴.

Princípios básicos da epidemiologia clínica

A epidemiologia clínica baseia-se em alguns fundamentos: a) no princípio das probabilidades, dado que as situações clínicas, que envolvem diagnóstico, tratamento e prognóstico são incertas e necessitam de uma estimativa numérica, que traduza cada situação; b) a melhor estimativa para um paciente individual, baseia-se na experiência anterior, com grupos similares de pacientes; c) as pesquisas clínicas podem ser afetadas por erros sistemáticos (viéses ou *biases*), originados tanto do investigador quanto do paciente, que podem invalidar suas conclusões; d) toda observação clínica está sob a influência do acaso (*chance*); e) o

Universidade Federal de Santa Catarina - Florianópolis
Correspondência: Mário Coutinho - Rua Percy João de Borba, 129 - 88036-200 - Florianópolis, SC
Recebido para publicação em 6/5/98
Aceito em 8/6/98

clínico deve orientar sua prática por observações baseadas em princípios científicos sólidos, que incluam o controle de vieses e a estimativa do papel do acaso sobre os resultados⁵.

População e amostra

População é um grupo de indivíduos que vive num determinado contexto (p.ex. habitantes do Brasil) ou que têm uma característica comum (p.ex. idade >60 anos). Ao estudarmos uma população, muitas vezes não é possível obter dados da sua totalidade, por isso, recorremos a amostras da população. A amostra pode ser obtida por conveniência (p.ex. pacientes internados em uma unidade hospitalar) ou de maneira aleatória (p.ex. hipertensos adultos de um bairro, detectados em inquérito de morbidade, precedido por um processo de amostragem aleatória). A amostragem aleatória, por ser composta ao acaso e não depender de critérios do investigador, é representativa da população. Por exemplo, se examinarmos uma amostra aleatória de 10% da população adulta do bairro, e encontrarmos 20% de hipertensos, temos razoável segurança em afirmar que, aproximadamente, 20% das pessoas adultas do bairro são portadoras de hipertensão. A escolha aleatória da amostra tira do investigador o poder de definir, de antemão, quem fará parte da amostra, evitando, assim, o viés de seleção (ver adiante). Se a amostra é de conveniência, não é possível se determinar quais foram os reais critérios de sua seleção, o que a torna suspeita de vies.

Números, probabilidades, risco e estatística

A Epidemiologia Clínica tem por objeto a pesquisa quantitativa e usa números e probabilidades para expressar desfechos clínicos, sintomas e incapacidade. Números são formas claras e objetivas de expressar e comunicar informações, eliminando a subjetividade e tornando possível as comparações. Predizer, com certeza, um desfecho clínico não é possível, motivo pelo qual usamos as probabilidades. Estas expressam a experiência prévia de pessoas semelhantes ao paciente em questão, observadas por um período. Por exemplo, a mortalidade no 1º ano em pacientes com IAM inferior não-complicado é menor que 1%, ao passo que em pacientes com IAM anterior e insuficiência ventricular esquerda há uma probabilidade cinco vezes maior de óbito no 1º ano. Estas estimativas são feitas com base em estudos prospectivos, e é a melhor estimativa disponível para estabelecer prognósticos na clínica.

As estimativas probabilísticas são expressas em decimais, que variam entre zero (impossibilidade) e um (certeza). Muitas vezes, transformam-se os decimais em porcentagens para melhor entendimento e comunicação.

Todos têm uma noção razoável de probabilidades e acaso, e o jogo é a forma mais popular de vê-los em ação. A probabilidade de tirarmos "cara" jogando uma moeda é de $1/2=0,5$ ou 50%. Da mesma forma, se apostarmos que, um dado honesto, ao ser jogado, mostrará o número 5, teremos $1/6=0,17$ ou 17% de probabilidade de acerto.

Na atividade clínica diária, lidamos com probabilidades, mesmo que, de maneira informal. Por exemplo, ao estimarmos o risco de um paciente desenvolver um infarto do miocárdio, com base em seus fatores de risco, ou ao avaliarmos os resultados de um teste diagnóstico no contexto dos dados clínicos do paciente. A estatística é uma disciplina que lida com dados numéricos coletados de maneira sistemática e os classifica, tabula e analisa. Através de procedimentos indutivos, ela generaliza os resultados de uma amostra para a população em estudo. Existem duas classes de procedimentos indutivos estatísticos: a estimação de parâmetros e os testes de hipótese. No 1º, procede-se a uma descrição dos dados e sua precisão (p.ex. média e desvio-padrão, ou taxa de eventos e intervalo de confiança). Os objetivos estatísticos na pesquisa clínica são: a descrição dos dados, a estimação de parâmetros, a exploração de associações entre as variáveis, a comparação de grupos e, finalmente, os modelos de regressão.

Na pesquisa clínica, o evento de interesse pode ser visto como uma resposta binária. Isto é, se estamos interessados em um dos seguintes desfechos no portador de doença aterosclerótica coronária: infarto do miocárdio ou óbito. Cada categoria de desfecho pode ser considerada uma resposta binária: infarto x não-infarto e óbito x não-óbito. Ou de um modo geral, sucesso ou insucesso. Com base nisto, podemos derivar taxas de risco do evento e medidas de precisão daquelas taxas (ver adiante). Conceitua-se risco como "a probabilidade de um indivíduo desenvolver uma mudança no seu status de saúde, ao longo de um determinado período de tempo".

Os conceitos de risco e probabilidade, assim como a sua aplicação prática, constituem instrumentos, que nos auxiliam na interpretação dos resultados dos estudos clínicos de intervenção e dos estudos observacionais. O domínio destes princípios estatísticos se constitui num dos pilares da epidemiologia clínica. Em artigos subsequentes desta série estes tópicos deverão ser detalhados.

Variáveis e medidas

A pesquisa clínica quantitativa implica na medida de variáveis, que constituir-se-ão nos dados a serem descritos e analisados, ao final do estudo.

Quando se considera o tópico mensuração, é importante se fazer a distinção entre os diferentes tipos de variáveis. Fundamentalmente, elas podem ser divididas em categóricas (morto/vivo, homem/mulher) e contínuas (pressão arterial (PA), altura, peso). Variáveis categóricas podem assumir apenas certos valores e são chamadas de discretas, ao passo que, as variáveis contínuas, em teoria, podem assumir um número infinito de valores.

Variáveis categóricas podem ser nominais ou ordinais. As variáveis nominais não obedecem a nenhuma ordem inerente. Por exemplo, sexo (feminino e masculino), tipo sanguíneo (A,B,AB,O) e hábito de fumar (sim ou não). Quando uma variável nominal só tem duas categorias ela é chamada de dicotômica ou binária. As variáveis ordinais,

	Droga A	Segundo IAM	
Segundo IAM	75 a	100 b	175 (a+b)
Sem novos eventos	925 c	900 d	1825 (c+d)
	1000 (a + b)	1000 (b + d)	2000 (a + b + c + d)

Fig. 1 - Cálculo das medidas de avaliação com dados hipotéticos. 1) Risco relativo (RR) = $(a/a+c)/(b/b+d) = (75/1.000)/(100/1.000) = 0,75/0,100 = 0,75$ ou 75%; 2) redução absoluta de risco (RAR) = $(a/a+c) - (b/b+d) = 0,100 - 0,075 = 0,025$ ou 2,5%; 3) redução relativa de risco (RRR) = $RAR/(b/b+d) = 0,025/0,100 = 0,25$ ou 25%; 4) número necessário de tratar (NNT) = $1/RAR = 1/0,025 = 40$; 5) odds ratio (OR) = $(a \cdot d)/(b \cdot c) = (75 \cdot 900)/(100 \cdot 925) = 0,73$ ou 73%.

por outro lado, possuem uma ordem inerente e podem mostrar uma gradação, por exemplo, pequeno-médio-grande, leve-moderado-severo. Porém, os intervalos entre as categorias são subjetivos, como no caso da classificação por “cruzes” do edema periférico, ou dos sopros cardíacos.

Variáveis contínuas têm ordem inerente e o intervalo entre valores sucessivos é constante. Como, por exemplo, a concentração de substâncias bioquímicas no sangue (sódio, creatinina, colesterol), peso corporal e PA⁵.

Medidas de benefício clínico

Medir eventos é a atividade cotidiana do pesquisador clínico. Os resultados de uma pesquisa clínica são expressos em número de eventos e taxas, onde o denominador representa o número de pessoas em risco (todos os membros do grupo) e o numerador representa o número de eventos ocorrido no grupo.

Por exemplo, em um ensaio clínico randomizado em pacientes pós-IAM, 75 pacientes num total de 1.000, do grupo, que receberam a droga A, sofreram um 2º IAM no 1º ano. Por outro lado, 100 em 1.000 do grupo placebo tiveram um 2º IAM naquele período.

A partir destes números, podemos derivar cinco importantes medidas de efeito clínico: a redução absoluta do risco (RAR), o risco relativo (RR), a redução relativa do risco (RRR), o número necessário de tratar (NNT) e a razão de chances (*odds ratio* (OR)) (fig. 1).

A RAR mostra a dimensão real e absoluta do efeito da intervenção e produz outra importante medida, o NNT. No exemplo da figura 1, o NNT de 40 significa que, se aplicarmos os resultados à prática, será necessário tratar 40 pessoas com a droga A, para evitar um 2º infarto no 1º ano. Este índice é útil, pois coloca, numa perspectiva clínica, os resultados da pesquisa. O RR estima a relação entre a magnitude do efeito da intervenção (droga A) e o placebo, sendo que, se <1, revela benefício da intervenção, e se >1, um efeito indesejável. OR ou razão de chances foi originariamente desenvolvida para estimar o risco em estudos de caso-contro-

le, porém hoje é usado como uma aproximação razoável, e algumas vezes equivalente, do RR, em especial, em doenças pouco frequentes (incidência <5%). OR e RR apontam sempre na mesma direção, e as diferenças numéricas, que possam existir, são, em geral, de pouco significado prático. Além disso, o cálculo do OR é mais simples do que o do RR⁷. A RRR é o complemento do RR (RRR = 1 - RR), isto é, ela nos dá a magnitude do efeito terapêutico da droga A em termos de benefício, ou seja, expressa em que proporção a droga A diminui a taxa de eventos em relação ao placebo.

Medida de precisão: intervalo de confiança

A precisão estatística de uma estimativa pontual é expressa pelo intervalo de confiança, usualmente, o intervalo de confiança de 95% em torno da estimativa. Sua interpretação é a seguinte: em um estudo livre de vieses, existe uma probabilidade de 95% de que o intervalo inclua o real efeito clínico da intervenção sob investigação. Quanto mais estreito for este intervalo, maior a probabilidade de que aquela seja a real magnitude do efeito. Por outro lado, intervalos muito amplos nos dão menos segurança na estimativa do efeito clínico da intervenção⁸. A precisão estatística aumenta com o poder estatístico do estudo, que por sua vez, depende do tamanho da amostra (ver adiante).

No exemplo da figura 1, houve uma RRR de 25% (0,25) na incidência de um 2º IAM com o uso da droga A. O intervalo de confiança de 95% calculado foi de 0,22 a 0,28 (22% a 28%), ou seja, com 95% de probabilidade, o efeito da droga A se situa dentro deste intervalo, diminuindo entre 22 e 28% a incidência de eventos. O intervalo de confiança contém informação similar à de um teste de significância estatística. Se o valor correspondente à ausência de efeito (diferença entre tratamentos igual a zero ou RR igual a 1) está fora do intervalo de 95%, o resultado é significativo para um nível de P igual a 0,05. Da mesma forma, se o intervalo contém o zero ou o um (conforme o caso), o resultado não é significativo. No exemplo acima, se o intervalo de confiança fosse de -0,20 a +0,45, teríamos, basicamente, duas informações importantes: 1ª, o intervalo contém o zero, ou seja, ausência de diferença entre os tratamentos e, portanto, não é, estatisticamente, significativo; 2ª, o intervalo de confiança é muito amplo, isto é, varia de menos 20% a mais 45%, refletindo que estamos lidando com uma amostra de tamanho pequeno. Além disso, o efeito clínico da intervenção, pode tanto ser benéfico, quanto ausente, ou mesmo, prejudicial, uma vez que o intervalo abrange números negativos e positivos. Portanto, o intervalo de confiança é uma forma simples de avaliar o efeito clínico em estudos experimentais, além de fornecer informações sobre a significância estatística e a magnitude daquele efeito.

Validade interna e validade externa

A validade interna define, até que ponto, os resultados de um estudo são corretos para a amostra de pacientes estudados. Chama-se interna porque se aplica às condições da-

quele grupo em particular, e não, necessariamente, a outros grupos. A validade interna é determinada pela qualidade do planejamento e da execução do estudo, incluindo adequada coleta e análise dos dados. Os erros sistemáticos e o acaso podem ameaçar a validade interna do estudo, e as medidas para seu controle devem ser previstas durante as fases iniciais de planejamento de um estudo de qualidade. A validade interna é uma condição necessária, mas não suficiente, para que um estudo clínico possa ser considerado de utilidade prática.

Validade externa diz respeito ao grau de aplicabilidade, ou de generalização, dos resultados de um estudo em particular, para outros contextos. Ao ler um relatório de um estudo, o clínico deve se perguntar, assumindo que os resultados sejam verdadeiros: “eles são aplicáveis aos meus pacientes?”. Se a resposta for positiva, teremos o que chamamos de validade externa. Este é um conceito subjetivo, e depende do julgamento do médico, que deve decidir se a amostra do estudo em questão, é tão parecida com os seus pacientes, a ponto dos seus resultados serem aplicáveis na prática⁵.

Confiabilidade e validade de um teste diagnóstico

Testes diagnósticos fazem parte da rotina médica. Entretanto, sua interpretação está condicionada à confiabilidade dos resultados. Um teste é confiável, quando seus resultados são reproduzidos, consistentemente, após a repetição do teste por uma ou mais vezes, em diferentes momentos e lugares. Em outras palavras, a medida realizada é reprodutível em diferentes situações, isto é, com diferentes observadores, ou com o mesmo observador em ocasiões diferentes, separados por um curto espaço de tempo. São sinônimos de confiabilidade: reprodutibilidade, precisão, fidedignidade e repetibilidade. Por exemplo, quando dosamos uma enzima miocárdio-específica (CK-MB) para detectar um IAM, os resultados devem ser similares, se o teste for repetido, na mesma amostra de sangue, por mais de uma vez.

A confiabilidade deve se diferenciada da acurácia do teste. A acurácia de um teste é medida pelo número de verdadeiro-positivos e verdadeiro-negativos em relação a todos os indivíduos testados. Poucos falso-positivos e falso-negativos refletem uma alta acurácia.

A verdade científica e suas determinantes

A ciência busca a verdade, esta definida como, “aquilo que está de acordo com a realidade”. A verdade científica é dinâmica e não-absoluta, pois, novas informações e formas de abordar um mesmo problema são propostos a cada dia. Assim, antigos dogmas dão lugar a novos conhecimentos. Isto nos alerta para o caráter transitório da verdade, e para uma permanente atitude crítica em relação a posições rígidas, baseadas na autoridade e na experiência.

Uma verdadeira revolução no tratamento e diagnósti-

co das doenças cardiovasculares ocorreu nas últimas duas décadas. Graças aos *mega-trials* (ISIS, GISSI, GUSTO e outros), a cardiologia dispõe de enorme volume de evidência sobre intervenções na doença coronária. Porém, antes dos ensaios randomizados controlados se tornarem o padrão de avaliação de intervenções terapêuticas, um longo caminho foi percorrido. Grande parte da prática médica era baseada em estudos observacionais, ensaios não-controlados, séries de casos e outros estudos metodologicamente frágeis. Esse panorama ainda não se modificou completamente e estima-se que apenas cerca de 15% dos procedimentos médicos, atualmente, são apoiados por evidência científica válida.

A busca da verdade envolve a aplicação rigorosa do método científico que, parte de uma hipótese de trabalho, testa-a num experimento e, finalmente, aceita-a ou refuta-a. Ao se testar uma hipótese, procura-se controlar todas as potenciais fontes de erros sistemáticos e aleatórios do estudo, para que, ao final, suas conclusões, possam ser consideradas válidas. Desta forma, podemos falar de aproximações da verdade, uma vez que a verdade absoluta é uma abstração. Portanto, esquematicamente, a verdade na pesquisa científica, pode ser definida como o resultado de uma observação empírica, controlados os erros sistemático e aleatório; **verdade = observado - (erro sistemático + erro aleatório)**.

Erro sistemático

O erro sistemático, ou viés (*bias*, do inglês) é definido como qualquer processo, em qualquer estágio da inferência, que tende a produzir resultados e conclusões, que diferem sistematicamente da verdade. Seu efeito é o de distorcer a estimativa de uma variável, por exemplo, aumentando a média de uma variável (PA) ou diminuindo a prevalência de uma característica (frequência de fumantes).

Basicamente, temos três grupos de vieses, embora mais de 40 já tenham sido catalogados⁹, são eles: a) viés de seleção; b) viés de aferição e c) viés de confusão (*confounding*).

O viés de seleção ocorre quando a amostra do estudo não é representativa da população. Ele é resultante da maneira como os indivíduos foram selecionados para o estudo. Por exemplo, dois grupos de pacientes diferem entre si, quanto aos fatores prognósticos para os desfechos clínicos finais do estudo. Imaginemos um estudo aberto (o investigador decide para qual grupo o paciente será alocado) com a droga A versus placebo em pacientes pós-IAM. Se o investigador tem a priori a convicção de que a droga A é pior do que o placebo, ele poderá alocar pacientes de melhor prognóstico (p.ex. com fração de ejeção do ventrículo esquerdo (FEVE) >0,40) para a droga A e os de pior prognóstico (FEVE <0,40) para o grupo placebo, evitando, assim, que pacientes de maior risco sejam expostos a uma droga que, segundo suas convicções, não funciona. Contudo, ao final do estudo, ele terá dois grupos heterogêneos, no que concerne a um fator prognóstico importante, fato que poderá invalidar os resultados do estudo.

O viés de seleção pode ser evitado ao introduzirmos o acaso na seleção dos pacientes para o estudo. Por exemplo, num ensaio clínico controlado, a alocação dos pacientes para o grupo da droga A e para o grupo placebo deve ser aleatória, processo que chamamos de randomização. Isto garante a mesma chance, a cada paciente individual, de ser alocado para um ou outro grupo. Assim, o investigador não interfere no processo, eliminando-se o viés de seleção.

Na organização dos grandes *trials*, o processo de randomização é centralizado e feito por via telefônica, por um sistema automatizado, no qual o investigador fornece dados básicos do novo paciente e, a seguir, recebe um código, que corresponde ao tratamento a ser administrado ao paciente (p.ex. droga A ou placebo)¹⁰.

Existem outros métodos para o controle do viés de seleção, entre eles estão: a) restrição; b) estratificação; c) ajustes estatísticos e d) análise de sensibilidade (melhor e pior cenário). Para maiores detalhes sobre eles, referimos o leitor para os textos especializados.

O viés de aferição (ou avaliação) se dá quando os métodos de medida diferem entre os grupos. Por exemplo, no estudo mencionado acima, se o desfecho final a ser avaliado, dependesse do julgamento do médico (p.ex. severidade da angina do peito), e este soubesse qual droga teria sido administrada ao paciente, fatores subjetivos poderiam levá-lo a avaliar, de forma distinta, pacientes que receberam, ou não, a droga A.

Para evitar o viés de aferição, três medidas podem ser tomadas: a) condição cega, ou seja, paciente e investigador não conhecem a que grupo o paciente pertence (duplo-cego); b) estabelecimento e aplicação rigorosa de normas rígidas do que seja um evento e c) agir de maneira uniforme e consistente na detecção dos eventos, em todos os grupos do estudo. Eventualmente, podemos ter um estudo triplo-cego, no qual o estatístico responsável pela análise, também ignora qual a intervenção administrada aos grupos a serem analisados. Os ensaios controlados randomizados e duplo-cegos são a melhor maneira de avaliar a eficácia e segurança de intervenções, pois incorporam no seu delineamento, medidas fundamentais para o controle dos vieses¹⁰.

O viés de confusão ocorre quando não há comparabilidade entre os grupos estudados. Isto acontece quando variáveis que produzem os desfechos clínicos, estão desigualmente distribuídas entre os grupos. Dois fatores estão associados (viajam juntos) e, o efeito de um deles é confundido ou distorcido pelo efeito do outro.

Por exemplo, num estudo sobre fatores de risco, uma associação entre o hábito de beber café e a doença coronária é detectada. Porém, se não for considerado o fato de que os fumantes bebem mais café do que os não-fumantes, pode se chegar à errônea conclusão de que o café é um fator de risco, independente para doença coronária, o que não corresponde à realidade. Neste caso, o café é um fator de confusão, e não um fator causal independente para a doença coronária.

O potencial de viés não significa que ele esteja sempre presente num estudo em particular. Tanto para o pesquisa-

dor, quanto para o leitor do relatório de pesquisa, a questão do viés demanda que, antes de tudo, se saiba onde e quando procurá-lo, e o que fazer para evitá-lo.

É importante, também, determinar a magnitude do viés, e se ele é suficientemente grande, a ponto de modificar as conclusões do estudo e sua aplicação na clínica. O ônus da prova, de que o viés não existe ou, se existente, não influenciou de maneira decisiva os resultados, é sempre do investigador. Por isso, ele deve se precaver de todos os potenciais vieses, tanto na fase de planejamento, como nas fases de coleta e análise do estudo, de forma a garantir a validade interna do estudo⁸.

Erro aleatório

O erro aleatório decorre, exclusivamente, do acaso, e pode ser estimado por testes estatísticos. Diferente do erro sistemático, o erro aleatório, porque é aleatório, varia de forma uniforme em torno do valor real, porém sem modificá-lo.

Uma premissa básica no uso de medidas de frequência das doenças em epidemiologia, é que se pode inferir sobre a experiência de toda uma população com base na avaliação de apenas uma amostra (inferência estatística). Observações clínicas são feitas em amostras da população (*samples*) e podem, eventualmente, não representar a real situação na população como um todo, apenas por obra do acaso (*play of chance*). Porém, se grandes amostras, repetidas e semelhantes, forem tomadas da população, os resultados tenderão a variar em torno do valor que representa a realidade.

A divergência entre uma observação feita na amostra, e outra feita na população total (se isto fosse possível), devido apenas ao acaso, é chamada de variação aleatória ou randômica (*random variation*). Se medirmos uma variável (p.ex. PA) repetidas vezes, teremos valores diversos entre si, mas que variam, uniformemente, em torno de um valor médio, que representa o valor real. Quanto maior o número de medidas, maior será o número de valores próximos da média e menor a variabilidade. Os valores dispersos ao redor da média são devidos a vários fatores ligados ao método (variabilidade biológica, largura do manguito, calibração inadequada, dificuldade auditiva do médico) ou, simplesmente, pela variação aleatória na medida. Neste último caso, as medidas tendem a se dispersar ao redor da média, ao passo que, no caso de erro sistemático, as medidas tendem a ocupar preferencialmente um dos lados da média real. Se, por exemplo, um esfigmomanômetro mal calibrado estiver superestimando a PA, logicamente, a maioria dos valores estará acima do valor real e vice-versa.

Há, portanto, uma necessidade clara de quantificar em que grau a variação aleatória pode ser responsabilizada pelos resultados de um estudo. Isto é feito através de testes de significância estatística (p.ex. qui-quadrado e teste t-student).

Se estamos estudando uma associação entre exposição a um fator de risco (p.ex. fumo) e a ocorrência de doença coronária, e encontramos uma frequência de coronariopatia entre os fumantes de 10 vezes maior em relação aos não-fu-

mantas, submetemos estas taxas a um teste estatístico, que leva em consideração a variabilidade da medida e o tamanho da amostra. De um modo geral, o resultado do teste é relatado em termos de valor de P, que indica a probabilidade de que um determinado efeito (no caso, alta frequência de doença coronária em tabagistas) possa ter ocorrido apenas pelo acaso, assumindo-se que não exista relação entre exposição e doença.

Portanto, se o valor de $P \leq 0,05$, significa que existe não mais do que 5% (ou uma chance em 20), de se observar um resultado tão extremo apenas pelo acaso, de onde se conclui que a associação entre exposição (fumo) e doença (doença coronária) é estatisticamente significativa.

A significância estatística é muitas vezes interpretada de maneira errônea. Em primeiro lugar, ela não deve ser vista como um corte (sim ou não) no que se refere às conclusões do estudo, mas sim, deve ser, meramente, um guia para a ação.

Da mesma forma, um resultado estatisticamente significativo não equivale dizer que o acaso não é responsável pelos resultados, mas sim que esta explicação é improvável. Em 2º lugar, a magnitude do valor de P não é uma medida da sua significância clínica, que é melhor representada pelo intervalo de confiança (ver adiante), mas sim uma medida da improbabilidade de que um resultado tenha ocorrido por mero acaso.

Finalmente, a significância estatística está intimamente ligada ao tamanho da amostra do estudo. Em estudos com grandes amostras, mesmo pequenas diferenças, muitas vezes sem significância clínica, podem ser detectadas pelos testes estatísticos e apresentarem resultados significativos. Da mesma forma, amostras pequenas têm pouca possibilidade de mostrar resultados significativos, mesmo que existam diferenças grandes, e clinicamente importantes, entre os grupos¹¹.

As duas fontes principais de erro - acaso e viés - não são mutuamente exclusivas. Muitas vezes, ambas estão presentes, e sua distinção auxilia na abordagem, que é diversa. Estes erros podem ser evitados ou controlados, se a investigação clínica for planejada e conduzida de maneira apropriada, e submetida a uma adequada análise estatística dos dados.

Reconhecer, evitar e minimizar os vieses são atribuições do pesquisador sério e do bom leitor de relatórios científicos. Quanto ao acaso, este não pode ser eliminado, mas sua influência pode ser minimizada pelo desenho apropriado do estudo, dando-se ênfase ao tamanho adequado da amostra, em conjunto com a adequada análise estatística.

O uso e o abuso da estatística têm transformado, freqüentemente, mentiras em verdades, especialmente para aqueles que, por ignorância, são seduzidos pelo poder dos números e dos gráficos bem elaborados. A rigor, a análise estatística não deveria ser aplicada a estudos mal delineados e metodologicamente fracos. Caso contrário, corre-se o risco de dar a respeitabilidade dos números a um trabalho sem mérito.

Conhecimentos de estatística básica aplicada às ciências biológicas são de fundamental importância para o en-

tendimento e a análise crítica da literatura médica. Sem este instrumental, a interpretação errônea dos relatórios de pesquisa pode resultar em conseqüências indesejáveis para o paciente e a sociedade.

Significância estatística e significância clínica

A significância clínica e significância estatística não são sinônimas. Sabe-se que diferenças de efeito clínico entre duas intervenções podem ser grandes e não serem detectadas, se a amostra for pequena. Ao contrário, em grandes amostras, diferenças de efeito, mesmo que muito pequenas, podem produzir resultados estatisticamente significativos.

Por exemplo, os resultados de um estudo clínico, com alguns milhares de pacientes, mostrou uma diferença significativa, na concentração do colesterol sérico, entre um grupo tratado com uma droga A e o grupo placebo. Se esta diferença bioquímica não se traduzir em modificações, por exemplo, na incidência de eventos cardiovasculares, seu significado clínico será nulo.

Assim, mais importante, do que o valor do colesterol sérico em si, é o significado que sua redução terá sobre o prognóstico do paciente. Portanto, a significância clínica é avaliada pelo impacto que os resultados do estudo produziram, se aplicados na prática. Quanto maior este impacto, maior a significância clínica.

Tamanho da amostra

Na fase de planejamento de um ensaio clínico, o cálculo do tamanho da amostra é fundamental, pois dele depende sua validade interna. Para tanto, são necessários níveis adequados de significância e poder estatístico que possam detectar diferenças entre grupos consideradas clinicamente relevantes.

Um dos grandes problemas relacionados aos ensaios clínicos, está no número insuficiente de pacientes envolvidos e, conseqüentemente, no baixo poder estatístico de detectar diferenças moderadas e pequenas (entre 10 e 20%) entre duas intervenções. As intervenções, atualmente sendo testadas em ensaios clínicos, apresentam, em geral, efeitos pequenos a moderados (raramente grandes, ou seja, acima de 50%) e, portanto, exigem amostras com alguns milhares de indivíduos.

O cálculo do tamanho da amostra leva, portanto, em consideração três fatores: o erro alfa, o erro beta (e seu complemento, o poder do teste) e, finalmente, a diferença a ser detectada, que seja clinicamente considerada relevante¹⁰.

O erro alfa (ou do tipo I) ocorre quando se detecta uma diferença devido apenas ao fator chance ou acaso; trata-se, portanto, de um falso-positivo. Na investigação clínica, o nível de significância estatística é pré-fixado em, por exemplo, 5% (0,05), o que quer dizer que a probabilidade de ocorrência pelo acaso de um resultado positivo é de um para 20.

Por outro lado, o erro beta (ou tipo II) é aquele no qual não se detecta uma diferença que, na realidade, existe. Este é, portanto, um falso-negativo. Nas pesquisas biomédicas,

aceita-se um valor de beta igual a 0,20, como uma probabilidade pequena, ou seja, duas chances em 10 de cometer um erro tipo II. Eventualmente, este número pode ser menor, dependendo da importância em se dar uma resposta definitiva à questão de pesquisa.

O poder estatístico da amostra é dado pelo complemento do erro beta (poder = $1 - \beta$) e representa a probabilidade de detectar um resultado significativo, se ele de fato existe.

Resumindo, o tamanho da amostra é inversamente proporcional ao erro alfa, ao erro beta e à magnitude da diferença considerada clinicamente significativa, e que, portanto, vale a pena ser detectada.

Programas estatísticos disponíveis (p.ex., EpiInfo ()) fornecem automaticamente o tamanho da amostra, uma vez alimentados com os dados mencionados.

Contaminação e co-intervenção

Nos estudos controlados randomizados (ECR) duas situações colocam em risco a validade interna. Uma delas refere-se à co-intervenção, que ocorre quando, após a randomização, os pacientes estão livres para receber outras intervenções, além daquela que está sendo avaliada. Se isto acontecer de forma desigual nos dois grupos (controle e intervenção), os resultados podem ser afetados de maneira imprevisível em magnitude e direção.

Um exemplo clássico é o da avaliação de uma terapêutica no pós-IAM, na qual, durante o seguimento, outros procedimentos, como revascularização cirúrgica e mudanças de hábitos de vida (dieta, controle do tabagismo) são introduzidos de maneira indiscriminada, de tal forma que, ao final do estudo, os resultados do tratamento original ficam diluídos e indiscerníveis das outras intervenções.

A segunda situação diz respeito à contaminação, onde pacientes, originariamente, alocados para o grupo B passam a receber a terapia do grupo A ou vice-versa. Isto pode atenuar, ou mesmo anular, eventuais diferenças existentes entre os grupos. Esse fato pode ocorrer quando uma ou mais das drogas do estudo se encontram disponíveis fora do seu âmbito.

Um exemplo muito citado é o do estudo ISIS-4, no qual no grupo alocado para o placebo-nitrato, metade dos pacientes recebeu o nitrato ativo, prescrito pelo médico assistente, o que certamente reduziu a chance de demonstrar um efeito positivo de maior magnitude daquela droga no IAM.

Concluindo, a contaminação só pode ocorrer no grupo controle, que recebe uma intervenção experimental, inicialmente, não-designada, ao passo que, tanto grupo controle quanto grupo de intervenção podem, eventualmente, receber uma co-intervenção.

Desfechos clínicos

Desfechos clínicos (*outcomes* ou *end-points*) são eventos considerados importantes, e objeto da hipótese do estudo. Eles são pré-definidos no protocolo, coletados e verificados no transcorrer do estudo, ou ao seu término.

Por exemplo, em um ensaio clínico com a droga A, a ocorrência de um IAM é considerada um evento primário. A sua detecção e registro dependem de um diagnóstico preciso e correto, baseado em critérios pré-estabelecidos no protocolo do estudo. A aferição dos desfechos depende, em muitos casos, do julgamento do investigador, e há, conseqüentemente, risco de viés. Se estamos lidando com desfechos, cuja aferição não dá margem a interpretações, como, por exemplo, mortalidade geral, chamamos este desfecho de “duro” (*hard*).

Porém, alguns desfechos dependem do julgamento do investigador e, portanto, são chamados de “moles” (*soft*). Um exemplo disto é o da causa de morte, que muitas vezes não pode ser determinada, por não se dispor de métodos diagnósticos ou de verificação de óbito apropriados. Estes desfechos, menos consistentes, podem ser também chamados de desfechos intermediários ou substitutos (*surrogate end-points*). Eles podem ser definidos como desfechos de um ensaio clínico, que seja baseado em um dado laboratorial ou em um sinal físico, usado para substituir eventos ou desfechos clinicamente significativos e que interessam mais, diretamente, ao paciente, tais como morbidade e mortalidade. Espera-se que as mudanças que ocorrem nos desfechos intermediários reflitam mudanças nos desfechos clinicamente significativos¹².

Na cardiologia, podemos citar alguns exemplos de desfechos intermediários, como supressão de arritmias, tolerância ao exercício na insuficiência cardíaca e controle de lípidos e da PA usando drogas.

Na prática, porém, nem sempre os desfechos intermediários substituem aqueles que realmente interessam ao paciente, por isso é recomendável que os ensaios clínicos, que usam variáveis laboratoriais e fisiológicas como desfecho, sejam vistos com reserva, e que seus resultados não sejam transpostos para a prática antes da realização de ensaios clínicos, que usem desfechos clinicamente relevantes.

O controle do viés de aferição é feito pela “condição cega”, na qual, investigador, paciente e eventualmente, estatístico, estão “cegos” quanto à alocação do paciente para os grupos de intervenção. Elimina-se, assim, a possibilidade de interpretação tendenciosa dos eventos, uma vez que se desconhece qual tratamento está sendo administrado ao paciente.

Regressão à média

Valores extremos de uma determinada distribuição tendem a voltar para valores, em média, menos extremos em medidas subsequentes. Este fenômeno é chamado de regressão à média e ocorre por razões puramente estatísticas. A razão é que qualquer teste observado é composto de duas partes: o valor real e um erro inerente.

Existem várias fontes de erro, incluindo variações dos instrumentos de medida, variação biológica intra-indivíduo, motivação, fadiga e erros no registro dos dados¹¹. Como é assumido que este erro é aleatório, algumas vezes ele atua aumentando, e outras, diminuindo, o valor real.

Se, por exemplo, dosarmos a glicemia de jejum de uma amostra de 1.000 pessoas, cuja média é 90mg/dL, cerca de 5% dos indivíduos terão valores acima de 140mg/dL e outros 5%, valores abaixo de 50mg/dL. Numa segunda dosagem veremos que, em média, muitos dos 100 indivíduos com glicemia anormal apresentarão valores mais próximos à média e dentro da faixa de normalidade.

A importância deste fenômeno está no fato que, ao medirmos uma variável contínua, por exemplo, PA, é fundamental que se tenha, pelo menos, uma segunda medida naqueles indivíduos que apresentam valores extremos, de forma a detectar o fenômeno da regressão à média. Por isso, para o diagnóstico de hipertensão arterial (um valor extremo na curva da normalidade da PA), é recomendado a média de três medidas da PA, num mesmo dia, e em três ocasiões diferentes.

Em estudos não-controlados, as modificações de uma variável podem se dever tanto à intervenção quanto à regressão à média. Porém, se existe um grupo controle, ambos os grupos estarão sob os mesmos efeitos aleatórios da

regressão à média, e portanto, os efeitos se anulam reciprocamente.

Considerações finais

Neste artigo procuramos abordar alguns conceitos básicos da epidemiologia clínica, colocando-os no contexto da especialidade cardiológica. A aplicação destes conceitos é múltipla, na prática clínica, e pretende tornar o ato médico mais objetivo e científico. Embora a vivência do clínico e sua experiência diagnóstica sejam imprescindíveis para uma boa prática da Medicina, somente estes atributos não são suficientes para oferecer o melhor do conhecimento médico, atualmente disponível. Bons médicos combinam a competência clínica com a melhor evidência externa científica disponível, na tomada de decisões sobre seus pacientes.

Portanto, a aplicação dos princípios epidemiológicos, no contexto clínico, torna o ato médico mais científico, e proporciona decisões, diagnósticas e terapêuticas, mais confiáveis, eficazes e seguras.

Referências

1. Sackett DL, Richardson WS, Rosenberg W, Haynes RB - Evidence-Based Medicine. New York: Churchill Livingstone, 1997.
2. Echt DS, Liebson PR, Mitchell B et al - Mortality and morbidity in patients receiving encainide, flecainide and placebo: the Cardiac Arrhythmias Suppression Trial. N Engl J Med 1991; 324: 781-8.
3. Packer M, Carver JR, Rodeheffer RJ et al - Effect of oral milrinone on mortality in severe chronic heart failure. The PROMISE Study Research Group. N Engl J Med 1991; 325: 1468-75.
4. Sackett DL, Haynes RB, Guyatt GH, Tugwell P - Clinical Epidemiology. A Basic Science for Clinical Medicine. 2nd ed. Boston: Little Brown, 1991.
5. Fletcher RH, Fletcher SW, Wagner EH - Clinical Epidemiology: The Essentials. 3rd ed. Baltimore: Williams & Wilkins, 1996.
6. Amberson JB, McMahon BT, Pinner M - A clinical trial of sanocrysin in pulmonary tuberculosis. Am Rev Tuberc 1931; 24: 401-35.
7. Davies HTO, Crombie IK, Tavakoli M - When can odds ratio mislead? Br Med J 1998; 316: 989-1.
8. Hennekens CH, Buring JE - Epidemiology in Medicine. Boston: Little Brown, 1987.
9. Sackett DL - Bias in analytic research. J Chronic Dis 1979; 32:51-63.
10. Friedman LM, Furberg CD, DeMets DL - Fundamentals of Clinical Trials. 3rd ed. St Louis: Mosby, 1996.
11. Hill AB - Principles of Medical Statistics. 9th ed. New York: Oxford University Press, 1971.
12. Fleming TR, DeMets DL - Surrogate end-points in clinical trials: Are we being misled? Ann Int Med 1996; 125: 605-13.