Ferramentas SUPPORT para a elaboração de políticas de saúde baseadas em evidências (STP)

16. O uso de evidências de pesquisa para avaliar os prós e contras das políticas

Andrew D Oxman[1*], *John N Lavis*[2], *Simon Lewin*[3] *and Atle Fretheim*[4]

Oxman AD, Lavis JN, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed Policymaking in health 16: Using research evidence in balancing the pros and cons of policies**. Health Research Policy and Systems 2009, 7(Suppl 1):S1. http://www.health-policy-systems.com/content/7/S1/S16

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; Health Systems Research Unit, Medical Research Council of South Africa
4 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

* 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, abordamos o uso de evidências para informar julgamentos sobre a avaliação dos prós e contras das opções de políticas e programas. Sugerimos cinco questões que podem ser consideradas ao fazer estes julgamentos. Estas são: 1. Que opções estão sendo comparadas? 2. Quais são os resultados potenciais mais importantes das opções sendo comparadas? 3. Qual é a melhor estimativa de impacto das opções sendo comparadas para cada resultado importante? 4. Qual o grau de confiabilidade dos formuladores de políticas e de outros nos impactos estimados? 5. É um modelo econômico formal susceptível de facilitar processos de tomada de decisões?

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

Você trabalha no Ministério da Saúde. O Ministro da Saúde pediu que você apresentasse um resumo dos benefícios esperados, prejuízos e custos de uma importante mudança na política de saúde que está sendo avaliada.

HISTÓRICO

Neste artigo, sugerimos cinco perguntas que os formuladores de políticas e aqueles que os apóiam podem fazer quando consideram como garantir que os julgamentos dos prós e contras das políticas de saúde e opções de programa sejam bem informados por meio de evidências de pesquisa. Tais questões podem ser feitas, por exemplo, em cenários como o descrito acima.

Não se tomam decisões somente com base em pesquisas [2]. Julgamentos são sempre exigidos, incluindo aqueles sobre quais evidências usar, como interpretar estas evidências, e nossa confiança nelas. E mais importante ainda, as decisões sobre opções exigem julgamentos para saber se as consequências desejáveis e antecipadas pesam mais do que as consequências indesejáveis. (consulte a Figura 1) [3]. Além de fazer julgamentos sobre o tamanho provável dos impactos, os processos de tomada de decisões exigem julgamentos sobre a importância dos impactos, os recursos que são exigidos para implementar a opção [4], e a extensão na qual a opção é uma prioridade com relação às outras formas nas quais esses recursos poderiam ser usados.

Seria simples tomar uma decisão se a opção tivesse diversos e amplos benefícios com poucas desvantagens e custos mínimos, se tivéssemos certeza sobre a evidência e a importância dos benefícios, e se a opção fosse realmente uma prioridade. Infelizmente,

isto quase nunca é assim. Frequentemente, os impactos esperados e os custos são incertos, e julgamentos complexos e difíceis devem ser feitos.

As questões que propomos aqui não reduzem a necessidade de julgamentos. Contudo, mais considerações e discussões sistemáticas destas questões poderiam ajudar a garantir que considerações importantes não fossem negligenciadas e que julgamentos fossem bem informados. Isto poderia também ajudar a resolver desacordos ou pelo menos ajudar a fornecer esclarecimentos. Se estes julgamentos fossem feitos com transparência, poderiam ajudar outros a entender a razão por trás das decisões políticas de saúde.

Preparar e usar um balancete (como explicado na Tabela 1 e abordada nas primeiras quatros questões discutidas abaixo) pode facilitar tomadas de decisões bem informadas. Às vezes, o uso de um modelo econômico formal, tal como uma análise da relação custo-benefício, pode ser útil também. Este último tema é abordado na quinta pergunta deste artigo. As considerações sugeridas aqui são baseadas no trabalho do GRADE Working Group [5]. Embora o foco do Group tenha sido principalmente nas diretrizes práticas clínicas, sua abordagem sobre decisões relativas a intervenções clínicas podem também ser aplicadas a políticas e programas [6].

QUESTÕES A SEREM CONSIDERADAS

As cinco perguntas seguintes podem ser usadas para guiar o uso de evidências para informar julgamentos dos prós e contras das opções de políticas e programas de saúde:

1. Que opções estão sendo comparadas?
2. Quais são os resultados potenciais mais importantes das opções sendo comparadas?
3. Qual é a melhor estimativa de impacto das opções sendo comparadas para cada resultado importante?
4. Qual o grau de confiabilidade dos formuladores de políticas e de outros nos impactos estimados?
5. É um modelo econômico formal susceptível de facilitar processos de tomada de decisões?

As primeiras quatro perguntas têm a intenção de orientar o uso do balancete no processo de tomada de decisões políticas. Responder a questão final pode ajudar a garantir que os recursos escassos usados em análises econômicas completas sejam aplicados onde mais necessário.

De preferência, os balancetes (e modelos econômicos) deveriam ser criados por pesquisadores ou pela equipe de apoio técnico juntamente com formuladores de políticas. Eles também deveriam se fundamentar em revisões sistemáticas pelas mesmas razões descritas em outro lugar que realcem a importância das revisões sistemáticas em geral [7]. Não consideraremos os diversos julgamentos detalhados que

devem ser feitos quando da construção do balancete, já que foram abordados em outro lugar [8]. Os formuladores de políticas raramente estão, se é que estiveram alguma vez, em uma posição na qual são exigidos a fazer, por conta própria, todos estes julgamentos. No entanto, mesmo em situações nas quais exista apoio técnico competente para preparar o balancete, é importante que os formuladores de políticas saibam o que procurar e quais questões perguntar. Isto garante que os balancetes possam ser usados de maneira judiciosa para informar as decisões pelas quais os formuladores de políticas são responsáveis.

1. Que opções estão sendo comparadas?

Quando um balancete é usado, como o mostrado na Tabela 2, a primeira consideração é a necessidade de identificar quais opções estão sendo comparadas. Geralmente, isto não é tão simples como parece (veja a Tabela 3, por exemplo). Aqueles que preparam um balancete devem decidir entre ambas as opções que estão sendo consideradas *e* a opção comparativa. Tipicamente, a comparação é o status quo. Contudo, o status quo pode variar entre cenários. Por isso, é preciso determinar quais características do status quo são:

- Fundamentais – de tal modo que pesquisas com comparações sem estas mesmas características sejam excluídas
- Importantes, mas não fundamentais – de tal modo que pesquisas com comparações sem estas mesmas características sejam excluídas, mas com menor certeza de que os resultados seriam os mesmos no cenário escolhido, e
- Irrelevantes – de tal modo que estaríamos confiantes de que os resultados seriam provavelmente os mesmos no cenário escolhido

Também é necessário fazer estes mesmos julgamentos sobre as opções sendo consideradas: quais das suas características são fundamentais, importantes ou irrelevantes em termos de como podem influenciar os impactos prováveis?

2. Quais são os resultados potenciais mais importantes das opções sendo comparadas?

Geralmente, os formuladores de políticas são motivados pelo desejo de servir às pessoas que representam e deveriam estar interessados primeiramente nos impactos das opções políticas e programas nos resultados que são importantes para aqueles atingidos (consulte, por exemplo, a Tabela 4). Estes incluem resultados de saúde, acesso a – ou utilização de – serviços de saúde, efeitos não planejados(prejuízos), e uso de recursos (custos ou economias) (consulte a Figura 1). Outras consequências importantes incluem a distribuição e equidade dos benefícios e custos [9], e os efeitos propagados a outros setores. Consequências éticas, como aquelas relacionadas à redução da autonomia das pessoas, podem ser igualmente importantes.

Ser explícito sobre quais resultados são importantes pode ajudar a garantir que as consequências importantes de uma opção não sejam negligenciadas. Isto também pode ajudar a garantir que consequências irrelevantes não sejam superestimadas. Isto é particularmente importante para os resultados secundários – por exemplo, resultados que não são importantes em si mesmo. Eles são considerados importantes porque acredita-se que refletem resultados importantes. Por exemplo, as pessoas, geralmente, não consideram a pressão arterial como uma preocupação importante. O que torna a questão da pressão arterial importante é sua associação a riscos de derrames, ataques cardíacos e morte, todos de *suma* importância para as pessoas. Portanto, quando se consideram opções dirigidas à hipertensão (ou outros fatores de riscos cardiovasculares), as decisões deveriam ser baseadas nos impactos destas opções nos resultados importantes (doença cardiovascular). A evidência isolada dos impactos na pressão arterial é somente uma forma de evidência indireta dos impactos nas doenças cardiovasculares.

3. Qual é a melhor estimativa de impacto das opções sendo comparadas para cada resultado importante?

Decidir se os impactos desejáveis de uma opção valem os impactos indesejáveis exige uma estimativa de quão importantes estes diferentes impactos serão (e suas consequências econômicas). De forma ideal, isto deveria ser como uma comparação entre o que se poderia esperar para cada resultado importante se uma opção *tivesse* sido implementada, e o que poderia ser esperado se *não tivesse sido* implementada– ou o que se poderia esperar se uma outra opção tivesse sido implementada (consulte a Tabela 5, por exemplo) . Também é útil saber quão precisa é cada estimativa – por exemplo, qual o intervalo de confiabilidade de cada estimativa (isto é explicado em mais detalhes na Tabela 6).

É importante que os tomadores de decisão reconheçam as diferenças entre as estimativas dos efeitos que são apresentados como efeitos *relativos* e aqueles que são apresentados como efeitos *absolutos*. Pacientes, profissionais de saúde e pessoas que tomam decisões sobre políticas e programas de saúde estão mais propensos a usar uma intervenção se seus efeitos são informados como relativos do que se são informados como absolutos [10] Por exemplo, um estudo informou que 61% de uma amostragem com profissionais de saúde na Austrália concordaram em implementar um programa de triagem de câncer colorretal que reduziria a taxa de mortalidade por câncer de intestino para 17% (a redução de risco relativo). Em comparação, somente 24% dos profissionais de saúde concordaram em implementar um programa que produziria uma redução absoluta de mortes por câncer de intestino de 0,4% (a redução de risco absoluto) [11]. Ambas as estimativas foram, de fato, do mesmo programa (para uma explicação da diferença entre efeitos relativos e absolutos, consulte a Tabela 4 no Artigo 10 desta série [9]).

4. Qual o grau de confiabilidade dos formuladores de políticas e de outros nos impactos estimados?

Seis fatores podem reduzir nossa confiança nas estimativas dos impactos de uma política ou um programa [12]:

- Um desenho de estudo inadequado
- Outras limitações de estudo
- Imprecisão
- Resultados inconsistentes
- Aspecto indireto das evidências
- Viés de publicação

Uma avaliação destes fatores é inevitavelmente técnica. Os formuladores de políticas não necessitam ter uma compreensão detalhada destes fatores ou saber como eles são avaliados. Mas tantos os formuladores de políticas quanto sua equipe de apoio técnico ainda podem se beneficiar se entenderem a importância de considerar estes fatores.

Estudos nos quais um programa é atribuído de forma randomizada reduzem o risco de diferenças desconhecidas ou não mensuradas entre os grupos sendo comparados. Isto dá uma maior certeza de que os impactos são atribuíveis ao programa e não a algum outro fator [13-15]. Os desenhos de estudos que não usam atribuições randomizadas podem explicar somente diferenças que são mensuradas. Por exemplo, um estudo no qual comunidades são atribuídas de forma randomizada a uma opção de programa ou política, tal como o licenciamento a vendedores de cigarro, forneceria evidências mais convincentes dos impactos de uma opção do que um estudo faria se comparasse comunidades que decidiram por si mesmas se implementariam uma determinada opção. Isto acontece porque as comunidades que decidem implementar uma opção tendem a se diferenciarem daquelas que não o fazem de formas que poderiam ter um impacto nos resultados de interesse (neste caso, a prevalência do consumo de cigarros). Por esta razão, seria impossível saber se as diferenças nos resultados foram devido à opção de política ou de programa ou devido àquelas entre as comunidades.

Outras limitações do estudo podem afetar tanto as avaliações de impacto randomizadas quanto as não randomizadas. Dados incompletos ou as medidas não confiáveis dos resultados, por exemplo, podem aumentar o risco de uma estimativa ser tendenciosa e, por isso, reduzir o grau de confiança nas estimativas derivadas.

Imprecisão (conforme indicado por um amplo intervalo de confiança) também diminui a confiabilidade na exclusão da chance como fator, formando algumas diferenças observadas nos resultados entre grupos comparados e, consequentemente, nossa confiabilidade no efeito estimado. (A Tabela 6 explica o conceito de intervalos de confiabilidade mais detalhadamente).

Se estudos diferentes da mesma opção de política ou programa tiverem resultados inconsistentes e não houver explicações convincentes de tais diferenças, também haverá menor confiabilidade no conhecimento dos impactos esperados decorrentes da implementação da opção.

Há várias formas nas quais estudos podem não ser diretamente relevantes a uma determinada questão e, por isso, gera menor confiabilidade nos resultados. Como observado acima, se um resultado indiretamente relevante (tal como, pressão arterial) for medido no lugar de um resultado importante (doença cardiovascular), haverá menor confiabilidade nos impactos dos resultados importantes (para os quais o resultado indireto é secundário). Se somente comparações *indiretas* forem fornecidas, a confiabilidade também será menor. Teremos menor confiabilidade nos estudos de uma opção que não apresenta comparações, por exemplo, entre a opção comparada a um controle (sem nenhuma intervenção) e estudos de uma opção diferente comparada a um controle. Outras formas na qual evidências podem incluir de forma indireta diferenças entre um estudo e o cenário de interesse:

- As características da população
- A opção sendo considerada, ou
- O status quo ou a opção de comparação

Estudos que descobrem efeitos estatisticamente significantes têm geralmente maior probabilidade de serem publicados do que aqueles que não [16]. Quando tal 'viés de publicação' parecer provável, a confiabilidade nas estimativas de estudos publicados isolados pode também ver-se reduzida. Os vieses de publicação deveriam ser considerados em casos nos quais houvesse um número de estudos pequenos, especialmente se patrocinados pela -indústria, ou se os pesquisadores fossem conhecidos por compartilharem outros conflitos de interesses similares.

Em resumo, avaliações da 'qualidade' ou robustez da evidência, e confiabilidade nas estimativas dos possíveis impactos das opções, dependem da considerações de *todos* os fatores observados acima. Embora não haja regras estabelecidas para avaliar estes fatores, os julgamentos relacionados à qualidade das evidências que abordam explicitamente cada fato ajudam a reduzir a probabilidade de fatores importantes serem negligenciados. Elas também ajudam a reduzir a probabilidade de avaliações tendenciosas da evidência (consulte a Tabela 7, por exemplo). Usando uma abordagem sistemática e transparente, tais como o método GRADE (consulte a Tabela 8), fica mais fácil verificar os julgamentos feitos [5].

5. É um modelo econômico formal susceptível de facilitar processos de tomada de decisões?

Modelos econômicos formais, tais como análises da relação custo-benefício e análises de custo-utilidade, podem ajudar a informar julgamentos sobre o equilíbrio entre as consequências desejáveis e indesejáveis de uma opção [17]. Os modelos econômicos

podem ser valiosos para processos decisórios complexos e para testar quão sensível uma decisão é no que se refere a estimativas ou suposições importantes. Um modelo, no entanto, só é tão bom quanto os dados nos quais se baseia. Quando estimativas de benefícios, prejuízos ou uso de recursos procedem de evidências de baixa qualidade, os resultados serão necessariamente altamente especulativos (a Tabela 9 fornece um exemplo disto).

É mais provável que um modelo econômico completo ajude a informar uma decisão quando:

- Haja uma grande diferença nos recursos consumidos entre as opções comparadas
- Grandes investimentos de capital sejam exigidos, como construção de novas instalações
- Haja incertezas sobre se os benefícios líquidos compensam os custos incrementais
- Haja evidências de boa qualidade com relação ao consumo de recursos

Um modelo econômico também pode ser usado para esclarecer necessidades de informações ao explorar a sensibilidade de uma análise em uma variedade de estimativas plausíveis.

Infelizmente, análises da relação custo-benefício publicadas, especialmente aquelas relacionadas a drogas, têm grande probabilidade de serem imperfeitas ou tendenciosas. Elas também são específicas a um cenário determinado, que pode diferir de importantes formas do cenário de interesse [18]. Os formuladores de políticas devem, portanto, considerar desenvolver seus próprios modelos econômicos. Para tanto, eles devem ter os recursos e a experiência necessários.

CONCLUSÃO

As decisões políticas são informadas por avaliações feitas do equilíbrio entre os prós e contras das opções. Como recomendamos, estas deveriam ser feitas de forma sistemática e transparente. Quando o benefício líquido (por exemplo, diferenças entre as consequências desejáveis e indesejáveis) é grande em relação aos custos, somos mais confiantes na decisão. Quando o benefício líquido é menor em relação aos custos, essa confiança diminui.

Geralmente, quanto menos confiança tivermos nos possíveis impactos de uma opção, menos confiantes seremos ao decidir o que fazer. Há exceções: primeiramente, nosso grau de confiança nos impactos de algo pode ser tão pequeno que torna-se fácil decidir não fazê-lo.

segundo, mesmo se houver pouca confiabilidade nos benefícios de uma determinada opção, poderá ser fácil decidir fazer algo simples porque haverá pouco ou nenhum risco de prejuízo, não custará muito e poderá apresentar benefícios. Muitos tipos de

informação relacionadas à saúde poderiam ser categorizadas como tais. Os formuladores de políticas, no entanto, devem ser cautelosos em pensar que políticas e programas aparentemente inofensivos não podem trazer prejuízos [19]. Mesmo algo tão simples quanto fornecer informações de saúde pode, na realidade, ser fatal [20]. Isto é demonstrado pelo conselho dado às mães em muitos países por quase 50 anos: que bebês deveriam dormir na posição prona. O conselho aparentemente inofensivo causou dezenas de milhares de mortes pela síndrome da morte súbita do lactente [21].

Finalmente, apesar de incertezas importantes sobre os possíveis impactos de uma política ou programa, pode ser fácil chegar a uma decisão de que algo promissor só deve ser feito no contexto de uma avaliação bem projetada de seus impactos [22].

Mesmo quando estamos confiantes sobre os impactos de uma política ou programa, pode não ser uma prioridade implementá-la. Nosso grau de confiança é um fator essencial no momento de decidir o que fazer e até que ponto fazê-lo é uma prioridade. Outros fatores adicionais (como aqueles descritos na Tabela 10) também podem determinar se a implementação de uma política ou um programa é prioridade ou não.

RECURSOS

Documentos úteis e leituras adicionais

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ, and the GRADE Working Group. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336:924-6

Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, and the GRADE Working Group. What is 'quality of evidence' and why is it important to clinicians? BMJ 2008; 336:995-8

Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Vist GE, Schunemann HJ, and the GRADE Working Group. Incorporating considerations of resource use. BMJ 2008; 336:1170-3

Links para websites

Sumários do SUPPORT: www.support-collaboration.org/index.htm – sumários concisos dos prós e contras das políticas e dos programas de saúde para países de baixa -e média renda com base em revisões sistemáticas

GRADE Working Group: www.gradeworkinggroup.org – o Grading of Recommendations Assessment, Development and Evaluation (ou GRADE) Working

Group desenvolveu uma abordagem para classificar a qualidade das evidências e a solidez das recomendações clínicas

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 Benjamin Djulbegovic e dois revisores anônimos 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. 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.

3. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al: **Going from evidence to recommendations.** *BMJ* 2008, **336:**1049-51.

4. Oxman AD, Fretheim A, Lavis JN, Lewin S: **SUPPORT Tools for evidence-informed health Policymaking (STP). 12. Finding and using research evidence about resource use and costs.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S12)**.

5. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al: **GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.** *BMJ* 2008, **336:**924-6.

6. The SUPPORT Collaboration: *How SUPPORT Summaries are prepared.* [http://www.support-collaboration.org/summaries/methods.htm]. 2009.

7. 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)**.

8. Guyatt G, Oxman A, Kunz R, Vist G, Brozek J, Norris S, et al: **GRADE guidelines: 1. Introduction – GRADE evidence profiles and Summary of Findings tables**. J Clin Epidemiol. 2009. In Press

9. Oxman AD, Lavis JN, Lewin S, Fretheim A: **SUPPORT Tools for evidence-informed health Policymaking (STP). 10. Taking equity into consideration when assessing the findings of a systematic review.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S10)**.

10. Akl E, Oxman A, Herrin J, Vist G, Terrenato I, Sperati F, et al: **Using alternative statistical formats for presenting risks and risk reductions**. Cochrane Database of Systematic Reviews. 2009. In Press

11. Adily A, Ward J: **Evidence based practice in population health: a regional survey to inform workforce development and organisational change.** *J Epidemiol Community Health* 2004, **58:**455-60.

12. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ: **What is "quality of evidence" and why is it important to clinicians?** *BMJ* 2008, **336:**995-8.

13. Kleijnen J, Gotzsche P, Kunz R, Oxman A, Chalmers I: **So what's so special about randomisation.** In *Non-Random Reflections on Health Care Research: On the 25th Anniversary Of Archie Cochrane's Effectiveness and Efficiency*. Edited by Chalmers I, Maynard A. London: BMJ Publishers; 1997:93-106.

14. Kunz R, Oxman AD: **The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials.** *BMJ* 1998, **317:**1185-90.

15. Savedoff WD, Levine R, Birdsall N: *When will we ever learn? Improving lives through impact evaluation*. Washington DC, Center for Global Development. 2006.

16. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K: **Publication bias in clinical trials due to statistical significance or direction of trial results.** *Cochrane Database Syst Rev* 2009, **1:**MR000006.

17. Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, *et al*: **Incorporating considerations of resources use into grading recommendations.** *BMJ* 2008, **336:**1170-3.

18. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL: **Evaluation of conflict of interest in economic analyses of new drugs used in oncology.** *JAMA* 1999, **282:**1453-7.

19. Macintyre S, Petticrew M: **Good intentions and received wisdom are not enough.** *J Epidemiol Community Health* 2000, **54:**802-3.

20. Chalmers I: **Invalid health information is potentially lethal.** *BMJ* 2001, **322:**998.

21. Gilbert R, Salanti G, Harden M, See S: **Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002.** *Int J Epidemiol* 2005, **34:**874-87.

22. Oxman AD, Lavis JN, Fretheim A, Lewin S: **SUPPORT Tools for evidence-informed health Policymaking (STP). 17. Dealing with insufficient research evidence.** *Health Res Policy Syst* 2009, **7 (Suppl 1:S17)**.

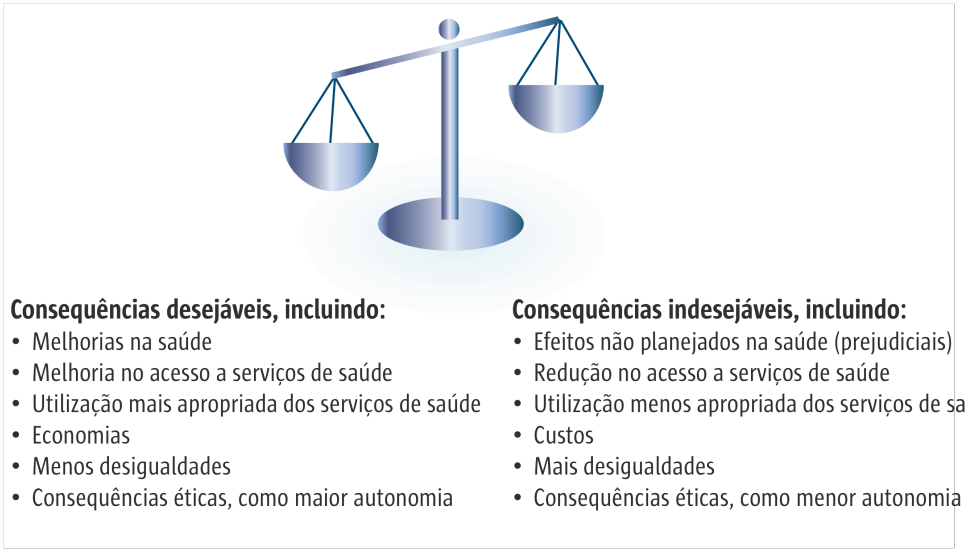
23. Eddy DM: **Comparing benefits and harms: the balance sheet.** *JAMA* 1990, **263:**2493, 2498, 2501.

24. Stead LF, Lancaster T: **Interventions for preventing tobacco sales to minors.** *Cochrane Database Syst Rev* 2005, **1:**CD001497.

25. Kapiriri L, Norheim OF, Martin DK: **Priority setting at the micro-, meso- and macro-levels in Canada, Norway and Uganda.** *Health Policy* 2007, **82:**78-94.

26. Aaserud M, Trommald M, Oxman AD, Innvaer S: **[Evaluation of reimbursement applications for new drugs].** *Tidsskr Nor Laegeforen* 2002, **122:**2619-23.

Figura 1. Ponderação dos prós e contras das políticas e dos programas de saúde*



* As decisões sobre as opções de políticas e programas de saúde exigem julgamentos sobre se as consequências desejáveis de uma opção compensam as consequências indesejáveis

Tabela 1. Os prós e contras de balancetes

Um balancete é uma maneira simples, mas poderosa de apresentar as vantagens e desvantagens de diferentes opções, inclusive opções de políticas [17,23]. Nesta seção descrevemos as evidências e os julgamentos necessários para preparar e usar um balancete, como o mostrado na Tabela 2. Também descrevemos as vantagens de usar um balancete comparado ao uso de julgamentos não sistemáticos e não transparentes de especialistas.

O objetivo de um balancete é ajudar os formuladores de políticas a desenvolver um entendimento preciso das consequências importantes das opções sendo comparadas. Os balancetes ajudam a conseguir isto de várias formas. Primeiramente, eles condensam as informações mais importantes possibilitando, deste modo, considerações eficientes. Segundo, os balancetes se concentram nos resultados mais importantes. Isso aumenta a probabilidade de que os tomadores de decisão obterão uma percepção precisa do que é conhecido sobre os impactos das opções sendo consideradas e suas consequências importantes. Terceiro, o ato de criar um balancete é uma mecanismo útil para organizar pensamentos, estruturar análises de evidências e focar discussões. Quarto, os balancetes podem ajudar a desenvolver julgamentos mais explícitos sobre quais são as consequências mais importantes das opções de políticas, as evidências implícitas, e os julgamentos subsequentes sobre o balanço entre as vantagens e desvantagem relativas das várias opções. E por último, os balancetes

podem fornecer a outros tomadores de decisão informações 'não processadas, ajudando-os, deste modo, a aplicar seus próprio julgamentos sobre conflitos entre as consequências desejáveis e as indesejáveis.

Mas há duas limitações importantes que também precisam ser consideradas quando se usam balancetes no processo decisório. Primeiramente, quando há difíceis conflitos entre os vários resultados, os julgamentos podem exigir um alto nível de informações processadas pelos formuladores de decisões. Segundo, quando se ponderam diferentes resultados, os julgamentos de valor utilizados pelos formuladores de decisões podem permanecer implícitos. O modelo econômico formal pode ajudar a lidar com estas limitações ao tornar quaisquer suposições implícitas (inclusive julgamentos de valor) mais explícitas. Isso permite o uso de análises de sensibilidade para explorar os efeitos tanto de incertezas quanto de suposições variantes nos resultados.

Tabela 2. O licenciamento a vendedores de cigarro deveria estar condicionado à não venda de cigarros a menores?

População: Menores (conforme definido pelo limite de idade legal)
Cenário: Europa
Intervenções: Licença para vendedores de cigarro + verificações de conformidade[1]
Comparação: Sem licenciamento ou verificações de conformidade

Resultados	Impacto			Número de estudos	Qualidade das evidências (GRADE)[2]
	Pessimista	Melhor estimativa	Otimista		
Redução do número de fumantes por ano	0	?	1.650 no país (população 4,5 milhões)	4	⊕◯◯◯ Muito baixa[3]
Anos de vida salvos por ano	0	?	9.240 no país (população 4,5 milhões)	4	⊕◯◯◯ Muito baixa[4]
Custo por ano	€10,5 milhões (3 controles por ano)	?	€7,2 milhões (1 controle por ano + controle interno)	0	⊕◯◯◯ Muito baixa[5]

[1] A lei de licenciamento proposta no país europeu em questão exigiria que os vendedores tivessem uma licença para vender cigarros. As opções de políticas que foram consideradas incluíram três verificações de conformidade por ano, e uma por ano juntamente com controle interno. As verificações de conformidade (mediante a tentativa de um adolescente de comprar cigarros) são feitas para garantir a não venda de cigarros a menores. A penalidade pelo não cumprimento é a perda da licença de venda. O controle interno exige que os próprios vendedores tenham rotinas para controlar a venda de cigarros a menores

[2] Consulte a Tabela 8

[3] A revisão sistemática usada como base para este sumário (que não foi usada no laudo pericial o qual nos referimos nas tabelas subsequentes), incluiu um ensaio randomizado relevante e três estudos controlados antes e depois com limitações importantes. Houve um alto risco de viés para os impactos estimados na prevalência do consumo de cigarros. Inconsistências importantes nos resultados não apresentaram explicações convincentes. Os estudos na revisão foram baseados nos Estados Unidos (2), no Reino Unido (1) e na Austrália (1), com diferenças nas intervenções e incertezas sobre se resultados similares seriam esperados no países nos quais esta política foi considerada. Dois estudos encontraram um efeito em grupos de baixa idade que não foi mantido em um estudo; dois estudos não encontraram uma mudança no comportamento do fumante. É difícil estimar, com base nestes estudos, qual seria a melhor estimativa do impacto do licenciamento a vendedores de cigarro com verificações de conformidade na redução do número de pessoas que fumam. Uma estimativa mais baixa seria a não existência de impactos desta intervenção. A estimativa mais alta é retirada de um laudo pericial (consulte as Tabelas 3-5)

[4] A estimativa mais alta de anos de vida salvos, que é retirada do mesmo laudo pericial, tem as mesmas limitações do que a estimativa do impacto no comportamento do fumante, já que se baseia naquela estimativa. Além disso, baseia-se em suposições sobre o que aconteceria mais além da extensão dos estudos que avaliaram impactos no comportamento do fumante também, e suposições sobre o impacto das alterações no comportamento do fumante em relação à mortalidade.

[5] As estimativas de custos da política foram tiradas do laudo pericial (descritas na tabelas subsequentes deste artigo). Estas tem como base uma estimativa de quantos vendedores venderam cigarros, uma suposição de custo do processo para cada licença, e uma suposição sobre os custos de cada verificação de conformidade.

Tabela 3. O que está sendo comparado? Exemplo de caso: o licenciamento a vendedores de cigarro

A redução dos adolescentes fumantes foi uma prioridade no Ministério da Saúde em um país europeu. Um relatório das opções de políticas para atingir isso foi comissionado pelo governo interessado e preparou-se um relatório por especialistas líderes em saúde pública. Uma das opções de políticas consideradas no relatório foi o licenciamento a vendedores de cigarro. A perda de tal licença foi proposta como penalidade para a venda ilegal de cigarros a menores. Esta opção foi comparada no relatório ao status quo, ou seja, a ausência de licenciamento a vendedores de cigarro. Os especialistas em saúde pública não realizaram nem usaram uma revisão sistemática, tampouco especificaram quais características da opção de política (ou comparador) consideraram como sendo cruciais ou importantes.

Várias questões importantes não foram consideradas no relatório. Podem ter existido, por exemplo, diferenças importantes entre o status quo das áreas nas quais os formuladores de políticas consideraram implementar a política e aquelas nas quais os estudos foram realizados. Tais considerações podem ter incluído outras políticas já adotadas para reduzir a venda de cigarros a menores. É possível que a legislação existente já tenha estabelecido a venda de cigarros a menores como ilegal, ou contenha outros métodos pelos quais a legislação poderia ser aplicada (por exemplo: mediante multas ou outras penalidades para a venda ilegal de cigarros, educação individual dos vendedores (informando-os sobre as exigências legais), ou campanhas na mídia (para conscientizar a comunidade). Também poderiam ter existido diferenças na facilidade com a qual menores conseguem cigarros de outras fontes (por exemplo: dos pais e amigos ou através de furtos).

Os especialistas consideraram de forma explícita duas opções de políticas para o licenciamento a vendedores de cigarro, instituindo três verificações de conformidades por ano (por um adolescente tentando comprar cigarros) para ter certeza de que os vendedores não estavam vendendo cigarros a menores, e uma verificação de conformidade por ano com controle interno (exigindo que os próprios vendedores controlassem a não venda de cigarros a menores). A penalidade para a não conformidade em ambos os casos era a perda da licença. Outras formas de fazer cumprir o licenciamento são possíveis, algumas delas foram avaliadas em outros estudos. Os especialistas que escreveram este relatório não abordaram explicitamente se as diferenças nas abordagens referentes à aplicação do licenciamento foram susceptíveis de gerar modificações importantes na efetividade da política.

Tabela 4. Quais são os resultados mais importantes? Exemplo de caso: o licenciamento a vendedores de cigarro

O resultado primário considerado pelo laudo pericial autorizado pelo governo interessado foi a prevalência do tabagismo. Isso foi reconhecido como sendo um resultado secundário das consequências do tabagismo. O impacto nos anos de vida salvos foi estimado com base no impacto estimado na prevalência do tabagismo e nos dados epidemiológicos vinculando tabagismo com mortalidade. Impactos na morbidade não foram considerados. Outros impactos que foram expressamente considerados pelos especialistas foram os custos administrativos e a aceitabilidade política e pública. Há vários outros resultados que o laudo pericial poderia ter considerado, inclusive:

- Os custos para os vendedores e possíveis prejuízos (por exemplo: aumento de furtos ou comércio de fronteiras)
- Quem pagaria pelos custos administrativos de tais esquemas
- As possíveis diferenças nos impactos da política sobre diferentes populações (por exemplo: menores em desvantagem sócio-econômica ou aqueles que moram próximos da fronteira do país (que poderiam atravessar a fronteira do país vizinho para comprar cigarros)
- Consequências éticas (por exemplo: aquelas relacionadas ao uso de um menor ou uma pessoa pretendendo ser um menor para verificações de conformidade, ou a equidade da política em relação aos impactos potencialmente diferentes nos distintos grupos de menores e vendedores)

Tabela 5. Quais são as melhores estimativas dos impactos? Exemplo de caso: o licenciamento a vendedores de cigarro

O laudo pericial sobre políticas para reduzir o tabagismo na adolescência autorizado pelo governo interessado estimou que outorgar licença a vendedores de cigarro resultaria em uma redução relativa de 10% no número de fumantes. Usando a atual prevalência de fumantes como referência, o efeito absoluto da política foi estimado como sendo uma redução de 1.650 fumantes por ano. Com base nos modelos epidemiológicos do aumento do risco de morte devido ao tabagismo, os especialistas estimaram que esta política pouparia 9.240 vidas por ano. Não foi fornecido intervalos de confiança, embora foi observado que o atual efeito era muito incerto e uma variedade de estimativas foi usada para calcular a relação custo-benefício do licenciamento a vendedores de cigarro. Os custos administrativos foram estimados, com base na estimativa de quantos vendedores venderam cigarro, uma suposição sobre quanto custaria processar cada licença, e uma suposição sobre quanto cada inspeção custaria (para verificar a conformidade com a exigência de não vender cigarros a menores).

Usando estas distintas suposições, o custo estimado custo total foi entre €7,2 milhões e €10,5 milhões por ano.

Tabela 6. Intervalos de confiabilidade

Um intervalo de confiabilidade (IC) é a variação aproximada de uma estimativa que exprime o grau de precisão de uma estimativa. O intervalo de confiabilidade é um guia que representa o grau de certeza sobre a quantidade na qual estamos interessados (por exemplo: o efeito da opção de uma política em um resultado de interesse). Quanto menor for a variação entre os números mais altos e mais baixos do intervalo de confiabilidade, mais precisa será a estimativa, e maior certeza haverá com relação ao valor real. Quanto maior a variação, mais será a incerteza. A amplitude, ou variação, do intervalo de confiabilidade reflete a extensão na qual a chance pode ser responsável por uma estimativa observada (intervalos mais amplos refletem maior probabilidades de a chance ser um fator). Um IC de 95% significa que podemos estar 95% confiantes de que o tamanho real de um efeito está entre o limite de confiança mais alto e mais baixo. Opostamente, há 5% de probabilidade de que o efeito verdadeiro esteja fora desta variação.

Tabela 7. Quão confiantes estamos nos impactos estimados? Exemplo de caso: o licenciamento a vendedores de cigarro

O laudo pericial autorizado pelo governo interessado concluiu que as bases empíricas para o licenciamento a vendedores de cigarro foram “sólidas”, mas a base para este julgamento não foi clara. Os especialistas não conduziram, nem mencionaram, a revisão sistemática que é indicada na Tabela 3, ou qualquer outra revisão sistemática como a base para suas estimativas, embora uma revisão sistemática estivesse disponível [24]. Em contraste com o julgamento não explicado dos especialistas, uma avaliação da evidência resumida na revisão sistemática usando a abordagem GRADE sugere que a qualidade da evidência era muito baixa para todos os resultados importantes (consulte a Tabela 8 para informações adicionais sobre o sistema de avaliação GRADE). A tabela 1 resume as descobertas do relatório dos especialistas na forma de balancete para esta decisão política e mostra uma avaliação de qualidade da evidência para as três estimativas usando a abordagem GRADE.

Os autores da revisão sistemática (que incluiu uma variedade mais ampla de intervenções e desenhos de estudos) concluíram: “Intervenções com vendedores podem levar a uma grande diminuição nos números de lojas de varejo vendendo cigarros a jovens. Entretanto, algumas das comunidades estudadas nesta revisão atingiram níveis

sustentados de alta conformidade. Isto pode explicar porque existe pouca evidência para um efeito da intervenção nas percepções dos jovens sobre a facilidade de acesso a cigarros e sobre o comportamento do tabagismo." As estimativas 'pessimistas' dos benefícios na Tabela 1 são consistentes com as descobertas da revisão sistemática e não foram consideradas no laudo pericial.

Tabela 8. O sistema GRADE para avaliar a qualidade das evidências

Avaliar a qualidade das evidências requer julgamentos rigorosos sobre até que ponto se pode ter certeza de que uma estimativa de efeito é correta. O GRADE fornece uma abordagem sistemática e transparente para tornar os julgamentos de cada resultado de uma decisão importantes [12]. Os julgamentos se baseiam no tipo de desenho do estudo (avaliações randomizadas versus estudos de observação), o risco de viés (limitações do estudo), a consistência dos resultados nos estudos, e a precisão da estimativa geral em todos os estudos. Com base nestas considerações para cada resultado, a qualidade da evidência é classificada como alta, moderada, baixa ou muito baixa, usando as seguintes definições:

⊗⊗⊗⊗ **Alta**	• Confiante que o efeito verdadeiro está perto daquele da estimativa do efeito
⊗⊗⊗○ **Moderada**	• O efeito verdadeiro pode estar perto da estimativa do efeito, mas há uma possibilidade de que seja substancialmente diferente
⊗⊗○○ **Baixa**	• O efeito verdadeiro pode ser substancialmente diferente do da estimativa do efeito
⊗○○○ **Muito baixa**	• Bastante incerto sobre a estimativa

Tabela 9. Um modelo formal econômico pode ajudar? Exemplo de caso: o licenciamento a vendedores de cigarro

O laudo pericial autorizado pelo governo interessado incluiu uma análise econômica. Esta conclui que o custo por ano de vida salvo por licenciar vendedores de cigarro e conduzir verificações de conformidade, esta aproximadamente entre €900 e €92.000, com a melhor estimativa de €8,000. Os autores observaram que estavam substancialmente incertos sobre suas estimativas e sugeriram enfocar na variação de estimativas em vez de na melhor estimativa. Contudo, eles relataram estimativas exatas (com base em suposições que fizeram) e concluíram que a base empírica para

recomendar o licenciamento a vendedores de cigarro era sólida. Como resultado, os formuladores de políticas que não leram este relatório poderiam fazer uma conclusão crítica (errônea, em nossa opinião) de que o relatório forneceu evidências de alta qualidade de que o licenciamento a vendedores de cigarro era tão efetivo em custo quanto (ou mais efetivo em custo que) uma ampla variedade de serviços preventivos clínicos pagos pelo governo. Uma revisão mais sistemática das evidências implícitas [24] e uma sumário das descobertas que incluíram julgamentos mais sistemáticos e transparentes da qualidade das evidências (conforme mostrado na Tabela 1), teria fornecido uma melhor base para o processo decisório.

Tabela 10. Fatores que podem determinar a importância da implementação de políticas e programas de saúde

O seguintes fatores podem, algumas vezes, ser considerados de forma independente (ou em combinação) como critério para definir prioridades na implementação de políticas e programas de saúde:

- Quão sério é o problema – q uanto mais sério for o problema, maior será a probabilidade de que a política ou o programa que o aborda seja uma prioridade
- O número de pessoas que são afetadas pelo problema – quanto maior for o número de pessoas afetadas, maior será a probabilidade de que a política ou o programa que o aborda seja uma prioridade
- Benefícios – quanto maior for o benefício, maior será a probabilidade de que a política ou o programa seja uma prioridade
- Efeitos adversos – quanto maior for o risco de efeitos indesejáveis, menor será a probabilidade de que a política ou o programa seja uma prioridade
- Uso de recursos (custos) – quanto maior for o custo, menor será a probabilidade de que a política ou o programa seja uma prioridade
- Relação custo-benefício – quanto menor for o custo por unidade de benefício, maior será a probabilidade de que a política ou o programa seja uma prioridade
- Impactos sobre equidade – políticas ou programas que reduzem desigualdades podem ser mais prioritários do que os que não as reduzem (ou aqueles que as aumentam)

As decisões sobre prioridades deveriam apoiar-se nos critérios compartilhados ou nas razões, como as ideias mostradas acima. Elas também deveriam ser abertas a controles e deveriam poder fazer parte das considerações feitas pelas partes envolvidas. A regulamentação deveria garantir o cumprimento destas três condições [25]. Quando critérios, como os acima, são usados implicitamente em vez de explicitamente, é difícil avaliar se os critérios ou as decisões foram adequados [26].