

# 知证卫生决策工具之十六

## ——使用研究证据权衡政策的利弊\*

### SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in balancing the pros and cons of policies\*

Andrew D Oxman<sup>1</sup>, John N Lavis<sup>2</sup>, Atle Fretheim<sup>3</sup>, Simon Lewin<sup>4</sup>

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

**摘要** 本文是为卫生政策及规划的决策者及其支持者写的系列文章之一。本文提出使用证据以判断所选政策和卫生规划的利弊关系。作出判断时我们建议考虑以下五个问题：① 进行比较的选择是什么？② 所比较的选择最重要的可能结果是什么？③ 对所比较的选择每一项重要结果而言，其影响的最佳估计是什么？④ 对所估计的影响，决策者及其他相关人员有多大的把握度？⑤ 有无正规的经济学模型帮助决策？

## 1 关于STP

本文是为卫生政策卫生规划决策者及其支持者写的“知证卫生决策工具”系列文章之一。该系列文章旨在确保他们了解当前可得最佳研究证据之后进行决策。SUPPORT 工具及其使用方法在本系列文章的导论中已有详述<sup>[1]</sup>。每篇文后均附有本系列文章的所有术语(见附件1)。SUPPORT 网站上可找到本系列论文西班牙文、葡萄牙文、法文和中文翻译的链接([www.support-collaboration.org](http://www.support-collaboration.org))。欢迎通过邮箱 STP@nokc.no 对如何完善该工具提出反馈意见。

## 2 案例情景

你在卫生部工作，卫生部长要求你就正在考虑的某卫生政策的重要变革提供一份关于其预期效益、危害及成本的总结。

## 3 背景

决策者及其支持者在考虑如何确保对政策和规划决策的利弊权衡是基于研究证据时，本文建议

其可提出的五个问题，如在上述案例情景中可提出的问题。仅靠研究不能形成决策<sup>[2]</sup>，还需要对使用何种证据、如何解释该证据及我们对该证据的把握度等作出判断。更重要的是，决定作出何种决策需判断所期望的结果是否大于所不期望的结果(见图1)<sup>[3]</sup>。除判断影响大小的可能性外，决策过程中还需判断影响的重要性、决策实施所需的资源<sup>[4]</sup>及与该资源的其他利用方法相比，该利用方式的优越程度。

如果某选择的预期效益很大，几乎无不利方面且成本很小，而我们对证据及利益的重要性有很大把握且该选择有明确的优越性时，作出决策很简单。遗憾的是这种情况很少出现。大多数情况下，预期影响与成本均不确定，必须作出复杂而艰难的抉择。

本文提出的问题不能减少对判断的需要。但对这些问题更系统的考虑和讨论有助于确保不忽略重要问题且知晓证据后作出判断，也有助于解决分歧

期望的结果包括：

- 健康改善
- 卫生服务可及性提高
- 卫生服务利用更合理
- 节约支出
- 不公平性降低
- 伦理结果如增加自主权

不期望的结果包括：

- 对健康的非预期的影响(危害)
- 卫生服务可及性降低
- 卫生服务利用更不合理
- 增加成本
- 不公平性增加
- 伦理结果如降低自主权



图1 权衡卫生政策和规划的利弊\*

\* 卫生政策或规划决策要求判断相对于不期望的结果，期望的结果是否值得。

\* 本文英文原文发表于 *Health Research Policy and Systems*, 2009, 7(Suppl 1):S16 doi:10.1186/1478-4505-7-S1-S16

作者简介: Andrew D Oxman, Email: [oxman@online.no](mailto:oxman@online.no)

表 1 平衡表的利弊

平衡表很简单,却是陈述不同选择利弊的一种有效方式,包括政策选择<sup>[17,23]</sup>。此部分描述了准备和使用平衡表(如表 2 所示)所需的证据和判断。同时描述与使用非系统和非透明的专家判断相比,使用平衡表的优点。

平衡表的目的在于帮助决策者准确理解所比较的决策选择的重要结果。该表通过一系列方式达到此目的。首先,凝练最重要信息使有效思考成为可能;其次,平衡表关注最重要的结果,提高决策者准确把握下述问题的可能,即对所考虑决策的影响我们已经知道什么及有哪些重要结果;第三,构造平衡表为组织思路、构建证据分析和聚焦争议提供了有用的机制;第四,平衡表有助于更准确地判断某政策选择最重要的结果及潜在证据,及随后对不同选择相对优势和劣势的权衡;最后,平衡表可向其他决策者提供“原始信息”,帮助他们应用于权衡期望和非期望结果的判断。

但平衡表用于决策还需考虑两个重要的局限性。首先,当权衡多种结果时,决策者作出的判断需进行高水平的信息处理;其次,当权衡多种结果时,决策者所采用的价值判断可能是隐含的。正规的经济学模型通过使任何潜在假设(包括价值判断)变得更加外显而有助于解决这些局限。并能用敏感性分析来探讨不确定性及对结果的不同假设。

表 2 可否将禁止向未成年人出售烟草列为发放烟草许可证的条件?

	悲观	预期最佳影响	乐观	研究数目	证据质量 (GRADE) <sup>#</sup>
每年减少的吸烟者	0	?	1 650 (总人口 450 万)	4	很低 <sup>#</sup>
每年挽救的生命年	0	?	9 240 (总人口 450 万)	4	很低 <sup>§</sup>
每年的成本	1.05 亿欧元 (每年控制 3 次)	?	0.72 亿欧元 (每年控制一次+内部控制)	0	很低 <sup>**</sup>

\* 欧洲国家所提议的执照法案,要求零售商拥有售卖烟草的许可证。该政策包括每年三次的依从性检查及每年一次的内部控制。通过检查依从性(青少年试图购买烟草)确保零售商未向未成年人出售烟草。不依从者可吊销许可证。内部控制需要零售商控制自身不向未成年人出售烟草。

## 见表 8。

# 系统评价是此总结的基础(下表中提及的专家报告中未使用)包括 1 个相关的随机试验和 3 个存在重要局限的前后对照试验。对吸烟率影响的估计存在偏倚的风险很高。对结果的重要不一致性缺乏有说服力的解释。评价基于美国(2 个研究)、英国(1 个研究)和澳大利亚(1 个研究),其干预措施不同,且若该政策在本国实施,不确定能否期待相似的结果。其中两个研究发现该政策对低年龄组有影响,而另一研究并未发现;两个研究未发现吸烟行为的改变。基于这些研究,很难确定烟草零售商许可证和依从性检查对减少烟民数量影响的最佳估计。估计的下限是此干预措施无效,上限见专家报告(见表 3 ~ 5)。

§ 从同一专家报告中得到的对每年挽救生命年上限的估计与对吸烟行为影响的估计一样具有局限性,因为都是基于同一估计。另外,这一估计基于超过评价吸烟行为影响研究时期很长时间将会发生什么,以及吸烟行为改变对死亡率影响的假设。

\*\* 该政策的成本估计来自专家报告(后面的表格中阐述)。成本的估计基于烟草零售商的数量、每一发放许可证过程的成本及每次依从性检查的成本。

或至少可澄清问题。如果决策透明,则有助于其他人理解卫生政策决策背后的理由。

准备并使用一份平衡表(见表 1 及下文提出的前 4 个问题)有助于很好地知证决策。有时采用正规经济模型如成本-效果分析也会有帮助。本文提出的第 5 个问题用于讨论后者。此处我们建议考虑的问题基于 GRADE 工作组的工作<sup>[5]</sup>。尽管工作组关注的重点是临床实践指南,他们用于临床干预决策的方法同样适用于卫生政策及规划的决策<sup>[6]</sup>。

#### 4 考虑的问题

以下 5 个问题用于指导证据使用从而为判断卫生政策及规划决策的利弊提供信息。

- (1) 进行比较的选择是什么?
- (2) 所比较的选择最重要的可能结果是什么?
- (3) 对所比较的选择每一项重要结果而言,其

影响的最佳估计是什么?

(4) 对所估计的影响,决策者及其他相关人员有多大的把握度?

(5) 有无正规的经济模型帮助决策?

前 4 个问题用于指导决策中平衡表的使用。回答第 5 个问题有利于确保全面经济分析中的稀缺资源用于最需要的地方。

理想情况下,应由研究者或技术支持者与决策者一起构建平衡表(和经济模型)。一般而言,同其它地方描述的强调系统评价重要性的理由一样;此处也应以系统评价为基础<sup>[7]</sup>。我们不考虑构建平衡表时必须作出的诸多详细判断,因为这已在其他地方交代<sup>[8]</sup>。即便有,也很少需要由决策者自己做出所有的判断。甚至在有足够技术支持来准备平衡表的情况下,决策者知道查找什么及提出什么样的问题这一点也很重要。这可以确保平衡表被明智地用

表 3 比较什么? 案例: 烟草零售商许可证发放

欧洲国家卫生部关注的首要问题是减少青少年吸烟。政府相关部门要求作出如何减少青少年吸烟的决策选择报告,于是公共卫生权威专家准备了一份报告。报告中考虑到的一项政策选择即向烟草零售商发放许可证。建议对非法向未成年人销售烟草处以吊销许可证的惩罚。报告中以该决策与现状进行比较,该现状即不存在烟草零售商许可证制度。公共卫生专家并未实施或使用系统评价,也未指出所考虑决策(或对照决策)的哪些特点是他们认为决定性的或重要的。

报告中未考虑许多重要的问题。如决策者考虑实施政策的区域的现状与研究开展地区可能存在重要差异。基于上述考虑,应该包括减少向未成年人销售烟草的其他现有政策。现行法规可能已规定向未成年人销售烟草违法,或包括了其他方法来增强立法[如,对违法销售烟草处以罚款或其他惩罚,面对面的零售商教育(告知他们法律要求),或媒体传播(以提高公众知晓率)]。此外,未成年人从其他途径(如父母、朋友或偷窃)获得烟草的难易程度可能也存在差异。

专家明确考虑了烟草零售商许可证的两种政策,即每年检查依从性三次(通过青少年试着去购买烟草的方式)确保零售商未向未成年人销售烟草及每年一次依从性检查加内部控制(要求零售商自我监督确保不向未成年人销售烟草)。两种情况下不依从的惩罚都是吊销许可证。还存在可能的加强许可的方法,其中有些已被其他研究评价过。专家撰写此报告时未明确指出不同的加强许可的方法是否可能存在政策效果的重要差异。

表 4 最重要的结果是什么? 案例: 烟草零售商许可证发放

政府相关部门委托的专家报告考虑的主要结果指标是吸烟率。这是吸烟后果的替代指标。对挽救生命年效果的估计基于对吸烟率效果的估计及将吸烟与死亡相联系的流行病学数据。未考虑对发病率的影响。专家明确考虑到的其他影响还包括行政管理成本、政策可接受性和公众可接受性。专家报告中还应包括却未涵盖的其他结果如下:

- 零售商成本和可能的危害(如增加偷窃或跨境购买)
- 该体系中谁承担行政管理成本
- 政策对不同人群影响的可能差异[如社会经济水平较低的未成年人或居住在边境区域者(可能跨越至邻国购买烟草)]
- 伦理问题(如那些有关依从性检查时用未成年人或假装未成年人者;或,与对不同未成年人群体或零售商群体可能存在不同影响的有关政策的公平性问题)

以为决策者作出负责任的决策提供信息。

#### 4.1 进行比较的选择是什么?

当使用如表 2 所示的平衡表时,首先需要明确所比较的选择是什么。这通常不像其听起来那样直观(如见表 3)。准备平衡表的人必须决定所考虑的选择和用于比较的选择。通常,现状是用作比较的选择。但现状很可能因地而异,因此必须决定现状的特点是什么:

- 决定性的——所用对照与现状无共同特点的研究将被排除;
- 重要但非决定性的——所用对照与现状无共同特点的研究将被纳入,但在选定环境中得出相同结果的可信度较低;
- 不重要的——我们很有把握地认为在选定环境中很可能得出相同的结果。

对所考虑的选择同样需要作出这样的判断:就作用于可能的影响而言,他们的哪些特征是决定性的、重要的或不重要的?

#### 4.2 所比较的选择最重要的可能结果是什么?

一般而言,决策者受其服务所代表人群的愿望所激励,首先应该对卫生政策和规划对结果的影响感兴趣,而这种结果对被影响的人群很重要(如,见表 4)。包括健康结果、卫生服务的可及性或利用、不期望的效果(危害)和资源利用(成本或节约)(见图 1)。其他常见重要结果还包括:利益与成本的分配与公平性<sup>[9]</sup>、对其他部门的溢出效果;伦理方面

的后果如削弱民众的自主权也同样重要。

明确哪些结果是重要的有助于确保某决策的重要后果不被忽略,同样有助于确保不给不重要结果赋予不相称的权重。这对替代结果尤为重要,即那些自身不重要,且同类相关也不重要的结果。他们被认为重要是因为人们认为其反映了重要结局。如人们通常并不认为血压重要,使血压变得重要的是其与中风、心脏病发作和死亡相关,后者对人们都很重要。因此当考虑以高血压(或其他心血管危险因素)为对象的决策时,应基于相关决策对重要结局(心血管疾病)的影响。仅影响血压的证据只是影响心血管疾病证据的一种间接形式。

#### 4.3 对所比较的选择每一项重要结果而言,其影响的最佳估计是什么?

决定期望效应是否与非期望的效应价值相当,需估计这些效应(及经济影响)的差异有多大。理论上应采取下面的形式:比较实施或不实施或实施另一选择时每一重要结果的期望值(如,见表 5)。明确每一估计值的精确性也很有用——如每一估计值的可信区间(表 6 作进一步解释)。

对决策者而言,认识到以相对效应形式呈现的估计与以绝对效应形式呈现者之间的不同很重要。病人、卫生职业人员及卫生政策与规划决策者更可能决定采纳以相对效应形式报告的干预措施而非以绝对效应形式报告者<sup>[10]</sup>。如某研究报道澳大利亚卫生职业人员样本中 61% 的人同意实施可降低

表 5 影响的最佳估计是什么? 案例: 烟草零售商许可证发放

由政府委托的关于降低青少年吸烟政策的专家报告中估计,实施烟草零售商许可证可相对减少 10% 的青少年烟民。参考当前的吸烟率,估计每年可绝对减少 1 650 个青少年吸烟者。基于因吸烟造成死亡风险增加的流行病学模型,专家们估计该政策每年可挽救 9 240 个生命。报告未给出可信区间,尽管指出了真实效应很不确定并采用区间估计来计算了向烟草零售商发放许可证的成本效果。估计了行政管理成本,该估计基于对烟草零售商数量的估计,对每次许可证发放过程成本的假设,以及每次督导成本的假设(检查不向未成年人销售烟草的依从性)。基于这些不同的假设,总的估计成本在每年 720 万至 1 050 万欧元之间。

表 6 可信区间

可信区间(CI)是包括估计值的一个范围,传递了估计精确性的信息。可信区间代表我们对感兴趣的量可能的确信程度(如某决策对所关心结果的影响)。可信区间范围越窄,估计精度越大,可能接近真值的可信程度越高。可信区间越宽则可能的确信程度越低。可信区间的宽度或范围反映了机遇可解释观察值的程度(可信区间越宽反映了机遇作为影响因素的可能性越大)。95%CI 意味着我们有 95% 的把握认为效应的总体参数(真值)落在该范围,换言之,真值落在该区间之外的概率为 0.05。

结肠直肠癌死亡率 17% (相对危险度减少率 RRR) 的结肠直肠癌筛查。相比而言,如果报道死亡率绝对降低 0.4% (绝对危险度减少率 ARR)<sup>[11]</sup>,则仅有 24% 的卫生职业人员同意实施该筛查。事实上,这两个数据均来自同一卫生项目(该系列文章之 10 的表 4 解释了相对效应和绝对效应的不同<sup>[9]</sup>)。

#### 4.4 对所估计的影响,决策者及其他相关人员有多大的把握度?

以下 6 个因素会降低我们对政策或规划效应估计的把握度<sup>[12]</sup>:

- 研究设计不佳;
- 其他研究局限;
- 不精确;
- 结果不一致;
- 间接证据;
- 发表偏倚。

对这些因素的评价必然是技术性的。决策者无需详细了解这些因素及怎样进行评价。但决策者及其技术支持团队仍可以从理解为何考虑这些因素具有重要性中获益。

对某规划进行随机分配的研究可降低比较组间存在未知或未测量差异的风险。这为将研究结果归因于该规划而非其他因素提供了更高的可信度<sup>[13-15]</sup>。未采用随机分配设计的研究仅能说明测量到的差异。如将社区随机分配到接受某规划或政策(如烟草零售商准入)的研究,与由社区自己决定是否实施某特定政策的研究相比,对决策效果所提供的证据更具说服力。这是因为,决定实施某计划的社区与不决定采用该计划的社区相比,在对所关注结果(在此即吸烟率)产生作用的方式方面可能存在差异。从而不可能知道结果的差异归因于该政策或规划还是由社区间其他差异所引起。

研究的其他局限可能影响到随机和非随机设计的效果评估。例如不完整数据或结果测量不可靠,

可增加作出有偏倚的估计的风险从而降低所得估计的可信度。

不精确(以很宽的可信区间作为指征)同样降低可信度,即降低了我们排除比较组间观测到的结果差异由机遇造成的把握度,从而降低了估计效应的可信度(表 6 进一步详细解释了可信区间的概念)。

若关于同一政策和规划的多个研究结果不一致,且没有对这种差异有说服力的解释,同样只有较低的可信度认为所期望的效应源于该政策或规划的实施。

很多方式使得进行的研究可能与某特定问题不直接相关,从而降低结果的可信度。如上所述,测量非直接相关的结果(如血压)而不是测量重要的结果(如心血管疾病),则对重要结果指标影响估计的可信度较低(因为间接结果是一个替代指标)。如果仅提供了间接比较,可信度也会降低。当我们探讨的某选择缺乏直接比较的研究时,则对其把握度较低;如某选择与对照(无干预),而另一选择与其对照相比较(这时两种选择不同,其对照也不同——译者注)。非直接证据的其他方式包括研究及我们所感兴趣环境间存在的下述差异:

- 人群的特征;
- 考虑的政策选则;
- 现状或用于比较的选择。

有统计学意义效应的研究比没有者更易发表<sup>[16]</sup>。若存在这种“发表偏倚”可能,则仅基于已发表文章作出估计的可信度降低。当存在许多小样本研究尤其是厂家赞助者,或我们知道研究者存在其他类似利益冲突时,应考虑发表偏倚。

总之,证据“质量”或其稳健性评价及对某决策的可能效应估计值的确信程度有赖于对上述所有因素的全面考虑。尽管尚无评价这些因素的特定规则,与证据质量相关的判断明确指出所针对的每一因

表 7 我们对影响估计的把握度多大? 案例: 烟草零售商许可证发放

政府委托的专家报告总结道: 烟草零售商许可证决策的经验基础强度较高, 但该评判的基础尚不明确。专家们未实施或引用系统评价(表 3 中已提及的或其他系统评价)作为其估计的基础, 尽管存在相关的系统评价<sup>[24]</sup>。与专家未解释的判断相比, 用 GRADE 方法进行系统评价的证据评价总结指出, 所有重要结果的证据质量均很低(与 GRADE 评价系统有关的详细信息见表 8)。表 1 以决策平衡表的形式总结了专家报告的结果, 并展现了使用 GRAED 方法评价三个估计值的证据质量。

系统评价的作者(该系统评价包括更多的干预措施和研究设计)总结道: “对零售商的干预可减少向青少年零售烟草的数量, 但研究社区几乎无一能维持较高的依从性。这也许可以解释, 为什么下述证据有限: 青少年对易于获得烟草的感知度和吸烟行为的有关干预效果的证据。”表 1 中对效益的“悲观”估计与该系统评价结果一致, 但专家的报告中并未考虑。

表 8 评价证据质量的GRADE系统

评价证据质量需要对认为某效应估计正确的把握度作出判断。GRADE 系统为我们对某决策的各重要结果作出判断提供了系统而透明的方法<sup>[12]</sup>。这些判断基于研究设计类型(随机试验与观察性研究)、偏倚的风险(研究局限)、研究间结果的一致程度及各研究总估计值的精确性。对每一结果, 基于上述考虑, 证据质量分为高、中、低和极低四级, 各级定义如下:

高	很有把握认为真实值与估计值接近
中	真实值可能与估计值接近, 但存在二者完全不同的可能性
低	真实值可能与估计值完全不同
很低	估计值很不确定

表 9 正规经济学模型有帮助吗? 案例: 烟草零售商许可证发放

政府委托的专家报告中包括了经济学分析。报告总结了实施烟草零售商许可证发放及依从性检查成本, 为每挽救一个生命年近 900 ~ 92 000 欧元(最佳点估计值为 8 000 欧元)。作者指出这些估计存在很大的不确定性, 并建议关注估计值范围而不是最佳点估计值。尽管如此, 他们还是报告了确切的估计值(基于他们所作的假设), 并总结说推荐使用烟草零售商许可证制度的经验基础很坚实。结果, 未批判性地阅读此文的决策者可能下结论说(在我们看来是错误的): 该报告提供的高质量证据显示: 实施烟草零售商许可证发放与政府付费的更多临床预防服务相比, 具有相同(或更高)的成本效果。对基础证据更进一步的系统分析<sup>[24]</sup>, 及总结那些对证据质量进行更系统透明的判断(表 1 所示)原本可以为决策提供更好的基础。

素, 有助于降低忽略重要因素的可能性, 也有助于降低对证据作出有偏倚评价的可能(如, 见表 7)。使用系统而透明的方法如 GRADE 分级(见表 8), 使得较易于检查所作出的判断<sup>[5]</sup>。

#### 4.5 有无正规的经济学模型以帮助决策?

正规经济学模型, 如成本效果分析和成本效用分析, 有助于提供信息以权衡某决策期望和非期望的结果<sup>[17]</sup>。经济学模型对复杂决策及检验某决策对关键估计或假定值的敏感性很有价值。但经济学模型的结果取决于其用于建模的数据的质量。当估计效益、危害和资源利用的证据质量很低时, 则结果必然存在高度风险(投机性)(见表 9 中的例子)。

当存在以下情况时, 完整的经济学模型更可能有助于为决策提供信息:

- 所比较的决策之间资源消耗差异巨大;
- 需大量资金投入, 如建立新设施;
- 不确定产生的净效益是否值得所投入的增量成本;
- 资源消耗的证据质量好。

经济学模型还可通过探讨对一系列似乎合理的估计的分析的敏感性而用于阐明信息需求。

遗憾的是, 已发表的成本效果文章特别有关药物的文章, 很可能存在缺陷或偏倚。因其所实施的

特定环境与决策者感兴趣的环境在很多重要方面不同<sup>[18]</sup>。因此决策者可考虑构建自己的经济学模型。要做到这一点, 他们必须具备专业知识并拥有相关资源。

## 5 结论

通过权衡政策选择的利弊为决策提供信息。正如我们所推荐的, 这应做到系统化和透明化。当净效益(期望和非期望结果的差异)与成本相比很大时, 我们更有信心作此决策。相反, 当与成本相比净利益很少时, 则决策的信心不足。

一般来说, 我们对某决策选择可能的影响越没把握, 我们决定做什么时就越没信心。但也有例外: 首先, 我们对某决策非常没有把握时很容易决定不做。其次, 即使对某特定选择的益处几乎没有把握时仍容易选择做, 只是因为危害风险很小或不存, 且花费不多, 同时可带来些许益处。许多卫生信息可依此分类。但决策者应警惕这样的推论, 即貌似无害的卫生政策或规划不会带来危害<sup>[19]</sup>。事实上, 即使简单得诸如提供健康信息这样的事都可以是致命的<sup>[20]</sup>。近 50 年许多国家向妈妈们提出建议提倡婴儿俯卧睡, 这一事实作为例证证明了上述观点。这一貌似无害的建议造成了成千上万的婴儿死于婴

表 10 实施卫生政策及规划重要性的决定因素

有时可单独 (或联合) 考虑以下因素, 作为优先实施某卫生政策或规划的判定标准:

- 问题的严重性 —— 问题越严重, 针对该问题的某卫生政策或规划可能越应被优先考虑
- 受某问题影响的人数 —— 受影响的人越多, 针对该问题的某卫生政策或规划可能越应被优先考虑
- 效益 —— 效益越大, 该卫生政策或规划可能越应优先
- 不利效应 —— 不期望效果的风险越大, 该卫生政策或规划越不应优先
- 资源使用 (成本) —— 成本越高, 该卫生政策或规划越不应优先
- 成本效果 —— 单位效益的成本越低, 该卫生政策或规划越应优先
- 对公平性的影响 —— 减少不公平性的卫生政策或规划比其他不减少不公平性或增加不公平性者越应优先

决定优先策略应基于上述综合的标准或推理。其应能接受审查, 并能吸引利益相关者提出他们的考虑。规则应确保符合这三种情况<sup>[25]</sup>。未明确说明所使用的上述规则而是隐含地使用时, 很难判断该标准或决策是否合理<sup>[26]</sup>。

儿猝死综合征<sup>[21]</sup>。最后, 尽管某卫生政策或规划的可能影响存在重要的不确定性, 很容易作出这样的决策: 好的前景只能在对其影响进行良好设计的评价的背景下才能取得<sup>[22]</sup>。

即使当我们对某卫生政策或规划的影响很有把握时, 并不一定要优先实施该决策。我们的把握度是决定做什么以及做某事的优先程度的关键因素。其他附加因素 (如表 10 所述) 也是决定是否应优先实施某卫生政策或规划的因素。

#### 资源

##### 有用的文件和补充阅读文献

- 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.

##### 网址链接

- SUPPORT Summaries: <http://www.support-collaboration.org/index.htm> - Concise summaries of the pros and cons of health policies and programmes for low- and middle-income countries based on systematic reviews.
- GRADE Working Group: <http://www.gradeworkinggroup.org> - The Grading of Recommendations Assessment, Development and Evaluation (or GRADE) Working Group has developed an approach to grading the

quality of evidence and the strength of healthcare recommendations.

#### 附件

术语: <http://www.biomedcentral.com/content/supplementary/1478-4505-7-S1-S16-S1.doc>

#### 利益冲突、作者贡献及致谢

见原文。

#### 参 考 文 献

- 1 Lavis JN, Oxman AD, Lewin S, *et al.* SUPPORT Tools for evidence-informed health Policymaking (STP). Introduction. *Health Res Policy Syst*, 2009, 7(Suppl 1): 11.
- 2 Chalmers I. If evidence-informed policy works in practice, does it matter if it doesn't work in theory? *Evidence & Policy*, 2005, 1: 227-242.
- 3 Guyatt GH, Oxman AD, Kunz R, *et al.* Going from evidence to recommendations. *BMJ*, 2008, 336: 1049-1051.
- 4 Oxman AD, Fretheim A, Lavis JN, *et al.* 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, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 2008, 336: 924-926.
- 6 The SUPPORT Collaboration: How SUPPORT Summaries are prepared 2009 [<http://www.support-collaboration.org/summaries/methods.htm>].
- 7 Lavis JN, Oxman AD, Grimshaw J, *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, *et al.* GRADE guidelines: 1. Introduction - GRADE evidence profiles and Summary of Findings tables. *J Clin Epidemiol*, in press.
- 9 Oxman AD, Lavis JN, Lewin S, *et al.* 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, *et al.* Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database of Systematic Reviews*, in press.
- 11 Adily A, Ward J. Evidence based practice in population health: a re-

- gional survey to inform workforce development and organisational change. *J Epidemiol Community Health*, 2004, 58: 455-460.
- 12 Guyatt GH, Oxman AD, Kunz R, *et al*. What is "quality of evidence" and why is it important to clinicians? *BMJ*, 2008, 336: 995-998.
  - 13 Kleijnen J, Gotzsche P, Kunz R, *et al*. 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: 1185-1190.
  - 15 Savedoff WD, Levine R, Birdsall N. When will we ever learn? Improving lives through impact evaluation 2006 [[http://www.cgdev.org/files/7973\\_file\\_WillWeEverLearn.pdf](http://www.cgdev.org/files/7973_file_WillWeEverLearn.pdf)]. Washington DC, Center for Global Development.
  - 16 Hopewell S, Loudon K, Clarke MJ, *et al*. 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, *et al*. Incorporating considerations of resources use into grading recommendations. *BMJ*, 2008, 336: 1170-1173.
  - 18 Friedberg M, Saffran B, Stinson TJ, *et al*. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA*, 1999, 282: 1453-1457.
  - 19 Macintyre S, Petticrew M. Good intentions and received wisdom are not enough. *J Epidemiol Community Health*, 2000, 54: 802-803.
  - 20 Chalmers I. Invalid health information is potentially lethal. *BMJ*, 2001, 322: 998.
  - 21 Gilbert R, Salanti G, Harden M, *et al*. 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: 874-887.
  - 22 Oxman AD, Lavis JN, Fretheim A, *et al*. 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, *et al*. [Evaluation of reimbursement applications for new drugs]. *Tidsskr Nor Laegeforen*, 2002, 122: 2619-2623.

李幼平 总审校  
 陈群飞 陈尹 译  
 杨晓妍 审校  
 本文编辑: 蔡羽嘉