

ESTADÍSTICA



BIOESTATÍSTICA:

ARMADILHAS E COMO EVITÁ-LAS

Carlos Alberto Mourão Júnior*

INTRODUÇÃO

Neste artigo procuramos enfocar alguns pontos críticos referentes à bioestatística que devem ser levados em consideração na elaboração de trabalhos científicos visando a obtenção de resultados mais consistentes.

Inicialmente é importante se ter em mente que a estatística abrange os seguintes domínios: planejamento do estudo, coleta e apresentação dos dados, análise inferencial dos resultados e discussão dos achados do estudo. O primeiro equívoco que se pode cometer é procurar assessoria estatística somente quando já se têm os resultados do trabalho.

Como veremos, é justamente na fase de elaboração do projeto que alguns erros passam despercebidos, podendo tomar dimensões que venham a comprometer todo o estudo. Assim sendo, um planejamento adequado é condição indispensável para a obtenção de resultados confiáveis. A seguir passaremos a discutir alguns pontos-chave a serem levados em conta quando se pensa em iniciar algum tipo de pesquisa.

A PERGUNTA DO ESTUDO

Toda pesquisa deve nascer de uma questão ainda não respondida, entretanto, por mais estranho que possa parecer, em muitos casos não é possível identificar com clareza qual é a questão a ser respondida. Na verdade a viga mestra que sustenta toda a pesquisa é justamente a pergunta do estudo colocada de forma clara, concisa e precisa, pois a partir dela poderemos classificar as variáveis de forma adequada permitindo uma análise posterior consistente. De acordo com a questão do estudo podemos classificar as variáveis da forma a seguir.

VARIÁVEL RESPOSTA

É o evento que se pretende estudar (exemplo: ocorrência de câncer), sendo dependente das variáveis predictoras.

VARIÁVEL PREDITORA

É a variável que permite prever uma resposta (exemplo: no caso da variável resposta ser a ocorrência de câncer, uma variável preditora poderia ser tabagismo). É importante ressaltar que, dentro do possível, o ideal é a existência de uma única variável preditora, uma vez que em muitos casos, em estudos com mais de uma variável deste tipo, uma pode ter influência sobre a outra gerando uma situação conhecida como **covariância**.

Quando ocorre covariância, muitas vezes os métodos estatísticos não se mostram suficientemente seguros para permitir inferências, e isso pode comprometer seriamente os resultados da análise e a conclusão dos achados. Assim sendo, por mais tentador que seja incluir variáveis predictoras em um primeiro momento, isso deve ser evitado a fim de aumentar a validade da pesquisa.

VARIÁVEL DE CONFUSÃO

Este tipo de variável, também conhecida como variável interveniente, se caracteriza por estar associada à variável resposta ou à variável preditora, interferindo na relação de causa e efeito entre elas. Vejamos um exemplo: há alguns anos um pesquisador afirmou que o consumo de café aumentava o risco de câncer de pâncreas e que a análise estatística mostrava uma associação significativa (este estudo foi publicado no *New England Journal of Medicine* em 1981). Quando o trabalho foi re-analisado, concluiu-se que na verdade o fator de risco era o fumo. O que ocorreu foi que os fumantes tomavam mais café. Assim, o café era uma variável de confusão, ou seja, estava associada tanto ao fator de risco verdadeiro (fumantes tomam mais café) como à doença (o consumo de café entre os portadores de câncer de pâncreas era maior que na população geral); no entanto não havia relação causa-efeito entre consumo de café e câncer de pâncreas.

* Depto. Fisiologia – Universidade Federal de Juiz de Fora.

Como podemos perceber, a presença de uma variável deste tipo não só compromete, mas invalida os resultados. Na Figura 1 podem ser vistas as inter-relações entre os tipos de variáveis.

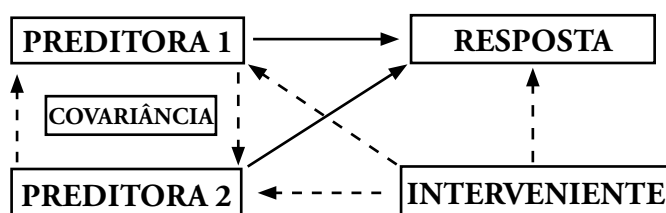


Figura 1: Relação entre variáveis.

Como já foi dito, a covariância pode e deve ser evitada restringindo-se a pergunta do estudo a uma única variável preditora. Mas como proceder com as variáveis de confusão?

Inicialmente, o mais importante é detectar a presença delas (lembre-se, elas são rasteiras e muitas vezes passam despercebidas). Uma vez detectadas, elas devem ser excluídas (se possível) ou **controladas**, isto é, distribuídas homoganeamente entre os grupos estudados (exemplo: idade, sexo). Assim se, por exemplo, o sexo for uma variável que pode afetar a variável resposta, podemos proceder de duas maneiras: ou estudamos somente elementos de um sexo (assim a variável deixa de existir) ou comparamos grupos com a mesma proporção masculino/feminino (neste caso a variável fica controlada).

Uma vez definida a pergunta do estudo e controladas as possíveis variáveis intervenientes, a próxima preocupação deverá ser evitar os erros (viés) que muitas vezes ocorrem sem que o pesquisador se dê conta do fato. A seguir trataremos dessa importante questão.

TIPOS DE VIÉS

Viés é sinônimo de erro sistemático, vício, tendenciosidade, desvio, deturpação, distorção ou *bias* (do original inglês). É um erro sistemático introduzido em qualquer fase de um estudo que o desvia da realidade prejudicando sua validade.

O viés é não-intencional, proveniente de questões conceituais ou metodológicas não adequadamente resolvidas, quer por impossibilidade prática, ou por não serem identificadas no devido tempo. Os três principais tipos de viés são: de Seleção, de Aferição e de Confusão (fórmula mnemônica = tríade SAC).

VIÉS DE SELEÇÃO

É caracterizado por distorções nos resultados de um estudo que resultam de falhas nos procedimentos utilizados para a seleção dos participantes e/ou de fatores que influenciam a participação na pesquisa. É também denominado viés da “amostragem”, da “amostra” ou da “população”.

Um exemplo deste tipo de erro ocorreu há alguns anos em uma pesquisa eleitoral para presidente da república nos Estados Unidos. Foram entrevistados um grande número de eleitores, escolhidos aleatoriamente a partir de uma lista telefônica, e estes informaram o candidato de sua preferência; ocorreu porém que o resultado das urnas foi o oposto daquele apontado pela pesquisa. O que ocorreu? Ao escolher a lista telefônica como fonte de amostragem, os pesquisadores, sem perceber, “selecionaram” os eleitores com maior poder aquisitivo. Como a maior massa de votantes estava em camadas sociais mais baixas e portanto não tinham telefone em casa, a pesquisa, apesar de todo o rigor, forneceu um resultado falso.

Para evitar o viés de seleção devemos:

- Planejar cuidadosamente a pesquisa e a população de referência;
- Analisar bem o tamanho da amostra (sua representatividade) e o método de amostragem;
- Estabelecer um esquema de controle de qualidade no acompanhamento.

VIÉS DE AFERIÇÃO

Refere-se a distorções nos resultados de um estudo decorrentes de erros de mensuração/aferição dos dados da exposição e/ou desfecho. É também denominado viés de “informação” ou de “observação”.

Um exemplo seria estudar a variável altura de pacientes que foram medidos com calçados, ou utilizando um instrumento de medida com precisão insatisfatória. Outro exemplo de viés de aferição seria a utilização de um teste estatístico inadequado para fazer inferência sobre os dados da amostra estudada. É fácil perceber que este tipo de viés torna os resultados encontrados duvidosos, comprometendo a validade interna e externa da pesquisa.

Para evitar o viés de aferição devemos:

- Utilizar instrumento e ambiente propícios para a mensuração dos dados, bem como se assegurar do treinamento padronizado dos observadores;

- Definir com clareza o evento a ser pesquisado ou verificado e o modo de coletar os dados;
- Padronizar os questionários e as escalas de mensuração;
- Estabelecer parâmetros fidedignos para a inclusão e exclusão de casos no estudo.

VIÉS DE CONFUSÃO

Erro sistemático que ocorre quando os resultados de uma associação entre dois fatores (causa-efeito) podem ser imputados, total ou parcialmente, a um **terceiro fator** não levado em consideração (variável de confusão ou interveniente). O conceito de variável de confusão já foi discutido anteriormente.

O que caracteriza o “confundimento” é a mistura de efeitos provocados por pelo menos duas variáveis sobre o desenvolvimento de uma doença (ou desfecho). Uma dessas variáveis é a exposição em estudo (principal) a outra é a variável de confusão que pode interferir no efeito da primeira.

Para evitar este tipo de viés devemos adotar medidas preventivas identificando as variáveis de confusão e dando a estas o tratamento adequado, conforme já foi discutido. No caso dessas variáveis serem identificadas após a conclusão da coleta dos dados, algumas técnicas de estatística multivariada (como por exemplo a regressão logística) podem eventualmente ser utilizadas visando tentar quantificar a interferências dessas variáveis no desfecho encontrado; porém é sempre bom lembrar que a estatística descreve resultados e faz inferências, porém não tem ferramentas para corrigir erros sistemáticos.

CONCLUSÃO

O maior objetivo de um pesquisador sério é produzir conhecimentos e divulgá-los para a comunidade científica visando contribuir para o ramo do saber para o qual dedica seu trabalho. Atualmente a aceitação de manuscritos em periódicos respeitados exige originalidade dos trabalhos e uma metodologia robusta que sustente a validade dos resultados encontrados.

A originalidade depende do talento criativo do pesquisador e dos recursos disponíveis, porém na metodologia muitos erros podem ser evitados se forem levados em consideração os aspectos que tentamos descrever neste artigo.

Podemos resumir o que foi dito nas seguintes sugestões:

- Defina com clareza a pergunta do estudo;
- Identifique as possíveis variáveis de confusão;
- Tenha sempre em mente a possibilidade de existir algum viés (triade SAC).

REFERÊNCIAS

Para um maior aprofundamento, sugerimos a leitura das seguintes obras:

ALTMAN, D.G. **Practical statistics for medical research**. London: Chapman & Hall, 1991.

BAILAR, J. C.; MOSTELLER, F. **Medical uses of statistics**. 2nd ed. Boston: NEJM Books, 1992.

MOTULSKY, H. **Intuitive biostatistics**. New York: Oxford University Press, 1995.

