

## Ensaio clínico, conferência de consenso e prática clínica\*

Alan D. Sniderman

O processo pelo qual introduzimos e autorizamos mudança na prática clínica passa por uma transformação fundamental. No passado, o que fazíamos na clínica, nós o escolhíamos a partir de uma prática individual. O que fazemos agora é cada vez mais ditado por uma variedade de grupos de “peritos”, alguns dos quais são indicados por entidades reconhecidas e competentes e alguns automeados. Essa transformação fundamental afeta profundamente pacientes e clínicos, e, portanto, deveria ser examinado, não simplesmente justificado.

As duas inovações responsáveis por muita, senão pela maior parte destas transformações, são o ensaio clínico aleatório e a conferência de consenso. Estes avanços interagem em sinergia, extraindo autoridade um do outro. Os ensaios clínicos aleatórios são tomados como “a melhor evidência” nas conferências de consenso; e as conferências de consenso são a arena na qual a importância prática dos ensaios clínicos aleatórios é determinada. Ambos são vistos como impessoais, desligados, e objetivos, e essas características tornaram-se parte essencial de seu capital intelectual e moral. Ou seja, elas integram o poder de persuasão deles.

Mas nenhum processo é perfeito. Nenhum processo está acima das pessoas envolvidas e nenhum processo pode fugir das limitações inerentes ao processo em si. Dada a extensão de seu impacto na medicina clínica, parece prudente refletir sobre o que podem ser essas limitações. Esse é o objetivo deste artigo, que terminará por fornecer algumas sugestões sobre como lidar com as limitações identificadas.

\* Publicado originalmente em *The Lancet*, July 24, 1999. Tradução de Clotilde Teixeira.

## Limitações e forças dos ensaios clínicos aleatórios

Os ensaios clínicos aleatórios são ferramentas poderosas para testar terapias e, quando bem desenhados e conduzidos, devem, sem dúvida, influir em nossa tomada de decisões. Mas nem mesmo seus mais fortes proponentes sugerem que não tenham limites. Essas limitações são de dois tipos: específicas e fundamentais. Algumas das limitações específicas são óbvias. Estão entre elas o fato de muito raramente os ensaios examinarem diferentes dosagens de agentes farmacológicos; de o pleno potencial de benefício ou dano que podem resultar de combinações entre terapias quase nunca ser definido, uma vez que a maioria dos ensaios estuda somente intervenções únicas. Além disso, os estudos têm limite de tempo e a extrapolação de seus resultados no porvir não é direta; no entanto, a extrapolação acontece rotineiramente. Feinstein e Horwitz notaram uma série de questões similares, incluindo o fato de que informações de ensaios aleatórios raramente estão disponíveis para responder questões sobre etiologia, diagnose e prognose, e que respondem a apenas parte das questões da clínica.

Há, no entanto, outro tipo de limitação ainda mais fundamental, porém menos examinada: o problema da aplicabilidade. Os ensaios podem ser divididos entre os que têm modelos de exclusão e os que têm modelos de inclusão. Nos ensaios de exclusão, os pacientes são selecionados com base em um critério de inclusão e exclusão explícita. Um ensaio sobre terapia para ataque cardíaco, por exemplo, pode selecionar pacientes com problemas cardíacos, baixa função sistólica e sem evidência recente de doença coronária instável. A faixa etária seria estabelecida, e pacientes portadores de várias doenças importantes seriam, normalmente, excluídos. A generalização automática de conclusões além do grupo estudado é, portanto, impossível. Uma vez que a generalização é limitada, os resultados serão de valor limitado no mundo clínico do dia a dia. Isto é, a aplicabilidade é limitada porque o desenho é de exclusão.

Este problema pode ser superado por um modelo inclusivo. No entanto, este modelo demanda estudos muito abrangentes, nos quais somente os resultados mais diretos e mais importantes – como, por exemplo, o efeito de uma terapia sobre a mortalidade – sejam levados em conta. Uma vez que no recrutamento se especificam poucos critérios de exclusão, esse tipo de modelo de ensaio é inclusivo em vez de exclusivo. Aqui, há um problema oposto de aplicabilidade: as limitações estão na interpretação dos resultados gerais para pacientes específicos. Todos os sub-grupos são beneficiados? Os benefícios são similares? Um resultado geral negativo não exclui a possibilidade de que alguns se beneficiaram enquanto um grande número de outros foram prejudicados pela terapia. Pode ser possível diferenciar beneficiados de prejudicados; mas para que isso aconteça é necessário que todas as determinantes potencialmente relevantes para o desfecho tenham

sido encontradas e especificadas em todos os participantes. Os mesmos problemas de aplicabilidade acompanham um resultado geral positivo ou negativo.

Assim, um modelo inclusivo ultrapassa a limitação de generalização que se aplica ao modelo exclusivo; mas, por sua vez, a aplicabilidade do estudo inclusivo é necessariamente limitada pela falta de especificidade. Ensaios inclusivos revelarão a magnitude geral do benefício ou dano de qualquer terapia, mas não revelarão as especificidades do equilíbrio entre o bom e o mau. Portanto, a aplicação dos resultados dos ensaios clínicos à prática clínica nunca poderá ser completamente direta. Isso não quer dizer que não tenham valor. Bem ao contrário, ensaios clínicos inclusivos ou exclusivos são instrumentos insubstituíveis para orientar a tomada de decisões clínicas. Porém, como seus proponentes reconhecem, o médico deve interpretar ativamente os resultados dentro do contexto clínico específico; e é seu dever não perder de vista a responsabilidade e a necessidade de fazê-lo. Parafraseando Tonelli, os ensaios clínicos são necessários, mas em muitos casos não suficientes para decisões clínicas.

Um motivo importante pelo qual se deu aos ensaios clínicos o papel de árbitro final da prática clínica foi que, apesar de organizados, executados, analisados e promovidos por indivíduos, eles são percebidos como coletivos impessoais. O acrônimo pelo qual os ensaios ficam conhecidos capturam essa identidade supra-pessoal.<sup>1</sup> Assim, os ensaios clínicos adquirem sua autoridade impessoal ao obscurecerem as conexões de seus autores individuais com aqueles que participam deles.

Não devemos, todavia, perder de vista o fato de que, como todos os outros atos científicos, os ensaios clínicos são produto de indivíduos que têm interesse na boa acolhida dos seus resultados, A expressiva influência dos ensaios clínicos na tomada de decisões determina um sucesso ou um fracasso, ou seja, recompensa profissional e material para os indivíduos envolvidos. Devemos nos lembrar, também, que a grande maioria dos ensaios clínicos são patrocinados por empresas farmacêuticas que desempenham um papel ativo no desenho do ensaio e na análise dos dados. Davidson relatou que os ensaios financiados por patrocinadores externos têm mais possibilidade de apresentar resultados positivos do que ensaios similares sem tais patrocinadores. Stelfox e seus colegas, disseram

1. O autor se refere às siglas pela qual os testes ficam conhecidos entre os médicos das especialidades. Por exemplo: o ensaio "Hearth and Estrogen/progestin Replacement Study", sobre efeitos da reposição hormonal em mulheres na pós-menopausa, tem o conveniente acrônimo HERS. Note-se que os autores do estudo obliteraram a inicial de "progestin" de forma a dar lugar à polissemia do acrônimo. Essa é prática comum no mundo da medicina baseada em evidências. (N. da T.)

que pesquisadores favoráveis ao uso de antagonistas de cálcio tinham uma probabilidade substancialmente maior de ter ligações financeiras com fabricantes dessas drogas do que os pesquisadores neutros ou críticos.

Mas existe outro ponto de relacionamento que não foi notado anteriormente. Muitos ensaios foram feitos usando agentes similares para problemas similares e, naturalmente, cada empresa tentará diferenciar seu produto do de seus concorrentes, mesmo quando o resultado final é semelhante. Da mesma forma, os pesquisadores tentarão diferenciar sua contribuição da de outros pesquisadores que tomam parte em ensaios similares. Nesse caso, os interesses das companhias e dos pesquisadores podem fundir-se imperceptivelmente. O resultado final é que debates científicos importantes podem se misturar a abordagens competitivas de *marketing* – e, algumas vezes, podem ser dirigidos por elas. O debate científico e a concorrência comercial são legítimos. Mas não há risco de o defensor da ciência se tornar, também, defensor da empresa, e vice-versa? Nós, os leitores; nós, os que prescrevemos; e nós, os consumidores, devemos nos lembrar da origem humana e corporativa dessas peças científicas, uma vez que nossos interesses nem sempre são idênticos aos de quem as produz.

### **Limitações das conferências de consenso**

Cada vez mais os resultados dos ensaios aleatórios levam às decisões das conferências de consenso e, cada vez mais, as decisões das conferências de consenso direcionam a prática clínica. Acontece, também, que os ensaios e terapias não aprovados por conferências de consenso enfrentam dificuldades quando se trata de obter reembolso. As recomendações das conferências de consenso podem começar a ser adotadas legalmente como padrão de decisão na prática médica. Mas uma conferência de consenso é um processo social e um processo científico. Aqui há um problema. A autoridade de uma conferência de consenso emana da aura de “científica” o que, nesse contexto, implica: primeiro, que todas as evidências relevantes foram identificadas e listadas, de maneira inteiramente objetiva e impessoal; segundo, que as conclusões derivadas daí são o produto simples e direto das evidências. No entanto, a identidade entre “conferência de consenso” e “atividade científica”, merece exame.

Pelo fato de mesmo a melhor evidência (os resultados de um ensaio clínico aleatório) ser necessariamente limitada, as conclusões de qualquer deles demandarão, sempre, julgamento e interpretação. “A” pode advir de “B”, mas não necessariamente em uma situação específica; os que estabelecem as regras e aqueles que devem segui-las precisam reconhecer sua falta de exatidão.

Por exemplo, muitas autoridades e muitos grupos que emitiram diretrizes, convenceram-se, com base em evidências disponíveis, que os estrogênios reduzem, em muito, a taxa de problemas coronários, e que esse benefício suplanta qualquer risco que possa estar associado ao uso. Apesar de as recomendações sempre terem sido associadas a alguma dúvida, a tendência final foi encorajar o uso de estrogênio em mulheres portadoras de doenças coronárias na pós-menopausa. No entanto, os resultados do único ensaio aleatório realizado até aqui contradizem essa visão, uma vez que nenhum benefício significativo foi observado no todo e, a curto prazo, o risco da terapia tende a ser maior que o benefício. Qualquer que venha a ser a resposta final, este exemplo ilustra o risco inevitável envolvido em previsões que se baseiam em informação insuficiente. Também salienta outra deficiência do processo – que as diretrizes valem até que sejam substituídas no próximo ciclo. Claramente, um método precisa ser desenvolvido para atualizá-las rapidamente, em aspectos específicos, quando necessário.

Outro problema com as conferências de consenso é que os participantes tendem a ser defensores de visões particulares. A participação na conferência dá, então, espaço para que o participante some o papel de juiz e júri ao papel de defensor. Advogar uma visão em ciência não é mau. De fato, é quase inevitável. Sendo este o caso, o remédio é assegurar que os membros das conferências tenham filiação intelectual ampla, e não limitada. Isso, no entanto, nem sempre é verdadeiro, pois ameaçaria uma das mais inovadoras características sociais das conferências de consenso – a unanimidade de sua autoria. Essa unanimidade é o que lhe dá autoridade. Porém, essa unanimidade nos deve alertar para o fato de que as conclusões representam soluções de compromisso e simplificações.

O compromisso não é, em si, ruim, nem deve necessariamente ser evitado. A simplificação, até certo ponto, também é essencial. Mas quando o compromisso e a simplificação acontecem, devem ser conhecidos, e esse conhecimento deve ser critério para a avaliação do mérito de qualquer processo específico. Da mesma forma, o modelo de unanimidade não precisa ser aplicado com tanto rigor. Médicos e pacientes podem lidar com opiniões divergentes e, na verdade, isso lhes dá grande satisfação. Nossas mais altas cortes não exigem unanimidade, e mudanças em nossa percepção da lei têm se servido da livre expressão de pontos de vista diversos.

Sugeri que a natureza social das conferências de consenso pode levar a recomendações científicas imperfeitas. Aqui está um exemplo em apoio a essa asserção. Há, agora, evidências persuasivas de que a diminuição das concentrações de colesterol LDL por meio de estatinas leva à diminuição da mortalidade e morbidez em pacientes com doenças coronárias. Há, também, forte evidência de que, diminuindo as concentrações de colesterol LDL com o uso de estatinas, pode-se evitar as primeiras crises. Portanto, as estatinas são bem-sucedidas nas manifestações iniciais e nas secundárias.

No entanto, o Programa Nacional de Educação sobre Colesterol, um programa que tem tido grande impacto tanto dentro dos E.U.A. como fora, tem recomendado estratégias que diferenciam a prevenção primária da secundária, e que favorecem uma terapia mais intensiva para a prevenção secundária do que para a prevenção primária da doença coronária. Mas, quão segura é essa abordagem?

Estruturalmente, a doença coronária começa bem antes que a crise ocorra. Além do mais, muitas, senão a maioria das ocorrências clínicas, se devem à ruptura de pequenas placas não-oclusivas, que não produzem sintomas clínicos antes da ruptura. Assim, muitas das pessoas selecionadas para prevenção primária que, clinicamente, não apresentam sintomas, já são portadoras de doença estrutural significativa; presume-se, portanto, que essas pessoas seriam as maiores beneficiárias da terapia com estatinas, dentro do tempo limitado dos ensaios de prevenção primária feitos até o momento. Qual seria, então, a evidência científica que apóia a recomendação de que a redução de colesterol LDL necessária para tratar coronárias anormais de pessoas sem doença sintomática deve ser menos acentuada do que a redução necessária para tratar as artérias anormais daqueles com doença sintomática? Na realidade, não há nenhuma evidência. A evidência disponível indica justamente o contrário. A distinção feita não se baseia em ciência porque não há ciência em que se basear. Sua origem está em outro lugar. Um dos motivos pode ser econômico. A redução acentuada das concentrações do colesterol LDL quase sempre demanda terapia farmacológica, o que envolve custos substanciais. Pacientes com doença sintomática correm alto risco individual de ocorrências adicionais e provavelmente fatais – portanto, os custos resultantes da terapia farmacológica aparecem como razoáveis. Por outro lado, muitos candidatos à prevenção primária não apresentam alto risco individual e, assim, um grande número de indivíduos deve ser tratado para evitar um único evento. Terapia farmacológica para todos esses indivíduos seria proibitivamente cara. Mas este é um problema econômico, não uma análise científica. O problema científico é que os critérios de julgamento de risco são fracos, e a evidência de sua fraqueza – sua imprecisão – são os grandes números que são ditos estar em grande risco em relação aos números pequenos daqueles que realmente irão desenvolver uma doença prematura. O objetivo científico deveria ser desenvolver maneiras melhores de identificar o risco.

No caso de doenças vasculares, há evidências de que é possível desenvolver abordagens melhores, mas isso exigirá a inclusão de variáveis outras que não as usuais ou, talvez, a substituição completa delas. Por exemplo, o Estudo Cardiovascular de Quebec, um estudo prospectivo em que foram avaliados tanto fatores de risco convencionais quanto novos, mostrou que a apolipoproteína B era o melhor indicador do risco de um evento cardiovascular entre os índices

de lipídeos tomados individualmente. A combinação de alta insulina em jejum, alta apolipoproteína B e densidade baixa de LDL estava presente em quase metade dos participantes do ensaio que tiveram eventos clínicos e estava associada a um risco 18 vezes maior. Por outro lado, a combinação de triglicérides alto e altas concentrações de colesterol LDL com baixa concentração de colesterol HDL estava associada a um risco apenas cinco vezes maior.

Não está em questão aqui se a apolipoproteína B e a insulina deveriam ser acrescentadas às variáveis, ou até substituí-las. O ponto é que o *status quo* não é adequado. Mas mudar não é simples. Pontos de vista se estabeleceram, e a percepção de ser necessária uma continuidade pode se tornar determinante. Por outro lado, devido ao custo da terapia farmacológica, é óbvia a necessidade de melhorar nossas tecnologias para identificar pessoas que apresentam alto risco. Se as deficiências de nosso conhecimento não forem destacadas e reconhecidas as limitações de nossas estratégias de prevenção, as soluções, infelizmente, podem ser retardadas indefinidamente.

## Recomendações

Ensaio clínico e conferências de consenso são instrumentos insubstituíveis e valiosos. A questão não é abandoná-los, e sim fortalecê-los. Seguem sugestões de como podemos fazer isso.

Os periódicos médicos já ajudam muito na interpretação de ensaios clínicos aleatórios, ao incluir editoriais que auxiliam no complexo processo de interpretação. Eles ajudam, também, ao publicar revisões, de tempos em tempos, dos princípios de interpretação de evidências e dos fundamentos das estatísticas. Esta prática deveria ser ampliada incluindo declarações e diretrizes de consenso.

Nos hospitais, muito pode ser feito. Em particular, a apresentação detalhada de um caso específico pode mostrar melhor como o tratamento de determinado paciente se liga ao estado geral do conhecimento. Esta prática diminuiu em muitos centros, mas merece ser reativada porque ela reafirma a todos que o objeto de nosso trabalho é o indivíduo.

O componente mais importante de uma conferência de consenso é a montagem do painel de especialistas. A diversidade de participantes é necessária. Além dos participantes conhecedores do assunto, pessoas de outras áreas deveriam ser incluídas. Em muitos casos, as conclusões têm amplas repercussões sociais, portanto, deveria haver uma tentativa para incluir a participação da sociedade em sua formulação. Além do mais, a cada ciclo, uma porção substancial de participantes deveria ser nova, para dissociar as pessoas do processo. Modelos

que incorporem muitas destas características foram bem-sucedidos e deveriam ser usados com mais freqüência.

Mecanismos de revisão periódica das diretrizes clínicas, em resposta a todos os avanços importantes na base de conhecimento, devem ser desenvolvidos como parte integral do processo de estabelecimento das diretrizes. A internet pode ser uma ferramenta poderosa. Pode haver demora entre o final da revisão científica e sua publicação. Portanto, cada conferência de consenso deve afirmar explicitamente quando parou de receber dados científicos e, da mesma forma como fazemos com os alimentos, deveria haver uma estimativa do prazo de validade das diretrizes em questão.

Acima de tudo, para melhorar as recomendações das conferências de consenso, não deveríamos temer a diversidade de pontos de vista. Ao contrário, o progresso da ciência pede essa diversidade. Face à incompletude do conhecimento, o único caminho para formular as melhores respostas possíveis é permitir que diferentes pontos de vista se expressem. Em *Assim falou Zarathustra*, Nietzsche escreveu: “Cansaço, que de um salto quer chegar até o último, de um salto mortal; um pobre cansaço ignorante, que nem mesmo querer quer mais”.<sup>2</sup> Nós, clínicos, não podemos, de tão fatigados, chegar ao ponto de abandonar a responsabilidade de pensar por nós mesmos.

96

## Referências

DAVIDSON, R.A. Source of funding and outcome of clinical trials. *J. Gen. Intern. Med.* n. 1, p. 155-58, 1986.

FEINSTEIN, A.R., HORWITZ, R. Problems in the “evidence” of “evidence based medicine”. *Am. J. Med.*, n. 103, p. 529-35, 1997.

STELFOX, H.T., CHUA, G., O’ROURKE, K., DESTSKY, D.A. Conflicts of interest in the debate over calcium channel antagonists. *N. Engl. J. Med.*, n. 338, p. 101-6, 1998.

TONELLI, M.R. The philosophical limits of evidence-based medicine. *Acad. Med.*, n. 73, p. 1234-40, 1998.

2. Nas palavras da tradução de Rubens Rodrigues Torres Filho. A citação está no décimo-terceiro parágrafo da seção “Ultramundanos”, na Primeira Parte de *Assim Falou Zarathustra*. O parágrafo completo: “Cansaço, que de um salto quer chegar até o último, de um salto mortal; um pobre cansaço ignorante, que nem mesmo querer quer mais: foi ele que criou todos os deuses e ultramundos”.