

An evident view of evidence-based practice in perinatal medicine: absence of evidence is not evidence of absence

Uma visão evidente da prática baseada em evidências na medicina perinatal: ausência de evidência não é evidência de ausência

Augusto Sola¹, Fernando Dominguez Dieppa², Marta R. Rogido³

Resumo

Objetivo: Proporcionar elementos valiosos e um pouco de humor nesta chamada era da "prática baseada em evidências" com o objetivo de ajudar os clínicos a fazer escolhas melhores no cuidado que eles provêem baseados em evidências, e não simples ou exclusivamente baseados em um ensaio clínico randomizado (ECR) ou meta-análise (o que pode não ser evidência).

Fontes dos dados: Livros e artigos com revisão por pares são citados e listados na bibliografia. Evidência de vida, aprendizado através de nossos próprios erros e muitos outros fatos evidentes que sustentam esta revisão não são citados.

Síntese dos dados: 1) "Ausência de evidência não é evidência de ausência" e "falta de evidência de efeito não significa evidência de nenhum efeito". 2) Os ECR com resultado "negativo" e aqueles com resultado "positivo", mas sem os resultados importantes, muitas vezes não podem concluir o que concluem. 3) Os ensaios clínicos não-randomizados e os estudos práticos podem ser importantes. 4) A pesquisa em busca de provas é diferente da pesquisa em busca de aperfeiçoamento. 5) A escolha clínica deve avaliar os efeitos nos desfechos importantes para os pacientes e seus pais. 6) A quantificação de desfechos adversos, do número necessário para causar dano e do número necessário para tratamento não é assim tão simples.

Conclusões: Desafios importantes inerentes à pesquisa em serviços de saúde devem ser correlacionados a possíveis aplicações clínicas usando ferramentas que permitam uma "visão mais clara da prática baseada em evidências" na medicina perinatal, lembrando que a ausência de evidência não é evidência de ausência.

J Pediatr (Rio J). 2007;83(5):395-414: Medicina baseada em evidências, número necessário para tratamento, ensaios randomizados, variáveis de desfecho, efeitos do tratamento, leitura crítica, significância estatística.

Introdução

A não equivalência da significância estatística e da importância clínica é reconhecida há muito tempo; esse erro de interpretação ainda é comum. Um resultado significativo

Abstract

Objective: To provide valuable elements and some humor in this so-called era of "evidence-based practice" with the aim of helping clinicians make better choices in the care they deliver based on evidence, not simply or exclusively based on a randomized clinical trial (RCT) or meta-analysis (which may not be evidence).

Sources: Books and peer-reviewed articles are quoted and listed in the bibliography. Evidence of life, learning from our own mistakes and many other evident facts that support this review are not quoted.

Summary of the findings: 1) "Absence of evidence is not evidence of absence" and "lack of evidence of effect does not mean evidence of no effect". 2) RCTs with "negative" results and those with "positive" results, but without outcomes that matter, often cannot conclude what they conclude. 3) Non-randomized clinical trials and practical trials may be important. 4) Research to prove is different than research to improve. 5) Clinical choice must assess effects on outcomes that matter to patients and their parents. 6) Quantifying adverse outcomes, number needed to damage and to treat is not that simple.

Conclusions: Significant challenges inherent to health service research must be correlated to possible clinical applications using tools to have a more "evident view of evidence-based practice" in perinatal medicine, recalling that absence of evidence is not evidence of absence.

J Pediatr (Rio J). 2007;83(5):395-414: Evidence-based medicine, number needed to treat, randomized trials, outcome variables, treatment effects, critical reading, statistical significance.

pode, algumas vezes, não ser clinicamente importante. Muito pior é a interpretação incorreta de achados "não-significativos". Outras interpretações incorretas comuns são a confusões entre "evidência de nenhum efeito" com

1. MD. MANA and Atlantic Neonatal Research Institute, Division of Neonatology, Atlantic Health System, Morristown, NJ, USA.
2. MD. Gonzalez Coro University Hospital, La Habana, Cuba.
3. MD. MANA and Atlantic Neonatal Research Institute, Division of Neonatology, Atlantic Health System, Morristown, NJ, USA.

Como citar este artigo: Sola A, Dieppa FD, Rogido MR. An evident view of evidence-based practice in perinatal medicine: absence of evidence is not evidence of absence. *J Pediatr (Rio J)*. 2007;83(5):395-414.

Artigo submetido em 06.07.07, aceito em 26.07.07.

doi 10.2223/JPED.1702

Tabela 1 - Erros do tipo I e tipo II e o poder de um teste

Erro do tipo I	Erro do tipo II	Poder do teste
A conclusão é de que há uma diferença quando, na verdade, não há nenhuma. A probabilidade associada é chamada de alfa.	A conclusão é de que não há diferença quando, na verdade, há uma no mundo real.	A probabilidade de se concluir que há uma diferença quando na verdade existe mesmo uma (geralmente estabelecida em 80%, derivado de 1- beta ou erro do tipo II).

“nenhuma evidência de efeito” e “ausência de evidência” com “nenhuma evidência de ausência”. Todos esses fatores têm um impacto na aplicação dos resultados da pesquisa clínica à prática clínica. Por isso, esta revisão é importante para a prática da neonatologia e pediatria.

Muitas pessoas nos instruíram e inspiraram em muitos ou todos os conceitos que compartilharemos nesta revisão.

Nosso objetivo é apresentar um resumo de alguns “conceitos evidentes da prática baseada em evidências” de uma forma simples e fácil de usar, com um pouco de humor, para lembrarmos que “um dos primeiros sintomas da aproximação de um colapso nervoso é a crença de que o seu trabalho é extremamente importante” (Bertrand Russell). Nossa esperança é de que alguns dos conceitos abordados nesta revisão se tornem de valor para os clínicos em sua prática diária e na prestação de serviços de saúde e, conseqüentemente, que tenham valor para uma criança, um dia, em algum lugar.

Conceitos básicos de inferência estatística

Há muitas falhas importantes na seleção e apresentação dos denominadores e testes estatísticos na literatura médica. Dois livros fazem essa descrição com muito humor e fácil leitura: “Bare Essentials of Biostatistics”¹ e “Biomedical Bestiary”².

Uma ‘doença intelectual’ moderna na literatura é a utilização inapropriada da significância estatística. Muitos autores, revisores e editores “morrem” pela significância estatística e pelo “ $p < 0,05$ ”. Entretanto, alguns deles não entendem bem o seu real significado nem compreendem inteiramente se o método estatístico apropriado foi aplicado corretamente. W. Castle disse: “A maioria dos pesquisadores usa a estatística da mesma maneira que um bêbado usa um poste de luz: mais pelo apoio do que pela iluminação”.

Os clínicos e a inferência estatística às vezes conflituam-se. As inferências estatísticas são usadas para determinar a probabilidade de uma conclusão baseada na análise de dados ser verdadeira e para quantificar o grau de imprecisão na estimativa. Então “jogamos um jogo”. É aceitável que se uma diferença ocorre mais que 5 vezes em 100, há uma grande probabilidade de que a diferença se deva apenas ao acaso. Por outro lado, se a probabilidade da diferença é devida ao acaso em menos de 5% das vezes, dizemos que a diferença “demonstra” que as duas amostras são mesmo diferentes.

Agora você já sabe o que significa $p < 0,05$, ou quase. Se não souber, leia novamente, e continue lendo, por favor. Em resumo, um p com valor $< 0,05$ significa que 95% das vezes, ou mais, os resultados encontrados não são devidos ao acaso. Se a probabilidade do acaso estar afetando os resultados for menor que 1 em 20, então podemos considerar o resultado como significativo, como disse Sir R. Fisher.

Para os erros do tipo I e do tipo II e para o poder estatístico de um teste, veja a Tabela 1 antes de continuar lendo. Quando se usa um nível de p de 0,05, aceitamos que em 5% das vezes podemos estar cometendo um erro do tipo I. Uma vez que os experimentos costumam ser feitos para demonstrar as diferenças, os estatísticos costumam se interessar pela probabilidade de detectar uma diferença real. Esses 5%, no entanto, não são um critério absoluto de verdade. Se, por exemplo, existir um efeito em exatamente $p = 0,049$, ele não desaparece de repente em $p = 0,051$. Rosnow disse de modo eloqüente: “Certamente, Deus adora o 0,06 tanto quanto o 0,05”.

Um erro freqüente na busca por um $p < 0,05$ é usar o teste t de Student ao analisar medidas ou variáveis repetidas. Se o desfecho fosse variação na pressão sanguínea ao longo do tempo e se fossem relatados quatro valores, fazendo comparações estatísticas *dentro* ou *entre* os grupos nesses momentos, repetir os testes t para cada uma das comparações aumenta as chances de diferenças estatísticas significativas, quando na realidade não existe nenhuma. Imagine jogar uma moeda uma vez após a outra. Vamos admitir que da primeira vez deu cara (coroa também tem uma probabilidade de 50%, como você sabe). Como da outra vez, no entanto, a possibilidade de dar coroa é maior em cada jogada sucessiva. Depois cada 3, 4 ou 5 jogadas, a possibilidade de dar coroa aumentou exponencialmente para 84-93% (não perca tempo com matemática; apenas acredite-nos desta vez). De forma semelhante, com a repetição sucessiva do teste t , a probabilidade de conseguir um p significativo, simplesmente por acaso, aumenta para aproximadamente 30% na quinta ou sexta vez, mesmo quando não existe uma verdadeira diferença. O teste t é bastante válido quando são comparadas as médias de dois grupos, mas em se tratando de mais de dois grupos ou quando as medidas forem repetidas em dois ou mais grupos, o teste t torna-se impreciso e não é válido estatisticamente. Fique atento e não aceite um “valor p ” por mais baixo que seja

Tabela 2 - Quando a - significância é insignificante

Fato	Consequência	Clínicos
Tamanho grande da amostra – número enorme de pacientes	Alta probabilidade de haver significância estatística	A magnitude da diferença é importante clinicamente?
$p < 0,05$; diferença de <i>pequena magnitude</i> (2-3 mg/dL) em 2.000 lactentes	Estatisticamente diferente "Parabéns!" Significância clínica prática mínima ou nula.	Suspeitam ou duvidam do estudo. Não mudem a prática ainda!
Necessidade de um maior número de estudos diferentes para apresentar diferença	A diferença real é pequena e provavelmente não é importante	Suspeitam ou duvidam da meta-análise. Não mudem a prática ainda!

J. Mortimer disse: "A falta de crença é um ato de fé: A única coisa da qual podemos ter certeza é a incerteza"

quanto você vir isso na literatura. A análise correta consiste na análise de variância (ANOVA; one-way ou fatorial) com uma de muitas comparações *post hoc*, mas isso é – desculpem – para ser tratado em um artigo de bioestatística.

Significância estatística e importância clínica

A significância que é insignificante

Os ECR que mostram uma diferença significativa entre os tratamentos comparados são comumente chamados de "positivos". *Significância estatística* é um termo probabilístico (a probabilidade de rejeitar uma hipótese nula quando ela é correta; o indício de que a diferença observada é de fato diferente de zero). Muitos clínicos tendem a estabelecer uma equivalência entre um p baixo (*significância estatística*) e algo com *importância ou significância clínica*. Contudo, um " $p < 0,0001$ " não tem nada a ver com a *magnitude ou importância* de uma diferença ou um efeito. Tal magnitude é chamada de *importância ou significância clínica*.

A indústria anuncia uma "descoberta" para a hipoglicemia. Os autores estudaram 2.000 recém-nascidos hipoglicêmicos através de um ensaio randomizado, prospectivo, multicêntrico, duplo-cego e controlado (ERPMDCC). Um grupo recebe o tratamento, o outro recebe placebo; ambos recebem glicose suplementar. A glicemia no grupo de tratamento aumenta a partir de um valor basal (média \pm dp) de 25 ± 8 mg/dl a 37 ± 6 , 46 ± 6 e 52 ± 4 em 30, 60 e 90 minutos. O grupo placebo vai de 26 ± 9 mg/dl a 35 ± 4 e de 43 ± 4 a 50 ± 3 . Os autores usam testes *t* de Student repetidos (evidentemente incorretos), relatando uma diferença em ambos os grupos em comparação aos valores basais ($p < 0,0001$) e uma resposta significativamente melhor no grupo tratado comparado ao grupo placebo em 60 e 90 minutos ($p < 0,001$). Antes de usar esse tratamento nos seus pacientes, e depois de haver lido os parágrafos anteriores, você escreveria para a revista e para os autores solicitando a análise de variância e os valores de p exatos, certo? Os autores agradecem (embora não muito satisfeitos, discutindo entre eles quem é você para mencionar publicamente os seus erros) e, com ressentimento, publicam uma errata com a análise de variância. Eles relatam um

$p = 0,048$. Agora você se sente confortável para usar esse tratamento nos seus pacientes. Espere! Antes de fazê-lo, gaste 2 minutos na Tabela 2 e também se pergunte se os autores mediram a glicose sérica, plasmática ou total. Eles nos disseram como manipularam as amostras e que método foi usado para as medições reais? Com base na Tabela 2 e nessas questões não respondidas, a maioria dos clínicos **não** vai querer expor crianças hipoglicêmicas à nova "descoberta" apesar desse abrangente, "evidente" ERPMDCC. Além disso, esse tratamento pode ser caro e pode-se descobrir, mais tarde, que ele produz efeitos colaterais infrequentes, mas importantes, não analisados a fundo pelo estudo. Para "os resultados clínicos que interessam" e efeitos colaterais infrequentes, mas importantes, por favor, continue lendo!

A não-significância que pode ser significativa

Até agora se sabe que o valor de p maior do que 5% ou $p > 0,05$ "não é estatisticamente significativo". Isso significa que qualquer valor entre 5,1% e 95% das vezes a diferença é devida ao acaso e as amostras não são diferentes. Veja a Tabela 3 para conceitos relacionados e as perguntas que um clínico deveria fazer em tais casos. Se ainda houver interesse no assunto, leia mais tarde sobre os denominadores corretos e os desfechos importantes.

Imagine agora que os autores dos exemplos acima acompanharam 300 das 2.000 crianças hipoglicêmicas até os 5 anos de idade. Usando avaliações cegas e detalhadas do desenvolvimento neurológico e análises cuidadosas dos potenciais fatores de confusão com regressão logística, eles descobriram que alguns dos testes de desenvolvimento e inteligência favorecem o grupo tratado em 7 a 10 pontos e que a incidência de paralisia cerebral (PC) é de 50% nas crianças tratadas (3% *versus* 5,8%). Os autores, ajudados por sua pergunta anterior e pela sugestão de alguns anos atrás, conseguiram obter excelentes análises estatísticas. Eles não relatam diferença estatística, apresentando o valor exato de p de 0,059 e uma razão de chances (RC) para a PC de 0,75 com intervalo de confiança (IC) de 0,67-1,03 que, por cruzar a unidade, não mostra diferença estatística (Figura 1). Além

Tabela 3 - Falta de diferença estatística NÃO é equivalente à falta de significância clínica

O dilema filosófico é que nunca se pode provar a “não existência” de algo

- 1) Quais foram as variáveis de desfecho estudadas e como foi definida a variável?
- 2) Quem foi excluído do estudo?
- 3) A amostra de tamanho “maior” é adequada para ser usada como denominador para essas variáveis de desfecho?
- 4) Qual é a incidência do problema no grupo controle?
- 5) Poderia haver uma amostra melhor para o estudo dessas variáveis de desfecho?
- 6) Um estudo dessa amostra poderia apontar alguma diferença?
- 7) Quando os resultados induzidos pela intervenção são ‘positivos’, esse desfecho é importante?
- 8) Caso positivo, quais os possíveis efeitos adversos para os desfechos importantes analisados pelos autores?
- 9) Se “há diferenças não significativas” nos desfechos ruins, isso significa segurança?
- 10) O que é considerado “importante” pela publicação e o que você, seus pacientes e suas famílias consideram “importante”?

Em questões de saúde pública, devemos ser céticos quanto à ausência de evidência de um resultado benéfico ser uma justificativa válida para a falta de ação. Em questões de saúde pública, devemos ser céticos quanto à ausência de evidência em efeitos adversos (infreqüentes, mas sérios) ser uma justificativa válida para a ação. Tenha cuidado com afirmações do tipo: “Esta prática demonstrou ser segura e deve ser implementada”.

disso, eles não mostram efeitos adversos do tratamento. A pergunta deveria ser: Essa significância não estatística é clinicamente significativa? Você terá que decidir! (Tabela 3). Pessoalmente, não abriríamos mão de 7-10 pontos do nosso QI, nem gostaríamos de ter um risco maior, quase o dobro, de PC. O que pode estar acontecendo aqui é que não deve ter havido cálculo do tamanho da amostra para a magnitude do efeito sobre esses desfechos e que o tamanho da amostra em 5 anos não é grande o suficiente para alcançar significância estatística para a magnitude do efeito encontrado (erro do tipo II ; veja Tabela 1).

No entanto, como clínicos, devemos decidir se os achados apresentam significância clínica e, caso positivo, oferecer o tratamento aos pacientes que confiam a nós os seus cuidados.

Você concorda que, diferentemente da lei e da justiça, uma publicação é “culpada” até que se prove o contrário? Até mesmo o manuscrito que você está lendo agora! Ao avaliarmos e criticarmos uma publicação científica com ceticismo saudável e críticas racionais, estamos criticando a publicação e NÃO os autores. Talvez você tenha outros motivos para criticar os autores, mas isso não tem importância nesta revisão.

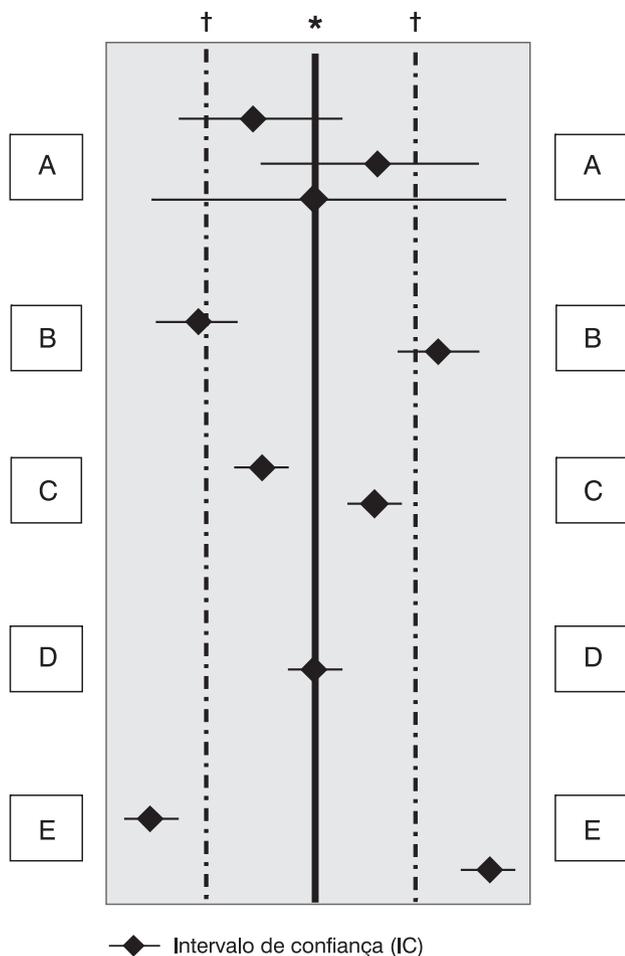
Denominador

A seleção do denominador é crucial em todos os estudos. A escolha de um denominador inadequado invalida parcial ou completamente os resultados, não importando o quão prospectivo, randomizado, controlado ou cego foi o estudo e não importando o quão sofisticadas foram as estatísticas. Os denominadores são essenciais para a incidência ou índices de

doenças, fatores de risco e para a magnitude do impacto ou efeito de uma intervenção. A escolha de um denominador correto está entre os fatores mais importantes nas mãos dos autores. Certamente, os revisores e os editores devem ser a “primeira linha de proteção” quando os autores escolhem os denominadores inadequados ou comparam erroneamente alguns deles. Contudo, os clínicos têm a obrigação de procurar cuidadosamente pelo denominador escolhido e pelos denominadores usados em todas as comparações e decidir o que eles significam, isso se realmente tiverem algum significado. Infelizmente, muitos artigos não fazem comparações válidas, “trocando” de denominadores e/ou usando denominadores incorretos. Veja um exemplo a seguir.

A taxa de câncer de próstata diminui com o passar dos anos em uma comunidade em oposição a muitas outras comunidades nas quais houve um aumento. Note que o denominador usado naquela comunidade foi a população inteira (crianças, homens jovens e mulheres!) Além disso, nos 5 anos anteriores, uma porcentagem de homens acima de 60 anos mudou-se da comunidade por motivo de aposentadoria e por razões climáticas!

Se você procurar cuidadosamente nos artigos que ler, algumas vezes encontrará erros significativos nos denominadores como esse aqui (esperamos que não tão atroz como esse). Então, pergunte-se: “Qual é o denominador? Qual deveria ser o denominador? Qual é o total da população em risco?” No caso do câncer de próstata, certamente não são as crianças, as mulheres ou os jovens adultos. Além disso, tome cuidado e preste bem atenção, uma vez que os autores “lhe aplicam” denominadores diferentes. A questão é identificar



* (Linha vertical sólida) Linha de ausência de efeito: risco relativo, razão de chance = 1; diferença de risco = 0.
 † (Linha vertical tracejada) Limites predefinidos de equivalência (Limites para diferenças importantes).
 Adaptado a partir de Alderson⁷⁰.
 A = O IC cruza a linha de ausência de efeito e um ou outro ou ambos limites de equivalência. Isso significa que há evidências insuficientes para confirmar ou excluir uma diferença importante.
 B = O IC não cruza a linha de ausência de efeito, mas cruza o limite de equivalência. Isso significa que há diferença estatística significativa, mas há incerteza se ela é importante para os pacientes.
 C = O IC não cruza a linha de ausência de efeito e está completamente dentro dos limites de equivalência. Isso significa que há diferença estatística significativa, mas que ela não é importante para os pacientes.
 D = O IC cruza a linha de ausência de efeito e está completamente dentro dos limites de equivalência. Isso significa que não há evidências de uma diferença importante. (Lembre-se: não é a mesma coisa que evidência de nenhuma diferença)
 E = O IC está inteiramente fora dos limites de equivalência. Isso significa uma diferença importante.

Figura 1 - Relação entre intervalo de confiança (IC), linha de ausência de efeito e limites para diferenças importantes (limites predefinidos de equivalência)

claramente sobre quem eles estão falando e comparar isso à população realmente em risco. Por exemplo, nos estudos de displasia broncopulmonar (DBP) e retinopatia da prematuridade (RDP) ou hemorragia intraventricular (HIV) grave, todos os nascidos vivos < 1.500 g são usados como denominador? Fazer isso pode ser semelhante ao exemplo do câncer de próstata acima. Se muitos recém-nascidos morrem antes da sexta

semana de vida ou não são avaliados detalhadamente através do exame oftalmológico ou ecografia cerebral, o denominador está errado e os índices serão falsamente baixos. Quando os denominadores corretos não são usados, a incidência do problema é provavelmente "melhor" do que na realidade! Os denominadores são extremamente importantes quando se faz comparações ou quando se decide mudar os tratamentos.

Numeradores

Você também precisa identificar como o desfecho ('numerador') foi definido. Você deve concordar que definir o DBP como O₂ para > 28 dias não é o mesmo que O₂ em casa, ou que a necessidade de ventilação por > 4 meses.

Número necessário para tratar (NNT) e número necessário para causar dano (NND), redução do risco absoluto (RRA) e redução do risco relativo (RRR)

O NNT é uma forma corriqueira de expressar o número de pacientes que deveriam ser tratados para prevenir um desfecho. Isso é calculado obtendo-se a recíproca de RRA (Tabela 4). Por exemplo, se em um ensaio de prevenção de DBP 13% daqueles tratados e 18% dos controles desenvolvessem DBP, a RRA seria de 5% (18% menos 13%). O NNT seria de 20 (1/5 x 100). Saber que teremos que tratar 20 crianças para prevenir um caso de DBP parece mais útil que uma razão de chances de 0,68 ou uma RRR de 28% (18% menos 13%/18%).

Graças ao NNT, você e os pais de seus pacientes podem entender melhor que um tratamento pode não beneficiar uma criança individualmente, já que para muitos tratamentos, o NNT é > 10. O clínico também deveria se perguntar a respeito do número necessário para causar dano (NND) (Tabela 4). No mundo ideal, adoraríamos a "certeza", ou seja, NNT de 1, NND infinito. Como isso não acontece, temos que tomar decisões baseadas na eficácia da intervenção em termos da redução do risco (Tabela 4) e de possíveis efeitos adversos. Se o NNT fosse relativamente baixo (digamos 6-10), a pergunta mais importante a fazer é "para qual desfecho?" Se o resultado for importante, então um NNT de 10 é sempre melhor que um NNT de 100. Se forem tratadas 100 crianças utilizando-se um NNT de 10, a probabilidade de não se observar nenhum benefício líquido é mais de 1.000 vezes mais baixa, ou seja, melhor do que com um NNT de 100. (Calculado pela distribuição binomial padrão). Não se preocupe, não iremos adicionar fórmulas para você decorar, estamos apenas tentando tornar o conceito claro. Você pode optar por não acreditar. Se você não acreditar, por favor, leia as referências a respeito³⁻⁸. Não podemos expandir o NNT ainda mais, mas dizer que seu enfoque é apenas no tratamento, negligenciando o tipo de placebo ou controle; isso tende a fazer com que o valor real nos ensaios controlados por placebo seja superestimado. Além disso, os NNTs são comumente apresentados como um valor separado, sem IC ou erro padrão (Tabela 4, Figura 1) e isso pode ser insuficiente para a obtenção da "evidência completa".

Tabela 4 - Reduções de risco absoluto e relativo, número necessário para tratar e intervalos de confiança.

Definições e conceitos	Exemplo 1	Exemplo 2
IGC: certifique-se de que é para a população em risco, com denominador bem definido.	IGC : 40%	IGC = 10%
IGT = verifique os denominadores novamente	IGT= 25%	IGT= 5%
RRR = (IGC-IGT)/IGC	RRR= 37,5%	RRR= 50%
RRA = IGC menos IGT ("importância" do tratamento)	RRA= 15%	RRA= 5%
NNT = 1/RRA x 100	NNT = 6	NNT = 20
A IGC no manuscrito é semelhante àquela da população que você trata?	Se a IGC é mais alta, você precisará de um NNT maior na sua unidade.	
Como a IGT se compara à sua população?	Se a IGT é semelhante ou mais alta que a incidência na sua população "não tratada": Não embarque!!	
Com a mesma RRR, se a IGC é baixa, o NNT será alto.	IGC = 0,9%	IGC = 90%
	IGT = 0,4%	IGT= 40%
	RRR = 55%	RRR = 55%
	RRA = 0,5%	RRA = 50%
	NNT = 200	NNT = 2
O NNT muitas vezes não é mostrado com o IC correspondente. Se a RRA para a morte é de 1,3%, o NNT é de aproximadamente 77.	Se o IC para a RRA varia entre 0,1-2,5%: o NNT vai variar entre 1.000 e 40 (!!!)	

RRA = Redução de risco absoluto; IGC = Incidência do problema no grupo controle; IGT = Incidência do problema no grupo de tratamento; NNT = Número necessário para tratar; RRR = Redução de risco relativo; Número necessário para causar dano (NND) = Um efeito colateral devido à exposição que não teria ocorrido se o tratamento não tivesse sido usado. Se a cada 40 crianças tratadas houver um efeito relatado, o NND é 40.

Sugerimos dar uma olhada em um artigo que você leu recentemente e ver se os autores relatam o NNT. Se não o relatam, você poderá fazê-lo sozinho usando a Tabela 4, se os dados apresentados nos resultados claramente descrevem a incidência nos grupos de controle e de tratamento. Se puder fazê-lo, terá uma "visão mais clara ou evidente". Se não puder fazê-lo, tenha cautela.

O quê não funciona e como mostrar isso

Como disse Alderson⁹, Cochrane sugeriu três perguntas básicas a respeito de qualquer intervenção em saúde: "Isso pode funcionar?" "Isso funciona na prática?" e "Será que vale a pena?" Seria maravilhoso se as respostas fossem sempre positivas, mas na vida real, as respostas possíveis seriam "sim", "não tenho certeza", e "não". As regras para optar por um "sim" são relativamente claras e conhecidas, mas pouco foi escrito sobre a decisão de que algo não funciona ou de que não causa efeitos adversos, mesmo que infreqüentes. Vamos dar uma olhada nas questões sobre intervenções e dilemas

relacionados à tentativa de decidir entre uma resposta "não tenho certeza" e "não", à avaliação do que são 'resultados importantes' e ao que fazer quando não temos certeza.

Tentar convencer o público de que um fator não tem efeito ou de que ele não apresenta risco é bastante difícil porque envolve "provar uma asserção negativa"¹⁰. Mesmo com todas as "evidências" que conseguimos juntar, muitas vezes ficamos incertos sobre a escolha do tratamento correto¹¹. É quase sempre impossível, e em muitos casos incorreto, dizer que não há diferença nos efeitos dos tratamentos. Pode-se ter mesmo a evidência da ausência de um efeito ou de uma diferença apenas se todos os estudos *suficientemente* amplos e bem estruturados mostrassem que um tratamento médico, uma exposição ou um não-tratamento não está associado a um desfecho¹². Independentemente de sua excelência, um ERPMDCC não consegue mostrar evidências de ausência. Na prática diária, devemos considerar "outros níveis de evidência" e não apenas o ERPMDCC (Tabela 5). Sempre haverá

Tabela 5 - É melhor obter uma resposta aproximada para a pergunta correta do que uma resposta exata para pergunta nenhuma.

Em sua "atividade diária"	"Níveis adicionais de evidência" sugeridos:
a) Formule a pergunta o mais claramente possível. Há respostas apenas para as perguntas formuladas. Uma pergunta que não foi feita não pode ser respondida.	1) Um ERPMC não é "evidência".
b) Procure pela evidência (Não apenas um ECR, não apenas uma conferência).	2) Ausência de evidência não é evidência de ausência.
c) Avalie a evidência de forma crítica.	3) Evidência de nenhum efeito não é o mesmo que ausência de evidência de efeito.
d) Decida de forma crítica se é aplicável na sua prática ou se deveria ser erradicado da sua prática.	4) Ensaio clínico não-randomizado: há um grupo de controle <i>concomitante</i> e avaliação uniforme da doença/condição em <i>ambos</i> os grupos.
e) Avalie de forma crítica seus próprios resultados.	5) Ensaio clínico prático.

incerteza acerca das estimativas dos efeitos do tratamento; diferenças pequenas, mas importantes nunca podem ser excluídas em um estudo¹³. Alegações de nenhum efeito ou de nenhuma diferença podem levar os clínicos a negar a seus pacientes intervenções com efeitos benéficos importantes ou a expô-los a uma intervenção com efeitos danosos sérios¹⁴. Portanto, alegações de nenhum efeito devem ser muito infrequentes, e quando forem feitas, devemos vê-las com ceticismo. Revisões ou estudos com "alegações de nenhum efeito" ou de "evidência de nenhum efeito" estão errados na maioria das vezes. Expressões do tipo "não reduziu", "não tem efeito" e "não é eficaz" não são geralmente justificáveis e não devem ser permitidas pelos revisores ou editores, já que o que foi mostrado foi "ausência de evidência de efeito" em oposição à "evidência de nenhum efeito". Frases aceitáveis nos artigos deveriam ser: "não foi detectada nenhuma diferença significativa" ou "não há evidências suficientes para apoiar ou refutar".

Ausência de evidência não é evidência de ausência

Há uma "evidente" concepção errônea nestes termos, que não são intercambiáveis. Interpretações errôneas de achados "não-significantes" podem se tornar um grande problema (Tabela 3). Em geral, estudos com um valor $p > 0,05$ costumam mostrar apenas a *ausência de evidência de uma diferença ou ausência de evidência de efeitos negativos*. Eles não podem ser considerados como "evidência de ausência", o que implica *erroneamente* que o estudo mostrou que não há diferença. Interpretar esses ensaios como evidência da falta de eficácia de um tratamento ou evidência de nenhum efeito adverso é "claramente errado e temerário"¹³. Poderíamos usar até mesmo termos mais fortes, mas poderíamos ser acusados de sermos "não-controlados" e conseqüentemente "não evidentes". Basta dizer: Há perigos na má interpretação de

resultados não significativos (Tabela 3). Um exemplo dramático é o tratamento fibrinolítico para prevenção do reinfarto após infarto do miocárdio (IM), que um de nós pode precisar ou já precisou. Dezenove de 24 ERPMDC não "mostraram nenhuma diferença" (" $p > 0,05$ "), levando a um "atraso estatisticamente significativo" antes que o valor verdadeiro da estreptoquinase, o qual realmente existe no mundo real, fosse reconhecido. Nossas desculpas! Nos confundimos e usamos erroneamente o termo!! ...Os estudos não mostraram "nenhuma diferença" como incorretamente descrevemos; eles apenas mostraram ausência de evidência de uma diferença. Mais tarde, a meta-análise mostrou uma redução significativamente alta (22%) na mortalidade. Graças a Deus, cientistas clínicos sérios se deram conta disso; alguns de nós não vamos morrer por causa do reinfarto apesar dos "estudos padrão-ouro!"

Em resumo, quando problemas de saúde pública causam preocupações, devemos olhar com ceticismo a ausência de evidência como justificativa válida para a ação ou falta de ação, já que quando os riscos são pequenos, ou o tamanho da amostra é pequeno, ou o tamanho da amostra é grande, mas o denominador correto para o desfecho em questão não é usado no estudo, os valores p "negativos" possivelmente levarão ao engano.

Em estudos desse tipo (*erroneamente* chamados de "estudos negativos") amplos intervalos de confiança (IC) muitas vezes contam a história ou iluminam a ausência de evidência¹⁵. Em casos como o da estreptoquinase, os IC são possivelmente mais amplos, indicando "incerteza evidente." Portanto, nos casos de ausência de evidência nos ERPMDC, devemos conhecer o IC e fazer avaliações detalhadas dos desfechos importantes desses ensaios. Além disso, precisamos avaliar a magnitude do efeito e o que é importante em qual

Tabela 6 - Desfechos compostos em neonatologia

O desfecho comparado não é um desfecho isolado	Exemplos /conceitos
Estratégia para o controle de riscos concorrentes ou de efeitos precoces que podem concorrer com efeitos tardios	Ter um dos desfechos individuais é suficiente para estar no grupo com o problema.
A variável de desfecho principal é "composta" por dois ou mais desfechos (morte, convulsões, paralisia cerebral)	Você sabe que estar morto ou ter convulsões não é a mesma coisa. (Veja os resultados de tais estudos!)
Quanto mais baixo o índice de sobrevivência, menor o número de crianças com RDP, TPL e efeitos adversos.	Morte prévia e DBP; Morte prévia e RDP. Morte prévia e desfecho de desenvolvimento neurológico adverso
Exemplo de "desfecho composto" em lactentes a termo pode ser encontrado na Referência 16 sobre asfixia e hipotermia Por favor, veja os desfechos individuais incluídos no desfecho composto.	No "mesmo resultado" o autor inclui déficit mínimo de audição e convulsões persistentes. Esses desfechos são bem diferentes. Um NNT de 6 é relatado no estudo para "desfecho composto", sem muitos detalhes, causando incerteza.

situação, observando se os autores descrevem os limites de equivalência decididos antecipadamente. Como mostrado na Figura 1, se o IC está entre esses limites de equivalência, um efeito é considerado pequeno demais para ser importante. Por favor, observe a Figura 1, onde tentamos esclarecer esses conceitos. Acredite; não é fácil. E também não é fácil avaliar claramente o quão importante é uma redução na incidência da síndrome de aspiração de mecônio (SAM) grave e fatal ou na persistência do canal arterial (PCA) grave levando à DBP. Também é difícil dizer quem decide quando tais reduções são importantes. Claro, se os pais de um bebê desnecessária e gravemente afetado tivessem que decidir, eles diriam que a "diferença de 1" seria grande o bastante.

Outras questões que "complicam a evidência" são o quão abrangente é a exposição e qual é a evidência do caso-controle ou estudos epidemiológicos prévios ou subsequentes, o que não deve ser ignorado. Outros fatores que aumentam a incerteza dos resultados no ERPMDCC (ou simplesmente ECR) são as falhas na hora de seguir o protocolo e as perdas não randômicas de seguimento. Como clínicos, não costumamos entender muito bem todos os jargões estatísticos recém mencionados. Autores e revistas precisam relatar resultados incertos de forma clara e deveríamos aumentar nosso ceticismo, tentando incorporar na nossa prática diária alguns dos conceitos descritos, para que não sejamos enganados quando o artigo deixa a impressão de que há evidências de nenhum efeito ou de nenhuma diferença, quando na verdade não há.

Resultados importantes que fazem a diferença

"As coisas mais importantes nunca devem ficar à mercê das menos importantes" (Goethe)

Como definir "desfechos importantes?" A escolha do tratamento deveria ser determinada pelos efeitos nos desfechos que são importantes para os pacientes e seus pais. Alguns até mesmo afirmam que eles também deveriam ser

importantes para a sociedade como um todo. Os desfechos importantes são, felizmente, 'freqüentemente infreqüentes', tais como morte, derrame, efeitos de longo prazo no desenvolvimento cerebral, visão, audição e outros. Portanto, para que haja uma diferença estatística, é necessário uma amostra maior ou um denominador correto da população realmente em risco, ou ambos. É por isso que desfechos "importantes" são comumente pouco estudados. Então, quando da leitura de artigos "evidentes", tente descobrir o que os autores escolheram como principal variável de desfecho. Ela é de importância clínica? Ela é biologicamente confiável? Os efeitos adversos raros e infreqüentes, mas graves, (como morte) foram bem analisados e relatados?

Em resumo, fique atento a resultados "positivos" de "pouca" importância clínica, mesmo se $p < 0,0001$, e a achados "negativos" de um tratamento se nada é claramente mostrado sobre os resultados importantes (mortalidade, morbidade grave). O problema com os "desfechos compostos" é apresentado na Tabela 6¹⁶.

Pesquisa para provar é diferente de pesquisa para melhorar: um ECR não é evidência

O padrão-ouro para a prática baseada em evidências é o ECR. Mas, às vezes, o ouro em um ECR não brilha ou é de baixo quilate. Como disse Jorge Luis Borges "sistemas racionais levados ao extremo de sua racionalidade transformam-se em pesadelos". É claro que ele não estava se referindo especificamente a um ECR; até onde sabemos nenhuma de suas histórias eram randomizadas.

Há coisas que sabemos que sabemos e que são evidentes sem ECRs. É "evidente" para você que a penicilina cura a faringite estreptocócica? Bem, nenhum ECR comprova isso. É evidente que se uma pessoa pula de um avião de uma altura de 1.000 metros e se o pára-quedas não abre, ela não estará bem de saúde depois que bater no solo? (Nenhum ECR evidenciou isso também). É evidente para você que o risco de ocorrer um

acidente com danos para si mesmo ou para outros é mais alto quando se está dirigindo um carro e os freios não funcionam ou funcionam inadequadamente? Se você não sabe que sabe disso e gostaria de fazer um ECR, fique à vontade.

Os ECRs são essenciais para avaliar a eficácia de intervenções clínicas se a cadeia causal entre o agente e o desfecho for relativamente pequena e simples e quando os resultados podem ser seguramente extrapolados para outras situações. Contudo, não deveria ser novidade para ninguém que o “primeiro artigo é intrigante, com os próximos três há uma crescente preocupação, talvez até um pouco de confusão e, depois disso, o que realmente queremos saber é a resposta verdadeira”. Devemos ser cautelosos e não aceitar resultados de um único experimento ou ECR, fazendo uma investigação mais abrangente sob diferentes condições, em outros lugares e em outros períodos.

Não importa o quão bem os ECRs comprovem ou contrariem as alegações terapêuticas, o quão forte seja sua credibilidade em se tratando da busca por evidências, eles são limitados em se tratando da garantia de cuidados adequados¹⁷⁻³⁰. Além disso, os ECRs geralmente são caros e sempre artificiais, realizados em um grupo seletivo e restrito, com critérios de exclusão. Em resumo, os ECRs nunca serão perfeitos porque são realizados por humanos..., para humanos; e são realizados em humanos afetados por condições médicas e patologias que são inevitavelmente heterogêneas. Os ECRs são como bons perfumes franceses, bons para cheirar, mas não para ser ingeridos; ou como vinhos Malbec argentinos: ótimos para se degustar e beber um pouco, mas não para se deixar intoxicar por eles.

Alguns pesquisadores na área da saúde defendem a posição de “adicionar quilates ao padrão-ouro” utilizando um ensaio clínico não randomizado (ECNR) e um ensaio clínico prático (ECP) (Tabela 5)¹⁹. Nos ECRs, os resultados estão sujeitos a alterações nos efeitos em diferentes populações. Por isso, a validade interna e externa dos achados de um ECR pode ser amplamente melhorada através de estudos observacionais com um delineamento de estudo que envolva adequação ou plausibilidade. Além disso, em intervenções de saúde pública e de grande escala, os estudos com delineamentos de plausibilidade são geralmente a única opção possível e podem fornecer evidências válidas de impacto quando os ECR não são adequados²⁰. Há também uma necessidade urgente de ECP que sejam relevantes para os clínicos e tomadores de decisão. Tunis et al. abordaram muito bem como avaliar o valor desses ensaios²¹ e Glasgow et al. fizeram recomendações e deram exemplos de como os ECP devem ser realizados para melhorar a validade externa sem comprometer a interna²².

Em resumo, desenvolver uma base de evidências em saúde pública e decisões práticas requer o uso de dados de estudos de avaliação randomizados e não randomizados. Estudos individuais e estudos em sínteses de pesquisas quantitativas requerem relatos transparentes do estudo, com detalhes e clareza suficientes para perceber as diferenças e

semelhanças entre os estudos realizados na mesma área. O CONSORT (*Consolidated Standards of Reporting Trials*) apresenta diretrizes para relatos transparentes de ECRs. Também há o TREND (*Transparent Reporting of Evaluations with Non-randomized Designs*). Essas diretrizes enfatizam o relato das teorias usadas, o delineamento da pesquisa e descrições de intervenção e condições de comparação, bem como métodos de ajuste para possíveis vieses em estudos não randomizados¹⁹⁻²². Na maioria das vezes, os ECNR e ECP são bem superiores aos ECR.

Deveria ser evidente para todos os clínicos que alguns ECRs não vão ao encontro das expectativas, mesmo com um tamanho da amostra grande, pois muitas vezes não são suficientes para fornecer “evidências evidentes” para a prática diária. Além disso, alguns levam a uma confusão desnecessária e à produção de danos humanos desnecessários, mesmo que infreqüentes ou raros. Nós somos os responsáveis pelos cuidados que proporcionamos, pelos cuidados que nossos pacientes recebem, e não um ECR ou uma “autoridade superior” em uma sociedade pediátrica. Opções de intervenção prática, delineamentos alternativos de pesquisa e representatividade ao nível do paciente são importantes para tratar das implicações clínicas e de políticas a fim de ajudar a reduzir a lacuna entre a pesquisa e a prática.

Exemplos da vida real

Os ECR “evidentes” podem levar a mudanças na prática devido a análises apressadas e ao fato de se “dar as costas” para evidências anteriores. Recomendamos não pular esta seção e passar algum tempo nos exemplos e tabelas como uma forma de auto-reflexão, identificando fatos relacionados nesse resumo, apresentados por muitos acerca da prática baseada em evidências. Muitos demonstram um extremismo contumaz em neonatologia. Nós, nossas crianças, professores, mentores, amigos, cães e gatos e a maioria de nossos inimigos consideram que alguém pode ganhar discernimento na vida (e na prestação de cuidados) com exemplos “evidentes”, aprendendo com eles e com os próprios erros. Mas isso não é evidente uma vez que não há ECR sobre o assunto! Como disse Popper: “*Se respeitarmos a verdade, aprenderemos com os próprios erros através das críticas racionais e da autocrítica*”.

Prevenção da prematuridade

É um objetivo extremamente importante! A Tabela 7²³ é usada para exemplificar isso. Lembre da analogia com o vinho Malbec. Então você poderá decidir com “mais evidências” o que fazer como clínico consciente.

Esteróides neonatais

Esta é uma história bem conhecida e triste para muitos bebês²⁴. Muitos casos vieram à tona na última década. A Tabela 8 mostra os riscos de generalizar a administração de terapias sistêmicas, melhorando, de forma milagrosa, os efeitos de curto prazo em um órgão em um ECR sem uma análise

Tabela 7 - Redução da prematuridade com 17 α Hidroxiprogesterona (17-OH-P) em um ECRPMDC ²³

Perguntas que devemos fazer	Respostas	Comentários
Prevenir a prematuridade é um objetivo extremamente importante?	Sim	Sem dúvida. Contudo, esse é um problema bem abrangente; nem toda a prematuridade é igual.
Como é definida a prematuridade no ECR?	< 37 semanas	O que eles estão dizendo em suas conclusões?
Diminuir o número de nascimentos na 36ª -37ª semana é um desfecho importante "que faz a diferença"?	Talvez	Se os custos, resultados ou efeitos adversos não forem altos.
Sobre quem os autores estão falando?	Alto risco	Se você vir o IGC, você perceberá o alto nível de risco
Qual é a pergunta feita pelos autores? O que eles concluem?	Continue lendo	Se estiver muito interessado, leia a referência 23 para acompanhar o que segue.
Qual foi a incidência de prematuridade na população de alto risco antes do início do estudo?	36%	O índice parece muito, muito alto. Essa população é semelhante à que tratamos?
Qual foi a incidência do problema (IGC)? O que eles dizem?	55%	Muito mais alta que 36% (estranho) Efeito significativamente benéfico para diminuir a prematuridade
Qual foi o índice de prematuridade no grupo tratado (IGT)?	36% (!!)	Alguma surpresa?
O que foi usado como placebo?	Óleo de mamona	O óleo de mamona pode induzir ao trabalho de parto!
É verdade que a 17-OH-P evita a prematuridade ou que o placebo (óleo de mamona) aumentou a prematuridade?	???	A decisão é sua
Qual é a dose de progesterona? Qual foi o número total de injeções semanais?	???	Essa e outras perguntas não foram abordadas no manuscrito!
Qual é o NNT para que a 17-OH-P evite a prematuridade (< 37 semanas)?	NNT 6-7	"Parece bom"
Qual é o NNT se usarmos o IC relatado?	NNT 50-75	Para cada 50-75 mulheres, seria evitado um nascimento <37 semanas NESSA população
Qual é o NNT para evitar uma criança com peso ao nascer < 1.500 g?	NNT 250	Acima de 250 mulheres teriam que ser "tratadas" para evitar um bebê com peso < 1.500 g
E quanto aos efeitos adversos da 17-OH-P?	"nenhum p < 0,05"	Uma ruptura uterina, 5 abortos, 6 mortes fetais, 1 torção testicular. NND?
A 17-OH-P é segura?	Quem sabe	Ausência de evidência (NÃO o contrário)
Qual é a eficácia da 17-OH-P?	Você decide	Agora, como um clínico consciente, você tem "mais evidências" antes de decidir
Você usaria a 17-OH-P para todas as mulheres em risco?	Você decide	Você decide com "mais evidências" o que fazer, como um clínico consciente, antes de usá-la

completa dos desfechos importantes que fazem a diferença para os pacientes e seus familiares.

Dióxido de carbono: o bom, o mau e o feio.

A Tabela 9 apresenta resumidamente algumas das fascinantes histórias acerca do CO₂ neonatal. O extremismo inicialmente tendia para o lado baixo, depois para o lado alto. Há muito tempo atrás, baseados em evidências emergentes, optamos por tentar evitar tanto a hipocarbica como a hipercarbica; recentemente, foi demonstrado que não ser

"extremistas" em relação ao CO₂ pode ser bom para os lactentes⁴⁴. A Tabela 9 deve ser divertida.

CPAP

Os problemas são apresentados resumidamente na Tabela 10. Este tópico poderia ocupar sozinho um artigo. Algumas referências para os interessados.^{27,28,45-47} A prática clínica é a ciência do particular ou dos indivíduos e a filosofia é a ciência das questões das populações. Consideramos que na prática é fundamental avaliar cada criança de acordo com as

Tabela 8 - Esteróides neonatais, doença pulmonar e prática clínica.

Problema	Motivo
História bem conhecida e triste para muitos bebês ²⁴	Os ECRs publicados apresentavam falhas no delineamento ou os desfechos principais apresentavam significância clínica questionável.
Desfechos de curto prazo ('não importantes')	Complacência pulmonar, resistência das vias aéreas, entre outros
Desfechos importantes "que fazem a diferença" não abordados	Não sabemos por quê. Você sabe?
"Nenhum problema" com infecção, hipertensão, hiperglicemia, hemorragia gastrointestinal, crescimento insatisfatório, alterações no metabolismo do cálcio e osteopenia.	"Ausência" de efeitos adversos que não foram relatados ou que não puderam ser detectados devido ao tamanho da amostra (erro do tipo II)
"Nenhum problema" com desfechos importantes a longo prazo	Não citado pelos autores e ignorado pelos clínicos
Foi ignorado problema prévio importante no sistema nervoso central ³¹⁻³³	Menos DNA no cérebro e cerebelo ³¹ , diminuição do perímetro cefálico e leucomalácia periventricular em neonatos humanos ³² e outros ³³
Menos dias em ventilação intermitente, raios X melhores, complacência e resistência das vias aéreas.	Efeitos benéficos dos esteróides neonatais nas inflamações do pulmão, entre outros
"Uso clínico amplo (ineficaz de esteróides neonatais para a prevenção ou tratamento de DBP	Todas acima, os ECRs ³⁴⁻³⁶ e recomendações na literatura e conferências para a prática da "medicina baseada em evidências"
"Devido ao uso de esteróides neonatais, não observamos mais DBP" (diferentes formas, doses, períodos de tempo)	Negação; busca de soluções rápidas; não fazendo as perguntas mencionadas nesta revisão
Quais resultados são importantes para o paciente e para a família?	Muitos. Alguns mais importantes do que outros
NND: 7-9 para obtenção de uma criança com sérias seqüelas do SNC: (PC, desenvolvimento neurológico, microcefalia) ²⁵	Uso amplo de esteróides neonatais com boas intenções e de muitos ECR
"Pouco uso" de esteróides neonatais (2-4% dos lactentes) na University of California San Francisco (UCSF) e Hospital Gonzalez Coro, em Havana.	"Experiência evidente": grupos desenvolvidos que revisaram de forma crítica a literatura "mais antiga" e os ECRs "modernos". Evidência evidente evitou o uso generalizado.
Não repetamos erros semelhantes com outras terapias	O bem estar dos pacientes sob nosso cuidado

"Todos os que bebem desse tratamento se recuperam em um curto período de tempo, exceto aqueles para quem o tratamento não adianta, que acabam morrendo. É óbvio, portanto, que ele falha apenas em casos incuráveis." (Galeno) (não no caso dos esteróides, é claro)

suas necessidades e não generalizar, mesmo se os conceitos parecerem "muito lógicos" ou "evidentes" ou forem repetidos por 'gurus em neonatologia'. Mas não realizamos um ECR.

Síndrome de aspiração de mecônio

Evidentemente (?) desobstruir as vias aéreas é o primeiro passo para a reanimação. A Tabela 11 mostra um ECR que revelou "ausência de evidência" de um efeito benéfico da desobstrução das vias aéreas superiores em neonatos com aspiração de mecônio⁴⁸. Antes de mudar a prática universalmente pode ser valioso fazer algumas perguntas importantes (Tabela 11). É necessário precaução na maioria dos ensaios (chamados inadequadamente de) "negativos" antes da universalização da prática. Conforme mencionado, o "primeiro artigo é intrigante, nos próximos três, há um aumento da preocupação, e, depois disso, o que alguém realmente gostaria de saber é a resposta verdadeira". Então, como tem sido recomendado, a prestação de cuidados não deve mudar com base apenas em um ECR¹⁷⁻²².

Ferro e oxidação

A Tabela 12 trata de um ECR recente, um editorial e um comentário sobre esse estudo^{29, 49,50} no mesmo volume de

uma revista de julho de 2007. Braekke relatou um estudo (chamado inadequadamente de) "negativo". O estudo não avalia nenhum desfecho clínico de longo ou curto prazo, mas os efeitos nos desfechos clínicos que importam para os pacientes e para os pais são os efeitos que devem guiar as escolhas para a prática. As evidências científicas básicas mostram que o excesso de ferro no neonato é muito prejudicial^{30, 51,52} (Tabela 12). Prevenir a deficiência de ferro não significa que devemos usar terapias que possam induzir ao "excesso de ferro". Os extremos em neonatologia não são geralmente bons (Tabela 12).

Persistência do canal arterial (PCA)

Afeta lactentes prematuros, extremamente pequenos. Nenhuma comunicação oral e nenhum ECR "sugerem" nenhuma intervenção. Se decidir não tratar, boa sorte para algumas das crianças não tratadas até que os dados sobre os desfechos que importam sejam disponibilizados. Veja a Tabela 13 para assuntos relacionados.

Clampamento do cordão

Um problema de saúde pública que "afeta" todos os recém nascidos (veja acima sobre ECRs para problemas de saúde

Tabela 9 - Dióxido de carbono no cuidado neonatal

Perguntas a serem feitas	Respostas	Comentários
Os investigadores ou profissionais da saúde foram cegados para o ECR em 1999 ²⁶ ?	Não	Continue lendo.
Houve uma pergunta clara?	Sim	"Ele reduz a duração da ventilação?"
Qual foi o tamanho da amostra?	49	De 114 candidatos
A randomização foi sistemática?	???	65 excluídos (24 devido à "ventilação curta"; 5 "por neonatologistas").
O que <i>não</i> foi considerado como ventilação assistida?	CPAP	Veja abaixo
A CPAP foi usada após a extubação?	Sim	De acordo com a "indicação clínica"
Estes fatores sugerem possível viés sistemático para você?	Você responde	Se não o sugerem, é provável que nada o faça
A pergunta (hipótese) foi respondida?	Sim	Os autores concluíram: a duração da ventilação não foi reduzida
Houve menos DBP ou redução nos dias de internação?	Não	É necessário ter uma amostra maior.
Os desfechos secundários foram analisados nos resultados?	Sim	(Sem cálculo do tamanho da amostra)
Ventilação assistida, frequência respiratória, menor pico de pressão inspiratória no grupo tratado?	Sim	(Lembra do "critério" para a CPAP pós-extubação?)
Qual foi a necessidade de reintubação em < 24 h?	17% tratados <i>versus</i> 28% controles	(Lembra do "critério" para a CPAP pós-extubação?)
Qual a importância dos dois desfechos acima para você, para o paciente e para a família?	Você decide	(Algum possível efeito adverso para a obtenção de tais desfechos?)
Alguma tabela com índice de reintubação total?	Sim	67% no grupo tratado <i>versus</i> 48% no grupo controle
Índice de reintubação para apnéia?	Sim	21% no grupo tratado <i>versus</i> 12% no grupo controle
Os efeitos indesejáveis importantes, como HIV, enterocolite necrosante, RDP, e o acompanhamento a longo prazo são diferentes?	Não	Você sabe ou pode imaginar quantos bebês seriam necessários para que haja uma diferença?
A ausência de evidência é evidência de ausência?	Não	Lembre também do erro tipo II
Um maior estudo foi realizado ³⁷	Sim	Não mostrou melhora em DBP
Optamos por não resumir devido a regras editoriais acerca do número de palavras	Sim	Na vida e nos cuidados clínicos, temos sempre que fazer escolhas!
Você usou "hipercarbia permissiva", seja lá o que isso significa?	"Provavelmente"	Não sabendo o que isso significa, podemos responder sim ou não?
Ao utilizá-la, que resultados foram medidos?	"mmm"	Sem comentários
"Ao usarmos "hipercarbia permissiva", notamos um TPL bem menor"	"mmm"	Sem comentários
Como você pode ter certeza de que não estava usando "hipercarbia iatrogênica"?	Hipercarbia iatrogênica??	Alto CO ₂ com ventilação alveolar inadequada e maior potencial para dano pulmonar
O que todas as evidências disponíveis até hoje dizem?	"mmm"	A hipercarbia não melhora os desfechos importantes
O que alguns estudos básicos selecionados dizem sobre os efeitos adversos do alto CO ₂ ?	Muitas coisas	Alto CO ₂ e alterações no cérebro e olhos em desenvolvimento ³⁸⁻⁴⁰ , entre outros
O que os estudos evidentes recentes mostram sobre os efeitos indesejados importantes ^{41,42} ?	Nada de bom	Maior necessidade de sedação, maior hemorragia intracraniana, piores desfechos a longo prazo
A hipocarbia é ruim!	Sim	Sabemos disso há > 25 anos ⁴³
Devemos ser "extremistas" e usar a hipercarbia para evitar a hipocarbia?	Não fomos antes e não devemos ser agora	Foi demonstrado recentemente que não ser "extremista em termos de CO ₂ " pode ser bom para os lactentes ⁴⁴

Alguém disse uma vez que em ECR a randomização não deve ser "aleatória", não deve ser arbitrária, não sistemática ou deixada ao acaso. O mesmo se aplica a essa prática. (Veja Galen, Tabela 8)

Tabela 10 - Questões (e confusões) relacionadas à - CPAP

Concorda ou discorda?	Nossa resposta	Comentários
A CPAP é uma ferramenta extremamente útil e de eficácia comprovada para os neonatos	Sim	Pergunte a Gregory et al, 1971
A CPAP funciona para quem precisa dela	Sim	* Esperamos que concorde
A CPAP falha ou não é suficiente em casos graves e em muitos bebês extremamente pequenos	Sim	Mesmo em centros com maior experiência (veja a literatura)
A CPAP não é necessária para muitos bebês	Verdadeiro	Veja*
A CPAP é utilizada excessivamente em muitos bebês que não necessitam dela	Verdadeiro	Veja*
Mitos, pesquisas, conferências e crenças um tanto "tendenciosas e de cunho religioso" são uma praga em CPAP	Achamos que sim	Veja*
A "evidência" (ECR) é muito clara com o uso precoce de CPAP "profilático"	Sim	Até o momento: Sem diferença e sem melhora nos desfechos importantes
A evidência não é tão clara com entubação + surfactante e da extubação à CPAP	Achamos que sim	Até o momento: Sem melhora nos desfechos importantes
A evidência acerca dos efeitos adversos e riscos da profilaxia precoce inclui tratamento para bebês que não necessitam dela, com pneumotórax e RDP	Sim	Há mais casos, mesmo que infreqüentes, de: uso tardio de surfactante, cabeças achatadas, narizes machucados, entubação de emergência e reanimação cardiopulmonar, dinheiro para os cuidados com a saúde gasto desnecessariamente.
A CPAP não é o melhor tratamento para o pneumotórax; na realidade, ela aumenta o seu índice	Sim	É "evidente", leia a literatura, dê uma olhada nos seus próprios dados.
Você ouviu ou disse (em relação aos esteróides e ao CO2): "Usamos a CPAP precocemente e bastante, vemos muito menos TLP e nós não temos problemas com a CPAP"	Ouvimos isso repetidamente	Não sabemos o que os neonatologistas vêem ou não vêem e não estamos preocupados se os <i>neonatologistas</i> têm mais ou menos problemas com a CPAP, pelo menos não neste artigo, nem no cuidado clínico neonatal
Use a CPAP, é claro, mas o faça com cautela	Sim	Use se indicado, não porque "nós usamos"
Uma CPAP é igual a qualquer outra CPAP	Não	Bubble, ventilador, Benviste, Aladdin, etc
"Bubble CPAP" é "o melhor"	Não há certeza	Evidências recentes sugerem o contrário
Podemos usar a CPAP com 100% de oxigênio	Nunca	Deve-se misturar, umedecer e aquecer os gases
Fluxo e pressão com CPAP são simples	Não	Não há espaço aqui para escrever a respeito. Seja cauteloso!
A melhor opção é nunca ter que usar um tubo endotraqueal, ou CPAP, ou surfactante.	Concorda?	Nós concordamos
Com a melhoria no cuidado pré-natal, não haverá mais nascimentos prematuros, nem tubos, CPAP ou surfactantes.	Ha!	Evidente alegação "feita pelo homem": se uma alegação for repetida várias vezes, ela "acaba se tornando verdadeira"
Profissionais da saúde criteriosos concordariam com a não entubação ou uso de CPAP quando desnecessário e com o não uso de surfactante tardio quando necessário	Sim	Devemos avaliar cuidadosamente cada lactente, individualmente, e ter muito cuidado para decidir quem precisa do que.

Tabela 11 - Um ensaio randomizado multicêntrico sem cegamento sobre líquido amniótico meconial (LAM)

Perguntas a serem feitas	Resposta	Comentários
A SAM é uma condição heterogênea com amplo espectro de gravidade?	Sim	A população em alto risco para desfechos ruins: LAM espesso - particulado
Qual foi o tamanho da amostra? Onde?	Amplo e multicêntrico	Continue lendo
Quem foi estudado? O que foi usado como denominador?	Todos com LAM	Em aproximadamente 2/3 dos casos o LAM é fino e fluido.
Qual foi a principal variável de desfecho e como a SAM foi definida?	Desconforto respiratório com necessidade de O ₂ > 12 horas	A SAM grave equivale à insuficiência respiratória importante que leva à ventilação intermitente obrigatória, ventilação oscilatória de alta frequência, ECMO e, infelizmente, a algumas mortes
Qual foi a incidência de SAM, conforme definida, que foi usada para calcular o tamanho da amostra?	7%.	A SAM grave ocorre após LAM espesso
Para o que o estudo teve poder estatístico?	diferença de 20%	Em SAM, como definida previamente
Para o que o estudo não teve poder estatístico?	Para diferenças em SAM grave ou em mortalidade	Ambos são resultados importantes e ocorrem após LAM espesso
No tamanho de amostra maior, quantos tiveram LAM espesso?	12%	61% tinham LAM fluido, o qual "evidentemente" não está associado à SAM grave ou à morte
Qual foi o IGC da SAM, conforme definido?	4% (e não 7% como esperado)	Alerta; veja o cálculo do tamanho da amostra.
O que pode acontecer quando o IGC é muito menor do que a incidência usada para o cálculo do tamanho da amostra?	Com pouco poder estatístico para a principal variável de desfecho; erro do tipo II	Tamanho da amostra para 20% de diferença com IGC de 4% pode ser estimado em aproximadamente 4.000
Qual foi a incidência de SAM grave na população estudada?	2%	Para haver diferenças, o tamanho total da amostra foi estimado em aproximadamente 7.000
Qual foi a incidência de mortalidade na população estudada?	0.6%	Para haver diferenças, o tamanho total da amostra foi estimado em > 14.000.
O que pode acontecer quando a incidência de um "desfecho secundário" é pequena?	Com pouco poder estatístico para esses desfechos; erro do tipo II	Incerteza. A amostra é insuficiente para o desfecho importante "que faz diferença"?
Quantas mortes no estudo?	13	Todas no grupo com LAM espesso
Alguma diferença na mortalidade entre os grupos com LAM espesso?	Impossível responder	Tamanho da amostra insuficiente.
Será que o grande denominador usado foi, na verdade, insuficientemente grande?	Talvez(Erro do tipo II)	Pode não haver indivíduos suficientes da população real em risco para desfechos ruins
O que pode ser feito em casos de amplo espectro e heterogeneidade da doença?	Alternativa para tamanho da amostra grande (potencialmente pequeno)	Use como denominador um tamanho da amostra suficiente de lactentes em maior risco
Eu devo mudar a prática com base em um ECR?	Achamos que não	Referências no texto.
O quão seguro seria fazê-lo?	Não há certeza	NND desconhecido
A desobstrução das vias aéreas superiores não é necessária em casos de LAM espesso?	Você responde	Achamos que sabemos a resposta.
Você pode concluir que tal prática não é eficaz para todos os lactentes?	Você responde	Achamos que sabemos a resposta.
Este ECR mostrou a "evidente" ausência de efeitos indesejados, importantes?	Você responde	NND desconhecido (tamanho da amostra insuficiente ou denominador incorreto)
Este estudo está realmente mostrando evidência de NENHUM efeito?	Você responde	Não há simplesmente "nenhuma evidência de efeito"?
Este estudo está mostrando evidência de ausência?	Você responde	Apenas "ausência de evidência"?
Nem todo o bebê necessita sucção antes da passagem de seus ombros, correto?	Sim, e diferentemente de...	"Não é necessária a sucção de nenhum bebê antes da passagem de seus ombros"
A desobstrução das vias aéreas em todos os lactentes com LAM é uma medida "extrema"?	Verdadeiro	A não desobstrução das vias aéreas em nenhum lactente é outra medida extrema
Os neonatologistas são "extremistas" que oscilam de um extremo ao outro?	Você decide	Quando uma oportunidade é perdida, pode ter sido para sempre. Quando o mecônio entra definitivamente nos pulmões, a oportunidade para a prevenção também foi perdida para sempre.
É hora de mudar isso?	Você decide	
"Podemos pisar dentro do mesmo rio duas vezes?"	Você decide	

Tabela 12 - FERRO e oxidação em um ensaio randomizado

Perguntas a serem feitas	Resposta	Comentários
Quem está sendo estudado?	Lactentes prematuros >6 semanas de idade, saudáveis, com leite materno exclusivo, com vitamina E	Sem fatores oxidativos adicionais.
O estudo mostra <i>evidência de nenhum efeito</i> da alta dosagem de ferro?	Não	Mostra “nenhuma evidência de efeito” da alta dosagem de ferro nos componentes oxidativos da urina em um período de tempo bem curto (uma semana apenas) em uma população muito seleta
A “não evidência do efeito” da alta dosagem de ferro nos componentes oxidativos da urina é um “desfecho importante” que realmente importa?	Não	Isso não significa que nada mudou no estado de oxirredução intracelular presente ou futuro.
Uma semana é tempo suficiente para ver alguma alteração na concentração componentes oxidativos da urina?	Pode não ser	
Os achados podem ser generalizados para todos os lactentes sob nossos cuidados?	Na verdade, não; mesmo que os desfechos fossem realmente importantes.	Este estudo não diz nada sobre os componentes da urina após o tratamento com ferro em lactentes mais jovens ou mais doentes, lactentes expostos a outros estresses oxidativos ou a outras dietas.
O estudo avalia a oxidação ou dano oxidativo?	Não	Veja acima. O dano oxidativo não foi estudado neste ECR
<i>A alta dosagem de ferro em lactentes é bem tolerada sem indicação de aumento do dano oxidativo. Isso é verdade?</i>	Incorreto, incompleto e não-evidente.	Esta afirmação é feita no comentário sobre o artigo original.
Estamos preparados para prover “cuidado evidente” administrando precocemente ferro via oral para lactentes prematuros extremamente pequenos, com base em 3 (!!) artigos da mesma revista?	Não estamos	Se você estiver, quanto você administrará e quando começará?
O que publicações relacionadas dizem sobre componentes oxidativos, ferro neonatal, dano oxidativo e toxicidade celular <i>de longa duração</i> ?	Não é somente a produção de componentes oxidativos que é responsável pelo dano oxidativo e pela toxicidade celular <i>de longa duração</i>	A reação dos componentes com ferro, alterando o estado de oxirredução, pode ser a principal responsável pelo dano e pela toxicidade <i>de longa duração</i>
O que mostram os estudos animais e básicos?	A elevação nos níveis de ferro neonatal produz um teor mais elevado de ferro na substância negra adulta, perda celular a longo prazo, exacerbação do dano oxidativo.	O ferro ferroso reage com a H ₂ O ₂ produzindo radicais de hidroxila, danificando as proteínas, os ácidos nucleicos, e os fosfolípidos da membrana.
Este estudo avalia efeitos clínicos de curta e longa duração que importam?	Não	Ele fornece informações importantes, mas não para mudar a prática universalmente.
Que efeitos nos desfechos clínicos devem guiar nossas escolhas práticas?	Aqueles que são importantes para os pacientes e para os pais	Quando nossa prática está baseada nesses desfechos, nos tornamos médicos melhores.
A prevenção dos estados de deficiência de ferro é extremamente importante?	Sim. Não fazê-lo leva a desfechos anormais de desenvolvimento	Realmente evidente! E tem sido assim há décadas
A prevenção da deficiência de ferro significa que devemos usar tratamento que possa induzir ao “excesso de ferro”?	Não	Evitar tal excesso pode ser tão importante quanto evitar a deficiência de ferro.
Lembra das três fases do metabolismo do ferro em neonatos a termo e prematuros?	Você responde. Espaço insuficiente nesta seção.	Se precisar de ajuda e quiser nossa colaboração, por favor, entre em contato.
O quanto de ferro devemos administrar a bebês a termo e quando devemos começar a sua administração?	1 mg/kg/d ; depois de estarem completamente alimentados	A partir de qualquer fonte, pelo menos durante o primeiro ano
O quanto de ferro devemos administrar a bebês prematuros e a partir de quando?	2-6 mg/kg/d. Não começar antes de 4 semanas de vida, nem depois de 8 semanas.	A partir de qualquer fonte, após serem completamente alimentados; mecanismos antioxidantes mais bem desenvolvidos
O que parece sensato?	Evitar a deficiência de ferro e tratamentos que podem causar excesso de ferro	Evitar desfechos de “extremismo oxidativo” potencialmente anormais a longo prazo

Tabela 13 - Questões sobre PCA e clampamento do cordão

Questão	Tratamento tardio ou não tratamento para PCA	Clampamento tardio do cordão
Muito “barulho” sobre	Não tratar uma PCA sintomática em um lactente prematuro doente.	Clampar o cordão umbilical “tardiamente” (1-2 minutos -?) ⁵³⁻⁵⁴
Baseado em que?	“Comunicação oral”/ relatos recentes (Nenhum ECR) “Evidências” suficientes para alterar a prática?	ECR. Evidências suficientes para alterar a prática em problemas de saúde pública?
Com que propósito?	Evitar efeitos piores (não comprovados) dos tratamentos disponíveis <i>versus</i> nenhum tratamento	Evitar a anemia ferropriva (um objetivo importante). Transfusões em prematuros.
Desfechos que importam?	Não é possível achar nenhum cuidadosamente analisado e relatado	Apenas valores hematológicos, ferritina, depósito de ferro, hematócrito e menos anemia.
Qual o tamanho da amostra nos estudos com “nenhum efeito adverso”?	Mínimo	Varia (40-300 lactentes), mas pode não ser suficiente para efeitos adversos raros e sérios.
Evidência de ausência?	Não! ausência de evidência	Não! ausência de evidência
Meta-análise	Não	Sim, várias
NND para desfechos adversos?	Não avaliado cuidadosamente	Não avaliado cuidadosamente
Efeitos adversos menos importantes e potencialmente associados	Dias em CPAP, oxigênio e furosemida, dias em nutrição parenteral e fluidos, hiperbilirrubinemia direta; osteopenia.	Policitemia, hiperbilirrubinemia, desconforto respiratório, volume e carga de ferro
Efeitos adversos mais importantes e potencialmente associados	Ventilação intermitente, DBP grave, semanas de desnutrição extrauterina, perímetro cefálico, RDP grave, enterocolite necrosante grave.	Necessidade de exsanguinotransfusão parcial, enterocolite necrosante, derrame cerebral e hemorragia
“Evidências” anteriores?	Sim, ignoradas	Sim, ignoradas
Alguma preocupação?	Talvez? Não há certeza? Ausência de evidência não é evidência de ausência.	Talvez? Não há certeza? Ausência de evidência não é evidência de ausência.
Pergunta	Você está preparado para “esperar” por uma PCA sintomática em um lactente prematuro? (O bebê está?)	Você está preparado para “esperar” para clampar o cordão? (O bebê está?)
Caso positivo	Quantifique cuidadosamente e compare os índices de todos os possíveis efeitos adversos nos lactentes “tratados tardiamente” com os lactentes com PCA não sintomática ou com aqueles tratados precocemente.	Quantifique cuidadosamente e compare os índices de todos os possíveis efeitos adversos nos lactentes “tratados tardiamente” com outros lactentes
Sugestão/reflexão	PCA persistentemente prolongada (“PCAPP”) pode estar associada a anormalidades pulmonares e de desenvolvimento persistentemente prolongadas (uma “PPDA” diferente, mas importante)	Existem outras maneiras de evitar a deficiência de ferro e o desfecho insatisfatório a longo prazo? Elas estão associadas a um menor potencial de danos em “circunstâncias inocentes”?

Você terá que decidir por conta própria acerca desses dois, como nós tivemos que fazer. Fique ligado!

Boa sorte para os lactentes não tratados ou tratados tardiamente até que os dados sobre os desfechos que importam (mesmo que raros) sejam disponibilizados. Muitas vezes, os problemas ou um problema importante em um lactente são devidos às soluções propostas.

pública). Apenas por diversão, falando sério, “comparamos” alguns aspectos breves dessa intervenção à não intervenção em PCA (Tabela 13). Você tem que decidir sozinho acerca desses dois. Fique ligado aos efeitos adversos raros, mas importantes!

Midazolam

Você consegue encontrar benefícios evidentes para neonatos? O midazolam é “evidentemente” um veneno para

Tabela 14 - Argumentos usados para justificar várias práticas e seus significados reais

Argumento	Possível significado real
1) "Na minha experiência isto funciona"	Sucessiva repetição de erros
2) "Não vemos nenhum problema com este procedimento ou tratamento"	Não estão analisando corretamente os seus dados disponíveis. (Eventos raros, ainda que importantes, são difíceis de quantificar)
3) "Tenho feito isto repetidamente e 'caso após caso' com ótimos resultados"	Vi dois casos, talvez três, e/ou tenho seletiva capacidade de observação. (A negação é freqüente)
4) "Funciona bem para nós"	Ninguém se importa sobre como isto funciona para os provedores de cuidado ("nós")
5) "Nós não temos problemas com este procedimento ou tratamento"	Novamente, alguém se preocupa apenas se os pacientes têm o problema (não os provedores de cuidado)
6) "Nunca vi este problema"	Ou ele /ela não trabalha, não presta atenção, não trata lactentes de risco ou está de férias quando os problemas ocorrem.
7) "O fazemos por causa disso ou daquilo (normalmente um neonatologista "renomado" diz isso)"	Eu não consigo pensar por conta própria e eu sigo opiniões e conferências como "discursos e sermões"
8) "O fazemos porque é o que eles fazem nesse lugar ou universidade (de prestígio)"	E daí? Muitas coisas que não são corretas são feitas por um monte de gente, incluindo presidentes e outros.
9) "É assim que fazemos as coisas aqui, temos feito isso há mais de 10 anos e temos bons resultados"	Não me confunda com os fatos, eu já me decidi. (Além disso, o que dizemos que fazemos acontece com todos os pacientes?)
10) "Não fazemos isto aqui"	Veja acima.
11) "Não há evidências suficientes e mais ECR são necessários"	Parece impressionante (Muito freqüentemente, não sabe o que é evidência evidente e ouviu isso em alguma conferência. Veja também o argumento 12.
12) "Os estudos não são claros"	Leu apenas resumos ou artigos superficialmente (geralmente de 10-15 anos atrás. Veja também o argumento 11.
13) "Onde eu trabalhei antes eles faziam desta maneira"	E daí? Veja argumentos 8-10 acima para o significado real.
14) "Onde está a evidência?"	O que é evidência para aqueles que usam este argumento de forma simplista? Com que freqüência eles se envolvem em discussões críticas?
15) "Ele /ela faz desta forma, mas alguns outros modificam isso quando estão de plantão ou em serviço"	Anarquia. (O que ele /ela faz é mais importante do que o que o paciente necessita?)
16) "Vamos deixar que nosso especialista no tópico se manifeste"	Veja o argumento 9. E o que é um especialista para você? (A opinião de especialistas não é o mesmo que opinião especializada)
17) "Tentamos muito garantir que os pacientes sob nossos cuidados recebam o melhor e mais eficaz tratamento disponível"	Bom! (veja a "evidência" criticamente, analise seus dados, quantifique os efeitos adversos, tente evitar danos)

recém-nascidos, com efeitos colaterais graves no sistema nervoso, além de hemorragia intracraniana e morte ou comprometimento do desenvolvimento neurológico, levando os neurônios a "cometer suicídio"⁵⁴⁻⁵⁶. Se você continuar a administrar o midazolam, você provavelmente não está entre os provedores de cuidados neonatais que tentam, com toda a incerteza e ignorância que temos, utilizar a prática baseada em evidências para melhorar os desfechos dos lactentes, um bebê de cada vez.

Os argumentos que geralmente ouvimos a favor e contra o uso de algumas práticas ainda são, infeliz e evidentemente,

simplistas demais. Com um pouco de humor, seus significados reais possíveis são apresentados na Tabela 14.

Comentários finais e conclusões

Resumimos tópicos importantes relacionados aos achados da pesquisa clínica e sua incorporação aos nossos encontros diários com os pacientes. Esperamos ter esclarecido melhor alguns desafios importantes inerentes à pesquisa nos serviços de saúde. Correlacionamos as principais idéias a muitas aplicações clínicas possíveis e utilizamos exemplos da vida real para enfatizar alguns pontos. Fornecemos ferramentas

para proporcionar uma visão mais “evidente da prática baseada em evidências” e reforçamos o fato de que a ausência de evidência não é evidência de ausência. A interpretação incorreta de um ensaio que não encontrou nenhum efeito significativo como se “não houvesse nenhum efeito” é um dos muitos problemas que surgem no âmbito mais geral da aplicação de evidências obtidas a partir da pesquisa clínica nos cuidados baseados em evidências. A maioria dos ensaios em neonatologia é pequena demais para descartar efeitos de um tamanho amostral que poderiam ser clinicamente importantes e a maioria também pode não apresentar evidências reais dos efeitos adversos. Além disso, as diferenças estatísticas podem não apresentar significância clínica ou se referir a desfechos que realmente não têm muita importância.

Não é fácil encontrar evidências de que alguma terapia ou intervenção é, na verdade, para o melhor, que é eficaz e apresenta (apenas) os efeitos desejados⁵⁷⁻⁷⁰. É mais difícil encontrar evidências de que alguma intervenção não é na verdade necessária em nenhum caso, é ineficaz para todos os pacientes, e que sua falta não tem importantes efeitos adversos, embora raros. Um dos objetivos do *Cochrane Effective Practice and Organization of Care* é ampliar os tipos de estudos incluídos, além dos ECR, e otimizar a validade, generalizabilidade e evidências “do que funciona” sem causar nenhum efeito adverso desnecessário importante: melhorar a prática profissional e a provisão de serviços de saúde eficazes.

A pergunta sobre o que é um bom cuidado não pode ser respondida com o uso de ECR tão somente, independentemente de quão impressionantes sejam suas evidências. A resposta depende do caráter do efeito desejado, do que é mais importante¹⁸. Se há diferenças boas e más em questão, um julgamento de valor é necessário e isso requer uma avaliação cuidadosa, listando e pesando os prós e contras. A tarefa dos pesquisadores ainda seria a de fornecer as evidências que constituem o pano de fundo no qual as escolhas podem tomar forma.

Na prática médica, somos responsáveis por cada um dos nossos pacientes, e não a ciência, os ECR, os autores, os desenvolvedores de diretrizes ou o que um especialista diz ou faz. Provedores de cuidados têm que decidir o que fazer usando evidências, incluindo as incertezas e certamente os desfechos que realmente importam para os pacientes, familiares e sociedade. O que faz cada um de nós mais ou menos responsáveis no cuidado que provemos não é apenas o que nós decidimos e/ou aceitamos fazer, mas também o que nos recusamos a fazer. É nossa responsabilidade como clínicos que cuidam de pacientes manter uma prática que esteja de acordo com esse entendimento.

É necessário criar uma cultura que seja confortável com a estimativa e discussão das incertezas. Esperamos que no futuro mudanças na prática ocorram através de estudos com denominadores corretos e precisão na interpretação e linguagem. Com isso e com a incorporação de *todas* as pesquisas disponíveis (não apenas um ou o último ECR), a incerteza pode

ser reduzida. Como isso vem ocorrendo com maior frequência, poucas diretrizes que são desenvolvidas e implementadas “universalmente” serão consideradas erradas mais tarde, aumentando assim os desfechos importantes que fazem diferença, bem como o bem-estar de mais lactentes.

Esperamos que, após ler (e reler) este manuscrito, você se sinta mais capacitado a fazer escolhas certas para os seus pacientes e que se torne evidente que algumas opiniões especializadas (e não opinião de especialistas) possam se tornar suas também. Para finalizar, desejamos o melhor para a prestação de cuidados que você dá aos seus pacientes, um bebê de cada vez, nesta complexa era da “prática baseada em evidências”.

Agradecimentos com gratidão evidente

Drs B. Griffin e F. Bednarek, Universidade de Massachusetts, mentores “por excelência”, influenciaram positivamente nossa abordagem acerca dos cuidados intensos.

Drs R. Phibbs, G. Gregory e J.A. Kitterman, Neonatologia, CVRI na Universidade da Califórnia, São Francisco, servindo como mentores, ensinando e prestando cuidados na prática diária, influenciaram positivamente nossa abordagem acerca da prática.

Drs J. Comroe, J. Clements, A. Rudolph e S. Glantz, CVRI na Universidade e da Califórnia, São Francisco, contribuíram para o nosso aprendizado sobre cuidados e estatísticas através de seus conselhos, palestras e publicações.

Dr E. Dueñas, Havana, Cuba, Inspirador na organização dos cuidados e melhoria dos desfechos.

Dr J. Sinclair, Universidade McMaster, Canadá, oficinas, publicações palestras. Inspirador.

Dr A. Brann, Universidade Emory, incansável clínico e garimpador de informações.

Drs G.R. Norman e D. Streiner, Canadá, oficina e livro *Biostatistics: the bare essentials*.

Dr P. Alderson, UK Cochrane Centre, seus escritos como modelo.

Outros, muitos, escritos, palestras, abordagem acerca da prática, amizade

Se você leu vagarosa e atentamente, tudo deve ter ficado evidente!

Referências

1. Norman GR, Streiner DL. *Biostatistics: the bare essentials*. 2nd ed. Hamilton: B.C. Decker; 2000.
2. Michael M, Boyce TW, Wilcox AJ. *Biomedical bestiary: an epidemiologic guide to flaws and fallacies in the medical literature*. Boston: Little, Brown; 1984.
3. Alderson P, Deeks JJ. *Concerning: 'The number needed to treat needs an associated odds estimation'*. *J Public Health (Oxf)*. 2004;26:400. Author reply in: *J Public Health (Oxf)*. 2004; 26:401.

4. Aino H, Yanagisawa S, Kamae I. [The number needed to treat needs an associated odds estimation](#). J Public Health (Oxf). 2004;26:84-7. Erratum in: J Public Health (Oxf). 2004;26:218.
5. Artalejo FR, Banegas JR, Artalejo AR, Guallar-Castillo, P. [Number-needed-to treat to prevent one death](#). Lancet. 1998;351:1365.
6. Cook RJ, Sackett DL. [The number needed to treat: a clinically useful measure of treatment effect](#). BMJ. 1995;310: 452-4.
7. Saver JL. [Deriving number-needed-to-treat and number-needed-to-harm from the Saint I trial results](#). Stroke. 2007; 38: 257.
8. de Craen AJ, Vickers AJ, Tijssen JG, Kleijnen J. [Number-needed-to-treat and placebo-controlled trials](#). Lancet. 1998;351: 310.
9. Alderson P, Groves T. [What doesn't work and how to show it](#). BMJ. 2004;328: 473.
10. Hunt A, Millar R. AS science for public understanding. Oxford: Heinemann, 2000.
11. Chalmers I. [Well informed uncertainties about the effects of treatments](#). BMJ. 2004;328: 475-6.
12. Joffe M. ["Evidence of absence" can be important](#). BMJ. 2003; 326: 1267.
13. Altman DG, Bland JM. [Absence of evidence is not evidence of absence](#). BMJ. 1995;311: 485.
14. Alderson P, Chalmers I. [Survey of claims of no effect in abstracts of Cochrane reviews](#). BMJ. 2003;326: 475.
15. Altman DG, Bland JM: [Confidence intervals illuminate the absence of evidence](#). BMJ. 2004;328:1016-7.
16. Shankaran S, Laptook AR, Ehrenkranz RA, Tyson JE, McDonald SA, Donovan EF, et al. [Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy](#). N Engl J Med. 2005; 353:1574-84.
17. Sinclair JC. [Weighing risks and benefits in treating the individual patient](#). Clin Perinatol. 2003;30: 251-68.
18. Mol A. [Proving or improving: on health care research as a form of self-reflection](#). Qual Health Res. 2006;16:405-14.
19. Des Jarlais DC, Lyles C, Crepaz N; TREND Group. [Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement](#). Am J Public Health. 2004;94:361-6.
20. Victora CG, Habicht JP, Bryce J. [Evidence-based public health: moving beyond randomized trials](#). Am J Public Health. 2004; 94: 400-5.
21. Tunis SR, Stryer DB, Clancy CM. [Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy](#). JAMA. 2003;290: 1624-32.
22. Glasgow RE, Magid DJ, Beck A, Ritzwoller D, Estabrooks PA. [Practical clinical trials for translating research to practice: design and measurement recommendations](#). Med Care. 2005; 43: 551-7.
23. Meis PJ, Klebanoff M, Thom E, Dombrowski MP, Sibai B, Moawad AH, et al. [Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate](#). N Engl J Med. 2003; 348: 2379-85.
24. Finer NN, Craft A, Vaucher YE, Clark RH, Sola A. [Postnatal steroids: short-term gain, long-term pain?](#) J Pediatr. 2000; 137: 9-13.
25. Mariani G, Cifuentes J, Carlo WA. [Randomized trial of permissive hypercapnia in preterm infants](#). Pediatrics. 1999;104: 1082-8.
26. Carlo WA, Stark AR, Wright LL, Tyson JE, Papile LA, Shankaran S, et al. [Minimal ventilation to prevent bronchopulmonary dysplasia in extremely-low-birth-weight infants](#). J Pediatr. 2002;141: 370-4.
27. Verder H, Robertson B, Greisen G, Ebbesen F, Albertsen P, Lundstrom K, et al. [Surfactant therapy and nasal continuous positive airway pressure for newborns with respiratory distress syndrome](#). Danish-Swedish Multicenter Study Group. N Engl J Med. 1994;331: 1051-5.
28. Vain NE, Szyld EG, Prudent LM, Wiswell TE, Aguilar AM, Vivas NI. [Oropharyngeal and nasopharyngeal suctioning of meconium-stained neonates before delivery of their shoulders: multicentre, randomised controlled trial](#). Lancet. 2004; 364: 597-602.
29. Braekke K, Bechensteen AG, Halvorsen BL, Blomhoff R, Haaland K, Staff AC. [Oxidative stress markers and antioxidant status after oral iron supplementation to very low birth weight infants](#). J Pediatr. 2007;151: 23-8.
30. Hutton EK, Hassan ES. [Late vs early clamping of the umbilical cord in full-term neonates: systematic review and meta-analysis of controlled trials](#). JAMA. 2007;297: 1241-52.
31. Howard E. [Reductions in size and total DNA of cerebrum and cerebellum in adult mice after corticosterone treatment in infancy](#). Exp Neurol. 1968;22:191-208.
32. Regev R, de Vries LS, Noble-Jamieson CM, Silverman M. [Dexamethasone and increased intracranial echogenicity](#). Lancet. 1987;1: 632-3.
33. Noble-Jamieson CM, Regev R, Silverman M. [Dexamethasone in neonatal chronic lung disease: pulmonary effects and intracranial complications](#). Eur J Pediatr. 1989;148: 365-7.
34. Stark AR, Carlo WA, Tyson JE, Papile LA, Wright LL, Shankaran S, et al. [Adverse effects of early dexamethasone in extremely-low-birth-weight infants](#). National Institute of Child Health and Human Development Neonatal Research Network. N Engl J Med. 2001;344:95-101.
35. Murphy BP, Inder TE, Huppi PS, Warfield S, Zientara GP, Kikinis R, et al. [Impaired cerebral cortical gray matter growth after treatment with dexamethasone for neonatal chronic lung disease](#). Pediatrics. 2001;107: 217-21.
36. Barrington KJ. [The adverse neuro-developmental effects of postnatal steroids in the preterm infant: a systematic review of RCTs](#). BMC Pediatr. 2001;1: 1.
37. Gu XQ, Xue J, Haddad GG. [Effect of chronically elevated CO2 on CA1 neuronal excitability](#). Am J Physiol Cell Physiol. 2004; 287:C691-7.
38. Martin C, Jones M, Martindale J, Mayhew J. [Haemodynamic and neural responses to hypercapnia in the awake rat](#). Eur J Neurosci. 2006;24: 2601-10.
39. Fritz KI, Zubrow A, Mishra OP, Delivoria-Papadopoulos M. [Hypercapnia-induced modifications of neuronal function in the cerebral cortex of newborn piglets](#). Pediatr Res. 2005;57: 299-304.
40. Kaiser JR, Gauss CH, Pont MM, Williams DK. [Hypercapnia during the first 3 days of life is associated with severe intraventricular hemorrhage in very low birth weight infants](#). J Perinatol. 2006; 26: 279-85.
41. Thome UH, Carroll W, Wu TJ, Johnson RB, Roane C, Young D, et al. [Outcome of extremely preterm infants randomized at birth to different PaCO2 targets during the first seven days of life](#). Biol Neonate. 2006;90: 218-25.

42. Sola A, Spitzer A, Morin FC 3rd, Schlueter M, Phibbs RH. [Effects of arterial carbon dioxide tension on newborn lamb's cardiovascular response to rapid hemorrhage](#). *Pediatr Res*. 1983;17:70-6.
43. Fabres J, Carlo WA, Phillips V, Howard G, Ambalavanan N. [Both extremes of arterial carbon dioxide pressure and the magnitude of fluctuations in arterial carbon dioxide pressure are associated with severe intraventricular hemorrhage in preterm infants](#). *Pediatrics*. 2007;119: 299-305.
44. Gregory GA, Kitterman JA, Phibbs RH, Tooley WH, Hamilton WK. [Treatment of the idiopathic respiratory-distress syndrome with continuous positive airway pressure](#). *N Engl J Med*. 1971; 284: 1333-40.
45. Finer NN, Carlo WA, Duara S, Fanaroff AA, Donovan EF, Wright LL, et al. [Delivery room continuous positive airway pressure/positive end-expiratory pressure in extremely low birth weight infants: a feasibility trial](#). *Pediatrics*. 2004;114:651-7.
46. Sandri F, Ancora G, Lanzoni A, Tagliabue P, Colnaghi M, Ventura ML, et al. [Prophylactic nasal continuous positive airways pressure in newborns of 28-31 weeks gestation: multicentre randomised controlled clinical trial](#). *Arch Dis Child Fetal Neonatal Ed*. 2004;89:F394-8.
47. Morley C; COIN Trial Investigators. COIN Trial Abstract at PAS Meeting, Toronto, Canada, May 2007.
48. Jobe AH. Iron for preterm infants. *J Pediatr*. 2007;151:A3.
49. Kling, P.J. [Iron supplementation in prematurity: how much is too much?](#) *J Pediatr*. 2007;151:3-4.
50. Zecca L, Youdim MB, Riederer P, Connor JR, Crichton RR. [Iron, brain ageing and neurodegenerative disorders](#). *Nat Rev Neurosci*. 2004;5: 863-73.
51. Kaur D, Peng J, Chinta SJ, Rajagopalan S, Di Monte DA, Cherny RA, Andersen JK [Increased murine neonatal iron intake results in Parkinson-like neurodegeneration with age](#). *Neurobiol Aging*. 2007;28:907-13.
52. Peng J, Peng Li, Stevenson FF, Doctrow SR, Andersen JK. [Iron and paraquat as synergistic environmental risk factors in sporadic Parkinson's disease accelerate Age-related neurodegeneration](#). *J Neurosci*. 2007;27:6914-92.
53. Chaparro CM, Neufeld LM, Tena Alavez G, Eguia-Líz Cedillo R, Dewey KG. [Effect of timing of umbilical cord clamping on iron status in Mexican infants: a randomised controlled trial](#). *Lancet*. 2006;367:1997-2004.
54. Anand KJ, Barton BA, McIntosh N, Lagercrantz H, Pelausa E, Young TE, et al. [Analgesia and sedation in preterm neonates who require ventilatory support: results from the NOPAIN trial](#). *Neonatal Outcome and Prolonged Analgesia in Neonates*. *Arch Ped Adol Medicine*. 1999;153:331-8.
55. Waisman D, Weintraub Z, Rotschild A, Bental Y [Myoclonic movements in very low birth weight premature infants associated with midazolam intravenous bolus administration](#). *Pediatrics*. 1999;104:579.
56. Olney JW, Young C, Wozniak DF, Jevtovic-Todorovic V, Ikonomidou C. [Do pediatric drugs cause developing neurons to commit suicide?](#) *Trends Pharmacol Sci*. 2004;25: 135-9.
57. Sinclair JC. [Evidence-based therapy in neonatology: distilling the evidence and applying it in practice](#). *Acta Paediatr*. 2004; 93: 1146-52.
58. Sinclair JC, Houghton DE, Bracken MB, Horbar JD, Soll RF. [Cochrane neonatal systematic reviews: a survey of the evidence for neonatal therapies](#). *Clin Perinatol*. 2003;30: 285-304.
59. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. [Evidence based medicine: what it is and what it isn't](#). 1996. *Clin Orthop Relat Res*. 2007;455: 3-5.
60. Sacket DL. [Patients and therapies: getting the two together](#). *N Engl J Med*. 1978;298: 278-9.
61. Sackett DL. [Superiority trials, noninferiority trials, and prisoners of the 2-sided null hypothesis](#). *ACP J Club*. 2004;140:A11.
62. Chalmers I. Proposal to outlaw the term "negative trial". *BMJ*. 1985;290:1002.
63. Chalmers I. [Well informed uncertainties about the effects of treatment](#). *BMJ*. 2004;328:475-6.
64. Fretheim A, Schunemann HJ, Oxman AD. [Improving the use of research evidence in guideline development: 15. Disseminating and implementing guidelines](#). *Health Res Policy Syst*. 2006;4:27.
65. Mainland D. [Is the difference statistically significant? An innocent request for quack treatment](#). *Clin Pharmacol Ther*. 1969; 10: 436-8.
66. Mainland D. Statistical ward rounds --14. "How many patients would we need?" *Clin Pharmacol Ther*. 1969;10:272-81.
67. Mainland D. [Medical statistics-thinking vs arithmetic](#). *J Chronic Dis*. 1982;35: 413-7.
68. Mainland D. [Statistical ritual in clinical journals: is there a cure?--II](#). *Br Med J (Clin Res Ed)*. 1984;288:920-2.
69. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, et al. [Need for expertise based randomised controlled trials](#). *BMJ*. 2005;330: 88.
70. Alderson P. [Absence of evidence is not evidence of absence](#). *BMJ*. 2004;328: 476-7.

Correspondência:

Augusto Sola MD
 MANA and Atlantic Neonatal Research Institute, Division of
 Neonatology, Atlantic Health System,
 Morristown, NJ
 100 Madison Ave
 Morristown, NJ 07960
 Tel.: (+1) 973-971-8985
 E-mail: augustosolane@gmail.com