

PRÁTICAS CIRÚRGICAS BASEADAS EM EVIDÊNCIAS: APENDICECTOMIA LAPAROSCÓPICA *VERSUS* A CÉU ABERTO

EVIDENCE-BASED SURGICAL PRACTICES: LAPAROSCOPIC VERSUS OPEN APPENDECTOMY

Bráulio dos Santos Júnior¹; Carlos Alberto Guimarães²;
Grupo de Estudos de Revisão Sistemática do Rio de Janeiro

RESUMO: Os cirurgiões são instados a não somente ler os artigos dos periódicos, mas também a compreendê-los e analisá-los criticamente quanto à validade. Eles cuidam melhor dos seus pacientes se são capazes de analisar criticamente a literatura e aplicar os resultados a sua prática. Este é o primeiro artigo de uma série - Práticas Cirúrgicas Baseadas em Evidência - que tem por objetivo apresentar tópicos de Cirurgia Baseada em Evidências (*Rev. Col. Bras. Cir.* 2008; 35(1): 056-060).

Descritores: Medicina baseada em evidências - Educação médica continuada - Ensaios clínicos controlados aleatórios - Apendicectomia - Laparoscopia.

INTRODUÇÃO

Os cirurgiões tomam decisões terapêuticas baseadas geralmente na experiência pessoal, nas recomendações de outros cirurgiões ou nos conhecimentos das ciências básicas cirúrgicas. Com essa prática, intervenções sem efetividade ou até mesmo danosas têm sido aceitas como tratamentos-padrão.

A prática cirúrgica baseada em evidências enfatiza a necessidade da avaliação da eficácia das intervenções e terapêuticas, antes de aceitá-las como padrão.

Os relatos de casos e as opiniões dos especialistas representam o mais baixo nível de evidência. Com um nível um pouco mais alto, estão os estudos retrospectivos de série de casos, forma comum de publicação em Cirurgia.

Os estudos prospectivos não randomizados, tais como coortes concorrentes de pacientes, não têm as limitações da coleta de dados retrospectiva e da comparação com “controles” históricos, mas, mesmo assim, estão sujeitos a múltiplos vieses.

Somente um ensaio clínico randomizado, com desenho e condução adequados, fornece um alto nível de evidência para suportar as decisões terapêuticas cirúrgicas. Finalmente, quando há vários ensaios clínicos randomizados, uma revisão sistemática desses estudos prove o nível mais alto de evidência. Idealmente, os cirurgiões deveriam avaliar criticamente as evidências publicadas, a fim de adequar a elas as suas práticas diárias.

O objetivo desta série denominada Práticas Cirúrgicas Baseadas em Evidências é ajudar os cirurgiões nesse último processo.

CENÁRIO CLÍNICO

Um homem de 20 anos se apresenta no Serviço de Emergência com dor abdominal de início súbito há doze horas, localizada no quadrante inferior direito. O exame físico mostrou dor à palpação nessa área, com alguma defesa voluntária e sem dor à descompressão; a contagem de leucócitos era de 11.000/mm³. A tomografia computadorizada do abdome e pelve evidenciou apêndice dilatado com edema mesentérico adjacente. O cirurgião-chefe solicitou ao residente que providenciasse uma laparotomia, mas esse último indagou se não seria melhor para o paciente se a apendicectomia fosse feita por videolaparoscopia. O cirurgião-chefe respondeu que desconhecia evidências científicas da superioridade da apendicectomia laparoscópica e deu ao residente a tarefa de buscar a literatura para que voltassem a discutir o assunto no dia seguinte.

BUSCA BIBLIOGRÁFICA

O residente, que já havia realizado um breve treinamento sobre busca em bases bibliográficas, resolveu acessar a base MEDLINE pela interface PubMed (<http://www.pubmed.gov>). Na base MeSH, ele selecionou os descritores relevantes e acabou elaborando a seguinte estratégia: (“Appendectomy”[MeSH] OR “Appendicitis/surgery”[MeSH]) AND “Laparoscopy”[MeSH]. Com essa estratégia e limitando a busca a estudos em humanos publicados nos últimos dez anos, ele chegou a 727 artigos. Como havia muito pouco tempo para o levantamento, ele decidiu restringir ainda mais a busca, usando as “Clinical Queries”,

1. Médico da Coordenação de Ensino e Pesquisa (Unidade de Epidemiologia Clínica) do Instituto Nacional de Cardiologia, Ministério da Saúde.

2. TCBC-RJ, Professor Titular de Clínica Cirúrgica do Curso de Graduação em Medicina da Fundação Educacional Serra dos Órgãos (Teresópolis, RJ); Professor Adjunto do Departamento de Cirurgia – UFRJ.

Recebido em 12/11/2007

Aceito para publicação em 20/12/2007

Conflito de interesses: nenhum

Fonte de financiamento: nenhuma

que são filtros metodológicos elaborados para várias categorias de perguntas clínicas. Aqui, ele repetiu a sua estratégia de busca (e seus limites) com o filtro específico para estudos sobre terapia e obteve 45 artigos. Dos artigos que poderiam responder a sua pergunta, apenas um (Katkhoua 2005) era identificado como de acesso livre que foi obtido no endereço <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1357752&blobtype=pdf>.

Na verdade, um outro artigo (uma revisão sistemática da Colaboração Cochrane) também estava disponível gratuitamente, mas infelizmente o nosso residente – como a esmagadora maioria dos profissionais de saúde brasileiros – não sabia que poderia acessar a Cochrane Library pelo portal da Biblioteca Virtual em Saúde (<http://cochrane.bireme.br>).

AVALIAÇÃO CRÍTICA DO ARTIGO ENCONTRADO

A avaliação crítica do artigo foi baseada na leitura do texto “Users’ guide to evidence-based surgery: how to use an article evaluating surgical interventions” publicado no Canadian Journal of Surgery em 2001.

Três principais perguntas devem ser respondidas quando se faz a avaliação crítica de um artigo que compara intervenções cirúrgicas: os resultados são válidos? Quais são os resultados? Os resultados se aplicam aos meus pacientes? Essas três questões serão respondidas, utilizando o artigo de Katkhouda et al. que compara a eficácia e a segurança da apendicectomia laparoscópica versus apendicectomia a céu aberto, por meio de um ensaio clínico randomizado duplo-cego.

OS RESULTADOS DO ESTUDO SÃO VÁLIDOS?

Os pacientes foram randomizados para receber o tratamento e houve sigilo de alocação?

Se um estudo que compara intervenções cirúrgicas não é metodologicamente consistente, devemos ter muita cautela em utilizar as suas recomendações na nossa prática diária. O principal conceito – mas não o único – da metodologia de um bom desenho de estudo é a randomização dos sujeitos da pesquisa.

Os desfechos clínicos são influenciados não somente pelas intervenções cirúrgicas específicas, mas também pela seleção dos pacientes, comorbidades, fatores prognósticos conhecidos e vários fatores desconhecidos. A randomização é a melhor maneira de se distribuir uniformemente, entre o grupo experimental e o grupo controle, esses determinantes dos desfechos clínicos. Além disso, o processo de randomização precisa ser “sigiloso”, isto é os pesquisadores que incluem os pacientes no estudo não podem ter acesso à tabela de alocação e, portanto, não podem saber a intervenção que será atribuída ao próximo paciente do estudo. A violação do sigilo de alocação permite a manipulação, mesmo que involuntária, dos grupos, resultando em vícios de seleção e perda do equilíbrio dos fatores de confusão do efeito do tratamento.

No estudo de Katkhouda et al. está descrito que a randomização utilizou números aleatórios gerados por um computador, os quais foram escritos sobre cartão fechado em um envelope opaco. Isso mostra que o ensaio clínico foi randomizado, mas não garante o sigilo da alocação, já que não se especifica quem teve acesso aos números aleatórios antes de serem lacrados.

Todos os doentes que entraram no estudo foram considerados nos resultados e conclusões?

Os leitores de um artigo de terapia devem observar se todos os pacientes alocados inicialmente foram analisados nos resultados. Nos ensaios clínicos podem ocorrer perdas – pacientes saem do estudo – ou alguns pacientes podem ser excluídos pelos investigadores em virtude de violações do protocolo. Esses fatos precisam ser bem descritos, a fim de que o leitor ateste que isso ocorreu ao acaso e não de maneira sistemática. Temos que considerar ainda que as perdas podem estar relacionadas à morbidade do tratamento.

No artigo sobre as apendicectomias, todos os 134 pacientes randomizados para a apendicectomia aberta foram incluídos na análise. No entanto, dos 134 randomizados para a apendicectomia laparoscópica, somente 113 foram analisados: 11 pacientes foram excluídos (dez se recusaram a submeter-se ao procedimento laparoscópico e uma mulher estava grávida) e em dez pacientes houve perda de dados. É estranho que todos as dez ocorrências de perda de dados tenham acontecido no procedimento laparoscópico: o fato sugere alguma diferença sistemática na coleta ou registro de dados entre as duas intervenções.

Um nível de perdas elevado é, em geral, um indicador de cuidado inadequado no desenho e/ou condução do estudo. Mas mesmo perdas pequenas podem ser inaceitáveis de acordo com a diferença entre as taxas de evento nos grupos experimental e controle. No estudo em questão, os dez pacientes perdidos poderiam ter influência considerável no resultado, dada a pequena diferença dos desfechos entre as intervenções testadas.

O seguimento foi completo?

Se os pacientes não são avaliados em um tempo definido após a intervenção ou a duração do seguimento é inadequada, podemos ter uma visão incompleta de todos os desfechos clínicos possíveis. Os pacientes podem apresentar complicações em várias épocas após o tratamento, logo, o seguimento precisa ser rigoroso e consistente.

O seguimento do estudo foi de duas semanas, o que permitiu avaliar apenas as complicações precoces após a alta hospitalar, porém, permitindo mais facilmente analisar a atividade e a qualidade de vida da maioria dos pacientes.

Os doentes foram analisados nos grupos em que foram randomizados (princípio da “intenção de tratamento”)?

Não é infreqüente que os sujeitos de pesquisa, randomizados para um determinado tratamento, não sejam submetidos a nenhuma terapia ou sejam submetidos ao tratamento do outro braço do estudo. Há três formas de se lidar com esse cruzamento entre os grupos: exclusão dos pacientes

que violaram o protocolo, análise de acordo com a intervenção realmente realizada (análise por protocolo) e análise de acordo com a intervenção à qual os pacientes foram alocados pelo processo de randomização, não importando se a receberam ou não (análise por intenção de tratamento).

Há duas razões principais para se preferir a análise por intenção de tratamento às outras alternativas. Em primeiro lugar, o princípio da intenção de tratamento é consistente com o da randomização e preserva o equilíbrio dos fatores de confusão conhecidos e desconhecidos entre os grupos. O outro motivo é que, na análise por intenção de tratamento, o viés introduzido pelo cruzamento tem direção conhecida: a mistura das intervenções com a manutenção dos grupos originais tende a diluir as diferenças entre os braços do estudo, resultando em estimativas mais conservadoras (exceto, é claro, nos ensaios de equivalência, nos quais a redução das diferenças produz estimativas menos conservadoras).

No artigo que estamos discutindo, os pacientes que tiveram laparoscopias convertidas para laparotomia foram corretamente analisados nos seus grupos originais. No entanto, dez pacientes que foram randomizados para a laparoscopia se recusaram a ser submetidos a esse procedimento e foram excluídos da análise. Levando-se em conta que indivíduos não cooperativos habitualmente apresentam resultados piores que os demais, a exclusão desses dez pacientes tende a favorecer de forma espúria a intervenção laparoscópica.

Houve mascaramento no tratamento?

Nos ensaios clínicos que estudam fármacos, os sujeitos de pesquisa, os investigadores e os avaliadores são geralmente “mascarados” em relação ao tratamento que os pacientes recebem. O mascaramento geralmente não é exequível nos ensaios cirúrgicos. Para minimizar o viés, quando o mascaramento não é possível, a avaliação da terapia cirúrgica deve ser feita por um grupo de profissionais separados, sem interesse nos desfechos analisados.

No estudo em tela, houve mascaramento por meio de três curativos e uma faixa abdominal, de modo que a natureza da intervenção era desconhecida do paciente, da enfermagem, da equipe médica e dos avaliadores independentes.

Os grupos foram similares no início do estudo?

Em um grande ensaio clínico randomizado, a distribuição dos pacientes com seus vários fatores prognósticos deve estar balanceada nos grupos de tratamento e controle. No entanto, nos ensaios menores, há sempre o risco de se comparar, por acaso, dois grupos não similares de pacientes. Assim, os leitores precisam estar seguros de que a dessemelhança ao acaso entre os dois grupos não tenha ocorrido.

No artigo de Katkhouda et al., não houve diferenças entre os grupos quanto à idade e à leucometria. A proporção de homens foi maior na apendicectomia a céu aberto (78%) do que na laparoscópica (69%), o que poderia influir na estimativa de pelo menos um dos desfechos (dor pós-operatória).

Exceto a intervenção, os grupos foram tratados igualmente?

O leitor precisa estar também atento se houve diferenças, exceto a intervenção cirúrgica, no tratamento dos dois grupos. Essas diferenças são denominadas co-intervenções. Por exemplo, quando se comparam as operações minimamente invasivas com as operações a céu aberto, se houver analgesia peridural em apenas um dos braços do estudo, seria uma co-intervenção importante que certamente enviesaria os resultados.

No artigo analisado, os pacientes foram tratados identicamente, exceto pelas intervenções experimentais.

QUAIS SÃO OS RESULTADOS DO ESTUDO?

Qual a magnitude do efeito do tratamento?

Os ensaios cirúrgicos randomizados comumente medem desfechos ou eventos dicotômicos, tais como morte, recorrência do câncer e complicações cirúrgicas. Esses eventos dicotômicos ou acontecem ou não e, assim, o artigo reporta geralmente a proporção dos pacientes que tiveram o evento de interesse. A análise e a apresentação dos resultados se transformam em um processo de comparar proporções.

Essas comparações podem ser feitas e expressas de duas formas básicas:

1. Diferença de risco: diferença entre os riscos dos grupos experimental e controle.
2. Risco relativo: o risco de eventos no grupo experimental dividido pelo risco no grupo controle. Riscos relativos abaixo de um indicam que a intervenção é protetora; se acima de um, a intervenção aumenta o risco.

As duas formas são complementares: o risco relativo é uma medida de associação enquanto a diferença de risco é uma medida de impacto. Por exemplo: se o risco basal de eventos é de 2%, um tratamento com risco relativo de 0,25 o reduziria para 0,5%. Se em uma outra doença, o risco de eventos é de 50%, um tratamento com risco relativo de 0,8 reduziria o risco para 40%. Portanto, tratamentos muito eficazes (RR=0,25) podem ter impacto insignificante (1,5%) enquanto tratamentos com eficácia modesta (RR=0,8) podem ter um grande impacto (10%) sobre o desfecho, dependendo do risco basal de eventos.

No artigo sobre a apendicectomia, as complicações pós-operatórias existiram em 18,6% dos pacientes submetidos à apendicectomia laparoscópica e em 17,2% dos pacientes submetidos à apendicectomia a céu aberto. A diferença de risco foi de 0,014 (0,186 – 0,172). Em outras palavras, a apendicectomia laparoscópica resultou em um aumento de 1,4 pontos percentuais das complicações pós-operatórias. O risco relativo de complicações pós-operatórias com a apendicectomia laparoscópica foi de 1,08 (0,186/0,172).

No ensaio também foram analisados desfechos numéricos como a avaliação da dor pós-operatória por uma escala analógica visual (que também não diferiu entre as duas intervenções).

Qual a precisão do efeito do tratamento?

Os ensaios cirúrgicos randomizados estudam um tratamento em um pequeno grupo de pacientes e fornecem uma estimativa do efeito do tratamento que seja aplicável a uma população maior de pacientes.

Isso levanta uma importante questão: quão precisa é a estimativa do efeito do tratamento? Os pesquisadores podem indicar essa precisão de várias maneiras. Uma delas é fornecer o tradicional valor de “p”, ou seja, a probabilidade de se obter a diferença encontrada entre as intervenções na amostra estudada, dado que essa diferença inexistia na população de onde se retirou a amostra.

Por exemplo, um “p” de 0,01 (1%) significa que, se a intervenção experimental é inefetiva com relação à intervenção controle (isto é, a diferença entre suas medidas de efeito é zero), a probabilidade de obtermos a diferença encontrada na amostra estudada é de 1%. Como essa probabilidade é baixa, é mais verossímil que exista realmente alguma diferença entre os tratamentos.

Outro modo de mostrar a precisão da estimativa do efeito é calcular o intervalo de confiança. Se, de uma população onde uma variável tem o valor “a”, retirarmos um número infinito de amostras e em cada uma delas estimarmos o valor da variável (chamaremos estas estimativas de “b”) e calcularmos intervalos de confiança de x% para essas estimativas, x% dos intervalos de “b” incluirão “a”. Por exemplo, se em uma população com altura média de 170 cm, retirarmos um número muito grande de amostras e calcularmos intervalos de confiança de 99% para a média da altura em cada amostra, 99% destes intervalos incluirão 170 cm.

No artigo, o risco relativo de complicações pós-operatórias com a apendicectomia laparoscópica é de 1,08, com um intervalo de confiança de 95% de 0,63 a 1,85. Portanto, temos 95% de confiança de que, na população de onde a amostra do estudo foi retirada, o risco relativo se encontre em algum ponto entre 0,63 e 1,85, ou seja, não é possível inferir se o tratamento experimental é benéfico, indiferente ou maléfico para o paciente.

Os intervalos de confiança e os valores de “p” são essencialmente formas diferentes de expressar o mesmo conceito: se o “p” de uma medida de efeito for inferior a 5%, o intervalo de confiança de 95% não passará pelo valor nulo da medida (zero, se for uma diferença, um se for uma razão). Se o “p” for maior do que 10%, o intervalo de confiança de 90% necessariamente incluirá o valor nulo da medida, e assim por diante. No entanto, a informação veiculada pelo intervalo de confiança é mais facilmente interpretável. Por exemplo, vamos supor duas situações:

a) risco relativo de 0,3 com intervalo de confiança de 95% de 0,12 a 0,98;

b) risco relativo de 0,3 com intervalo de confiança de 95% de 0,25 a 0,38.

Em ambos os casos o valor de “p” será inferior a 5% (nenhum dos intervalos de confiança de 95% inclui 1, o valor nulo do risco relativo). No entanto, no primeiro caso é verossímil que o efeito do tratamento seja clinicamente irrelevante (RR=0,98) enquanto no segundo caso estamos razoavelmente confiantes de que o tratamento seja efetivo.

OS RESULTADOS IRÃO AJUDAR NO CUIDADO DOS MEUS DOENTES?

Os paciente do estudo são semelhantes aos meus pacientes?

Os cirurgiões devem fazer-se essa pergunta antes de incorporar as recomendações de um ensaio clínico as suas práticas diárias. Se um ensaio randomizado tem critérios muito rígidos de inclusão dos pacientes, os resultados podem não ser aplicados a uma população maior de pacientes cirúrgicos. A idade e a capacidade funcional são critérios de inclusão que podem causar problemas para a generalização dos resultados. Por exemplo, se um ensaio randomizado só recruta pacientes jovens com boa capacidade funcional, suas recomendações para uma estratégia de terapia com muitos efeitos colaterais pode não ser aplicável a pacientes idosos ou com má capacidade funcional.

Os cirurgiões precisam examinar os critérios de inclusão para verificar se os resultados do estudo podem ser aplicados aos seus pacientes.

No artigo de Katkhouda et al., os critérios de inclusão para se considerar um paciente com apendicite foram: idade igual ou superior a 16 anos; história de dor no quadrante inferior direito ou dor periumbilical com migração para o quadrante inferior direito com náuseas e/ou vômitos; febre acima de 38° C e/ou leucócitos acima de 10.000 células/mm³; dor e defesa à palpação do quadrante inferior direito. Os resultados podem assim ser aplicados ao paciente do nosso cenário clínico.

Todos os desfechos clínicos importantes foram avaliados?

Antes de aceitar que um tratamento seja superior a um outro, os cirurgiões precisam questionar a importância clínica das medidas dos principais desfechos. Por exemplo, os ensaios randomizados comparando operação a céu aberto e por laparoscopia utilizam frequentemente o tempo de internação como uma medida de um desfecho principal. Embora o tempo de internação seja importante, não pode ser o desfecho clínico mais relevante para os cirurgiões e para os pacientes.

No artigo analisado, o tempo de anestesia e o tempo operatório pele a pele foram comparados. Esses desfechos não são tão importantes quanto outros como as complicações pós-operatórias. Um desfecho importante que não foi incluído entre as complicações foi a necessidade de conversão da laparoscopia para a cirurgia a céu aberto.

A minha experiência cirúrgica é semelhante a dos cirurgiões do estudo?

Os cirurgiões devem considerar as suas próprias habilidades antes de aplicar os achados de uma pesquisa a sua prática. Eles precisam perguntar-se se são tão proficientes tecnicamente quanto a equipe dos cirurgiões investigadores. A proficiência da equipe cirúrgica com uma determinada operação é obviamente um componente muito importante na avaliação da efetividade do tratamento.

Se a evidência publicada favorece, sem sombra de dúvida, uma operação que um cirurgião não realiza ou não realiza bem, esse cirurgião tem três escolhas: realizar outra operação, referir o paciente a um colega ou procurar treinamento naquela operação.

A falta de experiência cirúrgica vai enviesar um ensaio randomizado a favor do procedimento mais simples entre duas operações. Por outro lado, um ensaio randomizado, envolvendo uma complexa operação realizada por cirurgiões extremamente habilitados, pode não ser aplicável a uma comunidade cirúrgica maior.

RESOLUÇÃO DO CENÁRIO CLÍNICO

O estudo não forneceu evidências da superioridade da apendicectomia laparoscópica sobre a cirurgia a céu abert-

to; não houve diferença entre os grupos quanto à ocorrência de complicações ou à intensidade da dor pós-operatória. Na verdade, como alguns vieses do estudo (exclusão das violações do protocolo, não inclusão da conversão da laparoscopia para laparotomia) tenderam a favorecer o grupo laparoscopia, pode-se mesmo argumentar que a cirurgia a céu aberto poderia mostrar alguma superioridade se o estudo fosse melhor desenhado. De qualquer forma, como os resultados do estudo foram inconclusivos, o cirurgião-chefe e o residente concordaram que não havia motivos para modificar a sua prática habitual.

ABSTRACT

Surgeons are told that they need not only read journal articles, but also understand them and make a critical appraisal of their validity. They offer better care if they are able to appraise critically the original literature and apply the results to their practice. This is a first article of a series – Evidence-Based Surgical Practices – which focus on issues of Evidence-Based Surgery.

Key words: *Evidence-Based Medicine – Education, Medical, Continuing - Randomized Controlled Trials - Appendectomy – Laparoscopy.*

REFERÊNCIAS

1. Katkhouda N, Mason RJ, Towfigh S, Gevorgyan A, Essani R. Laparoscopic Versus Open Appendectomy: A Prospective Randomized Double-Blind Study. *Ann Surg.* 2005 Sep;242(3):439-50.
2. Urschel JD, Goldsmith CH, Tandan VR, Miller JD; for the Evidence-Based Surgery Working Group. Users' guide to evidence-based surgery: how to use an article evaluating surgical interventions. *Can J Surg.* 2001 Apr;44(2):95-100.
3. **Grupo de Estudos de Revisão Sistemática do Rio de Janeiro:** Alexander Magno Cordeiro, Programa de Pós-Graduação em Medicina, UFRJ; Bráulio Santos Jr., Coordenação de Ensino e Pesquisa (Unidade de Epidemiologia Clínica) do Instituto Nacional de Cardiologia, MS; Carlos Alberto Guimarães, Departamento de Cirurgia da Faculdade de Medicina, UFRJ; Fernando César David Silva, Hospital Geral de Jacarepaguá, MS; Giovanni Antonio Marsico, Instituto de Doenças do Tórax, UFRJ; Glória Maria de Oliveira, Hospital do Andaraí, RJ; José Eduardo Ferreira Manso, Departamento de Cirurgia da Faculdade de Medicina, UFRJ; José Henrique Cunha Figueiredo, Psicologia Médica e Saúde Mental do Hospital Universitário Clementino Fraga Filho, UFRJ; Juan Miguel Renteria, Departamento de Cirurgia da Faculdade de Medicina, UFRJ; Márcio Maranhão Costa, Hospital de Força Aérea do Galeão, MD; Sonia Maria Ramos de Vasconcelos, Núcleo de Catálise do Programa de Engenharia Química da COPPE, UFRJ.

Endereço eletrônico: GERS-Rio@yahoo.com