

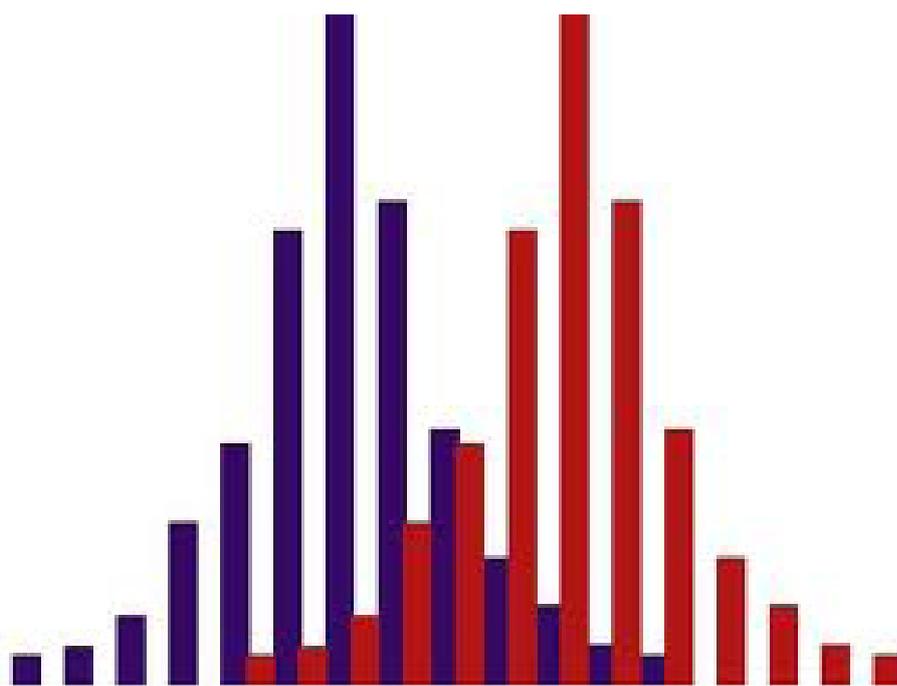
# PRÁTICA BASEADA EM EVIDÊNCIA EM SAÚDE MENTAL NA INFÂNCIA E ADOLESCÊNCIA

**John Hamilton & Füsun Çuhadaroğlu-Çetin**

Edição em Português

Tradutora: Ana Teresa Prata

Editora: Ana Angélica de Araújo Pontes  
(in memoriam)



John Hamilton MD, MSc

Diretor clínico, The Permanente Medical Group, Inc., Sacramento, California, USA

Conflito de interesses: não relatado.

Füsun Çuhadaroğlu-Çetin MD

Professor of Psiquiatria da infância e adolescência, Hacettepe University Medical School, Ankara, Turkey

Conflito de interesses: não relatado.

Agradecimentos: Professor Eric Youngstrom, Universidade da Carolina do Norte, Professor Peter Szatmari, Universidade McMaster, e Dr Eric Daleiden, PracticeWise LLC, foram revisores de grande ajuda na correção dos rascunhos anteriores.

Esta publicação é dirigida para profissionais em formação ou especialistas em saúde mental e não para o público em geral. As opiniões expressas são da responsabilidade dos autores e não representam, necessariamente, as visões do Editor ou da IACAPAP. Esta publicação procura descrever os melhores tratamentos e práticas baseados em evidência científica disponível no momento em que foi escrita e que foi avaliada pelos autores podendo mudar como resultado de novas pesquisas. Os leitores precisam aplicar este conhecimento em pacientes de acordo com as diretrizes e leis de prática do seu país. Algumas medicações podem não estar disponíveis em alguns países e os leitores devem consultar informações específicas sobre os fármacos já que nem todas as dosagens e efeitos indesejáveis são mencionados. Organizações, publicações e endereços electrónicos são citados ou associados para ilustrar conteúdos ou como uma fonte mais aprofundada de informações. Isto não significa que os autores, o Editor ou IACAPAP endossam o seu conteúdo ou recomendações, o que deve ser criticamente avaliado pelo leitor. Os endereços electrónicos podem também mudar ou deixar de existir.

©IACAPAP 2019. Esta é uma publicação de livre acesso sob a Creative Commons Attribution Non-commercial License. O seu uso, distribuição e reprodução em qualquer meio são permitidos sem permissão prévia desde que o trabalho original seja adequadamente citado e que o uso seja não-comercial. Envie comentários sobre este livro ou capítulo para [jmreyATbigpond.net.au](mailto:jmreyATbigpond.net.au)

Citação Sugerida: Hamilton J, Çuhadaroğlu-Çetin F. Evidence-based practice in child and adolescent mental health. In Rey JM, Martin A (eds), *JM Rey's IACAPAP e-Textbook of Child and Adolescent Mental Health*. (edição em Português; Dias Silva F, ed) Geneva: International Association for Child and Adolescent Psychiatry and Allied Professions 2019.

Este capítulo pretende mostrar como usar os dados de estudos com menos vieses e erros aleatórios na tomada de decisões clínicas corretas, no sentido de ajudar os jovens que procuram tratamento clínico para problemas de saúde mental. Concretamente, este capítulo pretende mostrar como:

- Considerar vieses e erros aleatórios como indesejados.
- Compreender os conceitos estatísticos mais relevantes para o alinhamento da prática clínica com a boa evidência.
- Procurar bases de dados online acessíveis gratuitamente para qualquer indivíduo com ligação à internet.
- Usar técnicas de entrevista diagnóstica adequadas para verificar se o jovem cumpre ou não critérios para síndromes bem definidas.
- Encontrar os tratamentos mais prováveis de serem eficazes para essas síndromes.

Em resumo, este capítulo explica aos profissionais de saúde como alinhar aquilo que fazem com *aquilo que funciona*: medidas que irão promover uma proporção maior de jovens com melhoria significativa do seu funcionamento de forma mais substancial e mais rápida do que com medidas alternativas. Mas para chegar à verdade sobre aquilo que funciona, primeiro é necessário lutar contra dois grandes inimigos da verdade: *viés e erros*.

## VIÉS E ERROS ALEATÓRIOS

### O problema do viés

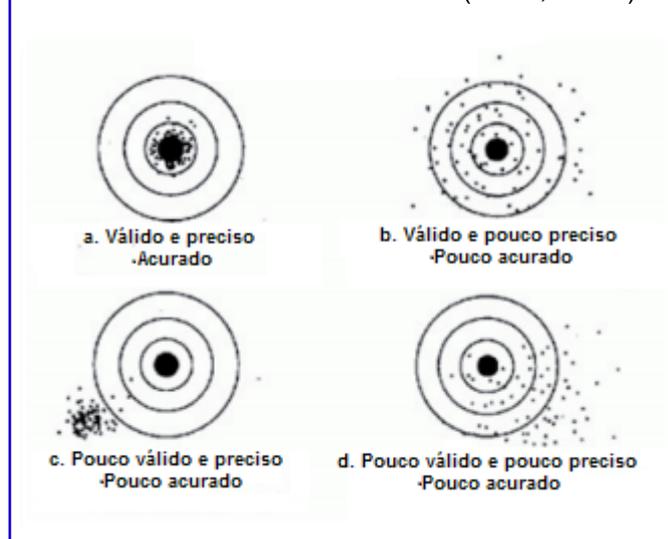
O viés interfere com a validade – o grau com que o estudo consegue medir aquilo a que se propõe medir – ao criar desvios sistemáticos (não aleatórios) da verdade subjacente (Guyatt & Rennie, 2012). O viés é diferente dos *erros aleatórios*. Os erros aleatórios diminuem ao aumentar a amostra, enquanto o viés pode permanecer numa amostra de qualquer tamanho. Considere os alvos (a) a (d) na Figura A.6.1 onde o alvo do centro marca o verdadeiro efeito. No estudo (a), existe pouco ou nenhum desvio do verdadeiro efeito (pouco ou nenhum viés), assim

A responsabilidade final pelo manuscrito, incluindo os erros, é de responsabilidade dos autores. O autor pode ser contactado através do email [john.hamilton@kp.org](mailto:john.hamilton@kp.org)

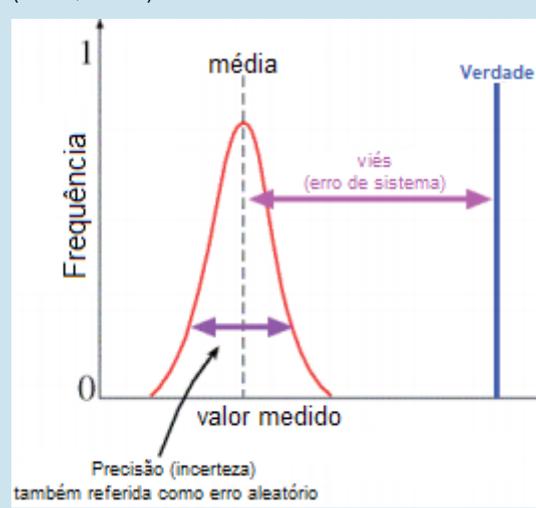
### Erro

Todas as experiências ou decisões clínicas estão sujeitas a erros aleatórios, erros sistemáticos ou ambos. *Erros aleatórios* são devido a alterações desconhecidas ou inesperadas nos instrumentos de medição ou nas condições do ambiente. *Erros sistemáticos* (ou viés), por outro lado, são imprecisões que são consistentes em qualquer direção. Podem ocorrer por algum problema no instrumento de medição (p.e., uma escala de medição, critérios de diagnóstico) ou porque o instrumento é usado de forma errada pelo investigador (p.e., um médico que sistematicamente sobrestima ou subestima o risco de suicídio).

**Figura A.6.1 Exemplos de viés e de erros aleatórios em amostras de dados (Kabai, 2011a).**



**Figura A.6.2 Representação esquemática de erros aleatórios e viés (Kabai, 2011b).**



como pouca dispersão (erros aleatórios). No estudo (b), também existe pouco viés, mas muitos erros aleatórios e no (c) muito viés e poucos erros aleatórios. No (d) existe muito viés e muitos erros aleatórios.

### Encontrar evidência com pouco viés e pouco erro aleatório

A evidência mais útil na orientação de decisões clínicas é a evidência *com o mínimo de viés e de erros aleatórios*. A Figura 6.A.2 mostra uma representação esquemática de erro aleatório e viés. A Tabela A.6.1 mostra fontes comuns de viés e estratégias para reduzir cada uma delas e a Tabela A.6.2 mostra fontes comuns de erros aleatórios e estratégias para sua redução.

**Tabela A.6.1 Fontes comuns de viés e estratégias para sua redução.**

Fontes de viés	Estratégias para reduzir a fonte de viés
<ul style="list-style-type: none"> <li>Doentes do grupo de tratamento e controle com prognósticos diferentes</li> </ul>	<ul style="list-style-type: none"> <li>Alocação randomizada do tratamento ou controle</li> </ul>
<ul style="list-style-type: none"> <li>Efeitos placebo do tratamento</li> </ul>	<ul style="list-style-type: none"> <li>Doentes cegos para o tratamento ativo ou controle</li> </ul>
<ul style="list-style-type: none"> <li>Cuidadores acrescentam outros tratamentos</li> </ul>	<ul style="list-style-type: none"> <li>Cuidadores cegos para o tratamento ou controle</li> </ul>
<ul style="list-style-type: none"> <li>Avaliadores dos resultados viesados</li> </ul>	<ul style="list-style-type: none"> <li>Avaliadores dos resultados cegos para o status de tratamento ou controle dos casos que avaliam</li> </ul>
<ul style="list-style-type: none"> <li>Perda de seguimento</li> </ul>	<ul style="list-style-type: none"> <li>Seguimento de todos ou de alta proporção de casos incluídos no estudo</li> </ul>

**Tabela A.6.2 Duas fontes comuns de erros aleatórios e estratégias para sua redução**

Fontes de erros aleatórios	Estratégias para sua redução
<ul style="list-style-type: none"> <li>Instrumentos de medição imprecisos</li> </ul>	<ul style="list-style-type: none"> <li>Melhoria da precisão dos instrumentos</li> </ul>
<ul style="list-style-type: none"> <li>Pequeno tamanho da amostra</li> </ul>	<ul style="list-style-type: none"> <li>Aumento do tamanho da amostra</li> </ul>

## CONCEITOS ÚTEIS NA PRÁTICA BASEADA EM EVIDÊNCIA

### Média

Tudo o que precisa são alguns conceitos estatísticos base para compreender muitas ideias chave usadas em abordagens baseadas na evidência. Em primeiro lugar, é preciso lembrar o *normal* ou curva de Gauss, que descreve a distribuição de vários aspectos na natureza, e que pode ser matematicamente derivado usando a teoria probabilística. O valor mais comum é o mesmo que o valor da *média*, representado aqui pela letra grega *mu* ( $\mu$ ).

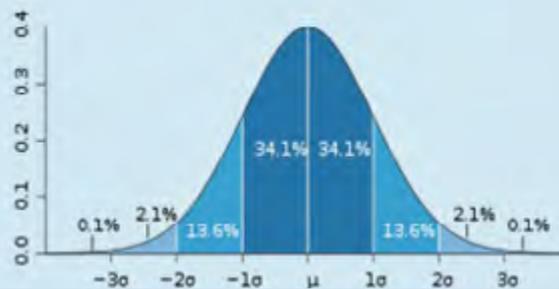
### Desvio padrão (DP)

O DP é uma medida de variabilidade: reflete quanta variação ou “dispersão” existe do valor médio (a média). Um pequeno DP relativo à média indica uma

### Viés

O Viés é um risco sempre presente em investigação, mesmo nos estudos controlados randomizados (ECRs). Por exemplo, resultados positivos são reportados muito mais frequentemente em experiências financiadas pela indústria farmacêutica (78%), comparativamente com aqueles sem patrocínio da indústria (48%), ou aqueles financiados por um concorrente (28%) (Kelly et al, 2006).

**Figura A.6.3** Gráfico da distribuição normal (Gaussiana) com o valor da média representado pela letra grega  $\mu$  e o desvio padrão  $\sigma$ , representado por cada banda colorida. Observe que um pouco mais de um terço (34,1%) da população desta distribuição encontra-se um desvio padrão acima da média; um pouco mais de um terço também (34,1%) se encontra até um desvio padrão abaixo da média. Apenas cerca de 2% da população encontra-se entre 2 ou mais desvios padrão acima da média, e apenas cerca de 2% se encontra entre 2 ou mais desvios padrão abaixo da média (Wikipedia).



distribuição em que os pontos dos dados se encontram muito juntos. Um grande DP relativo à média indica uma distribuição em que os pontos dos dados se encontram espalhados numa extensão muito grande. O DP é muitas vezes descrito como o *valor típico que difere os casos uns dos outros*. Espera-se que dois em cada três casos se encontrem dentro da variação de um ou menos um DP da média.

### Tamanho do efeito

O tamanho do efeito de um trabalho de investigação mede o quão grande é o efeito de um tratamento, relativamente às diferenças entre os casos no geral. É a diferença da média estandardizada entre dois grupos num estudo. Tecnicamente, é calculado pela razão da diferença causada pelo tratamento e o que se chama de desvio padrão *combinado*. *Combinado* refere-se à combinação de valores quer do grupo controle quer do grupo experimental para criar um desvio padrão combinado.

Pode encontrar-se em sites de estatística (e.g., <http://www.leeds.ac.uk/educol/documents/00002182.htm>) a forma de calcular o desvio padrão combinado. O tamanho do efeito coloca a ênfase no aspecto mais importante da intervenção em vez do seu significado estatístico, que é determinado tanto pelo tamanho do efeito como pelo tamanho da amostra.

#### Exemplo de tamanho do efeito

Um estudo (McCracken et al, 2002) pesquisou se um antipsicótico atípico (risperidona) era mais eficaz que o placebo em reduzir a agressividade, birras ou comportamentos autolesivos em crianças com autismo com idades compreendidas entre 5 e 17 anos. Os valores após 8 semanas numa escala de irritabilidade foram de 11.3 (DP 7.4) para o grupo experimental e de 21.9 (DP 9.5) para o grupo controle. Qual o tamanho do efeito?

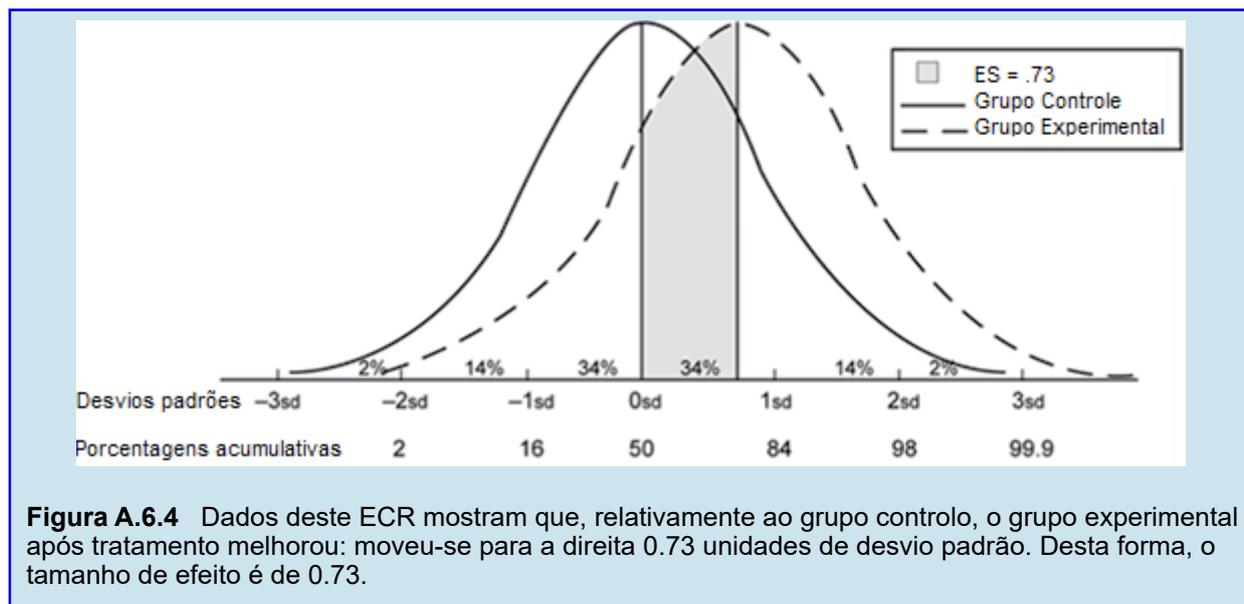
Tamanho de efeito =  $(11.3 - 21.9) / 9.5$ , ou seja, 1.1. Este é um tamanho de efeito grande e altamente significativo. Um tamanho de efeito acima de 0.5 é geralmente considerado clinicamente significativo, mas efeitos menores podem ainda ter significância clínica se o tratamento pode ser facilmente aplicado a grandes populações (p.e., o uso de aspirina na prevenção de infartos agudos do miocárdio, onde o tamanho de efeito é pequeno, mas quando ocorrem em grandes populações, as consequências na saúde pública são significativas).

$$\text{Tamanho de efeito} = \frac{[\text{média do grupo experimental}] - [\text{média do grupo controle}]}{[\text{desvio padrão combinado}]}$$

A Figura A.6.4 mostra uma representação típica de dados de um tratamento experimental, em psiquiatria infantil, com um tamanho de efeito de 0.73. Repare que as duas curvas se sobrepõem consideravelmente.

### Redução absoluta de risco (RAR)

Redução absoluta de risco é uma medida usada para comparar duas alternativas diferentes: quanto é que um tratamento diminui o risco de um resultado negativo



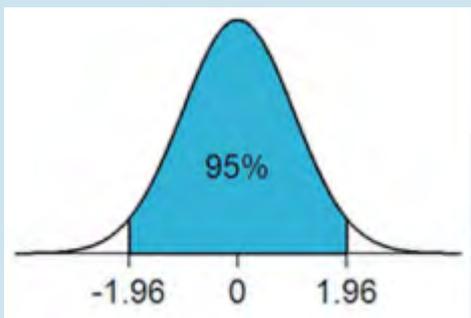
específico (como um jovem deprimido manter-se deprimido) comparativamente a um tratamento alternativo ou placebo? Por exemplo, no Estudo de Tratamento de Adolescentes com Depressão (TADS), às 12 semanas, 65.2% dos adolescentes que tomaram placebo não foram classificados como com melhorias ou com muitas melhorias, comparativamente com apenas 39.4% dos adolescentes em tratamento com fluoxetina. Assim, a RAR para a não melhoria dos jovens com transtorno depressivo maior que tomaram fluoxetina, comparativamente com placebo, foi de  $65.2 - 39.4 = 25.8\%$ . Em outras palavras, a fluoxetina reduz consideravelmente o risco de não melhorar: 25.8%, ou cerca de 1 em cada 4.

A redução absoluta de risco é útil por duas razões. Em primeiro lugar, conforme explicado abaixo, a RAR pode ser usada para calcular o *número necessário para tratar* (NNT). A fórmula é simples:  $NNT = 1/RAR$  com a solução arredondada ao próximo número inteiro. Em segundo lugar, a RAR não exagera o efeito do tratamento como faz a *redução relativa de risco* (RRR), que se calcula simplesmente através da divisão do risco do grupo experimental pelo risco do grupo controle. No estudo TADS, quando se compara os resultados entre os jovens que realizam tratamento com fluoxetina e os jovens que usam placebo, a RRR é o risco dos sintomas depressivos serem classificados como não significativamente melhorados no grupo de tratamento com fluoxetina dividido pelo risco para os jovens de não terem uma melhoria significativa no grupo placebo, ou  $39.4 / 65.2 = 0.60$ , geralmente expresso como 60%. Observe que o valor é bastante diferente do valor da RAR, como calculado no parágrafo anterior (25.8%). No geral, a RAR é um valor estatístico mais útil que a RRR.

### Intervalo de confiança (IC)

Em vez de começar com uma definição, um exemplo específico pode ser mais útil para compreender a importância deste conceito. Imagine que a Ministra da Saúde lhe diz que está preocupada com os adolescentes deprimidos e lhe pede que determine o *nível médio* (i.e., a média) de sintomas depressivos em crianças de 13 anos na cidade onde vive. Você aceita o desafio e inicia o projeto selecionando adolescentes de 13 anos da sua cidade para completarem um inventário de sintomas depressivos. O valor médio é de 39.8 de 49 questionários preenchidos. Calcula o

**Figura A.6.5** Um fato importante associado com o “1.96” na fórmula para o cálculo do intervalo de confiança de 95%: 95% da área de uma distribuição normal está entre 1.96 unidades de desvio padrão da média.



desvio padrão médio e é de 9.0. A Ministra da Saúde quer que calcule, baseando-se nos dados dos seus 49 questionários respondidos, quão perto é a sua média do valor *verdadeiro*, também conhecido como o valor da *população*. Ou seja, se fosse possível avaliar *todos* os adolescentes de 13 anos da cidade com esta escala de depressão, o valor calculado seria presumivelmente muito perto do valor real. Como se pode ver, o valor *verdadeiro* ou da *população* é, portanto, um conceito útil, mas na realidade com frequência não é prático avaliar todas as crianças. Apenas por acaso pode ter encontrado uma proporção maior de crianças deprimidas ou um valor não habitual de crianças felizes na sua amostra presente na cidade, mesmo que consiga evitar ser enviesado na amostra (por exemplo, escolher apenas estudantes da mesma escola ou apenas alunos de uma só turma).

Voce dirá à ministra que o valor da média na sua amostra é de 39.8, como uma estimativa aproximada. Ela, no entanto, quer saber mais, e pergunta “quão aproximada é a estimativa ? Qual é o valor real mais alto? E o mais baixo ?” A sua resposta “Sim, posso dar-lhe uma variação ou intervalo, mas depende em quão frequente quer que eu esteja certo. Quer que eu acerte em 80% das vezes que lhe der estas estimativas ? Ou 95% ? Ou 99% ? Quanto mais insistir que a resposta que lhe dê contenha a resposta correta para a população, mais largo terá que ser o intervalo”. A ministra diz “Ok. Quero que as suas respostas, neste tipo de estimativa, sejam corretas 95% das vezes.” A ministra acabou de lhe pedir um intervalo de confiança de 95% da sua média calculada.

Um estatístico pode ajudar nesta altura. Ele poderá nos dizer que uma boa estimativa para calcular intervalos de confiança em distribuições normais é a seguinte, em que N é o número de indivíduos de onde se originam os dados (neste caso, 49) e RQ significa raiz quadrada:

$$\text{Intervalo de confiança, limite superior} = \text{Média calculada} + 1.96 \times DP / RQ N$$

$$\text{Intervalo de confiança, limite inferior} = \text{Média calculada} - 1.96 \times DP / RQ N$$

Portanto, os intervalos de confiança são os seguintes:

$$IC_{\text{superior}} = 39.8 + 1.96 \times 9.0 / RQ 49 = 39.8 + 1.96 \times 9.0 / 7 = 39.8 + 2.52 = 42.32$$

$$IC_{\text{inferior}} = 39.8 - 1.96 \times 9.0 / RQ 49 = 39.8 - 1.96 \times 9.0 / 7 = 39.8 - 2.52 = 37.28$$

Agora voce já tem a resposta para a Ministra.

Em resumo, ICs são representados com três números, sendo o primeiro a média, a nossa melhor estimativa do valor real. Os outros dois números são, como é óbvio, os números que se referem ao limite superior e ao limite inferior do

Intervalos de confiança podem fazê-lo um mago dos dados





Ao calcular o NNT, quão alta devem os investigadores definir a barreira a ultrapassar para que considerem que um jovem tenha “melhorado”?

Barreiras mais altas irão aumentar o NNT porque menos jovens irão ultrapassá-las.

intervalo. Quando mais certos queremos ser, maior será o intervalo de confiança: intervalos de confiança maiores *garantem* que o valor correto está por lá algures. *Os intervalos de confiança são importantes porque a análise estatística de estudos é baseada na ideia central de que, apesar das observações serem realizadas numa amostra limitada de indivíduos, a verdade inferida do estudo aplica-se à população total desses indivíduos.* O principal objetivo do intervalo de confiança é indicar a precisão ou imprecisão da amostra do estudo como estimativa do valor real da população. Os intervalos de confiança são, portanto, úteis e até necessários sempre que é feita uma inferência de resultados de um estudo para uma população maior. No primeiro exemplo, apenas 49 crianças tinham preenchido o questionário, mas a Ministra espera poder generalizar os resultados para a população de adolescentes de 13 anos da cidade. Observe também que mesmo que escolha a sua amostra de forma adequada – p.e., de escolas muito diferentes em diferentes partes da cidade, para evitar viés – o problema da variação randomizada no nível de sintomas de depressão mantém-se. Apenas por acaso, pode-se cruzar com um grupo de crianças com níveis particularmente mais elevados ou mais baixos de depressão que os existentes na população completa de crianças daquela idade na cidade, apesar de não ter existido viés sistemático significativo no método de amostragem.

### Número necessário para tratar

O número necessário para tratar, com frequência abreviado para NNT, pode ser usado para sumarizar num único número quão eficaz é um determinado tratamento em comparação com o placebo. NNT é definido como o *número de pessoas que precisamos tratar para prevenir um resultado negativo adicional.* Um *resultado negativo* é definido por critérios específicos. Assim, NNT baixos (como 3 ou 4) indicam um tratamento eficaz, porque é comum para o tratamento *converter* um doente para um resultado positivo que não melhoraria sem esse tratamento. Os resultados positivos mais típicos são melhoria do humor ou melhoria do funcionamento, geralmente definidos através de um valor *cutoff* de uma escala.

Observe que, no entanto, o NNT depende muito da *barreira* que necessita ser ultrapassada para se considerar “melhorado” ou “curado”. *A definição mais alta ou mais baixa desta barreira vai ter muito impacto no NNT.* Logo, é importante identificar sempre qual a barreira definida pelos autores quando se lê um estudo e tomar nota do NNT.

NNT elevados (como 25, 30 ou até maior) sugerem que o tratamento não é muito eficaz ou que o resultado esperado é difícil de alcançar. Por exemplo, ao se exigir que um resultado positivo num adolescente deprimido é não ter qualquer

Mann e colaboradores (2014) realizaram uma metanálise para calcular a eficácia do acamprosato em alcançar a abstinência em alcoolistas. Referiram que nos estudos, 36.1% dos participantes que usavam acamprosato haviam atingido a abstinência em 6 meses, comparativamente com 23,4% dos que estavam a usar placebo. Qual é o NNT? Neste caso,  $NNT = 1 / (0.234 - 0.361) = 7.9$ , isto é, oito doentes (é regra arredondar para o número inteiro mais próximo) precisam ser tratados com acamprosato para um doente adicional ficar abstinente de álcool aos 6 meses em comparação aos que não fazem o tratamento.

### Número necessário para causar dano (NNH)

O NNH é igual a um a dividir pela taxa no grupo de não expostos ao tratamento menos a taxa no grupo de expostos.

No grupo de trabalho do estudo Tratamento de Adolescentes com Depressão (TADS, 2004) os autores reportaram que 11.9% dos participantes tratados apenas com fluoxetina revelaram pelo menos um episódio autolesivo, comparativamente com 5.4% para o placebo. Qual é o NNH?

Neste caso,  $NNH = 1 / (0.119 - 0.054) = 15.4$ . Ou seja, quinze adolescentes com depressão têm de ser tratados com fluoxetina para que um doente adicional tenha um episódio autolesivo atribuível à medicação.

sintoma depressivo, então será muito difícil alcançar este resultado e o NNT correspondente será maior do que se escolhesse um resultado mais modesto.

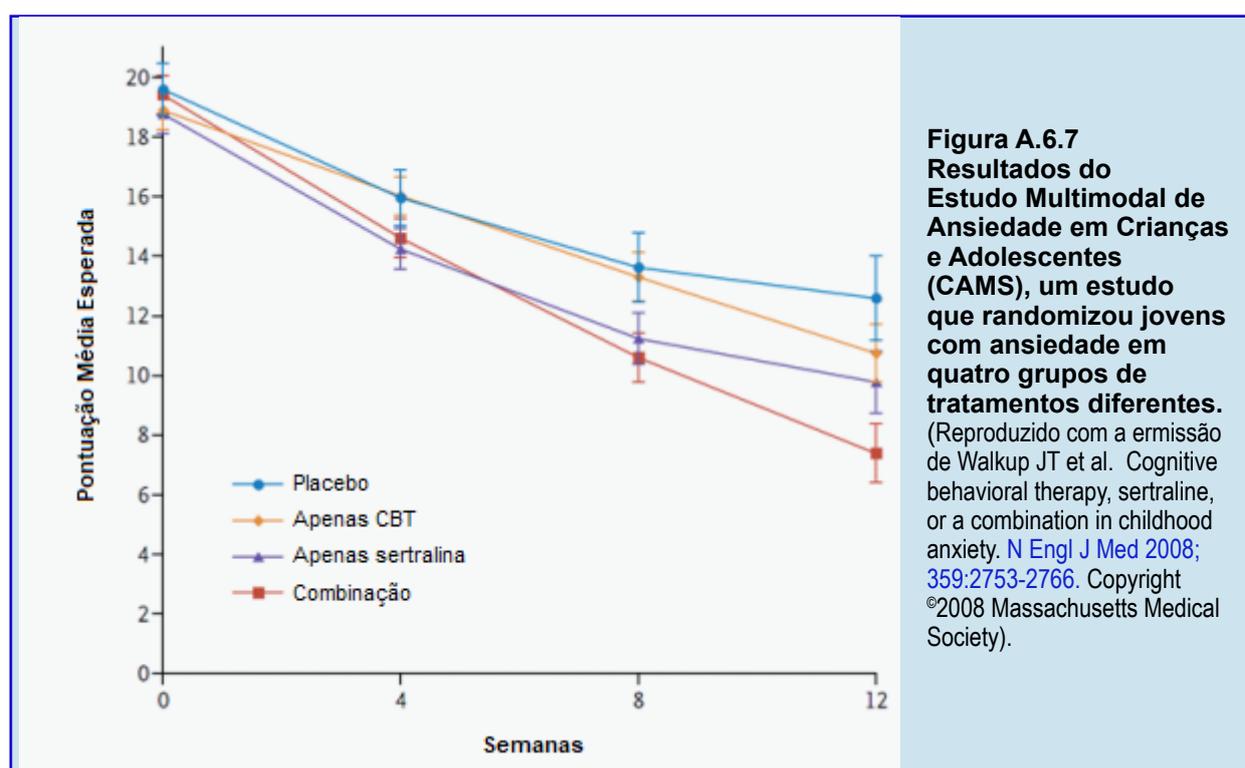
### Número necessário para causar dano (NNH)

O número necessário para causar dano é abreviado por NNH (*number needed to harm*) e resume o a frequência com que determinado efeito adverso ocorre num tratamento específico. O NNH é definido como o número de pessoas que é preciso tratar para que uma delas seja lesada pelo tratamento, que de outra forma não seria se fize Irarrázaval sse apenas um tratamento placebo. Por exemplo, com medicações, um efeito indesejável ou efeito adverso ocorre quando não ocorreria se o doente estivesse a fazer uma medicação placebo. Assim, um NNH elevado indica um tratamento seguro: muitos indivíduos são necessários a fazer tratamento para que um deles seja lesado pelo tratamento. As decisões relativas à recomendação de um tratamento específico exigem, portanto, um balanço entre os benefícios (NNT) e os riscos (NNH).

## COMPREENDER UM ESTUDO ATRAVÉS DA ANÁLISE DE SUA ESTATÍSTICA BÁSICA

Considere agora os resultados de um estudo que avaliou as respostas de crianças e adolescentes com transtorno de Ansiedade usando estes conceitos nucleares da estatística. Com o nome Estudo Multimodal de Ansiedade em Crianças e Adolescentes (CAMS), este estudo recrutou crianças e adolescentes de diversos locais com transtornos de ansiedade comuns – transtorno de ansiedade generalizada, fobia social ou transtorno da ansiedade de separação – e randomizou este grupo em quatro estratégias terapêuticas possíveis: apenas TCC, apenas sertralina, TCC e sertralina combinadas ou placebo.

O gráfico da Figura A.6.7 mostra os resultados da Pediatric Anxiety Rating



Scale (PARS) para 4 grupos, cada qual a receber um tratamento diferente. Valores superiores a 13 são consistentes com um transtorno de ansiedade. Com base nestes dados, a combinação de sertralina e TCC é o tratamento mais eficaz. As *barras de erro* nos valores da PARS mostram a *dispersão* que existe nesses números. O topo de uma barra de erro é mais facilmente entendido como a média mais o erro padrão da média (calculado através do desvio padrão / raiz quadrada do número de participantes), e a base da barra de erro é a média menos o erro padrão da média. Existem também outros métodos mais complexos de calcular o *erro padrão* representado pelas barras de erro.

Neste caso, a magnitude clínica do impacto do tratamento nos resultados foi avaliada calculando o tamanho do efeito de 3 grupos com intervenção e um grupo placebo: na PARS, quão diferentes eram os 3 grupos com intervenção comparativamente com o grupo placebo? Com base nas melhorias da escala que os autores usaram para medir a ansiedade nos jovens, a PARS, o tamanho do efeito às 12 semanas foi de 0.86 para a terapia combinada, 0.45 para o tratamento apenas com sertralina e 0.31 para o tratamento apenas com TCC.

Os autores também calcularam o número necessário para tratar. Recorde que para calcular o NNT temos de escolher uma “barreira” para cada indivíduo “ultrapassar”. Barreiras mais altas são mais difíceis de atingir e, portanto, geram NNT mais elevados; barreiras mais baixas geram NNT inferiores. Os autores escolheram como barreira que o jovem fosse classificado como melhorado ou muito melhorado na PARS no fim do tratamento de 12 semanas. Usando este critério, os autores calcularam o NNT para o tratamento apenas com sertralina de 3.2 e para TCC apenas de 2.8, e para a combinação dos dois tratamentos o NNT foi de 1.7. Este NNT baixo é encorajador porque demonstra que existem tratamentos eficazes para muitos jovens com transtornos de ansiedade.

Os autores também usaram as estatísticas para transmitir aos leitores o quão certos estão de que os seus resultados não são apenas um caso de sorte causado por eventos aleatórios que, em conjunto, resultaram na melhoria dos jovens que receberam tratamento com sertralina e TCC. Para averiguar esta possibilidade, os autores calcularam os intervalos de confiança para cada tamanho de efeito e para cada NNT. Os intervalos de confiança para 95% de certeza para o tamanho de efeito calculado de 0.86 foi de 0.56 a 1.15 para jovens que receberam a combinação de TCC e sertralina, em comparação com os jovens do grupo placebo, o que é um efeito bom e robusto em psiquiatria. Tamanhos de efeito de 0.8 ou superior indicam um grande efeito, 0.3 ou inferior indicam pouco efeito, de acordo com as convenções de Cohen (1988).

## CONFIABILIDADE E A VALIDADE DE MÉTODOS DIAGNÓSTICOS

### O FUNDAMENTO DA MEDICINA BASEADA NA EVIDÊNCIA

É importante alinhar os métodos diagnósticos da prática clínica com os métodos diagnósticos usados pelos investigadores. Desta forma, os jovens que na prática clínica sejam definidos como apresentando, por exemplo, transtorno depressivo maior, tenderão a ter uma constelação de sintomas depressivos semelhantes a dos jovens que os investigadores estudaram quando testaram tratamentos para o transtorno depressivo maior. Pode-se assim, transpor com

#### Fiabilidade

Fiabilidade é a consistência alcançada por um instrumento de medição (p.e., um teste ou questionário de diagnóstico) em todos os indivíduos, períodos de tempo e profissionais de saúde. A fiabilidade é medida usando a estatística kappa ( $\kappa$ ) para variáveis categoriais (p.e., diagnóstico) e coeficiente de correlação interclasse (CCI) para variáveis contínuas (p.e., valores de escalas). Valores comuns para ambos variam entre 0 (não melhor que o acaso) e 1 (concordância perfeita). Fiabilidade acima de 0.5 são aceitáveis e acima de 0.7 são excelentes.

#### Validade

A validade é a extensão pela qual um sistema de diagnóstico mede aquilo que diz medir. A avaliação da validade é geralmente indireta, confiando na capacidade de uma medida de concordar com outras medidas do mesmo constructo, ou de prever o curso, resultado ou resposta ao tratamento, etc. Coeficientes de correlação e análise fatorial são ambas formas comuns de quantificar pelo menos alguns tipos de validade.

**TABELA A.6.3 Exemplos de instrumentos de avaliação confiáveis\* (ver também Capítulo A.5)**

Patologia	Instrumento
Transtornos de ansiedade incluindo transtorno de ansiedade generalizada e transtorno de ansiedade de separação	<ul style="list-style-type: none"> <li>• <a href="#">Screen for Child Anxiety Related Emotional Disorders (SCARED)</a></li> <li>• Ver também os Capítulos F.1 e F.2</li> </ul>
Transtornos depressivos incluindo transtorno depressivo maior e distímia	<ul style="list-style-type: none"> <li>• <a href="#">Hamilton Rating Scale for Depression (HAMD-21)</a></li> <li>• <a href="#">Johns Hopkins Depression Checklist for Children (HDCL-C)</a></li> <li>• Ver também o Capítulo E.1</li> </ul>
Transtorno do déficit de atenção/hiperatividade	<ul style="list-style-type: none"> <li>• <a href="#">Swanson Scale for ADHD (SWAN ADHD)</a></li> <li>• <a href="#">The SNAP-IV Teacher and Parent Rating Scale</a></li> <li>• <a href="#">Vanderbilt ADHD Diagnostic Parent Rating Scale</a></li> <li>• Ver também o Capítulo D.1</li> </ul>
Transtorno obsessivo-compulsivo	<ul style="list-style-type: none"> <li>• <a href="#">Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS)</a></li> <li>• Ver também o Capítulo F.3</li> </ul>

\*Estes instrumentos foram desenvolvidos e testados na Europa e nos EUA. As síndromes em países distantes destas áreas geográficas podem ser substancialmente diferentes. Para além disso, é de lembrar que os resultados de um só instrumento são apenas um aspecto a ser levado em conta quando se formula uma conclusão, ou até uma hipótese inicial, sobre o diagnóstico; muitas outras fontes de dados, tal como a entrevista com o jovem, com os seus pais, assim como a história da doença atual e os registos médicos e psicológicos também são úteis.

confiança os resultados sobre o que ajuda aqueles jovens para a prática clínica.

Então, quão fiáveis são os procedimentos que se usam para identificar síndromes nos jovens? Fiável significa que, em primeiro lugar, quando o teste é usado de forma repetida produz resultados consistentes. Por exemplo, se uma entrevista de diagnóstico estruturada e bem definida produzir resultados iguais ou idênticos quando aplicada num curto período após a primeira aplicação, isto sugere que o teste é fiável (fiabilidade teste-reteste). Fiável também significa que produz os mesmos resultados ou resultados muito similares quando aplicados ao mesmo jovem por diferentes observadores (fiabilidade inter-observador)?

E quão válidos são os sistemas de diagnóstico que se usam para identificar síndromes nos jovens? Um sistema diagnóstico válido é um que mede realmente aquilo que se propõe a medir. Por exemplo, um sistema de diagnóstico válido para avaliar a gravidade da depressão é um que mede realmente a gravidade da depressão. A evidência que apoia a validade de um sistema de diagnóstico para a gravidade da depressão pode incluir dados que mostrem que relatórios de gravidade de depressão preenchidos pelos próprios jovens são altamente correlacionáveis com relatórios de observadores treinados que entrevistam os mesmos adolescentes usando aquele sistema de diagnóstico. A concordância diagnóstica entre o tipo de entrevistas de diagnóstico estruturadas usadas com frequência nos estudos de investigação e as avaliações clínicas variam bastante por patologia e é baixa a moderada na maioria das doenças (Rettew, 2008).

Concluindo, um sistema de diagnóstico confiável e válido é um fundamento essencial na prática clínica. Infelizmente, entrevistas clínicas não estruturadas têm

um registo fraco em termos de confiabilidade e validade no diagnóstico de jovens com doenças psiquiátricas. Entrevistas diagnósticas não estruturadas tendem a produzir conclusões precipitadas e geralmente falham na avaliação de comorbidades importantes e produzem resultados diagnósticos que não são confiáveis. Este é um problema importante, mas que felizmente pode ser resolvido.

### **Instrumentos de diagnóstico confiáveis e válidos**

Criar um sistema de diagnóstico confiável e válido é um desafio difícil. Os principais apoios na criação destes sistemas são instrumentos de diagnóstico específicos para classes específicas de transtornos. Apoios de diagnóstico estão geralmente disponíveis gratuitamente na internet, são muito úteis na definição de transtornos e na sua gravidade e, principalmente, servem de base na eventual aferição da mudança dos sintomas ao longo do tratamento. Alguns exemplos do que está disponível para algumas das patologias mais frequentes podem ser encontrados na Tabela A.6.1.

### **Um exemplo do uso de uma ferramenta diagnóstica para avaliar um adolescente**

Traduzir corretamente os sintomas descritos pelos jovens, por seus pais e professores para uma síndrome psiquiátrica ou, pelo menos, parte de uma síndrome (*condição subsindromica*) é importante. Este passo se baseia nas queixas únicas de uma família e uma criança individual, traduzi-las para um transtorno reconhecido tentando evitar o maior número possível de fontes de erro no processo. Considere, por exemplo, Sam, um jovem de 14 anos cujos pais se queixam de que está constantemente irritável e com dificuldades de sono. Há queda das notas escolares e está em isolamento social. A sua constelação de sintomas pode caber em uma ou mais síndromes. Os adolescentes podem tornar-se irritáveis quando estão ansiosos ou em oposição, mas a irritabilidade sugere também a possibilidade de um transtorno do humor; o seu isolamento social também pode ter múltiplas origens. Felizmente, várias ferramentas de avaliação largamente disponíveis, não dispendiosas e de espectro ampliado são úteis para encontrar o diagnóstico correto (ver Tabela A.6.1).

Enquanto que estes instrumentos nos ajudam a aproximar-nos e definirmos de forma mais precisa o nível de sintomas nos domínios sugeridos pela história clínica, também precisamos nos afastar e despistar no Sam outras síndromes menos óbvias tal como ser vítima de trauma ou abuso, uso de substâncias ou ter uma transtorno de comportamento que ele não nos tenha contado.

Neste caso, podem ser úteis questões de rastreio de uma avaliação compreensiva e bem reconhecida de psicopatologia. O Kiddie-SADS (K-SADS) é um procedimento de entrevista diagnóstica semiestruturada usada de forma alargada que se encontra generosamente em domínio público para uso sem fins lucrativos e investigação ética (Kaufman et al., 1997). De realçar que o K-SADS não é uma entrevista estruturada – ou seja, em que o entrevistador repete as perguntas escritas textualmente – mas sim uma entrevista semiestruturada em que o entrevistador pode escolher entre diversas perguntas para explorar a mesma área. Os entrevistadores não têm que fazer todas as perguntas, mas sim parar quando têm informação suficiente. Aqui estão algumas perguntas de rastreio do K-SADS desenhadas para explorar um possível consumo de álcool:

- *Que idade tinhas quando provaste a tua primeira bebida alcoólica?*
- *Qual é a tua bebida favorita?*
- *Tens um grupo de amigos com quem costumavas beber ou geralmente bebes sozinho/a?*
- *Onde é que costumavas beber? Em casa? Em festas? Na casa de um amigo? Na rua? Em bares?*
- *Há ocasiões especiais em que é mais provável beberes mais que os outros? Bailes da escola ou outras festas?*
- *Com que idade começaste a beber regularmente, digamos duas ou mais bebidas por semana?*
- *Nos últimos seis meses, houve pelo menos uma semana em que bebeste mais do que duas bebidas?*

### **O que é uma boa forma de rastreio de trauma?**

Aqui está um exemplo de como a entrevista semi-estruturada K-SADS pergunta sobre trauma, mas apenas após um período de “aquecimento” para coleta de informação básica para deixar o jovem à vontade. Repare que o uso do K-SADS exige treino e a lista seguinte inclui apenas algumas das perguntas sobre trauma (Kaufman et al., 1997).

#### **1. Eventos Traumáticos**

*Exploração:* Vou perguntar-te sobre um número de coisas más que acontecem com frequência a crianças da tua idade e quero que me digas se algumas destas coisas já te aconteceram. Confirma que me dizes se alguma destas coisas te aconteceu, mesmo que tenha sido apenas uma vez.

- a. *Acidente de Carro:* alguma vez estiveste num acidente de carro grave? O que aconteceu? Ficaste ferido? Mais alguém que estava no carro ficou ferido?
- b. *Outro Acidente:* Alguma vez estiveste em qualquer outro tipo de acidente grave? Por exemplo um acidente de bicicleta? Outros acidentes? O que aconteceu? Ficaste ferido?
- c. *Fogo:* Alguma vez estiveste num incêndio grave? A tua casa ou escola alguma vez tiveram um incêndio? Alguma vez pegaste fogo a alguma coisa que depois perdeu o controle? O que aconteceu? Alguém ficou com lesões físicas graves? Houveram muitos danos?

O K-SADS continua, a partir deste ponto, com perguntas sobre testemunhar um desastre como uma tempestade forte ou um crime violento, violência doméstica ou abuso físico ou sexual.

Os autores do K-SADS sugerem um patamar claro para aprofundar a avaliação: o jovem ou os pais referem que o jovem bebe duas ou mais bebidas por semana durante pelo menos 4 semanas. Os patamares explícitos do K-SADS para “contar” o sintoma como presente ou ausente são úteis, assim como as perguntas sugeridas. São claras e diretas o suficiente, no entanto oferecem ao entrevistador múltiplas formas de chegar à informação necessária. A especificidade das perguntas do K-SADS, os seus patamares claros, e a presença de perguntas de rastreio são úteis para os médicos que procuram melhorar a confiabilidade e validade das suas entrevistas clínicas. É necessário treino específico para o uso da entrevista completa e pode não ser realista em muitos contextos clínicos, mas as suas avaliações, perguntas sugeridas e patamares são úteis.

Assim que uma avaliação diagnóstica boa – confiável e válida – de um adolescente estiver completa, segue-se a tarefa de escolher o tratamento.

## **ESCOLHER UM TRATAMENTO**

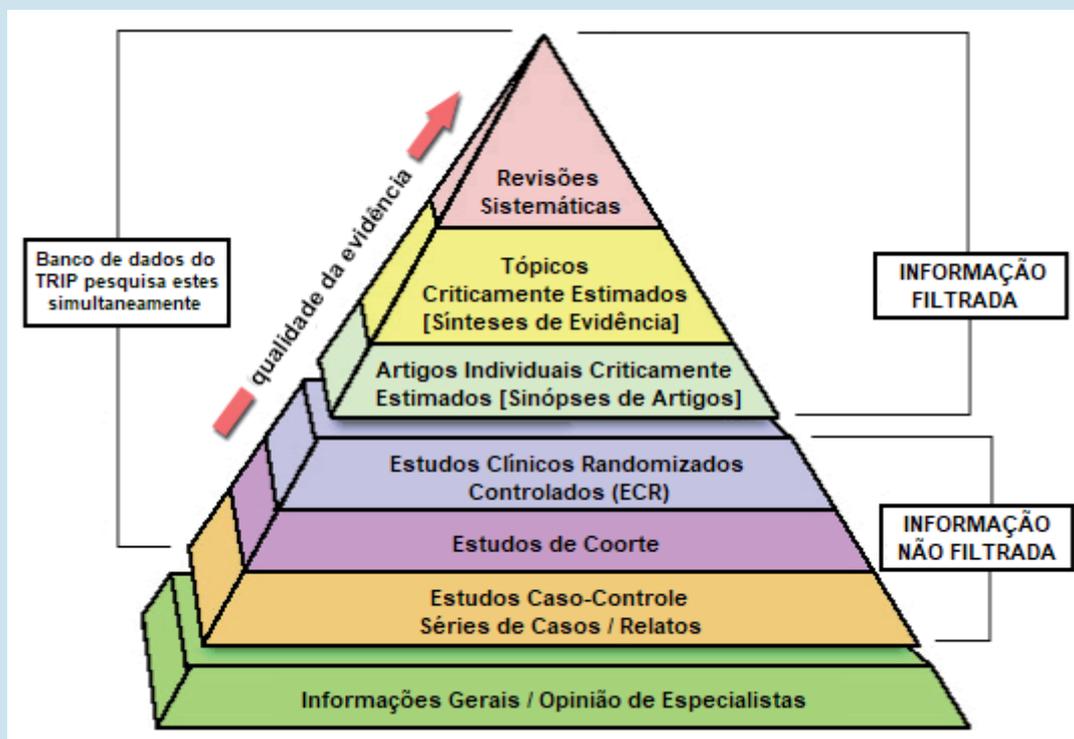
### **O conceito de pirâmide da evidência**

A *pirâmide da evidência* é uma forma útil de resumir o que é a melhor evidência – isto é, o mais perto de uma declaração precisa da verdade. Fontes de evidência com menos viés e menos erros aleatórios são colocadas mais para cima, em direção ao topo da pirâmide, como o mais plausível. As pirâmides são uma boa forma para representar a transição de centenas ou até milhares de fontes de “evidência de nível inferior” tal como estudos de caso, estudos não controlados ou até opiniões de especialistas, colocadas adequadamente na base da

As pirâmides são uma boa forma geométrica para representar diferentes tipos de evidência, visto que são largas na base e estreitas no topo. O nosso objetivo é chegar ao topo, onde a vista é boa, minimizando o viés e o erro aleatório que distorcem a imagem.



**Figura A.6.8** Uma pirâmide de evidência onde as posições mais para cima na pirâmide significam menos viés e erros aleatórios.



Fonte: Universidade da Califórnia, San Diego, sistema da biblioteca e a sua base de dados [TRIP](#) que usa “filtros” (programas de software) para selecionar a evidência publicada mais livre de viés e erros aleatórios.

pirâmide – até às relativamente escassas revisões sistemáticas de estudos controlados randomizados e meta-análises que sintetizam as evidências mais confiáveis – colocadas no topo da pirâmide. Visto que existem muitos relatos de caso e outros estudos que provavelmente contêm viés e erros aleatórios e relativamente poucos os que sintetizam de forma exaustiva toda a evidência crível, a forma da pirâmide com base alargada e o topo estreito é particularmente apropriada (Figura A.6.8).

### ***Informação de base***

Na base da pirâmide, a informação de base no assunto, tal como, por exemplo, a neurofisiologia da depressão ou autismo, e a opinião dos especialistas fornecem apenas fraca evidência. A informação de base apenas define o palco para aquilo que pode acontecer: pode acontecer vários desenlaces diferentes. Por exemplo, se estudos de imagem mostram alterações dos fluxos sanguíneos nas áreas corticais frontais na depressão, isto não nos ajuda necessariamente a encontrar o tratamento mais eficaz para a depressão.

### ***Opinião dos especialistas***

Da mesma forma, a *opinião dos especialistas* pode estar completamente errada. O *expert* pode estar contratado por uma empresa farmacêutica e sentir pressão para favorecer um determinado fármaco, ou pode sentir vergonha relativamente a resultados negativos, ou orgulho em resultados positivos e de forma subconsciente moldar os resultados numa direção positiva, ou apenas ser fiel a certo tratamento porque é com o qual está mais familiarizado. A questão principal é que uma única opinião individual não apoiada por dados sistemáticos coletados usando metodologias reconhecidas corresponde apenas a uma fraca evidência. De forma humorística podemos resumir esta conclusão como “Em Deus nós confiamos. Todos os outros, devem trazer dados.”

### ***Séries de casos clínicos***

Séries ou relatos de casos clínicos também se encontram perto da base da pirâmide de evidência. Um único profissional a avaliar um único caso, a aplicar um tratamento e a relatar os resultados é extremamente vulnerável a vieses. Séries de casos clínicos produzem dados fracos porque não existe grupo controle, apesar da sua informação poder ser útil a investigadores para gerar hipóteses em estudos futuros.

### ***Estudos de caso-controle***

Estudos de caso-controle podem ser úteis em mostrar uma associação entre a exposição a um fator de risco e um resultado. Nos estudos caso-controle, o resultado é conhecido, por exemplo, o suicídio na adolescência. O investigador cria então um grupo-controle de indivíduos sem o resultado final, mas com características similares aos indivíduos com o resultado conhecido e determina a proporção de exposição ao fator de risco nos controles em comparação com aqueles com o resultado conhecido. Usando este método, a existência e acessibilidade a uma arma em casa foi associada ao aumento de risco para suicídio na adolescência (Brent et al, 1993). Esta é uma evidência útil quando se aconselham pais relativamente a um adolescente potencialmente suicida.

### ***Estudos de coorte***

Estudos de coorte podem ser úteis em demonstrar de forma prospectiva

uma associação entre exposição a um fator de risco específico e o resultado. Aqui, a amostra é selecionada pela exposição e não pelo resultado. Por exemplo, sabe-se que crianças expostas a certos tipos de abuso têm um maior risco de desenvolver uma depressão. Alguns estudos de coorte usam bases de dados muito amplas, incluindo até todas crianças de um país inteiro, e por isso geram um grande poder estatístico (a capacidade de detectar até um pequeno efeito).

### ***Estudos controlados randomizados***

Estudos controlados randomizados (ECR) individuais são os principais blocos de construção dos níveis mais altos da pirâmide. Com as suas vantagens únicas de um grupo-controle escolhido ao acaso e geralmente com indivíduos, médicos e avaliadores, deliberadamente cegos para a atribuição da intervenção ativa e do grupo-controle, os ECR têm vantagens únicas. Níveis mais elevados da pirâmide são desenhados para diminuir as limitações nos dados dos ECR, tal como eliminar dados de ECRs mal desenhados e juntar dados de diversos ECRs bem feitos para gerar um número elevado de sujeitos estudados de modo a diminuir os erros aleatórios. Atenção deve ser dada para a exclusão de ECRs não publicados.

### ***Análise crítica***

A análise crítica de artigos individuais elimina dados de estudos com falhas significativas. Talvez um estudo assim tão defeituoso não tenha os avaliadores adequadamente cegos, por exemplo, ou tenha perdido uma grande proporção da amostra no follow-up. Se assim for, os resultados devem conter viés significativo e os estudos são excluídos. A análise crítica dos resultados de artigos individuais que abordam o mesmo tema, sabendo agora que se encontram livres de fontes óbvias de viés, podem então juntar-se para aumentar o tamanho da amostra para diminuir os erros aleatórios. Pondere usar listas convenientes específicas de critérios quando avaliar um estudo, tal como a lista do Consolidated Standards for Reporting Trials (CONSORT), se e quando tiver tempo para isso.

CONSORT significa Consolidated Standards of Reporting Trials (Padrões consolidados para reportar investigações).

O diagrama de fluxo da Figura A.6.9 (página 20) mostra como alguns passos críticos na execução de um ECR merecem ser lembrados quando se lê um artigo. Por vezes, por exemplo, um autor pode descrever o estudo como “randomizado” apesar de ser claro na secção dos métodos que a distribuição não foi realmente aleatória. Existem diversas fontes possíveis de viés ao realizar um ECR, mas a maioria dos clínicos não têm a vontade, o tempo ou a experiência estatística e de investigação para ler na totalidade um ECR publicado. Uma análise crítica implica ter uma visão crítica para cada etapa do estudo e assegurar que nenhum viés passa para os resultados finais. Frequentar um curso em Medicina Baseada na Evidência ou participar num *journal club* para pessoas interessadas pode ajudar a melhorar as competências de análise crítica.

### ***Síntese da evidência e revisões sistemáticas***

As sínteses da evidência juntam dados de ECRs individuais que tenham sido analisados de forma crítica para possíveis fontes de viés. A junção dos dados tende a diminuir os erros aleatórios e a produzir resultados mais precisos.

Por último, as revisões sistemáticas são um método especialmente organizado de síntese de evidência que se apoia em padrões bem definidos de síntese. Com esta metodologia, talvez melhor exemplificada pela biblioteca digital de Revisões

Sistemáticas da Cochrane Library, o método usado para combinar dados e desenhar conclusões está bem definido antes do início da revisão. Esta metodologia é também transparente, pelo que o leitor pode saber como os resultados foram obtidos. Uma meta-análise é um tipo de revisão sistemática que toma em consideração os dados de diversos estudos para chegar a uma estimativa única do tamanho do efeito ou NNT.

## USAR A INTERNET E A PIRÂMIDE PARA ENCONTRAR TRATAMENTOS EFICAZES

### Responder a questões relevantes para a prática clínica usando a Internet

Tente manter-se curioso quando faz o seu trabalho clínico: coloque questões relevantes para o cuidado dos doentes que vê nesse dia. Isto significa colocar questões relevantes para o caso específico e procurar responder usando a Internet para ver na literatura publicada, com foco na evidência sem viés ou erro aleatório.

O primeiro passo é colocar a questão. Para conseguir isto, comece a questão com “Qual é a evidência sobre...?”

Isto assegura que a questão pode ser respondida: ou descobre que não existe evidência, sendo que a resposta neste caso é “nenhuma”, ou encontra pelo menos alguma evidência.

O objetivo de uma pesquisa é um resultado prático, bom o suficiente para informar relativamente ao tratamento sem afirmar que é definitivo, dentro do tempo e recursos disponíveis. A tecnologia permite a pesquisa de resultados válidos e relevantes em bases de dados. Softwares tornaram a pesquisa sensível e específica, fácil no PubMed, o site da Internet da National Library of Medicine dos Estados Unidos. O PubMed é mantido pelo governo dos Estados Unidos sem nenhum custo para os utilizadores e sem associação necessária. PsycINFO, a página na Internet para estudos de Psicologia, é mantido pela *American Psychological Association* e disponibiliza os resumos sem qualquer custo para o utilizador. CENTRAL, a base de dados de estudos controlados da Cochrane Library, tem a sua base no Reino Unido. CENTRAL é a abreviatura usada para The Cochrane Central Register for Controlled Trials; contém resumos de mais de 700.000 estudos controlados quer da PubMed quer da EMBASE, uma base de dados especialmente forte em jornais europeus. EMBASE está acessível apenas para associados e é mantida pela Elsevier, uma corporação com a sua base em Amsterdam. PubMed é um bom local para começar a procurar questões relacionadas com a medicação, e inclui estudos além de ECRs. PsycINFO é bom para questões relacionadas com temas psicossociais. CENTRAL é bom para pesquisar especificamente resultados de ECRs.

### Terry: uma pesquisa de 10 minutos de uma solução para um rapaz com TDAH tipo desatento

Um psiquiatra da infância e adolescência está tentando ajudar um garoto de 10 anos com TDAH tipo desatento (TDAH-I). Baseado nos valores SWAN (ver Tabela A.6.3) do seu professor, ele precisaria de metilfenidato 15mg três vezes por dia para alcançar melhoria significativa (uma pontuação de um professor inferior a 9 nas primeiras nove questões SWAN, os itens da desatenção). Mas ele perdeu muito peso neste esquema, motivo pelo qual teve que ser interrompido. Agora

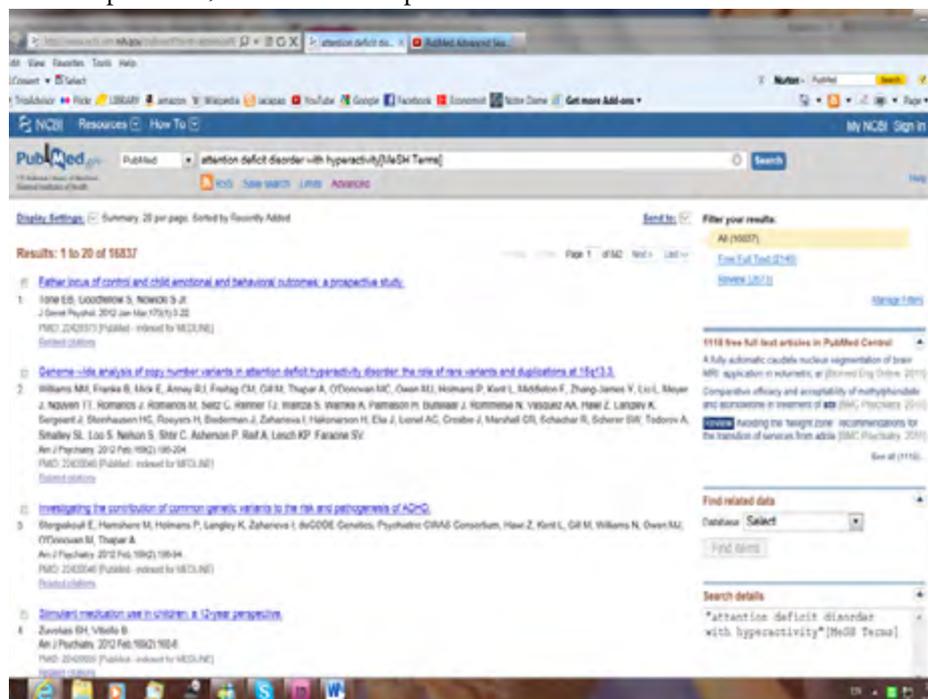


O logotipo da PubMed, uma fonte gratuita de evidência. Tem filtros para encontrar revisões sistemáticas, úteis como resumos de evidência de alta qualidade.

seus pais estão cada vez mais frustrados e à procura de uma solução: “não existe mais nada?” O psiquiatra da infância e adolescência quer um fármaco que não seja um estimulante ou a atomoxetina – ambos diminuem o apetite. O médico lembra-se do quão útil a guanfacina foi no passado para tratar garotos com TDAH com TOD e sintomas de hiperatividade/impulsividade. Mas é a guanfacina eficaz no TDAH-I? Qual é a evidência de que a guanfacina é eficaz em garotos de 10 anos com TDAH-I? Se é eficaz, qual é o tamanho de efeito?

1. Como local razoável para começar, vá ao [PubMed](#). Clique em “Advanced”, selecione no menu “MeSH term” (MeSH significa Títulos de Temas Médicos – Medical Subject Headings), escreva “attention deficit hyperactivity disorder” na caixa de pesquisa e carregue em “Search”. Os resultados devem aparecer conforme a figura abaixo.

Há demasiados resultados para serem lidos individualmente – mais de 16.000! Precisamos de uma maneira de limitar os resultados para um número manejável e escolher apenas aqueles resultados do topo da pirâmide, de modo a ver apenas a melhor evidência.



2. Comece no topo da pirâmide: pesquise por qualquer meta-análise publicada. Volte atrás para a página da pesquisa e, após repetir o passo anterior, escreva “guanfacine” na caixa da pesquisa “AND” e “All Fields” e selecione “Meta-Analysis” no novo menu e clique em “Search”; a pesquisa origina 2 resultados, mas este garoto não pertence à população estudada por nenhum dos artigos: um é referente a adultos, o outro, com crianças com TDAH e tics, e o nosso doente tem TDAH sem tics (ver Figura abaixo). Tornamo-nos demasiados seletivos e precisamos alargar a pesquisa.
3. Desloque-se para baixo na pirâmide para ECRs ao mudar a caixa “Limits”

## Pesquisar em PsycINFO

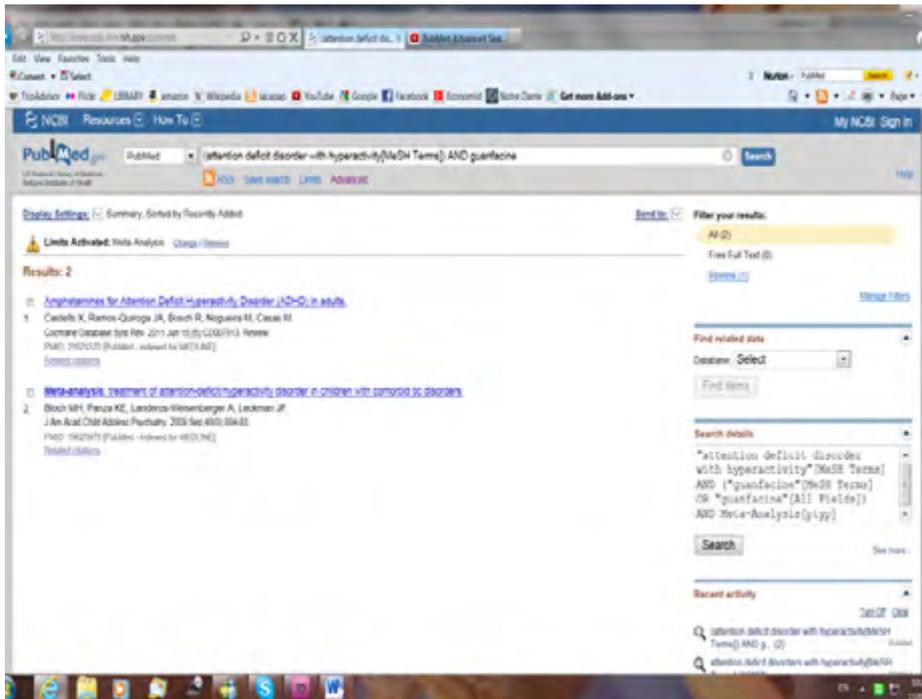
PsycINFO é uma excelente escolha de base de dados para pesquisas de construtos como “bullying” ou “identidade” porque tem uma lista de termos indexados. Pesquisar o termo “bullying”, por exemplo, mostra os seguintes termos indexados listados numa coluna à esquerda dos resumos, junto ao número de resultados para cada termo indexado.

Termo indexado:

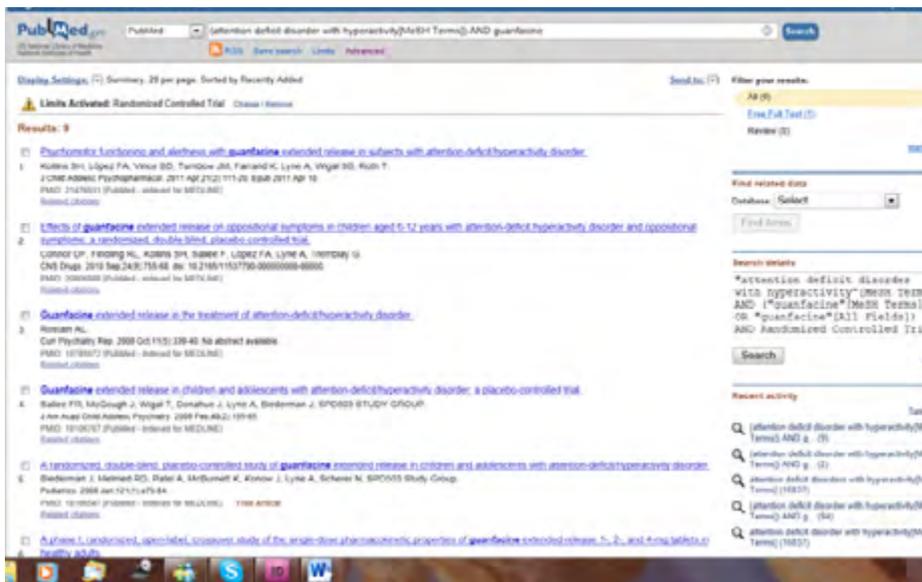
- Bullying (2,808)
- Victimização (1,175)
- Comportamento Agressivo (845)
- Relações entre Pares (543)
- Escolas (437)
- Mais 5...

Clicar na hipertexto da página leva aos resumos dos artigos que se encontram disponíveis, sem custos. Use a caixa “Advanced Search” para pesquisar “bullying” e escolha também “Only Show Content Where...” para incluir a opção do menu “Treatment Outcome/Randomized Controlled Trials” Isto reduz de 2.808 entradas para apenas 10 (Outubro 28, 2011).

Como exercício, tente encontrar todos os estudos randomizados que estudam o efeito do treino de relaxamento em adolescentes).



de “Meta-analysis” para “Randomized controlled trials” e aparecem 9 artigos (ver Figura em baixo).



Clique na imagem para ver um vídeo tutorial como fazer pesquisas usando o PubMed (2:27). Nota do Editor: Este vídeo foi legendado em português.

Os artigos 4 e 5 são os únicos diretamente relevantes para um garoto de 10 anos, e apenas o artigo 5 tem o texto completo disponível sem custos (Biederman et al, 2008). Este artigo foi financiado por empresas farmacêuticas, uma potencial fonte de viés, e aborda a guanfacina de liberação-prolongada, que é muito cara e não acessível para este garoto. Uma rápida pesquisa na Internet mostra, no entanto, que a semi-vida da guanfacina em crianças e adolescentes é de 13 a 14 horas, o que sugere a possibilidade de reproduzir a preparação de liberação prolongada com uma preparação mais barata – a própria guanfacina – tomando pelo menos duas vezes ao dia (Strange, 2008).

Quão grande é o efeito da guanfacina nos sintomas de desatenção em jovens com TDAH do tipo misto entre as crianças estudadas? Os autores notam que “O tamanho do efeito nos grupos de tratamento com GXR, calculado *post hoc* foram 0.58, 1.19 e 1.34 para grupos medicados com 0.05 a 0.08 mg/kg, 0.09 para 0.12 mg/kg, e 0.13 para 0.17mg/kg, respectivamente.” No entanto, estes tamanhos de efeito foram para mudanças nos valores totais de TDAH (sintomas de desatenção e hiperatividade/impulsividade). Felizmente, os autores clarificam “todos os grupos de crianças que tomaram GXR mostraram melhorias significativas em ambas as subescalas de hiperatividade/impulsividade e desatenção da TDAH-RS-IV [uma escala de medição] comparados ao placebo. Mudanças da média inicial da hiperatividade/impulsividade dos grupos placebo e GXR 2-mg, 3-mg e 4-mg foram -4.06, -6.94, -7.09, e -9.46, respetivamente. *Alterações da média inicial da desatenção foram -4.78, 8.46, -8.71, e -9.51, respectivamente* [itálicos adicionados]” Apesar de nenhum tamanho de efeito ser apresentado relativamente a mudanças apenas na desatenção, estes dados sugerem que o tamanho do efeito será pelo menos igual ao tamanho de efeito para sintomas de desatenção e hiperatividade/impulsividade combinados.

É razoável concluir que o paciente de 10 anos com TDAH tipo desatento pode responder, com uma redução significativa de sintomas, à guanfacina titulada até a uma dose de 1.0 a 1.5 ou 2.0mg, duas vezes por dia.

### **Estratégias na Internet para encontrar informação sobre uma área alargada de temas**

Em vez de procurar por evidência sobre uma questão clínica específica sobre um jovem, imagine que vc está escrevendo normas de orientação clínica para o seu país, sobre como tratar crianças e adolescentes com esquizofrenia. Uma pergunta relevante pode ser “Qual é a evidência que existe sobre tratamentos eficazes para jovens com esquizofrenia?”. Comece por encontrar os dados mais relevantes e atuais de estudos randomizados. Comece por pesquisar no PubMed por estudos de tratamento em crianças e adolescentes com esquizofrenia. Aqui fica como fazê-lo (de referir que o número de resultados vai variando ao longo do tempo consoantes novos estudos são adicionados).

1. *Encontre o título do tema médico para o seu tópico de pesquisa.* Escreva “Schizophrenia” na caixa de pesquisa da PubMed após usar o menu para mudar os parâmetros para “MeSH”. Isto significa que está agora a pesquisar nos Medical Subject Headings usados pela Biblioteca Nacional de Medicina dos Estados Unidos. Estes títulos do tema médico são usados para categorizar cada artigo publicado na literatura médica, tornando possível a recolha de todos os artigos sobre o mesmo tema. Neste caso, vários termos MeSH (12) foram identificados, comumente descrito como 12 resultados. Cada uma contém a palavra “esquizofrenia”, quer no título quer na descrição.
2. *Coloque os termos relevantes na caixa de pesquisa “Search Builder”, combinando-os com a palavra “OR”.* Apenas 6 parecem relevantes visto que as últimas 6 correspondem a uma proteína rara. Após selecionar estes 6, clique em “Add to Search Builder”.
3. *Adicione à pesquisa uma versão truncada (tw) do conceito central como*

### **Pesquisar na base de dados CENTRAL da Cochrane Library**

Ao procurar na CENTRAL por “bullying” encontramos 36 resultados, todos ECRs que estudam o bullying. Um destes, por exemplo, realizado na Austrália em 2009 com garotos adolescentes, parece promissor, mas não apareceu na pesquisa do PsycINFO para ECRs de “bullying” (Berry K and Hunt CJ (2009), Evaluation of an intervention program for anxious adolescent boys bullied at school. Journal of Adolescent Health 45:376-382). CENTRAL é útil em encontrar ECRs de todo o mundo e tornar os resumos disponíveis para todos sem custos.

uma palavra de texto com um “OR”. Neste caso, adicione “Schizophren\*” (tw) na caixa de pesquisa com o operador “OR”. Este passo irá recolher qualquer artigo que contenha qualquer palavra começada com as letras “Schizophren” no título ou resumo. Este termo adicional pode ajudar a encontrar artigos que podem não ter sido corretamente classificados em algum termo MeSH e que contenha informação sobre esquizofrenia, assim como artigos que usem apenas termos como “esquizofrênico” ou “esquizofrênicos”.

4. Procure no PubMed com esta linha de termos “OR”.

Surgem 125889 resultados. É o momento de tornar mais específico e começar a filtrar, desse conjunto grande de referências, apenas as mais relevantes.

5. Clique em “Limits” e assinale a caixa “Children 0 to 18”

Permanecem 22545 resultados.

6. Clique em “Limits” e assinale as caixas “Meta-analysis” e “Randomized controlled trial Click on “Limits” and check the boxes “Meta-analysis” and “Randomized controlled trial”.

Permanecem 945 resultados. Este pode ser um número razoável para ver com atenção um artigo importante para todo um país. Dois ou três indivíduos podem filtrar os títulos e resumos e escolher os mais relevantes. De forma alternativa, pode voltar a “Limits” e escolher a caixa “Published in the last 5 years” (publicado nos últimos 5 anos), para reduzir ainda mais o número de referências. Repare também que várias referências dizem respeito apenas adultos, e não crianças ou adolescentes, e podem ser imediatamente descartados.

Aqui fica uma alternativa se não tem tempo de avaliar tantos resultados:

1. Regresse ao passo 1 e assinale o termo MeSH “Schizophrenia, childhood”. Envie apenas o termo “Schizophrenia, childhood” [MeSH] para a caixa de pesquisa, ordenando ao motor de busca que procura apenas este termo MeSH. Aparecem 1434 resultados.
2. Agora aplique os limites “Meta-analysis” ou “Randomized controlled trial”. Permanecem 8 resultados. Estes resultados são de leitura obrigatória, mas os limites parecem ter eliminado demasiados resultados.
3. Regresse ao passo 2 com 1434 entradas e aplique o limite “Last 10 years” [últimos dez anos]. Mantêm-se 59 resultados. Uma avaliação destes resultados revela uma série de artigos muito relevantes com tópicos como “Efeitos da clozapina e olanzapina na espessura cortical na esquizofrenia com início na infância” ou “perda assimétrica é local e não global na esquizofrenia de início na infância”. Repare na diminuição significativa do número de resultados, o que aponta para que o termo MeSH “Schizophrenia, childhood” – como o site PubMed nos notifica – seja principalmente um termo histórico; no entanto, os 59 resultados identificados dos últimos 10 anos aparecem principalmente como muito relevantes e interessantes para um *position paper*.



**Mantenha-se atualizado na Medicina Baseada na Evidência**

Mantenha-se atualizado na Medicina Baseada na Evidência: São publicados mundialmente todos os dias mais de 5000 artigos biomédicos e os resultados de mais de 100 estudos de tratamento e os números continuam a subir. Acompanhar esta onda de informação é assustador. Para facilitar esta tarefa, pode inscrever, sem custos, à página [Evidence Updates](#) dirigida pelo grupo BMJ e pela Unidade de Informação de Investigação em Saúde da Universidade McMaster. Estas instituições colaboram para providenciar aos clínicos acesso à melhor e mais atual evidência da investigação, ajustada aos interesses pessoais relativos a cuidados de saúde, para apoiar decisões clínicas baseadas na evidência. Este serviço pesquisa mais de 120 jornais médicos do topo e mais de 50 000 artigos todos os anos, classificando-os de acordo com a sua validade e relevância, e envia regularmente aos seus subscritores um email com os artigos principais da sua especialidade.

A mensagem principal é *manter as estratégias de pesquisa transparentes e claras, e monitorizar sempre quantos resultados surgem; estes dados permitem ajustar continuamente a pesquisa aos seus objetivos, tempo e interesse.*

Mantenha-se atento a quantos resultados produz a estratégia de pesquisa no contexto dos objetivos da sua investigação: os resultados são manejáveis e adequados para o seu propósito? Se sente que não são manejáveis (demasiados resultados para muito pouco tempo), continue a apurar os resultados com mais limites. Se sente que não são suficientes, mude a sua estratégia de pesquisa inicial para se tornar mais sensível. Neste exemplo, adicionar o termo “Schizophren\*” (tw) adicionou mais de 20,000 resultados. Outra alternativa para aumentar a sensibilidade é clicar na hiperligação “Related citations” [citações relacionadas] por baixo das citações que parecem ser particularmente úteis; o software apresenta uma série de citações similares ao tema da citação que considerou útil.

### MBE CLÁSSICA E A ‘PERGUNTA RESPONDÍVEL’

A “MBE clássica” é a nossa expressão para a abordagem desenvolvida originalmente por David Sackett e seus colegas na Universidade McMaster no Canadá e mais tarde em muitos outros sítios. O seu trabalho foi fundamental para a abordagem aqui discutida. Distinguimos a nossa abordagem, no entanto, que é desenhada para encaixar na realidade de trabalhar com a saúde mental de crianças



Clica na imagem para ir ao site Centre for Evidence Based Medicine, que tem dados e formações em MBE

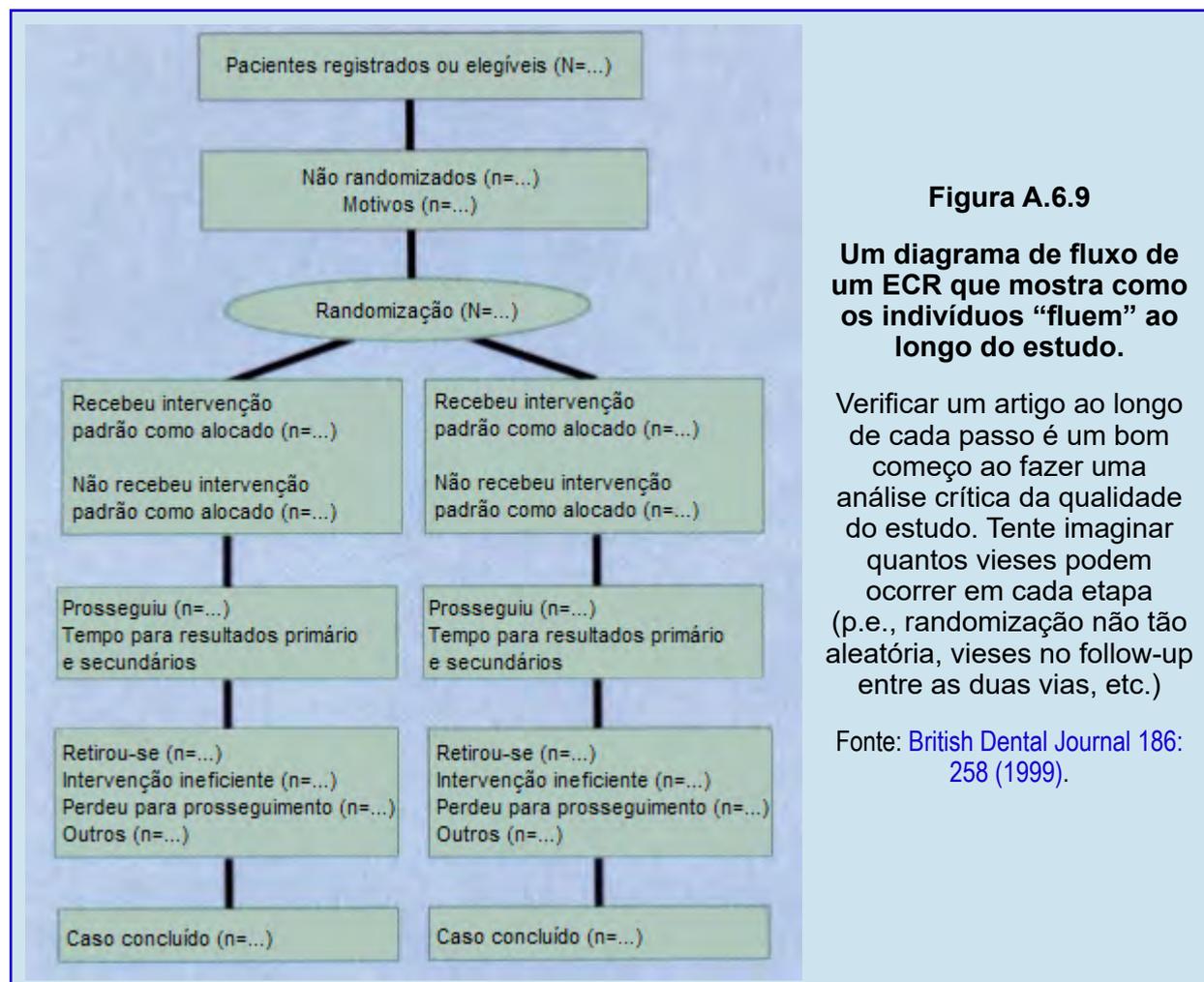


Figura A.6.9

Um diagrama de fluxo de um ECR que mostra como os indivíduos “fluem” ao longo do estudo.

Verificar um artigo ao longo de cada passo é um bom começo ao fazer uma análise crítica da qualidade do estudo. Tente imaginar quantos vieses podem ocorrer em cada etapa (p.e., randomização não tão aleatória, vieses no follow-up entre as duas vias, etc.)

Fonte: [British Dental Journal 186: 258 \(1999\).](#)

e adolescentes, da “MBE clássica”, que foi desenvolvida inicialmente no contexto da medicina interna. É importante conhecer a “MBE clássica” como o precursor de base daquilo que é aqui discutido.

A “MBE clássica” pode ser resumida num modelo de cinco passos:

1. Colocar uma questão “respondível” sobre uma situação clínica, geralmente começando com “Qual é a evidência de que...” um doente ou problema específico, tratado com uma intervenção específica em comparação com outra intervenção ou placebo produz em termos de resultados? As vezes são chamadas de perguntas PICO (Paciente/Intervenção/Comparação/Resultados [Outcomes]).
2. Pesquisar por evidência para responder à pergunta PICO.
3. Avaliar de forma crítica a evidência ao colocar questões sobre como os estudos foram feitos.
4. Integrar a experiência clínica e os valores do doente com a evidência para tomar uma decisão sobre o tratamento.
5. Avaliar os resultados.

A “MBE clássica” toma em consideração questões importantes como a sensibilidade e especificidade dos testes de diagnóstico e como aplicar a evidência (Sackett et al, 2000).

## PRÁTICAS BASEADAS NA EVIDÊNCIA E CULTURA LOCAL

A cultura – aqui deixada por definir de propósito – é um fator importante na forma como as ideias discutidas neste capítulo são recebidas por médicos assim como por doentes. Alguns componentes culturais que afetam a nossa prática incluem crenças e valores, práticas de parentalidade e educação infantil, o papel da família alargada, e o impacto da Internet.

### Crenças e valores

Crenças relacionadas com o tratamento de doenças mentais são importantes. É crucial que os médicos compreendam a cultura onde trabalham de modo a permitir a aceitação de intervenções eficazes. Em sociedades onde existe um maior estigma sobre a doença mental, pode ser muito difícil convencer uma criança e a família a comparecer sequer à entrevista psiquiátrica inicial. Mesmo que compareçam, pode ser muito difícil para eles aceitar o diagnóstico assim como o tratamento sugerido. Em algumas culturas, acredita-se que a doença mental é espiritual e, portanto, tratamentos espirituais são a primeira escolha. As famílias podem aderir ao tratamento apenas se o líder espiritual concordar. Num caso particular, um líder espiritual turco preparou um *muska* – um colar mágico de papel embrulhado e pano para afastar a doença – para um adolescente psicótico a quem o líder espiritual tinha enviado para o internamento. Durante o processo do tratamento, o adolescente abriu o *muska*. Encontrou apenas uma palavra rabiscada: *Haldo!*

No entanto, os clientes procuram com frequência curas imediatas e mágicas. É muito difícil por vezes explicar, e ser aceito, o tempo necessário para um tratamento eficaz ou a recomendação de usar uma medicação apesar dos seus efeitos adversos. Em culturas onde as pessoas perguntam primeiro aos vizinhos ou familiares sobre o problema e confiam nas suas experiências pessoais ou de

terceiros, os tratamentos baseados na evidência sugeridos por um médico podem não encaixar sempre nas sugestões dos vizinhos e com frequência adquirem um estatuto de “não usar” (o filho do vizinho melhorou com outro medicamento, este pode não ser o certo).

### **Pressupostos em educação infantil e parentalidade**

Pressupostos como “os rapazes normalmente apresentam elevada atividade motora” pode ser uma razão para que uma família recuse medicação para tratar TDAH, porque a família pensa que a medicação lentifica a criança e torna-a “anormal”. Outra suposição comum é a de que “as crianças não devem tomar medicamentos durante muito tempo”. Isto geralmente perturba a intervenção medicamentosa. Ou então, a avó diz que “ele é tal e qual o pai quando era pequeno e o pai está bem. Tem a certeza de que ele precisa de tratamento?”

### **Papel da família alargada**

Em culturas tradicionais, os avós ou irmãos mais velhos dos pais podem estar encarregados das decisões relativas ao tratamento. Podem recusar dar permissão aos pais para administrarem um tratamento sugerido. Em algumas situações, os pais que acreditam nos tratamentos baseados na evidência têm de lidar com o conflito de administrar o tratamento às escondidas. Membros da família também podem interferir com a dosagem prescrita, baseados nas suas opiniões pessoais.

### **Influência da Internet**

Pais ou adolescentes que pesquisam na Internet poderão recusar tratamento. Visto que a Internet pode desinformar para além de informar, os doentes e as suas famílias, com frequência encontram mais desinformação que orientações. Por exemplo, podem pesquisar no Google uma palavra e ler qualquer coisa que apareça, geralmente a informação encontrada enfatiza os efeitos adversos, ignora os benefícios, e opõem-se à medicação para a doença mental. Fornecer aos pais e adolescentes informação sobre páginas na Internet com menos vieses e informação mais corretas e completas pode melhorar a adesão ao tratamento.

Resumindo, o papel chave do clínico é entender o sistema cultural de crenças e valores dos seus pacientes, assim como as ideias fundamentais da epidemiologia e da investigação clínica que compõem a primeira parte deste capítulo. A combinação bem sucedida de ambos, ao ser aplicada na prática clínica ao doente e à sua família é a essência da prática baseada na evidência.

---

## REFERÊNCIAS

- Biederman J, Melmed R, Patel A et al (2008). A randomized, double-blind, placebo-controlled study of guanfacine extended-release in children and adolescents with attention-deficit/hyperactivity disorder. *Pediatrics* 121: e73-84
- Brent D, Perper J, Moritz G et al (1993). Firearms and adolescent suicide. A community case-control study. *American Journal of Diseases of Children*, 147:1066-1071.
- Guyatt GH, Rennie D (2002). *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* Chicago: AMA Press
- Kabai P (2011b). *Bias and Random Error*. [http://www.behav.org/ecol/wildlife/w\\_05\\_design.htm](http://www.behav.org/ecol/wildlife/w_05_design.htm)
- Kabai P (2011a). *More on Bias (systematic) and Random Error*. [http://apollo.lsc.vsc.edu/classes/remote/lecture\\_notes/measurements/bias\\_random\\_errors.html](http://apollo.lsc.vsc.edu/classes/remote/lecture_notes/measurements/bias_random_errors.html)
- Kaufman J, Birmaher B, Brent D et al (1997). *Kiddie-SADS-Present and Lifetime Version (K-SADS-PL)*. Pittsburgh: Department of Psychiatry, University of Pittsburgh School of Medicine.
- Kelly R, Cohen L, Semple R (2006). Relationship between drug company funding and outcomes of clinical psychiatric research. *Psychological Medicine*, 36:1647-1656.
- McCracken JT, McGough J, Shah B (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine*, 347:314-321.
- Sackett DL, Straus SE, Richardson WS et al (2000). *Evidence-Based Medicine: How to Practice and Teach EBM*, 2<sup>nd</sup> edition. New York: Churchill Livingstone.
- Strange BC (2008). Once-daily treatment of ADHD with guanfacine: patient implications. *Journal of Neuropsychiatric Disease and Treatment*, 4:499-506.
- Walkup JT, Albano AM, Piacentini J et al (2008). Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *New England Journal of Medicine*, 359:2753-2766
- Wikipedia (2011). *Standard Deviation*. [http://en.wikipedia.org/wiki/Standard\\_deviation](http://en.wikipedia.org/wiki/Standard_deviation)
-