

疾病治疗证据的分析和评价

范维琥

复旦大学附属华山医院

- George Washington had 2.4 liters of blood drained from him in the 15 hours prior to his death: he had been suffering from a fever, sore throat and respiratory difficulties for 24 hours.
- Some have stated that in this way Washington was murdered.



- A physiologic study demonstrated that β blockers result in a decline in ejection fraction and increases in end-diastolic volume in patients with prior myocardial infarction
- A meta-analysis of 18 RCTs and 3 large trials in patients with heart failure found a 32% relative risk reduction in death in patients receiving β blockers

循证医学

- 虽然循证医学的基础已经有了好几百年，但其基本概念的形成不过是近一二十年的事
- 循证医学是要解决临床问题的
- 最初，循证医学关注的焦点是，对某一特定的临床问题寻找和发现最佳的定量研究的客观证据，用于解决该临床问题
- 这一早期的聚焦点旨在弱化“将直觉、未经系统整理的临床经验和病理生理学原理作为临床决策的足够基础”，强调“检查来自临床研究的证据”

4

循证医学

- 以后的版本则强调，单有临床证据永远也不足以作出临床决策
- 研究证据本身很少告诉我们应该对个体病人作些什么，而是给我们提供了有用的信息，使我们能作出更符合这些信息的决策
- 临床医师必须结合病人的情况不断寻找和评价证据，然后对可能的得益与风险、费用和伴随的各种不便进行权衡

5

循证医学概念的发展

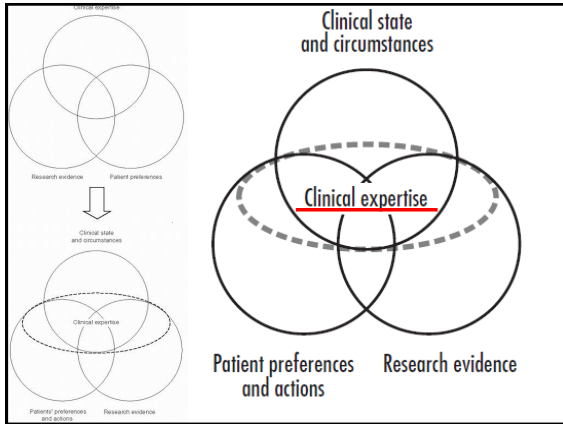
- “Evidence-based medicine is the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients”.

— Sackett DL, et al. *BMJ* 1996;312:71

- “Evidence-based medicine is the integration of best research evidence with clinical expertise and patient values”

— Haynes RB, et al. *ACP Journal Club* 2002;136:A11

6



循证治疗的步骤

1. 提出临床要解决的问题
2. 根据临床问题寻找最佳临床证据
3. 对临床证据进行评价
4. 根据治疗对象的个体情况和现有医疗条件、备选治疗，作出治疗决策
5. 实施决策
6. 后效的动态评估→循证调整治疗决策

8

1. 提出临床要解决的问题

- 目的
 - 治愈或根治疾病
 - 预防疾病或并发症的发生
 - 改善症状，提高运动耐量和生活质量
 - 延长寿命
- 循证的理由
 - 寻找对该病人最合适防治手段的临床证据
 - 临床疗效
 - 安全性
 - 经济
 - 其它

9

1. 提出临床要解决的问题

- 要解决什么问题：疗效？安全性？依从性？费用？
- 检验假设是什么
- 在此基础上确定
 - 研究对象
 - 治疗手段
 - 考核指标

10

2. 根据临床问题寻找最佳临床证据

- 旨在寻找解决临床问题的最佳证据
- 不仅包括阳性的结果的研究，也包括相反结果的研究
- 由于临床试验的发起和发表有不同程度的偏倚，因此，并非所有临床问题都能找到最佳证据
- 注意检索未发表的证据
- 医学实践的发展使得“最佳”证据都是相对
- 一个称职的临床医师必须从临床实践的需要出发，不断寻找和总结当前的最佳证据来指导自己的临床实践
- 临床实践中，也并非所有问题的解决都必须有临床试验的证据为依据

11

寻找临床证据：文献检索

- 一级证据来源
 - PubMed (美国)
 - EMBASE (欧洲)
 - 中文科技期刊数据库
 - 中国生物医学文献数据库 (CBM)
 - 中国期刊全文数据库 (CNKI)
 - 万方数据
- 二级来源证据
 - Cochrane图书馆
 - Best evidence
 - Evidence base medicine reviews

12

寻找临床证据：文献检索

- 临床指南
 - 美国国立指南数据库 (National guideline clearinghouse)
 - Guidelines (英国)
 - 各专业指南数据库
- 临床试验注册
 - Clinicaltrials.gov
 - 中国临床试验注册中心

13

3. 对临床证据进行评价

- 临床研究证据评估的目的：识别可靠、有用的临床研究证据
- 可靠：该研究的结果能反映真实的疗效和安全性（真实性/科学性）
- 有用：该研究的结果能为所研究临床问题的解决提供有用的信息

14

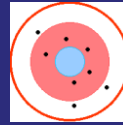
影响研究研究结果真实性的因素：两类误差

- 随机误差 (Random error) :
精确性

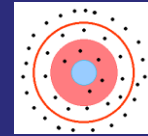


- 系统误差 (Systematic error), 偏倚 (Bias) :
准确性

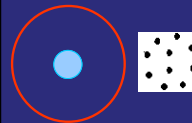
准确+精确



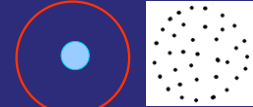
准确, 但不精确



精确, 但不准确



既不准确, 又不精确



证据评估的第一步：弄清基本问题

- 先通过文献阅读弄清研究证据是否切题
 - 首先看摘要
 - 理出
 - 该研究要解决的问题
 - 检验假设
 - 研究对象 (定义)
 - 考核的治疗手段 (定义)
 - 考核指标 (定义)

17

证据评估的第二步：识别系统误差

- 系统误差可以产生于临床研究中的任一环节
 - 研究设计
 - 研究实施
 - 结果分析

18

证据评估的第二步：识别系统误差

- 对象纳入误差：病例来源、受试者代表性、诊断正确性（诊断标准）
- 病情自然演变？向众数回归？
- 安慰剂作用？霍桑效应？
- 受试者对治疗方案的依从性和失访程度
 - ITT分析（意向分析）和PP分析，病例剔除及其原因

19

识别系统误差：正确设立对照组

- 是否按研究目的设置了正确的对照组？
 - 安慰剂对照：旨在平衡安慰剂效应
 - 阳性对照：旨在验证新治疗措施的疗效（安全性）是否更好
 - 不同剂量的对照：确定合适的剂量范围

20

识别系统误差：正确设立对照组

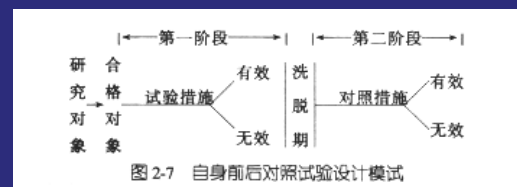
- 如果设立了对照组进行比较
 - 对照组设置是否合理（内容、方式）？
 - 随机分组的方法是否正确（包括实施）？
 - 各组间基线特征（影响疗效的因素）是否可比？
 - 各组间其它治疗措施有无不一致

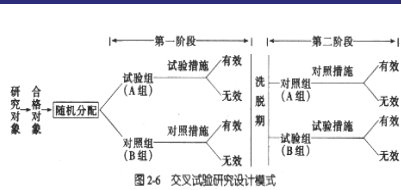
21

对理想的对照组的要求

对照组病例除了未接受所考核的治疗措施外，其病情特点和预后因素，以及同时接受的其它治疗措施均应与治疗组病例相同，亦即治疗组和对照组应均衡可比

22





25

识别系统误差：正确的随机化分组

- 将所有研究对象随机分配到治疗组和对照组，使每一对象都有同等的机会被分配到各组
- 其结果，是使各种预后因素（包括已知的和未知的）均匀地分布于各研究组，从而达到各组均衡可比的目的

26

识别系统误差：正确的随机化分组

- 使治疗组和对照组均衡可比的措施：随机化分组
 - 随机化：使受试者进入各组的机会相同
 - 正确的随机化方法
 - 简单随机化、区组随机化、分层随机化
 - 随机化的隐藏和破坏

27

什么是真正的随机化分组？

- 真正的随机化应符合下列原则：
 - 医生和病人都不能选择将进入的组别
 - 医生和病人都不能从上一个病人已经进入的组别推测出下一个病人将进入哪一组
 - 根据日历上单日进入研究的病人分配到治疗组，双日进入对照组的病人分配方法，不是真正的随机化分组方法

28

识别系统误差：盲法

- 减少来自研究人员和受试者的主观倚倚
 - 双盲、单盲和开放试验
 - 盲法评定终点
 - 安慰剂
 - 双盲双模拟
 - 仿盲？

29

Groups that can potentially be blinded in RCTs

- Participants
- Healthcare providers
- Data collectors
- Judicial assessors of outcomes
- Data analysts
- Data safety and monitoring committee
- Manuscript writers

DEVEREAUX EBM 2002 ;7:4

30

识别系统误差：疗效评价指标

- 选择具有临床意义的疗效评价指标
 - 延长寿命：生存率（急性期，远期）
 - 治愈或根治疾病：治愈率
 - 预防疾病或并发症的发生：发病率、心梗/脑梗发生率
 - 改善症状，提高运动耐量和生活质量：心绞痛发作次数、心功能分级、平板运动试验运动时间延长、生活质量量表、中医证候量表
 - 改善体征：神经功能评分、血压
 - 改善重要靶器官损伤
 - 改善实验室指标：血糖、血脂、肝肾功能

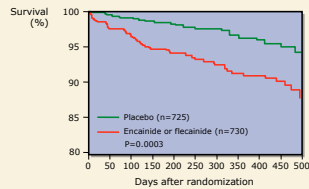
31

CAST（心律失常抑制试验）

- 1989年发表结果
- 背景：预后研究的结果显示心梗后频发室早者远期预后不良
- 目的：抑制心梗后频发室早可能改善预后
- 随机、先开放后双盲、安慰剂对照试验
- 1727例心梗后（6天~2年）频发室早者，在先开放阶段其早搏能为抗心律失常药控制者，被随机分到英卡氟或氟卡氟组（730例）、莫雷西嗪组（272例）和安慰剂对照组（725例）

32

All-cause mortality



CAST Investigators. *N Engl J Med* 1989;321:406-12.

CAST: Cardiac Arrhythmia Suppression Trial

	Placebo (n=725) No. (%)	Encainide/flecainide (n=730) No. (%)	Relative risk (95% CI)
Non-fatal cardiac arrest or death from arrhythmia	9 (1.2)	33 (4.5)	3.6 (1.7-8.5)
Other cardiac death	6 (0.8)	14 (1.9)	-
Non-cardiac or unclassified death or cardiac arrest	7 (1.0)	9 (1.2)	-
Total death or cardiac arrest	22 (3.0)	56 (7.7)	2.5 (1.6-4.5)
Average days of exposure	300	293	-

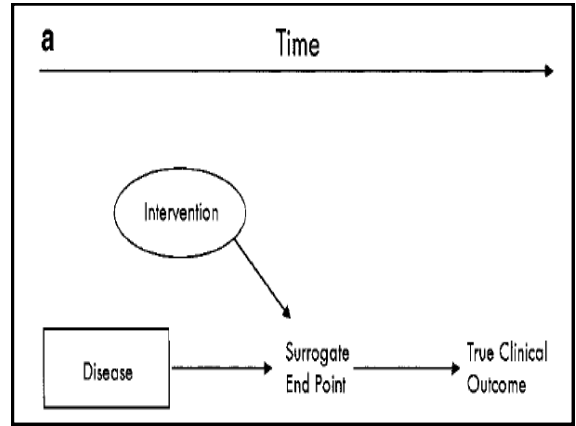
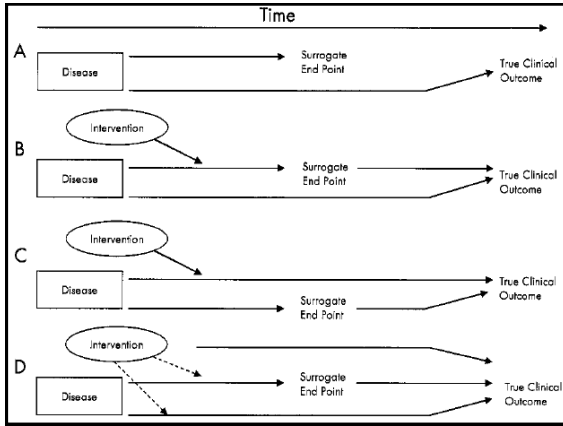
CAST: Cardiac Arrhythmia Suppression Trial

- In patients with asymptomatic or mildly symptomatic ventricular arrhythmias after MI, encainide or flecainide started at mean of 15 days after MI caused:
 - Excessive mortality risk
 - Excessive risk of death from arrhythmia

35

CAST: Cardiac Arrhythmia Suppression Trial

- 结果
 - 治疗10个月，英（氟）卡氟组心律失常病死率（4.5%）高于安慰剂对照组（1.2%）（RR 3.6，95% CI 1.7~8.5）
 - 英（氟）卡氟组总病死率（7.7%）高于安慰剂对照组（3.0%）。（RR 2.5，95% CI 1.6~4.5）
- 结论
 - 英（氟）卡氟不应该用于心梗后室性心律失常的治疗，不管其心律失常是否伴有症状，即使这些药物在开始服用时能有效地抑制室性心律失常



使用“替代终点”的危险性

治疗	疾病	替代终点	硬终点
英(氟)卡胺	心梗后频发室早	减少室性早搏	猝死↑
短效硝苯地平	原发性高血压	有效降压	心梗增多

识别系统误差：随访的完整性

- 失访的含义及解释
 - 如何判定失访对结果评定的影响
 - 失访资料的处理
 - ITT分析
 - PP分析

40

失访和不依从的处理

随机分组结果	A治疗		B治疗	
	未完成A治疗或改为B治疗	完成A治疗	完成B治疗	未完成B治疗或改为A治疗
	①	②	③	④

ITT 分析 (1+2~3+4)
PP 分析 (2+3)

两组患者的2年随访结果

随访2年结果	随机服药组		随机手术组	
	实际手术	实际服药	实际手术	实际服药
存活	48	296	354	20
死亡	2	27	15	6
合计	50	323	369	26

用两种分析方法计算的2年病死率

分析方法	服 药	手 术	卡 方	P 值
ITT 分析	29/373 (7.8%)	21/395 (5.3%)	1.9	0.17
PP 分析	27/323 (8.4%)	15/369 (4.1%)	5.6	0.018

研究结果的临床意义：拒绝缺乏临床意义的试验结果

- 相对危险度 (RR)
- 相对危险度的减少 (RRR)
- 绝对危险度 (AR) 和绝对危险度减少 (ARR)
- 需要治疗的病人数 (NNT)

44

证据评估的第二步：识别随机误差

- 随机误差：抽样误差，机遇
- 随机误差可以产生于临床研究中的许多环节：病例入选、样本大小、结果分析
- 以下重点讨论临床试验中治疗试验结果的随机误差问题，即所谓治疗结果的统计学意义

45

治疗结果的精确性估计：可信限

什么是可信限？

- 由于不可能让所有的该病患者都来接受同一临床试验，因此，治疗的总体效果不可能准确得知
- 所能得到的，是由其中某一部分具有代表性的病人样本经过严格的临床试验所提供的疗效估计值，即“点上的估计值”，总体真正的疗效值应位于该点附近
- 根据所研究样本的结果，可以计算出一个区间，总体的真正效果有 95% 的可能位于这一区间内，这一区间称为95%可信区间，或95%可信限

46

治疗结果的精确性估计：可信限

- 95%可信限的含义
 - 真正的疗效（如RRR）有95%的机会位于此范围内。
 - 或者，如果进行100次类似的临床试验，其中有95次试验的 95% CIs 中会包括总体真正的疗效值
- RRR CI 的估计
 - $CI = RRR \pm 2SE$
 - 根据P值估计: 如 $P=0.05$, 则RRR的95%CI的下限为0, 不能排除治疗没有作用的可能。如 $P < 0.05$, 则RRR的95%的可信限的下限 >0

47

治疗结果的精确性估计：可信限

- 95% 可信限的计算：
 - $95\% CI = RRR \pm 2SE$ (标准误)
 - 根据P值估计
 - 如 $P=0.05$, 则 RRR 95% CI 的下限为0, 不能排除治疗没有作用的可能
 - 如 $P < 0.05$, 则 RRR 95% CI 的下限 >0

48

治疗结果的精确性估计：可信限

- 可信限比 P 值更有用，因为它能告诉我们有关研究结果的论证强度，而P值仅对无效假设进行检验
- 临床试验的样本量越大，可信限越窄，越可以相信所观察得到的疗效值（如RRR）接近真正的疗效

49

治疗结果的精确性估计：可信限

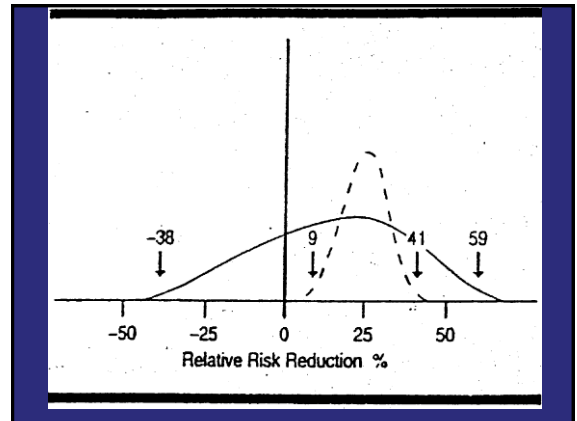
- 一项临床试验将200例病人随机分入治疗组和对照组，每组各100例
- 随访结果对照组死亡20例，治疗组死亡15例
- 对RRR的点上的估计为25%。计算得到RRR的95%CI为-38%至+59%
- 表示总体真正治疗效果有95%的可能是在“接受新疗法的病人较对照组病人的病死率高38%”至“接受新疗法的病人较对照组病死率低59%”这一范围内
- 显然，这一结果对医师决定是否要给予病人以新疗法毫无帮助

50

治疗结果的精确性估计：可信限

- 如上述研究中每组病例数为各1000例
- 而所观察到的病死率与上例中相同，即对照组死亡200例，而治疗组死亡150例
- 同样，RRR的点估计为25%
- 而计算得到的RRR的95%CI为+9%至+41%，全部大于 0
- 表示总体的真正治疗效果有95%的可能是在“接受新疗法的病人较对照组病人的病死率低9%至41%”的范围内

51



三种不同的临床试验

- 预先设定治疗组和对照组之间差别的**优效**和**非劣效**界限，并入选**足够数量**的研究对象，保证研究结果有足够的把握度
- 优效试验
- 非劣效试验
- 等效试验

53

治疗证据强度的分层

- 但是
 - 所有的证据都有价值
 - 所有的证据都有局限性

54

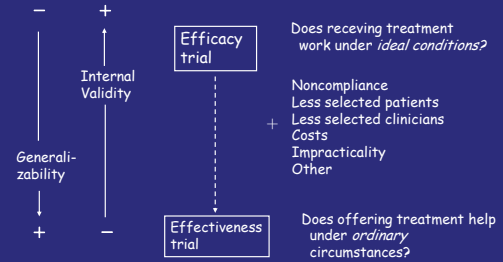
治疗证据强度的分层

- 多种来源证据的综合
- 若干设计良好的大型RCT的系统综述
- 单个大型 RCT
- 若干设计良好的较小RCT的系统综述
- 单个小样本 RCT
- 若干设计良好的观察性研究的系统综述
- 单个观察性研究
- 生理学研究
- 来自医师的非系统性观察
- 动物研究



55

Efficacy trial 和 effectiveness trial



临床问题

- 80岁男性高血压患者，发现房颤10个月
- 诊断房颤前一个月消化道出血住院，急症胃镜检查显示十二指肠溃疡出血，幽门螺杆菌，治疗后稳定
- 心超（食道）：心瓣膜及左心室功能正常，左房增大（65mm），未见血栓
- 复律未成功
- 患者十分担心会发生脑梗，因其妻曾于脑梗后2年去世
- 主诊医师最近有另一位服用华法林的病人发生消化道大出血，考虑到本病人的年龄和消化道出血史，因而十分顾虑其服用华法林的出血风险

58

临床问题

- 要解决的问题：华法林预防慢性房颤病人脑梗塞的疗效和出血风险
- 检验假设是什么：与安慰剂相比，华法林能减少慢性房颤病人脑梗塞的发生而具有可接受的出血风险
- 在此基础上确定
 - 研究对象：发生过上消化道出血的慢性房颤病人
 - 治疗手段：华法林
 - 考核指标：脑梗死，出血

59

房颤治疗的循证决策

- 文献检索证据
 - 无任何抗栓治疗，脑卒中8%/年，大出血1%/年
 - 华法林治疗，脑卒中3%/年，大出血8%/年。
 - 如愿接受INR自我监测，大出血降至4%/年
 - 阿司匹林治疗，脑卒中6%/年，大出血1.3%/年
- 患者知情，告知
 - 什么是房颤
 - 什么是脑卒中和大出血
 - 自我监测凝血酶原时间

60

房颤治疗的循证决策

- 患者表示愿意参加治疗决策
- 医师告知：对该个体患者来说，华法林治疗、阿司匹林治疗和无抗栓治疗的血栓事件/大出血风险
- 在仔细考虑这一信息之后，患者决定选择华法林治疗；他了解了华法林治疗会增加出血风险，但更看重避免发生脑卒中
- 在作出这一决策后，需要监测患者，以保证决策的实施

61