

Instrumentos de coleta de dados na Saúde Pública: resenha crítica da obra de Monteiro e Horta

Data collection instruments in Public Health: critical review of the Book from Monteiro and Horta

Carla Jorge Machado¹, Claudia Aguiar Pereira²

Validade e confiabilidade são conceitos fundamentais em Epidemiologia, e a obra “Pesquisa em Saúde Pública: Como Desenvolver e Validar Instrumentos de Coletas de Dados” traz à tona essa questão. O livro é dividido em seis capítulos, priorizando a teoria da medição.

O capítulo 1, “Introdução”, chama a atenção para o objetivo central da obra, que é trazer um texto introdutório sobre os temas da validade e confiabilidade das medidas, sem ser exaustiva nos tópicos abordados. Os autores tratam dos cuidados que devem ser tomados no planejamento de um estudo científico, com ênfase especial à fase da coleta de dados. Nesse aspecto, diferenciam os erros sistemáticos (ou constantes) dos aleatórios (casuais ou variáveis) e apresentam, didaticamente, os conceitos de validade e confiabilidade. Há ainda um curto relato histórico sobre a contribuição de Charles Spearman à teoria clássica das medidas — vale lembrar a implementação em *softwares* estatísticos da correlação não paramétrica rho de Spearman, muito utilizada em pequenas amostras — e uma definição do coeficiente de confiabilidade.

A “Teoria da Medição” é o capítulo 2. Após estabelecer o que é “medição”, são analisados os conceitos de “variável” e suas diversas derivações: “variável conceitual”, “variável operacional”, “variável quantitativa”, contínua ou discreta, “variável qualitativa”, nominal ou ordinal, entre outras. Os autores abordam a Teoria clássica da medição e os componentes do valor observado: valor verdadeiro e erro aleatório. No tópico “Escala de Medição”, separam os significados dos instrumentos de escalas de medição e definem-nas como os meios pelos quais se atribui um número a um objeto de estudo que se deseja mensurar. Os autores destacam que a obra não aborda os modos de operacionalizar as análises estatísticas voltadas para as escalas. O trabalho define em subtópicos, com exemplos, as possíveis estatísticas advindas de cada tipo de escala: nominal, ordinal, de intervalo e de razão. Os autores apresentam ainda um quadro que facilita a compreensão das diferenças entre os tipos de escala, relacionando-os a algumas categorias de aplicações funcionais. A obra destina, sabiamente, um tópico inteiro à Escala *Likert*, utilizada desde 1932. Neste os autores chamam a atenção para a possibilidade de expressar opinião, diferentemente das respostas dicotômicas, e é apresentado um exemplo prático. Por fim, tratam da obtenção de dados, com ênfase nas diferentes formas de observação (desde assistemática até em laboratório), exame de registro (busca documental) e entrevistas (com presença de entrevistador) ou questionários (quando o sujeito da pesquisa lê e responde as questões por si), relatando suas

Trabalho realizado na Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte (MG), Brasil.

¹Ph.D. Population Dynamics at Johns Hopkins University – Baltimore, USA; Professora Associada da UFMG – Belo Horizonte (MG), Brasil.

²Ph.D. Population Health at University of Wisconsin – Madison, USA; Pesquisadora do Departamento de Administração e Planejamento em Saúde da Escola Nacional de Saúde Pública Sergio Arouca da Fiocruz – Rio de Janeiro (RJ), Brasil.

Endereço para correspondência: Carla Jorge Machado – Avenida Augusto de Lima, 1376, sala 909 – CEP: 30190-003 – Belo Horizonte (MG), Brasil –

E-mail: carlajmachado@gmail.com

Conflito de interesses: nada a declarar.

diferentes apresentações, vantagens e desvantagens. Também diferenciam questões abertas e fechadas, apresentando as vantagens e desvantagens associadas a cada tipo.

Os capítulos seguintes podem ser considerados os mais importantes do livro por tratarem de validade e confiabilidade. O capítulo 3 é intitulado “Validade” e remete, inicialmente, à definição de que é o grau em que um instrumento mede o que se propõe a medir. Nesse caso, a ideia subjacente é a adequação entre o fenômeno estudado e o conceito teórico a ser medido. Assim, os autores apresentam uma variedade de possíveis definições para validade e trazem o relato histórico de uma das primeiras avaliações empíricas de validade, ocorrida em 1905. O relato histórico segue contando os avanços na década de 1950, com o surgimento de múltiplas linhas de “evidência de validade” e classificações de validade em quatro tipos: validade de conteúdo (cuja temática foi abordada em revisão integrativa recente na área de saúde¹), concorrente, preditiva e de construto. Em seguida, os autores passam ao conceito de “validade relacionada a critério”, que surgiu em 1966, a partir da combinação entre validade concorrente e preditiva. Passou-se, então, da concepção de “tipos” de validade para “aspectos” da validade. Adiante, os autores direcionam o capítulo para a abordagem aprofundada da validação do conteúdo, validação do construto e validação correlacionada a critério. Neste último, aprofundam o assunto em tópicos voltados para a acurácia dos testes (sensibilidade, especificidade e valores preditivos). Tratam ainda da curva *Receiver Operating Characteristics* (ROC), que é útil para a escolha do melhor ponto de corte, a partir da sensibilidade e especificidade do instrumento e do objetivo do estudo, bem como da razão de verossimilhança.

Uma ressalva nos tópicos voltados para acurácia dos testes (como especificidade e sensibilidade) é que tais conceitos não se aplicam apenas à relação doença e teste, como abordado pelos autores, mas também a inúmeros construtos e desfechos. Em Ciências Sociais, o conceito de validade aplica-se em medidas como inventários de personalidades, testes educacionais, *surveys*, questionários de atitudes, entre outros².

Os autores chamam a atenção para o fato de que a validade do instrumento deve estar relacionada com sua finalidade específica, pois um teste pode ser apropriado para um devido fim e não para outro. Dessa forma, a validade não deve ser considerada como ‘tudo ou nada’, válido ou não válido, mas o que se pode fazer é defender o uso de um teste em determinada

situação, reconhecendo que o que está validado não é o teste em si, mas seu uso para propósito particular.

O capítulo 4 foi denominado “Confiabilidade” e nele, assim como no capítulo anterior, os autores têm o cuidado de fornecer aos leitores várias possíveis definições da expressão abordada. Iniciam com o conceito mais simples, que é o da medição ser capaz de produzir os mesmos resultados em repetidas tentativas, ou seja, remete à consistência, independentemente de quem ou quando se está aferindo. Na sequência, abordam as fontes de erros e as formas de estimá-los: comparando o resultado obtido por um observador com outro igualmente treinado, repetindo a medida no mesmo sujeito e comparando o resultado do instrumento com o de outro similar. A partir daí, apresentam a confiança intraobservador, interobservador e a consistência interna. Desenvolvem, em seguida, uma seção sobre indicadores de confiabilidade e suas respectivas fórmulas: genérico; estimadores de consistência interna, incluindo correlação de Pearson; correlação item-total; métodos de bipartição ou divisão ao meio, também conhecidos como *split-half*; fórmula 20 de Kuder-Richardson (KR-20) e Alfa de Cronbach (coeficiente de alfa), que é o mais usado e conhecido indicador de confiabilidade interna de um instrumento de coleta de dados. O capítulo termina abordando os indicadores de confiabilidade intraobservador e interobservador, e fornecendo algumas considerações para melhorar a confiabilidade de um questionário ou entrevista.

No capítulo 5, “Desenvolvimento de Novos Instrumentos de Pesquisa”, os autores apresentam resumidamente alguns conselhos e instruções àqueles que desejam desenvolver novos instrumentos de pesquisa, como: definição ou identificação das informações necessárias, identificação do público a ser estudado, escolha do melhor método para atingir esse público, execução de um ou mais pré-testes, entre outros.

No último capítulo, “Considerações Finais”, os autores retomam questões que motivaram o livro e que haviam sido mencionadas no início, como o planejamento cuidadoso da pesquisa e da escolha adequada do instrumento de pesquisa.

À guisa de conclusão, trata-se de leitura que, apesar de contemplar conceitos quantitativos complexos do campo da medição, é leve e agradável. O livro representa uma contribuição à literatura nacional sobre o assunto e é recomendado para alunos de cursos de graduação e pós-graduação em Saúde Coletiva, de variados objetivos que abordem a qualidade na coleta de dados e aspectos relacionados.

REFERÊNCIAS

1. Alexandre NMC, Coluci MZ. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Cienc Saude Colet*. 2011;16(7):3061-8.
2. Sireci SG. The construct of content validity. *Soc Indic Res*. 1998;45:83-117.

Recebido em: 29/05/ 2014

Aprovado em: 09/12/2014