

# 我国《临床实践指南》 的现状和思考

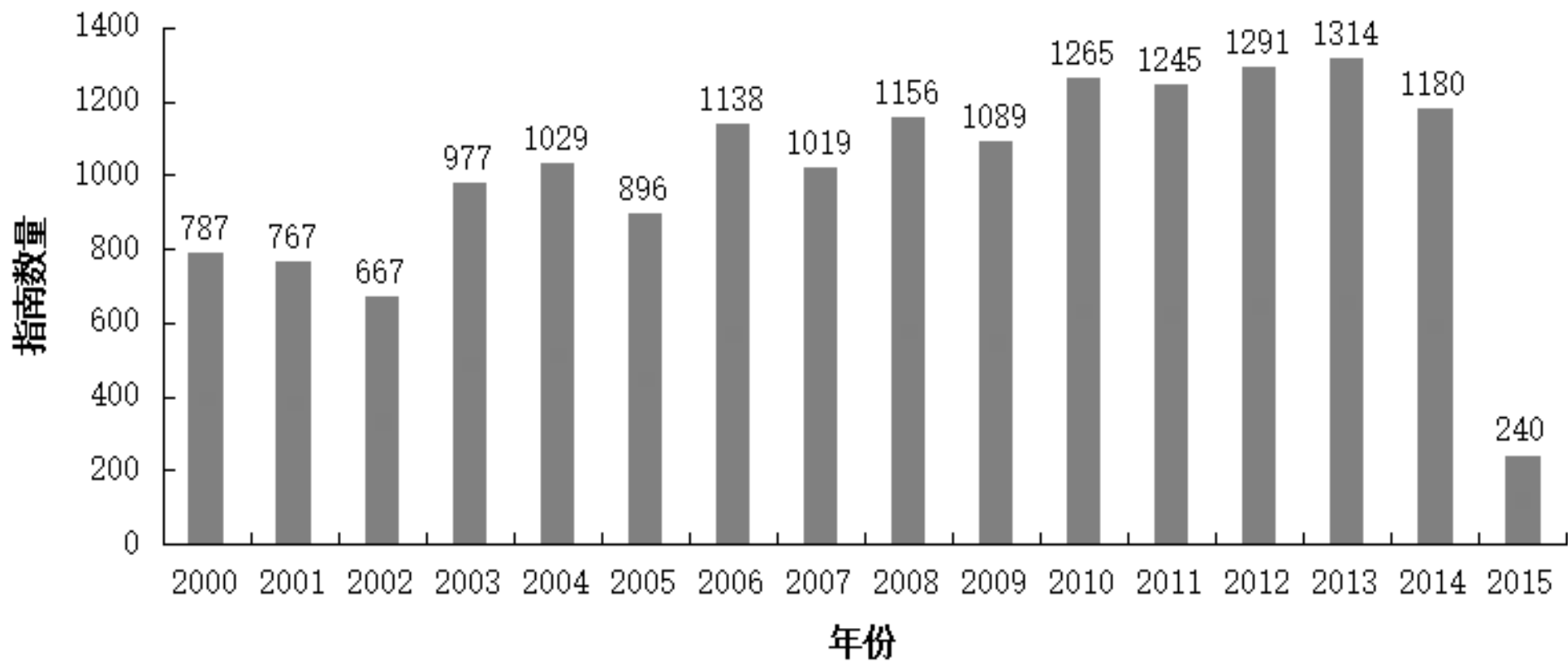
王吉耀

复旦大学附属中山医院

复旦大学循证医学中心主任

中国临床流行病学工作网主席

中华医学会临床流行病和循证医学分会名誉主委



2000-2015年国际临床实践指南的发布数量趋势图

注：检索截止时间为2015年9月18日，检索策略为"Practice Guideline" [Publication Type]

# 临床实践指南的定义

---

## 美国医学研究所 (Institute of Medicine, IOM)

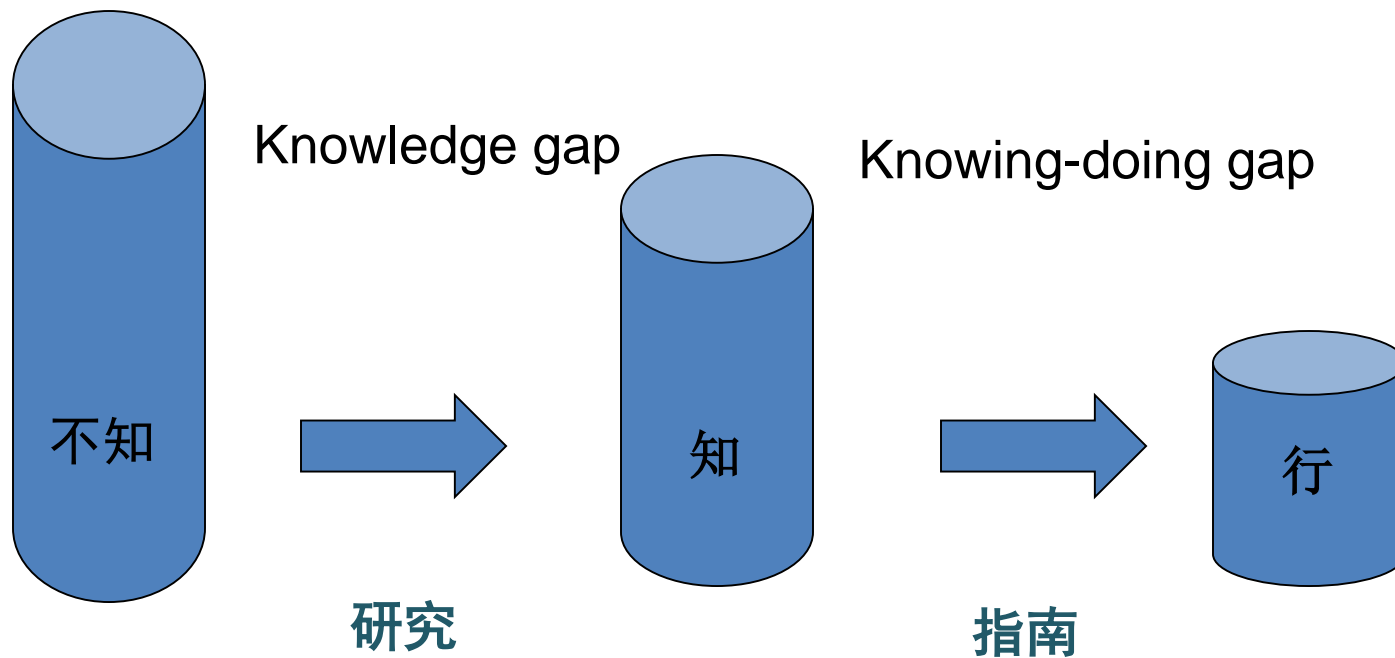
- 针对特定的临床问题，经系统研究制定后发布，帮助医生和患者作出恰当的判断的指导性意见，从而选择和决策适宜的卫生保健服务。
- 在对临床证据的系统性回顾以及对替代治疗手段风险和收益评估的基础上制定的推荐，该推荐旨在最大程度地优化患者的医疗护理过程。

# 临床实践指南的重要性

- ①改善临床结局，提高医疗质量；
  - ②以患者为中心、尊重患者价值观的临床决策；
  - ③减少临床实践中的不恰当差异，确保病人安
  - ④促进医疗资源的合理配置；
  - ⑤科学合理研究结果，使用临床决策清晰、透明
- 指导临床实践的指南与原始的临床研究和系统评价的区别在于：指南为临床医生提供具体的推荐意见指导医疗行为，是连接证据和临床实践的桥梁，更加贴近临床实践的需要。

# 指南：证据和实践之间的桥梁

CPG – the bridge between bench and bed



# 循证临床实践指南

- 随着循证医学的发展，循证临床实践指南（Evidence-based Clinical Practice Guidelines E-CPGs）成为解决上述问题的方法。E-CPGs针对某一特定的问题，特定的人群，由特定的组织和人员按照规范化的流程，集合当前最佳的证据，根据证据等级和推荐制定，用以指导临床医生从事预防、诊断、治疗、康复、健和管理工作。它增加了指南的的科学性、针对性、实用性。

# 循证临床实践指南

## (Evidence-based Clinical Practice Guidelines E-CPGs)

- ①清晰地界定的临床问题以及人群；
- ②规范的证据质量评价和分级：科学规范地获取评价证据、根据规范统一标准进行证据质量评价并给予分级。
- ③根据证据等级结合病人的价值观进行推荐，从而达到证据的临床转化。
- ④实时更新的特点：如果有新的证据出现，指南应做相应的调整。

# 存在问题

- 数量多、质量参差不齐
- 缺少国内高质量的研究证据
- 没有按照国际通用的规则制定指南
- 没有用GRADE对证据评价及推荐分级
- 没有充分考虑利弊的整体平衡，以对证据的解释提供建议；
- 没有考虑不同的价值观和偏好以形成最终推荐意见。
- 没有用国际通用的规则评价指南
- 受到医药公司的利益影响



- 过去几十年间，我国制定的临床指南的数量呈指数式增长，但这些指南的制定过程差别很大，导致许多指南质量不符合基本要求
- 临床医生、患者和一些利益相关者因此不得不面临众多质量参差不齐甚至内容相互矛盾的临床指南

## ■ 2006至2010年制定的327个指南

- 17.4%未对如何制定指南进行任何描述，所以无法获知指南的制定流程
- 1.8%为循证指南
- 95.9%指南未声明指南制定小组成员的利益冲突
- AGREE评价结果显示，有77个（28.5%）指南被推荐使用，但其中只有1个被强烈推荐，其它经修改后被推荐

## ■ 327个指南涉及多种疾病

- 消化系统疾病（17.1%）
- 循环系统疾病（16.8%）
- 肿瘤（11.0%）

## ■ 327个指南的更新情况

- 17%的指南为更新版本，更新距离上次制定时间的中位数为**4年**，其中最短的间隔时间为**1年**，最长的为**10年**。

# 指南制定规范

- 指南的修订需要科学化的过程，以保证修订内容的科学性和必要性
  - 国际指南联盟（Guidelines International Network, G-I-N）关于临床实践指南制定的国际标准
  - 美国医学研究所（Institute of Medicine, IOM）、WHO、NICE (National Institute for Health and Clinical Excellence)、苏格兰校际指南网络以及国家健康与医学研究理事会等许多医疗团体曾经为指南制定者提出制定的建议

# “临床实践指南”的制定过程

---

➤ 国际指南联盟（Guidelines International Network, G-I-N）委员会于2012年明确提出临床指南的制定应基于以下核心因素：

- 指南制定小组构成
- 决策过程
- 利益冲突
- 指南范围
- 方法学
- 证据综述
- 指南推荐
- 证据和推荐的评级
- 同行评审和利益相关者协商
- 指南的失效和更新
- 资金支持和资助机构

Amir Qaseem, Frode Forland, Fergus Macbeth, Gu ¨ nter Ollenschla ¨ ger, Sue Phillips, and Philip van der Wees. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med* .2012.

# 指南制定的过程：成立指南制定小组

---

## ➤ 成立指南制定小组

- 指南制定小组的构成：图

## • 指南制定小组的职责：

- 1) 确定指南要解决的PICO 问题；
- 2) 选择结局指标并对其排序；
- 3) 根据需要对指南范围的修改提供建议；
- 4) 评价支持指南的证据 (GRAD评价系统)；
- 5) 充分考虑利弊的整体平衡，以对证据的解释提供建议；
- 6) 考虑不同的价值观和偏好以形成最终推荐意见。

# 指南制定的过程：成立外部评审小组

---

## ➤ 成立外部评审小组

外部评审小组由对该指南主题感兴趣的人员组成，包括利益相关者和专业领域专家。各成员需要对指南制定过程的不同阶段进行评审。他们可能会在前期评审指南范围、草拟推荐意见及PICO 问题。在最终确定推荐意见时，他们还需要评审指南文件。

# 指南制定的过程：利益声明与处理

---

## ➤ 利益声明与处理

- WHO 对“利益冲突”的定义为：任何可能或被认为会影响到专家提供给WHO 建议的客观性和独立性的专家利益。
- 在指南外部评审小组和指南制定小组最终确定及发出邀请函之前，所有可能被纳入这些小组的成员都需要进行利益声明。



- 我国在制定**CPGs** 的指南制定小组构成往往只有临床专家，没有包括方法学的专家和社会学的专家，更没有患者参加。
- 没有采用**PICO** 模型构建问题；
- 没有用**PRISMA checklist** ([www.prisma-statement.org](http://www.prisma-statement.org)) 对系统评价/Meta分析的方法学质量进行评价
- 没有成立外部评审小组
- 缺少利益声明与处理的陈述。

表 牛津证据分级与推荐意见强度

推荐意见	证据水平	治疗/病因
A	1a	同质性RCT的系统综述
	1b	单一的RCT（可信区间较窄）
	1c	全或无效应（为治疗前所有患者均死亡或部分死亡，治疗后仅部分死亡或全部存活）
B	2a	同质性队列研究的系统综述
	2b	单个的队列研究（包括低质量的RCT，如随访率<80%）
	2c	结局研究(outcomes research)；生态学研究
	3a	同质性病例对照研究的系统综述
	3b	单个的病例对照研究
C	4	病例分析、低质量的队列和病例对照研究
D	5	没有严格评价的专家意见，或完全基于生理学和基础研究

# GRADE系统(Grading of Recommendations Assessment, Development, and Evaluation)

## “推荐意见的评定、发展和估价分级”

---

- ①其由有广泛代表性的国际指南研究小组担任分级。
- ②在证据质量与推荐强度之间有清楚的分割。
- ③对不同处理措施结局的重要性有明确的评价。
- ④对证据质量分级的上升和下调有明确的标准。
- ⑤从证据转化到推荐意见的过程透明化。
- ⑥对评估价值和选择有明确的说明。
- ⑦为临床医生、病人、政策制定者清楚的有程序地对推荐意见从强到弱进行解释。
- ⑧有利于系统综述、卫生技术测定和指南的制定。

# GRADE 系统证据质量及其定义

---

证据级别	定义
高质量 (A)	非常确信估计的效应值接近真实的效应值，进一步研究也不可能改变该估计效应值的可信度
中等质量 (B)	对估计的效应值确信程度中等，估计值有可能接近真实值，但仍存在二者不相同的可能性，进一步研究有可能改变该估计效应值的可信度。
低质量 (C)	对估计的效应值的 <u>确信程度有限</u> ：估计值与真实值可能大不相同。进一步研究极有可能改变该估计效应值的可信度。
极低质量 (D)	对估计的效应值几乎没有信心：估计值与真实值很可能完全不同。 对效应值的任何估计都很不确定

---

## 框 2-1 影响证据质量的因素

### 可能降低证据质量的因素及其解释

- 偏倚风险** 包括未正确随机分组；未进行分组方案的隐藏；未实施盲法(特别是结局指标为主观性指标且对其评估易受主观影响时)；研究对象失访过多、未进行意向性分析；选择性报告结果（尤其是仅报道观测到阳性结果的数据）；发现有疗效后提前终止研究。
- 一致性** 在排除了合理的原因外，不同研究间仍然出现了大相径庭的结果，可能意味着各种疗法的疗效确实存在差异。差异可能源于人群(如药物对重症人群的疗效可能相对显著)、干预措施(如较高药物剂量会使疗效更显著)或结局指标(如随时间推移疗效降低)。当结果存在异质性而研究者未能意识到并给出合理解释时，证据质量需降低。
- 间接性** 有两类：一是欲比较两种干预措施的疗效时，没有二者直接比较的随机对照试验，但可能存在均与同一安慰剂比较的随机对照试验，这样的试验可进行二者之间疗效的间接比较，但提供的证据质量比直接比较的随机对照试验要低。第二类间接证据包括人群、干预措施、对照措施、预期结局等。
- 精确性** 当研究纳入的患者和观察事件相对较少而致可信区间较宽时，将降低其证据质量。
- 发表偏倚** 若研究者未能发表研究(通常是阴性结果的研究)时，证据质量亦会减弱。典型情况是当公开的证据仅局限于少数试验而这些试验全部由企业赞助，此时应怀疑存在发表偏倚。

### 可能增加证据质量的因素及其解释

- 效应值很大** 当方法学严谨的观察性研究显示疗效显著或非常显著且结果一致时，将提高其证据质量。
- 偏倚相反** 当影响观察性研究的偏倚不是夸大，而是减小其效果时，可提高其证据质量。
- 有剂量-效应** 当干预的药量和引起的效应大小之间有明显关联时，可提高其证据质量。

表. GRADE 证据质量分级方法概要

研究设计	证据集群的初始质量	如果符合以下条件, 降级	如果符合以下条件, 升级	证据集群的质量等级
随机试验	高	偏倚风险	效应量大	高 (4 个“+”: +++) 中 (3 个“+”: +++ ○)
	中	-1 严重 -2 非常严重	+1 大 +2 非常大	
观察性研究	低	不一致性 -1 严重 -2 非常严重	剂量反应 +1 梯度量效证据	低 (2 个“+”: ++ ○○) 极低 (1 个“+”: + ○○○)
	极低			
		不精确 -1 严重 -2 非常严重	+1 如未观察到效应意味着是一种假效应	
		发表偏倚 -1 可能 -2 非常可能		

# GRADE证据概要表

证据评价							患者数量		效应值		证据质量	结局重要性
纳入研究数量	研究设计	偏倚风险	不一致性	间接性	不精确性	发表偏倚	干预组	对照组	相对效应值 (95% CI)	绝对效应值		
结局指标1												
结局指标2												
结局指标3												
结局指标4												
结局指标5												
结局指标6												
结局指标7												

备注：1) 结局指标最多纳入7个；2) 结局指标按其重要性分为3级，至关重要，重要和不太重要。

# 表 支持推荐意见形成的决策表

推荐意见		
人群		
干预措施		
因素	决策	解释
证据质量（证据质量越高，越可能做出强推荐）	高	
	中	
	低	
	极低	
利弊平衡与负担（利弊间的差别越大，越可能做出强推荐；净效益越小及利弊的确定性越低，越可能做出弱推荐）	利明显大于弊	
	利弊平衡	
	潜在危害明显大于潜在效益	
意愿和价值观（意愿与价值观的可变性或不确定性越大，越可能做出弱推荐）	无重要可变性	
	有重要可变性	
资源利用（干预的成本越高，即资源使用越多，越可能做出弱推荐）	资源耗费较少	
	资源耗费较多	
总体推荐强度（强或弱）		
研究差距		



## 推荐强度等级

推荐强度	具体描述
强推荐（1级）	明确显示干预措施利大于弊或弊大于利
弱推荐（2级）	利弊不确定或无论质量高低的证据均显示利弊相当

在形成推荐意见时，不仅考虑到证据的质量，还要权衡干预的利弊平衡与负担、患者偏好和价值观的不确定性或可变性，以及资源的合理利用、推荐措施的公平性与可实施性等。

# AGREE II评估工具

---

- 指南研究与评价工具( appraisal of guidelines research and evaluation , AGREE II) 的作用是为临床实践指南的质量审查提供一个框架。
- AGREE II评估工具可以用来评价地方、国家、国际组织或联合政府组织发行的指南。
- AGREE II包括6个领域和23个条目，每个领域针对指南质量评价的一个特定问题。

---

## 指南的范围和目的

1. 指南的撰写目的有明确说明
2. 指南包括的临床问题有详细的描述
3. 指南针对的目标人群有明确说明

---

## 利益相关者的参与度

4. 指南制定的小组成员包括所有相关方面的专家
5. 指南考虑了目标患者的意愿和喜好
6. 对指南使用人群有清楚的说明
7. 指南发表前已经在目标使用者中进行了预先测试

## 指南开发的严格性

8. 采用严格、系统的方法检索证据
9. 证据的纳入标准有清楚的描述
10. 形成推荐建议的方法有清楚的描述
11. 形成推荐建议时充分考虑了干预措施的获益、风险和副作用
12. 推荐建议和证据之间存在客观联系
13. 指南发表前接受了制定小组之外专家的同行评议
14. 提供了更新指南的程序

## 表述明确和清晰程度

15. 推荐建议专业而不含糊
  16. 针对不同临床状况提供了不同的治疗选择
  17. 关键的推荐建议可以很容易找到
  18. 提供了方便的应用工具（如：概要、教育工具、针对患者的宣传单等等）
- 

## 指南的可应用性

19. 讨论了应用该指南的过程中可能遇到的困难
  20. 讨论了应用该指南的过程中可能的费用
  21. 提供监督和审查指南执行情况的评估指标
- 

## 编辑工作的独立性

22. 指南编辑工作独立于其他经济实体
23. 指南制定小组成员的利益冲突在指南中有表述

- **李楠等用AGREE II对2012-2013年在我国大陆期刊发表的指南质量进行回顾性的评估，发现在制定指南的严谨性、编辑工作的独立性和利益相关者参与度三个很重要方面的评分低下，与国际上发表的指南有较大的差距**

李楠、姚亮、吴琼芳等：2012-2013年中国大陆期刊发表  
临床实践指南质量评价。中国循证医学杂志 2015； 15（3）： 259-263

# 受到医药公司的利益影响

利益冲突是一个非常重要并且广泛存在的问题,是指南制定过程中最常见的偏倚来源。为了保证指南制定的透明化,制定小组成员必须明确声明是否存在潜在的利益冲突。中国临床指南大多由行业协会制定,而这些指南大多都有厂家资助。但是到现在为止,中国几乎没有任何指南制定组织有利益冲突政策。我国发布的指南都没有利益冲突的声明。

# 指南的改编

- 我国缺少基于中国人群的数据，以往的指南大多数是以国外的指南为蓝本进行的改编。改编国际上高质量的指南应用于本地是一种避免重复工作以及提高应用性的方法。
- 由于国外的科学证据、专家共识主要依据他们国家的医疗体系和临床实践特征作为背景。而由于社会经济水平差异，特别是医疗资源的可及性、医保政策、新药应用等因素导致国外的指南和共识对我国不一定适用，因此需要改编。
- 但是，在改编时专家们没有用国际通用的规则进行指南的改编。



# ADAPTE指南改编方法(2009年)

改编的基本原则是：

- 尊重指南制定的循证原则；
- 用可靠的、统一的方法来保证改编指南的质量
- 让所有关键的利益相关者参加指南的改编，增加可接受性；
- 充分考虑当地的医疗实践实际情况；
- 透明报告以增加改编指南的可信性；
- 用灵活的形式来适应特殊的需要和环境；
- 改编的指南有责任说明原始指南的来源。

# 我国临床实践指南的未来

- 是完全跟风国外指南，还是应根据我们既

有的尚不完善的研究证据



# 面临的问题

- 我国缺少基于中国人群的数据，以往的指南基本是以国外的指南为蓝本
- 在欧美指南中，**只有不到15% 的建议是A 级证据**，大部分建议是B 级或C 级证据（专家们的共识）。
- 国外的科学证据、专家共识主要依据他们国家的医疗体系和临床实践特征作为背景。而由于社会经济水平差异，特别是医疗资源的可及性、医保政策、新药应用等因素导致国外的指南和共识对我国不一定适用

# 建议

1. 制定指南的指南
2. 规范指南改编的流程
3. 加强对指南制定者的方法学培训
4. 重视指南的传播、实施和评估
5. 坚守科学家的道德良知，不受商业利益干扰。
6. 加强我国自己的临床研究，尤其是实效研究，提高研究水平

# 重视指南的传播、实施和评估

指南制定后要宣传和传播到使用者，在使用指南本土化时，应该评估可能会有哪些障碍，应监查指南使用；对推荐意见中的主要的结局指标进行评估，维护指南使用。

- 指南不能替代临床医师的决策，而是建言和推荐，去帮助临床医师临床的决策，鼓励临床医师在运用指南时要因地制宜地结合患者自身情况进行诊治
- 各地应根据具体地区具体情况制定相应的标准。考虑国家或地区更多灵活性。在比较良好的医疗环境下，对指南的实践显得容易；而对于我们国家的边远地区，医疗条件尚比较落后，实践指南有时会困难重重。特别强调临床医师完全有权利根据具体情况作出相应的诊治

# 结语

- 临床诊疗指南是指针对特定的临床情况，系统制定出帮助临床医生和患者做出恰当处理的推荐意见。
- 制定指南为我国制定合格的《临床诊疗指南》提供适合国情、操作可及的开发工具和方法。
- 加强指南的实施与评价

谢谢！