

Resenhas

Trisha Greenhalgh 2008. *Como Ler Artigos Científicos: Fundamentos da Medicina Baseada em Evidências*. Porto Alegre: Editora Artmed. 255 pp. ISBN : 978-853-631-612-3. Tradução da edição de 2006, publicada por Blackweel Publishing.

O livro é dividido em 13 capítulos. A autora, Trisha Greenhalgh, tem uma licenciatura em ciências sociais e uma segunda licenciatura em medicina, sendo autora de 120 artigos e 8 livros sobre medicina e sociologia, com abordagens inovadoras interdisciplinares, baseadas em narrativa, métodos etnográficos e participativo e com ênfase em questões de política na área da saúde contemporânea. Este livro é dirigido para os profissionais de saúde, abordando os aspectos fundamentais para a correta análise estatística de um ensaio clínico, bem como a forma apropriada de orientar a investigação, nomeadamente como colocar em prática os conhecimentos e toda a informação adquirida no campo teórico, facilitando a sua análise. De forma incisiva, a autora explica como se deve efetuar essa análise, ponderando os riscos de benefício e de dano; refere os tipos de leitura que podemos assumir face a um determinado artigo e mostra também como avaliar a qualidade de um artigo. De igual modo, explica ainda como validar um artigo, possibilitando a sua publicação e ajuda na tomada de decisões sobre intervenções clínicas (revisões sistemáticas e metanálises), salientando a importância dos aspetos éticos. Neste sentido, pode ter também a vertente de orientar o leitor para possíveis pesquisas que possa realizar.

A medicina baseada em evidências começa por ser redefinida por Trisha Greenhalgh como 'o uso de estimativas ma-

temáticas do risco de benefício e de dano derivadas de pesquisas de alta qualidade sobre amostras populacionais para informar a tomada de decisões clínicas no diagnóstico, investigação ou no manejo de pacientes individuais'. Em consequência, surgem inúmeras perguntas como, por exemplo, 'quais os sintomas, sinais físicos ou diagnósticos', entre outras, para a qual é, então, feito um resumo de etapas essenciais que nos irão ajudar a responder a algumas questões; formular o problema; rastrear evidências; análise crítica; implementar resultados na prática clínica e, por fim, avaliar o desempenho.

Para os críticos, este é um argumento destinado a que ações relacionadas com a saúde não sejam realizadas por qualquer tipo de profissional desta área, sem que antes haja uma aprovação por um comité específico, baseado em vários ensaios clínicos com grande número de participantes. Na realidade, tem-se verificado, após alguns estudos, que o número de decisões clínicas tomadas com base em evidências tem vindo a aumentar, não obstante algumas limitações metodológicas. Neste sentido, são exemplos do que a medicina baseada em evidências não é: *i.* a tomada de decisão por relato de casos (é inevitável, enquanto ser humano, o médico ignorar a sua experiência pessoal quando está perante um novo paciente); *ii.* tomada de decisão por recorte de artigos (a abordagem a um determinado problema clínico não deveria ser feita apenas com base em algum artigo, sem que o médico seja capaz de responder a algumas questões como, por exemplo, sobre os métodos utilizados); *iii.* tomada de decisão à moda antiga (como, por exemplo, o uso da sangria para qualquer doença aguda, ou da talidomida para enjoo ma-

tinal no início da gestação); e, por último, *iv.* a tomada de decisão por minimização de custos (cada vez mais se tem em conta este fator).

A questão inicial é como colocar em prática os conhecimentos e toda a informação que adquirimos na dimensão teórica e, para isso, temos, então, que começar por formular um problema que se desenvolve em 3 etapas: 1) definir sobre que pessoa é a questão; 2) definir qual a manobra terapêutica que pode ser considerada para o paciente em questão; e, finalmente, 3) definir qual o desenlace. O risco-benefício e as prioridades pessoais de quem está a ser avaliado são dois fatores importantes que devem ser tidos em conta. Neste contexto, existem três tipos de leitura (superficial, busca de informações e pesquisa) que devemos considerar e, além disso, deve ser encontrado um meio-termo entre a ingenuidade de acreditar em tudo o que lemos e a crítica que podemos fazer. Para isso, temos de recorrer a uma base de dados e índices que nos ajudem a organizar informações dispersas e juntar artigos. A autora sugere, conseqüentemente, a expressão 'nível de evidência' que se refere à confiança que podemos ter num artigo.

Uma grande base de dados designada Medline, compilada pela National Library of Medicine (NLM) dos Estados Unidos, indexa mais de cinco mil periódicos de setenta países, mas apresenta, porém, segundo a autora, algumas limitações, destacando-se a falta de referência a algumas áreas como a psicologia, sociologia médica, saúde ocupacional e ambiental, entre outros. Por outro lado, temos os filtros que nos permitem obter níveis de evidência elevados, podendo ser simples ou complexos, consoante o tipo de pesquisa que queremos fazer; sensível, caso queiramos uma pesquisa mais ampla; ou específica, caso seja algo altamente limitado e focado. Além disso, existem ainda 'artigos pré-avaliados' que se encontram em bases de dados relativamen-

te pequenas, listando artigos que alguém avaliou e selecionou. Este tipo de artigos pode encontrar-se na Biblioteca Cochrane, mais especificamente em Cochrane Database of Systematic Reviews (CDSR) que é uma base de dados de evidências sintetizadas, ou seja, é a base onde autores de revisões sistemáticas atualizam periodicamente as suas contribuições. E o mesmo tipo de artigos pode ser encontrado em Database of Abstracts of Review of Effectiveness (DARE) que faz uma avaliação crítica de revisões sistemáticas publicadas em periódicos.

No caso de não ser possível uma pesquisa totalmente abrangente, precisamos de recorrer a um arquivo, como é o caso da UK National Research Register que permite evidenciar o que nos ficou em falta. Para complementar o nosso trabalho, ainda devemos rastrear algumas citações, de modo a acompanhar o estudo e determinar a importância deste, consoante o número de vezes que foi citado. É ainda importante salientar o contacto humano, dado que cada um tem a sua experiência, cada um tem os seus conhecimentos e pode ser uma boa ajuda nesta tarefa.

A maioria dos artigos publicados em periódicos médicos obedecem ao formato IMRAD: Introdução (por que?); Métodos (como?); Resultados (o que?) e Discussão (significado). Uma forma de saber se devemos ou não ler um artigo é compreender como o investigador recolheu e escolheu analisar os resultados (métodos) e não pelo impacto ou interesse que nos desperta. Uma forma de validar um artigo é enviá-lo para um revisor e este comentar a validade científica do mesmo (peer review). Grande parte dos artigos contém falhas e, daí, acabam por não ser publicados, salientando-se motivos como: não ser original; abordar um aspeto científico pouco importante; não testar a hipótese do investigador; tamanho da amostra inadequado; não ser controlado devidamente, entre outros.

Por outro lado, existem pontos fun-

damentais que auxiliam a interpretar um artigo. Assim, deve ser identificada, na introdução, de forma reduzida, a questão da pesquisa (metodologia ou discussão), por que o estudo foi realizado e devem ser apresentadas as bases da pesquisa. Por sua vez, a metodologia deve definir claramente a questão de pesquisa ou hipótese que os autores decidiram testar. Para definir o delineamento do estudo, é necessário compreender se estamos perante um estudo primário ou secundário, ou seja, o tipo de estudo que procura resumir ou tirar conclusões dos estudos primários. Estes últimos podem ser: experimentos (ambientes artificiais e controlados); ensaios clínicos (intervenção); ou levantamentos (em que algo é mensurado). Por seu lado, a pesquisa secundária pode ser dividida em revisões (sistemáticas, não sistemáticas ou metanálises); diretrizes (tiram conclusões de estudos primários); análises de decisão ou análises económicas. Posteriormente, é necessário compreender se o delineamento da pesquisa foi apropriado à questão.

Num ensaio clínico randomizado, os participantes são escolhidos aleatoriamente e formam-se dois grupos que são acompanhados (ex: tratamento com medicamento vs placebo). Em média, são idênticos, daí que qualquer diferença seja atribuída à intervenção. Os ensaios clínicos randomizados são os mais utilizados em pesquisa médica para questões de intervenção (tratamento ou prevenção), permitindo uma avaliação rigorosa de uma variável, delineamento prospetivo, realização de metanálises e raciocínio hipotético-dedutivo. No entanto, são caros e dispendiosos, podendo introduzir vieses ocultos.

O estudo de corte é realizado em dois ou mais grupos de pessoas que são selecionadas com base nas diferenças na exposição e, por outro lado, acompanhados para observar em quantos indivíduos se manifesta a doença. Geralmente,

é medido em anos e os participantes podem ou não desenvolver uma doença.

Quando os pacientes com uma doença são identificados com controlos, tratam-se de estudos de caso-controlo, relacionados com aquilo que causa a doença e não com o tratamento, sendo os dados coletados com bases de dados. Em estudos transversais, uma amostra é entrevistada, examinada ou estudada para obter respostas a uma questão clínica. Os relatos de caso descrevem um único caso de um paciente que, posteriormente, podem ser reunidos e formar uma série de casos.

Para tomar decisões sobre intervenções clínicas, o ideal é utilizar revisões sistemáticas e metanálises. Caso não seja possível, recorre-se a ensaios clínicos randomizados com resultados definitivos ou não definitivos, estudos de coorte, estudos de caso-controlo, estudos transversais ou relatos de casos.

Neste ponto, a autora refere que muitos investigadores dão pouca importância aos aspetos éticos, mas são considerações que devem ser sempre tidas em conta em qualquer estudo, salientando o consentimento informado.

Por outro lado, para avaliar a qualidade de um artigo é necessário ter em conta a robustez da metodologia utilizada. Em teoria, não há motivos para testar uma hipótese científica que alguém já testou, no entanto, a ciência não é tão exata assim. Desta forma, é necessário compreender se a nova pesquisa acrescenta alguma informação à pesquisa realizada anteriormente, o que ajuda a caracterizar a originalidade do estudo. Antes de aceitar os resultados de uma pesquisa, devem ser consideradas algumas questões como de que forma os sujeitos foram escolhidos e quem foi incluído e excluído do estudo. Para compreender se o delineamento do estudo foi adequado, é necessário, por outro lado, avaliar que procedimento/intervenção foi considerado, qual foi o resultado e como foi medido.

Outro critério para avaliar a qualidade da metodologia é compreender se o viés sistemático foi evitado ou minimizado. O viés é caracterizado como sendo qualquer coisa que influencia erradamente as conclusões. O objetivo é que os grupos sejam tão parecidos quanto possível, exceto no que está a ser examinado. Consoante o delineamento do estudo, as etapas para reduzir o viés sistemático são diferentes. Em ensaios clínicos randomizados, o viés sistemático é evitado pela seleção de uma amostra subdividida, aleatoriamente, em diferentes grupos. Nestes ensaios clínicos randomizados, existem viés de seleção (diferenças sistemáticas, randomização incompleta); viés de performance (diferenças além da intervenção avaliada), viés de exclusão; e viés de deteção (diferenças na mensuração de desfecho). Em estudos de caso-controlo, o diagnóstico de caso e a decisão de quando o indivíduo se tornou um caso é que são mais sujeitas a viés, e não a avaliação do desfecho.

A avaliação deve ser realizada sem que as pessoas que julgam o estudo saibam a que grupos pertencem os indivíduos que estão a avaliar. Assim, consegue-se imparcialidade na interpretação dos resultados.

Como em qualquer estudo, as questões estatísticas essenciais não devem ser esquecidas, tais como o tamanho da amostra, a duração do acompanhamento e a integralidade do acompanhamento. O tamanho da amostra prende-se, essencialmente, com duas situações: se o nível de diferenças entre dois grupos irá ter um efeito clinicamente significativo; e qual a média e o desvio-padrão da principal variável de desfecho. Note-se que, antes de iniciar o estudo, é possível descobrir, através de fórmulas, o tamanho ideal de uma amostra para que se detete uma verdadeira diferença entre os grupos. O acompanhamento deve ser realizado durante um período de tempo suficiente para que seja visível o efeito da intervenção. Quando os participantes abandonam uma pesquisa, a probabilidade de continuarem o trata-

mento é menor, assim como aumenta a probabilidade de faltarem as consultas periódicas. Neste contexto, existem inúmeras razões para abandonarem o estudo, entre as quais temos a inclusão incorreta do paciente no estudo, suspeita de reação adversa, perda de acompanhamento, perda de motivação, ou morte.

Para uma boa análise dos resultados, é indispensável saber qual o melhor teste estatístico para aplicar em problemas comuns, devendo o analista ser capaz de descrever o que o teste faz e em que circunstâncias é ou não aplicável. Nos ensaios clínicos comparáveis, é necessário ter em conta se os grupos em estudo são semelhantes, quanto à idade, sexo e outras variáveis relevantes. Se os grupos diferirem significativamente, devem ser feitos alguns ajustes, de modo a facilitar a sua interpretação. Todos os testes estatísticos são paramétricos (assumem que os dados foram retirados de forma particular de distribuição) ou não-paramétricos (verificam a ordem de classificação dos valores e ignoram as diferenças absolutas entre eles). É mais difícil demonstrar a significância estatística com testes não paramétricos, embora, por vezes, os dados assimétricos podem ser transformados, de modo a produzir um gráfico com distribuição normal.

Um intervalo de confiança permite fazer uma estimativa tanto nos estudos positivos, como nos negativos, quer o poder de influência seja forte ou fraco e quer o estudo seja ou não definitivo. Se um estudo for repetido inúmeras vezes, não se conseguia obter exatamente o mesmo resultado. No entanto, em média, conseguiria estabelecer-se um nível particular de diferença entre os estudos, o intervalo de confiança. Apesar da sua importância, o uso de intervalos de confiança não é muito comum e, quando são aplicados, algumas vezes são incorretamente interpretados.

Antes de se submeter um paciente a um tratamento com fármacos, o médico

deve identificar o objetivo final do tratamento para esse paciente, selecionar o tratamento mais apropriado e especificar o alvo do tratamento. A escolha de um tratamento tem de ser determinada pela evidência do que funciona e não pelo que parece ou deve funcionar. Antes que os resultados clínicos possam ser usados no marketing, a indústria deve justificar a utilidade dessas medidas, demonstrando uma ligação plausível e consistente entre o desfecho clínico e a progressão da doença.

Por outro lado, para avaliar a utilidade de um teste, é necessário definir uma amostra, fazer o teste de rastreio e o de diagnóstico (padrão-ouro) e, de seguida, compará-los para verificar se os resultados correspondem (estudo de validação). Se os valores da sensibilidade (verdadeiros positivos), especificidade (verdadeiros negativos), falsos positivos e falsos negativos se situarem dentro dos valores razoáveis, o teste foi válido. Assim, existem 10 questões essenciais a aplicar num artigo que pretende validar um teste de diagnóstico ou de rastreio. 1. O teste é potencialmente relevante? 2. O teste foi comparado com um verdadeiro padrão-ouro? 3. A amostra é apropriada? 4. O viés de investigação e de expectativa foram evitados? 5. O teste é reprodutível? 6. Quais as propriedades do teste determinadas deste estudo de validação? 7. Os intervalos de confiança para sensibilidade e outras propriedades foram fornecidos? 8. Uma variação normal sensível foi obtida com estes resultados? 9. O teste foi colocado no contexto de outros testes potenciais na sequência de investigação diagnóstica para o problema? 10. Depois de respondidas e analisadas as questões, poderá ou não ser validado o teste.

Nesta sequência, uma revisão sistemática é um resumo de estudos primários e realizada de acordo com uma metodologia explícita e reproduzível, contendo uma descrição dos objetivos, materiais e métodos. É necessário sempre haver um

especialista em revisões sistemáticas e outro no respetivo assunto do artigo, para uma avaliação cuidadosa das evidências. Quando rejeitamos um estudo, é porque os seus objetivos ou métodos de estudo não atenderam aos critérios de qualidade, independentemente dos resultados.

As revisões sistemáticas permitem, assim, chegar a conclusões mais acuradas e fiáveis. O atraso entre as descobertas científicas e a implementação das estratégias é reduzido. Os resultados de diferentes estudos podem ser comparados, permitindo a generalização que ainda aumentam a precisão do resultado global.

Em suma, numa revisão sistemática, deve-se conseguir encontrar a questão de partida, o autor tem de tornar claro de onde veio a informação e deve, ainda, fazer uma pesquisa extensa em várias bases de dados e outras fontes importantes, mencionando-as todas, incluindo pesquisas prévias, devendo, de igual modo, evitar o imperialismo linguístico. A qualidade metodológica é assegurada, prevenindo erros sistemáticos e de precisão, diminuindo erros aleatórios e aumentando a confiança nos resultados. Por último, deve ser preservada a validade externa dos estudos incluídos, isto é, a extensão com que os resultados são generalizáveis ou aplicáveis a uma população em particular.

Outra área significativa apresenta por esta autora é a metanálise que constitui uma síntese dos resultados de estudos independentes voltados para uma única. Uma metanálise é mais fácil de entender para o não-estatístico do que imensos artigos de pesquisa primária da qual resultou, não devendo ser confundida com uma revisão sistemática. Na linguagem da metanálise, heterogeneidade significa que os resultados de cada ensaio clínico randomizado individual são incompatíveis com outros. No entanto, é arbitrário chamar um conjunto de ensaios clínicos randomizados de heterogêneos com base na sobreposição dos intervalos de confiança, porque esta última é também

arbitrária. Para isto, é, então, usada uma variante do teste Qui-Quadrado.

Os protocolos são instruções sobre como agir em determinada situação. Estes incluem menos espaço para crítica individual, ao contrário das diretrizes que são afirmações para auxiliar nas decisões do médico sobre o cuidado de saúde apropriado para situações clínicas específicas. Por outro lado, facilitam a tomada de decisões clínicas, fornecem um guia para avaliar o desempenho profissional, educam pacientes e profissionais sobre a prática mais eficaz, melhorando a relação serviço-custo dos serviços de saúde e possibilitando também um melhor controle externo.

Apesar de tudo isto, há resistência dos médicos às diretrizes, devido, essencialmente, à interferência com a liberdade clínica, senso de competência e autonomia, uma vez que alguns agem de acordo com o que aprenderam há muitos anos, sem lerem um artigo desde então. Mas é necessário não esquecer, também, que uma medicina enviesada por evidências, isto é, baseada em grupos diferentes e com tratamentos apenas semelhantes, também não é o correto. Assim sendo, uma boa diretriz não deve forçar ao abandono do bom senso e do julgamento, estando o seu seguimento dependente da estratégia de desenvolvimento, divulgação e implementação, sendo maior a probabilidade de a diretriz ser efetiva, se esta for desenvolvida, interna e especificamente, para determinado paciente. Hoje, todos os profissionais de saúde devem prestar serviços, baseados nas melhores evidências de pesquisa disponíveis.

Uma análise econômica tem que ser colocada no contexto da qualidade geral e relevância do estudo, envolvendo o uso de técnicas analíticas para definir opções na distribuição de recursos. Assim, existem quatro tipos de avaliação econômica: análise de minimização de custos; análise de custo-efetividade; análise de custo-utilidade; e análise de custo-benefício.

Cada estudo possui apenas uma análise importante e, no caso da saúde, há muita controvérsia na comparação entre os custos e os benefícios das intervenções em saúde. Em alguns casos, as medidas específicas de bem-estar para a doença são mais válidas que as medidas gerais, porém, o economista da saúde tende a tomar decisões sobre grupos de pacientes em que as medidas da qualidade de vida específicas para cada paciente possuem pouca importância.

Há autores de instrumentos padrão que se dedicaram a analisar se uma intervenção melhora ou piora a saúde do paciente, e se a sua própria mensuração é válida e confiável. Várias medidas são usadas para analisar e mensurar a qualidade de vida e a sua variação com determinada intervenção clínica e, uma delas, relaciona os anos de vida com a qualidade (QALY), conduzindo à conclusão que qualquer medida de valores de preferência de estado de saúde é reflexo das preferências e dos preconceitos dos indivíduos que as desenvolveram e, além disso, os valores dependem da forma como as questões foram apresentadas às pessoas.

Desta forma, qualquer medida e valor atribuído devem ser examinados cuidadosamente, uma vez que este valor é, supostamente, um indicador do estado de saúde de uma pessoa. E também deve ser examinada a forma como cada medida pode diferir, de acordo com os diversos estados de doença. Todas são uma mais-valia, mas nenhuma é absoluta e unânime.

A pesquisa qualitativa começa com uma intenção de explorar uma área em particular, com uma recolha de dados a partir dos quais se geram ideias e hipóteses, sendo que estas últimas se podem ir alterando, consoante se progride na pesquisa, não deixando o pesquisador de ter bem claro o que está a investigar. Uma lista de verificação para avaliação crítica de pesquisas qualitativas é possível e é uma tarefa que tem sido trabalhada.

Embora não possa ser tão exaustiva e universalmente aplicável como os guias para avaliar pesquisas quantitativas, é possível estabelecer regras básicas. Um artigo qualitativo deve, assim, ser avaliado de acordo com as seguintes questões: 1. o artigo descreve um problema clínico importante, avaliado por meio de uma questão claramente formulada? 2. A decisão de fazer uma abordagem qualitativa foi apropriada? 3. Como foram selecionados a situação e os sujeitos? 4. Qual era a perspectiva do pesquisador e foi tida em conta? 5. Que métodos o pesquisador usou para recolher os dados? 6. Estes métodos são descritos com detalhes suficientes? 7. Que métodos o pesquisador usou para analisar os dados e que medidas de controlo de qualidade foram implementadas? 8. Os resultados possuem credibilidade e são clinicamente importantes? 9. Que conclusões foram tiradas? 10. As conclusões são justificadas pelos resultados? 11. Os achados do estudo são aplicáveis a outras situações?

A popularidade da pesquisa qualitativa nas ciências biomédicas tem vindo a aumentar, uma vez que os métodos quantitativos baseados em números não fornecem muitas vezes respostas, ou fornecem respostas incorretas, podendo ser enganosos, reducionistas e irrelevantes para os factos reais.

Um questionário, por sua vez, não é um método qualitativo nem quantitativo, mas sim um instrumento para recolher dados através das suas perguntas. Devemos, assim, ter em conta que dados de má qualidade levam a conclusões enganosas e, de facto, os erros metodológicos são comuns em pesquisas por questionário. Desta forma, há um conjunto de questões que devem ser colocadas para avaliar um artigo que descreve um estudo por questionário: *i.* Qual era a pergunta de pesquisa e se o questionário era adequado para respondê-la. *ii.* O questionário utilizado no estudo era válido e confiável? *iii.* Como parecia o questionário?

iv. Era adequado à população alvo? *v.* As instruções estavam claras? *vi.* O questionário foi adequadamente testado em um estudo piloto? *vii.* Qual foi a amostra? *viii.* Como foi aplicado o questionário e a taxa de respostas foi adequada? *ix.* Como foram analisados os dados? *x.* Quais foram os principais resultados? *xi.* Quais são as conclusões-chave?

As estratégias de cuidados em saúde, como o caso de novos tratamentos, levam, muitas vezes, anos a serem implementadas. Neste sentido, sendo a tentativa de redução desse tempo alvo de muitos estudos frequentemente inconclusivos, a implementação de melhores práticas é altamente complexa e depende das pessoas. Uma abordagem que resulta em determinado estudo pode, no entanto, não resultar noutra. Neste sentido, é geralmente aceite que uma organização assimila mais rapidamente um produto ou prática se for grande, funcionalmente diferenciada, se tiver recursos que possa canalizar para novos projetos e se as equipas puderem trabalhar de forma autónoma. No entanto, continua, mesmo assim, a ser difícil incluir as evidências na elaboração de políticas, porque estas são definidas em relação a um contexto específico, o que as torna ambíguas.

Em conclusão, o livro de Trisha Greenhalgh é muito útil, porque é claro, de fácil leitura e percepção, com todas as etapas que devem ser seguidas para a realização de um artigo, constituindo uma ajuda imprescindível na análise de diversos artigos. Um dos pontos muito positivos é o facto de dar exemplos quando é referida uma etapa daquilo que deve ser feito e, assim, ser mais fácil para investigadores, estudantes e profissionais de saúde cumprir todas as regras exigidas.

Margarida Pocinho

Escola Superior de Tecnologia da Saúde de Coimbra / Instituto Superior Miguel Torga