

17. Como lidar com insuficiente evidência de pesquisa?

Andrew D Oxman^{1}, John N Lavis², Atle Fretheim³ and Simon Lewin⁴*

Oxman AD, Lavis JN, Fretheim A, Lewin S: SUPPORT Tools for evidence-informed health Policymaking (STP). **17.**

Dealing with insufficient research evidence. Health Research Policy and Systems; 2009, 7(Suppl 1):S17

doi:10.1186/1478-4505-7-S1-S17.

<http://www.health-policy-systems.com/content/pdf/1478-4505-7-S1-s17.pdf>

1 Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway

2 Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, and Department of Political Science, McMaster University, 1200 Main St. West, HSC-2D3, Hamilton, ON, Canada, L8N 3Z5

3 Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway; Section for International Health, Institute of General Practice and Community Medicine, Faculty of Medicine, University of Oslo, Norway

4 Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs plass, N-0130 Oslo, Norway; Health Systems Research Unit, Medical Research Council of South Africa

* Autor responsável por comunicações (oxman@online.no)

Esta é a tradução de um artigo publicado no Health Research Policy and Systems, 2009; 7:Supplement 1 (www.health-policy-systems.com/supplements/7/S1).

O uso, a distribuição e a reprodução irrestritas por qualquer meio estão permitidas desde que a fonte seja citada. Podem ser encontrados links das traduções desta série para o espanhol, português, francês e chinês no website do SUPPORT (www.support-collaboration.org). Opiniões sobre como melhorar as ferramentas nesta série são bem-vindas e devem ser encaminhadas para: STP@nokc.no.

A série de artigos foi preparada como parte do projeto SUPPORT, apoiado pelo 6º Programa-Quadro INCO da Comissão Europeia, contrato 031939. A Norad (Norwegian Agency for Development Cooperation), a AHPSR (Alliance for Health Policy and Systems Research) e o Milbank Memorial Fund organizaram um encontro de revisão por pares no qual se discutiu uma versão prévia da série. John Lavis recebeu salário como Canada Research Chair in Knowledge Transfer and Exchange (Catedrático de pesquisa no Canadá para a transferência e troca de conhecimento). A Norad, o satélite norueguês do grupo EPOC (Cochrane Effective Practice and Organisation of Care), o Norwegian Knowledge Centre for the Health Services, a AHPSR, a CHSRF (Canadian Health Services Research Foundation), a EVIPNet (Evidence-Informed Policy Network) e a Organização Pan-Americana da Saúde apoiaram a tradução e difusão dos artigos. Nenhum dos financiadores atuou na elaboração, revisão ou aprovação do conteúdo.

Este artigo foi traduzido para o português por Ocean Translations e contou com o apoio da Canadian Health Services Research Foundation (CHSRF) <http://www.chsrf.ca/>; Centro Rosarino de Estudios Perinatales (CREP) www.crep.org.ar; e Organização Pan-Americana da Saúde (OPAS) (www.paho.org/researchportal).



Resumo

Este artigo faz parte de uma série escrita para as pessoas responsáveis pela tomada de decisões relacionadas a políticas e programas de saúde e para aqueles que dão apoio a estes tomadores de decisão.

Neste artigo, trataremos do problema de tomar de decisões em situações nas quais há evidências insuficientes disponíveis. Geralmente, os formuladores de políticas têm poucas evidências para realmente conhecer os impactos futuros de uma opção de política ou programa de saúde mas devem, ainda assim, tomar decisões. Sugerimos quatro questões a serem consideradas quando não houver evidências suficientes para ter certeza dos impactos da implementação de uma opção. Estas são: 1. Existe uma revisão sistemática dos impactos da opção? 2. A evidência inconclusiva foi mal interpretada como evidência sem efeito? 3. É possível ter certeza de uma decisão mesmo com a falta de evidências? 4. A opção é potencialmente prejudicial, ineficiente ou não compensa o custo?

SOBRE O STP

Este artigo faz parte de uma série escrita para as pessoas responsáveis pela tomada de decisões relacionadas a políticas e programas de saúde e para aqueles que dão apoio a estes tomadores de decisão. A série se destina a ajudar essas pessoas para assegurar que suas decisões sejam devidamente sustentadas pelas melhores evidências de pesquisa disponíveis. As ferramentas SUPPORT e como elas podem ser usadas estão descritas de maneira detalhada na Introdução desta série [1]. Um glossário para toda a série acompanha cada artigo (ver Arquivo adicional 1). Podem ser encontrados links das traduções desta série para o espanhol, português, francês e chinês no website do SUPPORT (www.support-collaboration.org). Opiniões sobre como melhorar as ferramentas nesta série são bem-vindas e devem ser encaminhadas para: STP@nokc.no.

CENÁRIO

O Ministério da Saúde está analisando algumas estratégias para selecionar e contratar profissionais em áreas rurais carentes. O Ministro da Saúde pediu sua opinião sobre tais estratégias. Você encontrou vários artigos que descreve estratégias usadas anteriormente em outros cenários, porém sem avaliações seguras de seus impactos [2].

HISTÓRICO

Neste artigo, apresentamos cinco perguntas que os formuladores de políticas e aqueles que os apóiam podem fazer ao considerar cenários nos quais possa haver evidências insuficientes para avaliar os impactos das opções de política e programa.

É irreal supor que os impactos de um programa ou política de saúde possam ser previstos com um grau de certeza absoluto. Muitos procedimentos administrativos, financeiros e de entrega não foram rigorosamente avaliados. Tampouco os programas, serviços e medicamentos para os quais tais procedimentos oferecem suporte. Contudo, os formuladores de políticas precisam tomar decisões, independentemente da disponibilidade (ou escassez) de evidências que as apoiem.

Neste artigo, nos concentraremos nas decisões tomadas em situações nas quais há poucas evidências disponíveis para saber se uma determinada opção terá os impactos pretendidos ou não pretendidos (e indesejáveis). Erros comuns cometidos quando há poucas evidências disponíveis incluem suposições sobre a evidência sem uma revisão sistemática, confundindo a falta de evidências com evidências sem efeito, supondo que a falta de evidências suficientes implica necessariamente em incertezas sobre uma

decisão e a suposição que fingir certeza é politicamente conveniente. Apresentamos neste artigo quatro questões que podem ajudar a evitar tais enganos.

QUESTÕES A SEREM CONSIDERADAS

As seguintes questões podem ser consideradas caso não haja evidências suficientes que garantam um grau de certeza sobre os impactos da implementação de uma política ou de um programa:

1. Existe uma revisão sistemática dos impactos da opção?
2. A evidência inconclusiva foi mal interpretada como evidência sem efeito?
3. É possível ter certeza de uma decisão mesmo com a falta de evidências?
4. A opção é potencialmente prejudicial, ineficiente ou não compensa o custo?

1. Existe uma revisão sistemática dos impactos da opção?

O primeiro passo para lidar com a falta de evidências é averiguar as evidências que *estão* disponíveis. É arriscado criar suposições sobre a disponibilidade de evidências sem consultar revisões sistemáticas. Considerações referentes à detecção e avaliação crítica de revisões sistemáticas podem ser encontradas nos Artigos 5 e 6 desta série [3,4].

Para muitas questões relacionadas a sistemas de saúde, não é possível encontrar revisões sistemáticas relevantes e atualizadas. É de conhecimento geral, por exemplo, que os profissionais da área de saúde são críticos na concretização dos Objetivos de Desenvolvimento do Milênio (MDGs) e de outras metas da área da saúde. Apesar disso, uma visão geral das revisões sistemáticas de opções que abordam recursos humanos na área da saúde encontrou apenas uma pequena quantidade de evidências de pesquisa sintetizadas de alta qualidade referentes aos efeitos de algumas opções para o aperfeiçoamento dos recursos humanos na área da saúde [5]. Outras visões gerais das revisões encontraram lacunas semelhantes [e.g. 6]. A falta de revisões sistemáticas não implica necessariamente em falta de evidências. Porém, em tais circunstâncias, torna-se difícil para os formuladores de políticas saber qual evidência encontra-se disponível (consulte a Tabela 1, por exemplo).

Talvez seja necessário realizar avaliações rápidas quando houver limitações de tempo ou recursos. Tais avaliações devem ser transparentes quanto aos métodos usados, assim como a quaisquer limitações metodológicas ou dúvidas relacionadas. Deveriam também tratar da necessidade e urgência da realização de uma futura revisão sistemática [7]. A comissão de novas revisões também deve ser levada em consideração, sempre que uma revisão relevante e atualizada de boa qualidade não esteja disponível. Processos adequados deveriam ser estabelecidos, inclusive a definição de prioridades para revisões sistemáticas [8]. Criar e estreitar laços de colaboração internacionais, tais como o Cochrane Collaboration (www.cochrane.org), podem evitar duplicações

desnecessárias de esforços que envolvam a produção de revisões sistemáticas e ajudar a garantir a pronta disponibilidade de revisões atualizadas.

2. A evidência inconclusiva foi mal interpretada como evidência sem efeito?

Outro erro comum encontrado em situações nas quais as evidências não são conclusivas é a má interpretação entre a falta de evidência de um efeito e a ‘evidência sem efeito’ [9]. É errado alegar que uma evidência não conclusiva demonstra que uma política ou um programa tenha sido ‘sem efeito’. ‘Significância estatística’ *não deve* ser confundida com importância.

Quando os resultados não são ‘estatisticamente significativos’ não que dizer que não houve impactos. Geralmente, um limite de 5% é usado para indicar significância estatística. Isso significa que os resultados são considerados ‘estatisticamente não significativos’ se a análise mostra que diferenças tão grandes quanto (ou maiores que) à diferença observada podem ocorrer por mero acaso *mais de* uma em vinte vezes ($p \geq 0,05$). Há, no entanto, dois problemas com tal suposição. Primeiramente, o limite de 5% é arbitrário. Segundo, os resultados ‘estatisticamente não significativos’ (geralmente considerados erroneamente como ‘negativos’), podem ser inconclusivos ou não. A Tabela 2 apresenta uma discussão mais detalhada desse ponto e a Figura 1 ilustra como o uso do termo ‘estatisticamente não significativo’ ou ‘negativo’ pode ser enganoso.

As tendências que são ‘positivas’ (por exemplo, em favor de uma opção) mas ‘estatisticamente não significativas’ são frequentemente descritas como ‘promissoras’, o que pode ser errôneo também. Tendências ‘negativas’ da mesma magnitude, por outro lado, não são tipicamente descritas como ‘sinais de alerta’.

Os formuladores de políticas deveriam saber que pesquisadores normalmente cometem esses tipos de erros. Para evitar enganos, eles deveriam ser cuidadosos quanto às interpretações equivocadas da significância estatística.

3. É possível ter certeza de uma decisão mesmo com a falta de evidências?

Alguns formuladores de políticas talvez concordem com Charlie Brown que afirmava: “Eu sempre estou certo se é uma questão de opinião”. Mas muitos concordariam que uma evidência de alta qualidade oferece uma base bem mais confiável na tomada de decisões. No entanto, deve haver bons motivos para se manter confiante sobre uma decisão mesmo na falta de evidências. Há pouquíssimas evidências de qualidade que demonstrem, por exemplo, que administrar aspirinas a crianças com gripe ou catapora pode causar a síndrome de Reye, doença rara, porém fatal [10]. Apesar das limitações dessa evidência, o US Surgeon General e outros desaconselharam o uso de aspirina nessas circunstâncias. Isso é devido à disponibilidade de paracetamol (acetaminofeno) como uma alternativa de efeito semelhante e de baixo custo que evita que as crianças

sejam colocadas em risco, mesmo na ausência de certeza do grau de risco propriamente dito. Por outro lado, talvez seja razoável acreditar que as políticas ou os programas de alto custo e efeitos potencialmente adversos *não* deveriam ser implementadas sem uma rigorosa avaliação de seus impactos.

4. A opção é potencialmente prejudicial, ineficiente ou não compensa o custo?

“As boas intenções e teorias plausíveis dos profissionais são insuficientes para selecionar políticas e práticas para a proteção, promoção e restabelecimento da saúde. Humildade e incerteza são condições prévias para avaliações imparciais dos efeitos de prescrições e proscições dos formuladores de políticas e profissionais a outras pessoas. Prestaremos serviço ao público de maneira mais responsável e ética quando pesquisas elaboradas para reduzir a probabilidade de vieses e casualidades se tornarem um elemento esperado da prática profissional e da formulação de políticas e não um complemento opcional.” (Iain Chalmers, Editor, the James Lind Library, apresentação no Norwegian Directorate for Health and Social Welfare em 1 de setembro de 2003. Para mais detalhes sobre esses comentários, consulte as Referências bibliográficas [11])

É arriscado não admitir ter incerteza por razões de conveniência política. Como pudemos observar no Artigo 1 dessa série [12], admitir a existência de dados incorretos na sustentação de políticas pode reduzir riscos políticos, pois isso permite que formuladores de políticas criem maneiras para alterar o curso caso tais políticas não obtenham o resultado esperado.

Conforme a citação acima sugere, boas intenções e teorias plausíveis não são suficientes na hora de selecionar políticas e práticas. Isso é verdade para sistemas de saúde e intervenções clínicas. Exemplos de intervenções clínicas consideradas relativamente ineficazes ou prejudiciais após terem sido inicialmente tidas como benéficas e amplamente usadas incluem:

- Soluções de reidratação osmolar altas em vez de baixas para crianças com diarreia [13]
- Diazepam ou fenitoína em vez de sulfato de magnésio para mulheres com eclampsia [14,15]
- Seis ou mais visitas pré-natal em vez de quatro [16]
- Corticosteróides para pacientes com trauma craniano grave [17]
- Albumina em vez de água salgada para reanimação de pacientes críticos [18]
- Terapia de reposição hormonal para reduzir o risco de doenças coronárias e AVCs em mulheres [19]
- Repelentes eletrônicos de mosquitos para evitar picadas de mosquitos infecção por malária [20]

Todas as intervenções citadas acima foram baseadas em teorias subjacentes, evidências indiretas, resultados secundários e estudos de observacionais: ensaios randomizados subsequentemente desaprovaram todas as suposições subjacentes. Tal fato sustenta a afirmação (acima citada) de que com avaliações rigorosas uma *expectativa* em vez de uma *opção* para apoiar decisões sobre a prestação de intervenções clínicas, o público poderá ser atendido de forma mais responsável e ética.

Tais preocupações se aplicam também a sistemas de saúde e intervenções de saúde pública. Exemplos de sistemas de saúde e intervenções de saúde pública amplamente usados e defendidos, porém com chances de serem ineficazes e mais prejudiciais do que benéficos são os seguintes:

- Intervenções na comunidade e em escolas para reduzir o risco de gravidez na adolescência [21]
- Tratamento diretamente observado no caso de tuberculose [22]
- Taxas de usuários por medicamentos essenciais [23]
- Hospitais privados com fins lucrativos em vez de sem fins lucrativos [24]
- Redução da má distribuição ao impor que os médicos trabalhem durante um número mínimo de anos em áreas carentes antes de se especializarem [2]
- Algumas formas de financiamento com base em resultados ou remuneração por desempenho [25]
- Contrato com o setor privado para prestação de serviços de saúde [26]

Recomenda-se muita cautela antes de investir recursos escassos em políticas ou programas que demandam grandes investimentos que não possam ser recuperados [27]. Em caso de grau significativo de incerteza sobre os impactos de tais opções, uma avaliação rigorosa (um estudo piloto, por exemplo) poderá evitar um potencial desperdício de recursos. E embora tais procedimentos pareçam apresentar atrasos desnecessários, Julio Frenk, ex-Ministro da Saúde do México, observou o seguinte: “Tanto politicamente, em termos de prestação de contas àqueles que financiam o sistema, quanto eticamente, em termos do melhor uso possível dos recursos disponíveis, a avaliação é absolutamente crucial” [28]. As decisões tomadas tanto a favor quanto em contra de uma determinada opção poderão, da mesma forma, ter consequências indesejáveis no caso de falta de evidências (consulte a Tabela 3 para ver um exemplo e obter explicações detalhadas). Sustentar a formulação de políticas testando uma opção proposta em uma avaliação bem desenhada oferece uma melhor abordagem.

Quando os julgamentos sobre os efeitos de opções se baseiam em teorias, resultados secundários, estudos limitados de observação, avaliações inadequadas dos impactos, analogias ou provas casuísticas, os formuladores de políticas devem ser cautelosos quanto à implementação destas (consulte exemplo na Tabela 4) [29].

E mesmo quando o grau de incerteza sobre os benefícios de uma opção for pequeno, poderá ainda existir uma incerteza significativa sobre *outras* consequências

potencialmente importantes, inclusive efeitos (prejuízos) e custos não intencionais (consulte exemplo na Tabela 5). Políticas ou programas com lógicas convincentes podem ser, na realidade, prejudiciais.

No caso de opções promissoras, mas para as quais não há evidências suficientes para indicar a probabilidade de danos, ineficácia ou de custo não compensador, uma avaliação bem projetada dos impactos é altamente recomendável. Tal procedimento pode ser realizado antes da implementação do programa ou da política ou ser integrado como parte da implementação propriamente dita. Considerações mais detalhadas referentes ao monitoramento e à avaliação encontram-se no Artigo 18 desta série [29].

CONCLUSÃO

A maioria das políticas e dos programas é complexo e tendem a gerar efeitos diversos. Algumas evidências estarão quase sempre disponíveis com base na experiência adquirida com políticas e programas similares em outros cenários. Entretanto, conforme os Artigos 8 e 9 dessa série, é de crucial importância que os formuladores de políticas considerem o grau de confiabilidade de tal evidência e avaliem a aplicabilidade dos dados encontrados em seu próprio cenário [4,30]. A incerteza quanto ao impacto de políticas e programas em resultados significativos é típica. Na existência de incerteza significativa, erros comuns, como aqueles descritos nesse artigo, deveriam ser evitados.

RECURSOS

Documentos úteis e leituras adicionais

Chalkidou K, Hoy A, Littlejohns P. Making a decision to wait for more evidence: When the National Institute for Health and Clinical Excellence recommends a technology only in the context of research. *J R Soc Med* 2007; 100:453-60.

<http://jrsm.rsmjournals.com/cgi/content/full/100/10/453>

Oxman AD, Bjørndal A, Becerra F, Gonzalez Block MA, Haines A, Hooker Odom C, et al. Helping to ensure well-informed public policy decisions: a framework for mandatory impact evaluation. *Lancet*. In press.

CONFLITO DE INTERESSES

Os autores declaram que não têm interesses conflitantes.

CONTRIBUIÇÕES DOS AUTORES

ADO preparou o primeiro esboço deste artigo. JNL, AF e SL contribuíram no esboço e na revisão do artigo.

AGRADECIMENTOS

Veja na Introdução desta série os agradecimentos àqueles que financiaram e aos que contribuíram com este trabalho. Além disto, gostaríamos de agradecer a Iain Chalmers e Malcolm Maclure pelos comentários úteis em uma versão prévia deste artigo.

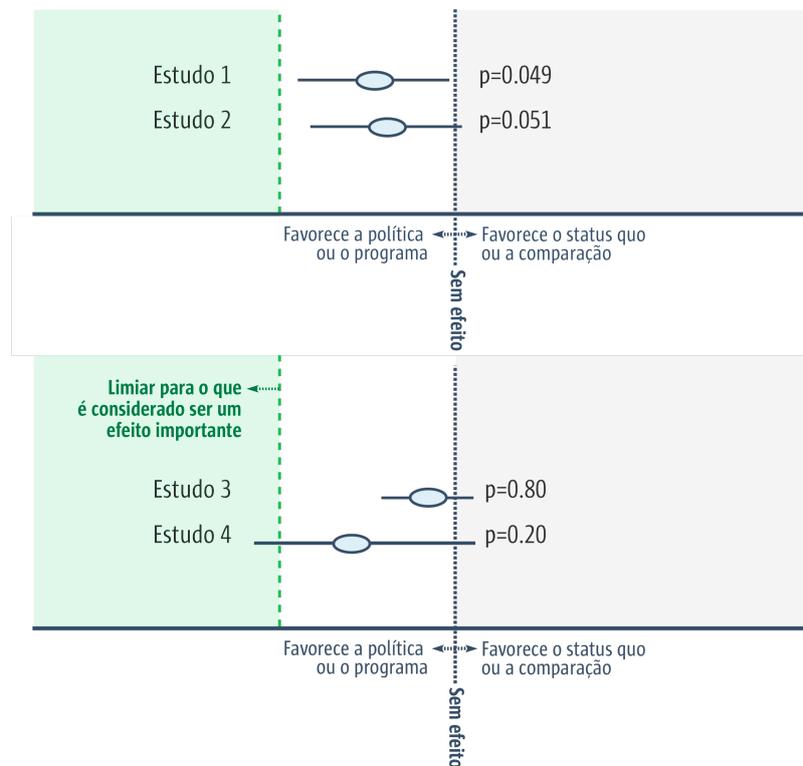
REFERÊNCIAS BIBLIOGRÁFICAS

1. Lavis JN, Oxman AD, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). Introduction.** *Health Res Policy Syst* 2009, 7 (Suppl 1:I1).
2. Grobler LA, Marais BJ, Mabunda S, Marindi P, Reuter H, Volmink J: **Interventions for increasing the proportion of health professionals practising in underserved communities.** *Cochrane Database of Systematic Reviews* 2009, 1:CD005314.
3. Lavis JN, Oxman AD, Grimshaw J, Johansen M, Boyko JA, Lewin S, *et al*: **SUPPORT Tools for evidence-informed health Policymaking (STP). 7. Finding systematic reviews.** *Health Res Policy Syst* 2009, 7 (Suppl 1:S7).
4. Lewin S, Oxman AD, Lavis JN, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 8. Deciding how much confidence to place in a systematic review.** *Health Res Policy Syst* 2009, 7 (Suppl 1:S8).
5. Chopra M, Munro S, Lavis JN, Vist G, Bennett S: **Effects of policy options for human resources for health: an analysis of systematic reviews.** *Lancet* 2008, 371:668-74.
6. Lewin S, Lavis JN, Oxman AD, Bastias G, Chopra M, Ciapponi A *et al*.: **Supporting the delivery of cost-effective interventions in primary health-care systems in low-income and middle-income countries: an overview of systematic reviews.** *Lancet* 2008, 372:928-39.
7. Oxman AD, Schunemann HJ, Fretheim A: **Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence.** *Health Res Policy Syst* 2006, 4:20.
8. Lavis JN, Oxman AD, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 3. Setting priorities for supporting evidence-informed policymaking.** *Health Res Policy Syst* 2009, 7 (Suppl 1:S3).
9. Alderson P, Chalmers I: **Survey of claims of no effect in abstracts of Cochrane reviews.** *BMJ* 2003, 326:475.
10. Centers for Disease Control (CDC): **Surgeon General's advisory on the use of salicylates and Reye syndrome.** *MMWR Morb Mortal Wkly Rep* 1982, 31:289-90.

11. Chalmers I: **If evidence-informed policy works in practice, does it matter if it doesn't work in theory?** *Evidence & Policy* 2005, **1**:227-42.
12. Oxman AD, Lavis JN, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 1. What is evidence-informed policymaking.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S1)**.
13. Hahn S, Kim S, Garner P: **Reduced osmolarity oral rehydration solution for treating dehydration caused by acute diarrhoea in children.** *Cochrane Database Syst Rev* 2002, **1**:CD002847.
14. Duley L, Henderson-Smart D: **Magnesium sulphate versus diazepam for eclampsia.** *Cochrane Database Syst Rev* 2003, **4**:CD000127.
15. Duley L, Henderson-Smart D: **Magnesium sulphate versus phenytoin for eclampsia.** *Cochrane Database Syst Rev* 2003, **4**:CD000128.
16. Villar J, Carroli G, Khan-Neelofur D, Piaggio G, Gulmezoglu M: **Patterns of routine antenatal care for low-risk pregnancy.** *Cochrane Database Syst Rev* 2001, **4**:CD000934.
17. Alderson P, Roberts I: **Corticosteroids for acute traumatic brain injury.** *Cochrane Database Syst Rev* 2005, **1**:CD000196.
18. Liberati A, Moja L, Moschetti I, Gensini GF, Gusinu R: **Human albumin solution for resuscitation and volume expansion in critically ill patients.** *Intern Emerg Med* 2006, **1**:243-5.
19. Farquhar C, Marjoribanks J, Lethaby A, Suckling JA, Lamberts Q: **Long term hormone therapy for perimenopausal and postmenopausal women.** *Cochrane Database Syst Rev* 2009, **2**:CD004143.
20. Enayati AA, Hemingway J, Garner P: **Electronic mosquito repellents for preventing mosquito bites and malaria infection.** *Cochrane Database Syst Rev* 2007, **2**:CD005434.
21. Guyatt GH, DiCenso A, Farewell V, Willan A, Griffith L: **Randomized trials versus observational studies in adolescent pregnancy prevention.** *J Clin Epidemiol* 2000, **53**:167-74.
22. Volmink J, Garner P: **Directly observed therapy for treating tuberculosis.** *Cochrane Database Syst Rev* 2007, **4**:CD003343.
23. Austvoll-Dahlgren A, Aaserud M, Vist G, Ramsay C, Oxman AD, Sturm H, *et al*: **Pharmaceutical policies: effects of cap and co-payment on rational drug use.** *Cochrane Database Syst Rev* 2008, **1**:CD007017.
24. Devereaux PJ, Choi PT, Lacchetti C, Weaver B, Schunemann HJ, Haines T, *et al*: **A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals.** *CMAJ* 2002, **166**:1399-406.
25. Oxman AD, Fretheim A: *An overview of research on the effects of results-based financing.* Report Nr 16-2008. Oslo, Nasjonalt kunnskapssenter for helsetjenesten. 2008.
26. Lagarde M, Palmer N: **The impact of contracting out on access to health services in low and middle income countries.** *Cochrane Library* 2009, **4**:CD008133.

27. Chalkidou K, Hoy A, Littlejohns P: **Making a decision to wait for more evidence: when the National Institute for Health and Clinical Excellence recommends a technology only in the context of research.** *J R Soc Med* 2007, **100**:453-60.
28. Moynihan R, Oxman A, Lavis JN, Paulsen E: *Evidence-Informed Health Policy: Using Research to Make Health Systems Healthier*. Rapport nr. 1-2008. Oslo, Nasjonalt kunnskapssenter for helsetjenesten. 2008.
29. Fretheim A, Oxman AD, Lavis JN, Lewin S: **SUPPORT Tools for evidence-informed health Policymaking (STP). 18. Planning monitoring and evaluation of policies.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S18)**.
30. Lavis JN, Oxman AD, Souza NM, Lewin S, Gruen RL, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 9. Assessing the applicability of the findings of a systematic review.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S9)**.
31. UK Cabinet Office: *Modernising government: presented to parliament by the prime minister and the minister for the Cabinet Office by command of Her Majesty*. Cm 4310. London, Stationary Office. 1999.
32. Macintyre S, Chalmers I, Horton R, Smith R: **Using evidence to inform health policy: case study.** *BMJ* 2001, **322**:222-5.

Figura 1. Dois problemas na classificação de resultados como ‘estatisticamente insignificantes’ ou ‘negativos’



Os pontos em azul na Figura acima indicam o efeito previsto em cada estudo e as linhas horizontais, os intervalos de 95% de confiabilidade. Um intervalo de 95% de confiabilidade significa que podemos ter 95% de certeza de que a verdadeira extensão do efeito está entre o limite de confiabilidade superior e inferior (as duas extremidades das linhas horizontais). Por outro lado, há 5% de probabilidade de que o efeito verdadeiro esteja fora desta variação.

Tabela 1. Uma investigação independente sobre as desigualdades na área da saúde – um exemplo da necessidade de revisões sistemáticas atualizadas para saber que evidências há

Em 1997, o primeiro governo trabalhista britânico estava interessado em reduzir as desigualdades na área da saúde. Para tanto, aconselhou-se com a comunidade de saúde pública sobre como reduzir as desigualdades, mas foram impostos limites claros sobre quais conselhos seriam considerados aceitáveis. O governo exigiu rapidez, mas determinou que tais conselhos fossem baseados em evidências para cumprir o desejo expresso do governo de ter uma política pública baseada em evidências [31]. A saúde pública e outras comunidades responderam com entusiasmo. Uma quantidade

considerável de material foi produzido pela e para a investigação e foram sugeridas várias recomendações [32].

Revisões subsequentes de tais recomendações, entretanto, encontraram pouca evidência da provável ou real efetividade de muitas das recomendações [32]. Houve também uma falta notável de pesquisas adequadas de evidências relevantes ou de tentativas para evitar vieses na maneira na qual as informações foram identificadas, avaliadas e usadas.

Isso não significa que governos não possam desenvolver ou implementar políticas apoiadas em evidências equivocadas. A falta de evidência dificulta, entretanto, que decidam prioridades. A rapidez com que os pesquisadores recomendam políticas tendo pouco conhecimento sobre a provável efetividade das mesmas torna esse processo ainda mais difícil.

Essa investigação no Reino Unido poderia ter sido mais fácil se revisões sistemáticas atualizadas estivessem disponíveis. Além disso, um sistema que garanta que as recomendações resultantes da investigação fossem revisadas regularmente como novas informações e evidências emergentes de revisões sistemáticas atualizadas teria ajudado a garantir que ajustes nas políticas tivessem sido realizados. Tal fato teria também ajudado a evitar dificuldades futuras e semelhantes quando investigações similares fossem realizadas, ou políticas análogas consideradas em outras jurisdições. Redes internacionais, como The Cochrane Collaboration (www.cochrane.org) (cujo foco é saúde) e a Campbell Collaboration (www.campbellcollaboration.org) (cujo foco é educação, crime e justiça e bem-estar social) têm estruturas para preparar e manter revisões sistemáticas atualizadas e assim facilitar o uso mais efetivo de evidências.

O investimento de recursos públicas em pesquisas primárias tem sido substancial e se mantido dessa forma. Contudo, o retorno permanece bem menor do que o esperado e os resultados dispersos em vez de sintetizados. Indivíduos envolvidos em tarefas e calendários semelhantes àqueles da investigação britânica teriam sido muito mais beneficiados se revisões sistemáticas atualizadas estivessem rapidamente disponibilizadas. Quanto ao desenvolvimento de políticas de saúde, não existem respostas inequívocas à questão “O que funciona?” Uma revisão sistemática é o melhor ponto de partida para descobrir o que é conhecido.

Tabela 2. ‘Estatística não significativa’

A Figura 1 ilustra dois problemas que surgem quando os resultados são classificados como ‘estatisticamente não significativos’ ou ‘negativos’:

1. *A classificação é baseada em um limite arbitrário.* Os resultados do Estudo 1, por exemplo, são ligeiramente diferentes dos do Estudo 2. Contudo, com o uso do limite convencional de $P < 0,05$, os resultados do Estudo 1 são classificados como ‘estatisticamente significativos’ e os resultados do Estudo 2 como ‘estatisticamente não significativos’.
2. *‘Resultados estatisticamente não significativos podem ser inconclusivos ou não.* Se a linha vertical curta de cor verde na figura abaixo indicar o menor efeito considerado importante, os resultados do Estudo 3 serão conclusivos, uma vez que um impacto importante é altamente improvável. Os resultados do Estudo 4 serão categorizados como ‘inconclusivos’ uma vez que um impacto importante não seria improvável (o intervalo de confiabilidade de 95% ultrapassa o limiar do que é considerado como um efeito importante). Ambos os resultados, entretanto, poderiam ser considerados como ‘estatisticamente não significativos’ ou ‘negativos’.

Tabela 3. Dizer “não” ou “sim” em vez de “apenas no contexto de uma avaliação” e suas consequências

Todos os países enfrentam limitações de recursos. Por esse motivo, no Reino Unido, por exemplo, o National Institute for Health and Clinical Excellence (NICE) tem oficialmente como princípio recomendar que, na existência de incertezas significativas sobre os efeitos de uma intervenção, tal intervenção deve ser usada somente no contexto de uma pesquisa [27]. 16 (aproximadamente 4%) das recomendações sobre a avaliação de tecnologias do NICE publicadas entre 1999 e início de 2007 aconselharam o uso de uma tecnologia apenas no contexto de pesquisas. Tomar decisões erradas dizendo “não” ou “sim” a uma tecnologia ignorando tal fato apresenta resumidamente as consequências abaixo (ver Referência bibliográfica [27] para mais detalhes).

Dizer “não” em vez de “apenas em pesquisa” e suas consequências

- O acesso a tecnologias potencialmente efetivas e promissoras é negado aos pacientes
- Ocorrem Atrasos na elaboração de evidências em áreas importantes tendo como resultado um impacto geral negativo na área da saúde

Dizer “sim” em vez de “apenas em pesquisa” e suas consequências

- O acesso a intervenções ineficazes ou potencialmente prejudiciais e improcedentes é promovido
- Qualquer pesquisa futura ou em curso é rigorosamente dificultada Talvez questões importantes sobre eficácia e relação custo-benefício não sejam nunca respondidas
- Poucos recursos serão desperdiçados

- Reverter um “sim” por qualquer evidência futura compromete a credibilidade e torna a implementação mais difícil

Tabela 4. Exemplo de uma intervenção potencialmente ineficaz ou prejudicial que foi altamente promovida com base em evidências insuficientes

Há medicação eficaz para tuberculose disponível desde os anos 40. Apesar disso, essa doença continua matando dois milhões de pessoas por ano, sobretudo em países de baixa renda. Indivíduos acometidos de tuberculose necessitam de tratamento de seis a oito meses. A dificuldade que muitos têm em completar o tratamento é a maior barreira para erradicar a doença. A baixa adesão ao tratamento pode causar infecção prolongada, resistência às drogas, reincidência ou mesmo morte. Consequentemente, o tratamento incompleto oferece um grande risco tanto para o indivíduo quanto para a comunidade como um todo.

O DOT (Directly observed therapy, tratamento diretamente observado) visa aumentar a adesão de indivíduos ao tratamento de tuberculose por meio de funcionários da área da saúde, membros da família ou da comunidade para que observem de modo direto os pacientes submetidos à medicação contra a tuberculose. O DOT é uma vantagem em potencial pois a probabilidade da adesão aumentar é maior em função de um monitoramento mais próximo das pessoas e do processo social envolvendo a pressão do grupo. Uma das desvantagens potenciais é que esse tratamento foge dos modelos de comunicação de adesão com ênfase na cooperação entre paciente e prestador de serviços, e retoma a abordagem médica tradicional na qual o paciente recebe conselhos e tratamento com passividade. Além disso, as implicações de recursos são substanciais para uma política desse tipo, particularmente em países de média ou baixa renda nos quais a carga pode ser maior. O DOT pode tornar a adesão ainda pior quando aplicado rigidamente em um cenário autoritário, ou quando os pacientes precisam viajar longas distâncias para a supervisão de seus tratamentos.

A Organização Mundial da Saúde e outras organizações vêm promovendo ativamente o DOT desde os anos 80, geralmente como parte de um programa de gerenciamento abrangente de atendimento à tuberculose conhecido como DOTS (directly observed therapy, short course, tratamento diretamente observado, curso rápido), uma estratégia composta de cinco elementos para o controle da tuberculose. Embora a estratégia como um todo pareça segura, há uma incerteza significativa quanto ao DOT ser um elemento chave do DOTS. Quando o DOTS foi lançado, a eficácia do DOT foi totalmente evidenciada através de estudos de observação e nenhuma avaliação randomizada dos impactos do DOT foi realizada. Subsequentemente, 11 ensaios randomizados compararam o DOT a uma autoadministração e descobriram que o DOT

não melhorava a adesão ao tratamento, apesar dos recursos substanciais exigidos e suas outras desvantagens [22].

Tabela 5. Exemplo de incertezas significativas sobre prejuízos importantes em potencial

Embora haja quase nenhuma dúvida de que os incentivos financeiros, quando suficientemente grandes, podem mudar comportamentos, eles podem também causar comportamentos inesperados. O custo tanto dos incentivos como de sua administração também pode ser substancial [25]. Os efeitos inesperados observados da remuneração por desempenho (pagamento pela obtenção de resultados bem definidos) incluem:

- *Comportamentos inesperados*: As transferências condicionais de dinheiro (CCT, conditional cash transfers) fizeram com que algumas mães mantivessem seus filhos desnutridos para manter a elegibilidade. Um aumento da fertilidade entre 2% e 4%, observado em outro estudo, pode ter ocorrido devido ao fato de apenas gestantes serem elegíveis para o subsídio CCT
- *Distorções*: Incentivos financeiros podem fazer com que os recebedores ignorem outras funções importantes
- *Jogos*: Os incentivos financeiros podem se tornar um mero jogo (mudanças em relatórios em vez de mudanças esperadas na prática)
- *Corrupção*: Os incentivos financeiros podem ser roubados ou utilizados incorretamente, se administrados inadequadamente
- *Escolha a dedo*: Os incentivos de desempenho para prestadores de serviços podem influenciar o acesso dos pacientes a tratamentos de saúde ao mudar o grau de empenho dos funcionários ou das organizações da área da saúde em cuidar de tais pacientes, populações carentes e pacientes mais difíceis
- *Aumentando a diferença entre ricos e pobres*: Os incentivos de desempenho para prestadores de serviços podem aumentar a lacuna de recursos existente entre organizações que atendem pacientes carentes e aquelas que não
- *Dependência de incentivos financeiros*: Contar com incentivos pode fomentar a dependência neles. Se o comportamento dos prestadores de serviços não for inerente, poderá diminuir ou desaparecer quando tais incentivos cessarem ou novos forem introduzidos
- *Desmoralização*: Os incentivos financeiros podem causar sentimentos de injustiça e desmoralização em situações nas quais, por exemplo, os profissionais com contratos de curto prazo recebam mais incentivos financeiros do que aqueles com práticas a longo prazo ou, quando houver favoritismo
- *Burocracia*: Os esquemas financeiros com base em resultados podem ter custos administrativos substanciais associados ao controle de desempenho e ao gerenciamento de pagamentos dos incentivos financeiros