

A arte de ler um artigo científico

Por: Luis Claudio Lemos Correia

Neste ano planejo uma série de postagens que denominarei de “*bússolas*” para a interpretação crítica de trabalhos científicos. Normalmente discuto interpretação científica no contexto de algum artigo ou assunto controverso. Pretendo agora revisar estes conceitos de forma estruturada, em postagens definidas por tipos de estudo e tipos de resultado.

Não gosto muito da conotação burocrática e simplista dos termos “guia” ou “checklist” de análise de uma evidência científica. A leitura de um artigo requer não apenas a “checagem” de detalhes. *Checklist* é para avião que vai decolar ou cirurgia que vai começar. Não basta *checklist* na interpretação científica. É necessária visão do todo, envolvimento com o assunto, percepção intuitiva trazida pela experiência, capacidade de reflexão, um equilíbrio entre ceticismo e pragmatismo. É como ler um livro ou assistir a um filme, nossa interpretação não se resume a um *checklist*.

Imaginem um crítico de cinema que analisa filmes pela forma estruturada de um *checklist*. Não chegaria ao seu objetivo. Por outro lado, na análise de uma evidência científica não podemos ter liberdade artística para interpretar como quisermos pois isso seria um convite ao viés de confirmação de crenças ou valorização seletiva de resultados. Precisamos controlar nossos vieses pessoais durante a leitura; precisamos saber como procurar componentes ilusórios dos trabalhos, perceber quando ruídos se confundem com sinais; precisamos seguir princípios científicos. Estes princípios devem nos nortear como uma *bússola*.

A leitura de um artigo não é um processo passivo. Mais do que uma leitura, é um processo de pensamento baseado em dados. Pensamos primeiro, depois procuramos a resposta no trabalho. E assim vamos construindo uma ideia mental do quanto aquela evidência deve influenciar nosso pensamento.

Com estas colocações não quero transmitir a ideia de que a interpretação de uma evidência é difícil ou laboriosa. Pelo contrário, a leitura de um artigo deve ser atividade leve, natural, e até divertida.

Como tenho proposto ao longo de múltiplas postagens, **a leitura do artigo começa antes do artigo**. Começa dentro de nós, com a procura de vieses pessoais que atrapalhem uma análise racional. Precisamos nos preparar mentalmente para a análise de qualquer evidência externa, pois esta pode vir de encontro a nossas perspectivas internas. Essa preparação passa pelo questionamento de quais são nossas evidências internas (crenças) a respeito do assunto que está exposto no título do trabalho de interesse. Esta proposta de aquecimento mental foi tema de [postagem prévia no Blog](#).

Prosseguindo com a reflexão pré-artigo, precisamos saber do que o artigo trata. Para isso teremos um contato inicial com o conteúdo do trabalho, pela leitura de apenas duas sentenças do resumo: objetivo e conclusão. Esses são os pilares de nossa análise. O objetivo mostra a hipótese testada ou a realidade descrita. A conclusão reflete o que o autor quer que você acredite como mensagem final do artigo. Neste momento exercitaremos o ceticismo, pois **a leitura do artigo deve ser uma tentativa de refutar a conclusão**. Uma rigorosa análise de ruído (viés, acaso) *versus* sinal (verdade). Uma baixa relação sinal/ruído trará dúvida quanto à veracidade da conclusão do trabalho.

A leitura da conclusão também nos permite um importante *insight* do quanto o autor é enviesado. Refiro-me ao diagnóstico de *spin*, quando o autor reconhece um resultado primário negativo, mas logo em seguida gera uma tendência positiva a partir da apresentação de um resultado secundário. Por exemplo, “na análise global não houve diferença entre os grupos quanto à fração de ejeção do ventrículo esquerdo [desfecho primário], porém houve melhora da troponina [desfecho secundário]”. *Spin* é um forte marcador de tendenciosidade. A detecção desde o início nos fará mais atentos.

Ainda na leitura pré-artigo, vale apenas observar se (1) o estudo é uma iniciativa do fabricante do produto testado, (2) se é realizado por grupo independente, porém recebe ajuda de custo do fabricante ou (3) se o trabalho não possui nenhuma relação de financiamento com o fabricante. Essa numeração representa uma escala do grau de envolvimento do

fabricante com o trabalho científico. Teoricamente, quanto menor o envolvimento, mais isento seria o estudo. Recente trabalho publicado no [Annals of Internal Medicine](#) demonstrou que quanto maior a vinculação com indústria farmacêutica, mais utilização de métodos estatísticos inadequados com objetivo de obter o resultado desejado.

Devemos também investigar se os autores possuem conflito de interesses, seja relação com indústria, seja conflitos assistenciais ou intelectuais. Conflito de interesses assistencial é aquele desejo do autor em provar uma hipótese que fortaleça sua atividade profissional. Esse pode ser um conflito maior do que o daqueles que recebem um dinheirinho da indústria. Por exemplo, trabalhos que testam acurácia de métodos diagnósticos ou eficácia de tratamentos podem ser feitos por profissionais cuja atividade principal é ligada ao método ou tratamento. Já conflito de interesses intelectual ocorre quando um autor faz parte de uma legião de estudiosos e crentes a respeito de um assunto. Por exemplo, no caso do [escândalo do estudo PREDIMED](#), os autores fazem parte de uma legião de crentes na dieta do mediterrâneo.

Não proponho aqui invalidar o estudo se este tiver fonte financiadora interessada ou conflito de interesses. Mas devemos aumentar nosso grau de atenção. [Revisão sistemática da Cochrane](#) mostra que trabalhos com conflito de interesses apresentam maior probabilidade de resultados positivos do que trabalhos sem financiamento do fabricante, um viés que não deveria existir. Interessante notar que, pela avaliação tradicional de risco de viés da Cochrane, o estudo não detectou maior risco de viés nos trabalhos com conflito de interesses. Isto ocorre porque estudos financiados pela indústria são "bem feitos" de acordo com uma avaliação metodológica superficial: seguem a metodologia padrão (tamanho amostral, randomização, cego). Os ruídos destes trabalhos usualmente são detectados pelo tipo de leitura que propomos, algo que vai além de um *checklist* presente em revisões sistemáticas.

Portanto, até aqui analisamos ***primeiro a nossa tendenciosidade como leitor, depois a tendenciosidade do autor.***

Em se tratando de um trabalho analítico, o próximo passo é a análise da probabilidade pré-teste da hipótese testada. O experimento científico é o teste, enquanto o pré-teste é o conhecimento que existe antes do trabalho. Há dois componentes da probabilidade pré-teste: a plausibilidade da hipótese (mecanismo lógico, conhecimento de ciência básica que suporta este mecanismo) e trabalhos prévios que testaram a hipótese.

Esta etapa é importante pois o valor preditivo do resultado do trabalho depende não só do trabalho, como também da probabilidade pré-teste da ideia sugerida pelo trabalho. Ao final da leitura de um artigo, queremos saber qual a probabilidade daquela conclusão ser verdadeira. Isso depende também da qualidade pré-teste da ideia.

Trabalhos cujo resultado propõe ideias sem sentido terão baixo valor preditivo positivo, independente da força da evidência. Assim como trabalhos negativos terão menor valor preditivo negativo quando a ideia for muito promissora. Esta é a perspectiva bayesiana que defendo para o processo de análise crítica de uma evidência. ***Uma evidência não deve ser analisada no vácuo, mas sim no contexto da ideia.***

Além de analisar qualidade da hipótese testada, devemos analisar a probabilidade pré-teste do resultado encontrado. Resultado improváveis possuem menor probabilidade pré-teste e menor valor preditivo positivo. ***Precisamos avaliar se estamos diante de um resultado "bom demais para ser verdade".***

A próxima etapa será a procura dos registros prévios do protocolo, com o intuito de saber: o protocolo foi publicado *a priori* (bases de protocolos ou revistas)? O protocolo *a priori* coincide com método publicado no artigo ou houve mudanças *a posteriori*? Respostas “sim” para a primeira pergunta e “não” para a segunda pergunta trarão credibilidade aos resultados. Não é frequente com grosseiras mudanças de protocolo, como exemplificadas por duas postagens prévias ([ISCHEMIA](#) E [SCOT-HEART](#)). Uma [recente varredura](#) nos registros do [clinicaltrials.org](#) demonstrou que 30% dos trabalhos modificam o desfecho primário definido *a priori*. Estas mudanças

reduzem substancialmente o valor preditivo do estudo, ou seja, sua credibilidade.

Chegamos finalmente à leitura do artigo. Agora, procuraremos por erros sistemáticos (vieses) e aleatórios (acaso) que promovam ruídos e resultados ilusórios. Esta leitura do artigo se resume a “métodos” e “resultados”.

Um parêntese: “introdução” e “discussão” não precisam ser lidas. Eu quase nunca leio. Não que seja ruim ler, é que minha preguiça não me deixa perder tempo com o desnecessário. Mas se lerem, não simplesmente acreditem (fé) nos argumentos do autor para a ideia do trabalho (introdução), nem na interpretação dos resultados (discussão). ***Não leiam acreditando, leiam duvidando. Mesmo que ao final concordem.***

Temos que saber o que procurar em métodos e resultados. O que procurar depende do objetivo do estudo: eficácia de tratamento, acurácia diagnóstica, acurácia prognóstica. A análise também depende de se o estudo for positivo ou negativo. ***A direção da leitura do artigo é diferente a depender da direção do resultado.*** Coisas diferentes causam falsos-negativos ou falsos-positivos.

Outra questão interessante é o diferente nível de atenção que devemos ter em métodos e resultados. ***Em métodos avaliamos predominantemente erros sistemáticos. Em resultados avaliamos erros aleatórios.*** P

Detectamos potenciais vieses quando notamos falhas no desenho do estudo descritos nos métodos.

Mas é nos resultados que veremos o grau de significância estatística ou a imprecisão dos intervalos de confiança, aspectos que representam o risco de erro aleatório. Nos resultados que perceberemos se a amostra cumpriu com as premissas do cálculo do tamanho amostral descrito nos métodos. Por exemplo, o cálculo (métodos) está perfeito e supõe uma certa incidência do desfecho, porém veremos em resultados que esta premissa não foi obedecida. Portanto, este seria um trabalho com maior risco de erro aleatório. Também será nos resultados que veremos desvios de integridade científica que predisõem a erro aleatório: o estudo foi interrompido precocemente pois o resultado aparentou positivo

(truncado), o autor começa a valorizar descrição de desfechos secundários ou análise de subgrupo (*reporting bias*).

Depois de toda essa análise, na ausência de mudanças de protocolo e quando observamos baixo risco de viés ou acaso, concluiremos que temos um estudo com alto nível de evidência.

Por fim, a resposta final será bayesiana: devemos calcular a probabilidade pós-teste. Quando temos uma hipótese moderadamente provável *a priori*, um estudo de baixo risco de viés e erro aleatório é suficiente para gerar um valor preditivo positivo alto. No caso de hipótese pouco provável, um bom estudo terá valor preditivo positivo moderado e requeremos um segundo estudo para confirmar a ideia (reprodutibilidade).

No caso do estudo negativo, quando desistir da ideia? Normalmente, se a hipótese não for promissora, temos uma tendência a julgar que não vale mais a pena insistir na ideia. Mas se a ideia for muito provável, precisaremos de mais de um estudo negativo para nos fazer desistir do conceito.

Ciência não prova inexistência, ciência prova existência de conceitos. Como é impossível provar que algo não existe, o que está em jogo na análise de um estudo negativo é se este deve nos fazer desistir de provar a ideia ou se devemos continuar tentando provar. Devemos avaliar quando estamos sendo teimosos ou quando abandonar a ideia. Isso depende do valor preditivo negativo do estudo, que depende da probabilidade pré-teste e da relação sinal/ruído do estudo.

Resumindo, devemos avaliar:

- Tendenciosidade do leitor (minhas crenças internas)
- Tendenciosidade intrínseca do autor (spin)
- Tendenciosidade extrínseca do autor (conflito de interesses)
- Probabilidade pré-teste da ideia
- Mudanças de protocolo a posteriori

- Relação sinal/ruído do estudo (métodos e resultados)
- E finalmente, calcularemos mentalmente o valor preditivo do trabalho

Após esta introdução genérica de como ler artigos científicos, trabalharei em uma série de postagens ao longo desse ano, com as especificidades de cada tipo de análise crítica:

- Como avaliar probabilidade pré-teste de uma ideia
- Como julgar se premissas do cálculo amostral foram obedecidas
- Como julgar trabalhos negativos sobre condutas médicas
- Como julgar trabalhos positivos sobre condutas médicas
- Como julgar trabalhos de acurácia diagnóstica
- Como julgar estudos de coorte para marcadores prognósticos
- Como julgar estudos de coortes para causalidade
- Como julgar estudos caso-controle para causalidade

Bem, não será exatamente nessa ordem, pois a inspiração vem de diferentes fontes, de trabalhos que surgem para ser analisados, de aulas que dou, de perguntas desafiadoras de alunos, de discussões interessantes com colegas e de casos clínicos que um provocam reflexões (*skin in the game*, como diria Taleb). Textos assim ficam bons mesmo quando surgem do sentimento de nossa “pele no jogo”, de nossa experiência clínica.

Portanto, não esperem nada previsível, pois esse blog é imprevisível, como já devem ter percebido.

Denominarei estes posts de “**bússolas**”, pois não há mapas precisos ou GPSs quando se trata do verdadeiro pensamento científico. Nossas “bússolas de análise crítica” estimularão um pensamento reflexivo, livre de dogmas, repleto de perguntas e com poucas certezas.

O que avaliar?

Onde avaliar?

