EBM簡介及臨床應用

高雄榮民總醫院 藥劑部
林佩津 總藥師

Outline

- 何謂實證醫學
  - 5A (ask, acquire, appraise, apply, audit)
  - 6S
- 實證醫學之臨床應用

Review of History

- 1960's: Dave Sackett: Nephrologist at USA.
- 1972: Archie Cochrane : Publish : Effectiveness and Efficiency : RCT.
- 1980's: Clinical Epidemiology & Biostatistics at McMaster University: Canada⇨實證醫療〔Evidence-based clinical practice〕
- 1992:
  - Gorden Guyatt in McMaster U. ⇨EBM
  - UK: Cochrane Collaboration by NHS(national health service) for review group. (1993)

Evidence-Based Medicine

The ability to track down, critically appraise, and incorporate this rapidly growing body of evidence into one's clinical practice has been named EBM

5

實證醫學五步驟

- 提出問題 (Question Formulation) – ask
- 搜尋證據 (Evidence Search) – access, acquire
- 嚴格評讀 (Critical Appraisal) – appraisal
- 恰當運用 (Evidence Application) – apply
- 衡量結果 (Outcome Evaluation) – audit

2012/9/14

6

1. Asking an answerable question

- Background question: ask for general knowledge about a certain condition, illness, or some aspect of health status
  - a question root (ie, who, what, when, where, why, how)
  - a disorder, test, or treatment (eg, hypertension, angiography, or exercise)

2012/9/14

Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. 3rd ed. London: Churchill Livingstone;2005

7

1. Asking an answerable question

- Foreground question

- PICO(D)
  - What type of **P**articipants?
  - What types of **I**nterventions?
  - What types of **C**omparison?
  - What types of **O**utcomes?

- D: design
  - Etiology
  - Diagnosis
  - Causation
  - Therapy or Prognosis

- Diagnosis Questions formulated : PPICO
  - **P**opulations of interest
  - **P**rior test(s) (if appropriate)
  - **I**ntervention
  - **C**omparisons (if appropriate)
  - **O**utcomes

2012/9/14

8

2. Acquire the evidence

- What evidence should we search?
  - For a clinical question?
    - A good-quality study my be enough
  - For a decision making (policy)
    - Several good-quality studies will be enough
    - Systematic review and meta-analysis

2012/9/14

9

Table 3. Clinical Outcomes

Study	Treatment	Daily Dose, mg	Study Duration	Clinical Outcome: Severity	Clinical Outcome: Resolution Rate According to Defined Time or Average Time to Resolution
Single-agent, open-label studies					
Breitbart[25]	Olanzapine	Initial: 3.0 ± 0.14	7 days	–	MDAS ≤10: 45% by Time 2 (Day 2–3), 76% by Time 3 (Day 4–7)
		Final: 6.3 ± 0.52			
Parellada[26]	Risperidone	Initial: 2.6 ± 1.3	7 days	–	DRS <13 by Day 3: 91%
		Final: 1.5 ± 0.8			
Sasaki[27]	Quetiapine	Mean: 44.9 ± 31	"Until resolution"	–	DRS-J < 12: 4.8 ± 3.5 days
	Olanzapine				
	Control	Final: haloperidol 3.4 ± 1.62, olanzapine 2.78 ± 1.85			
Kim[32]	Haloperidol	Initial: haloperidol 2.67 ± 2.71, risperidone 0.97 ± 0.67	7 days	–	DRS-R-98 < 15.25: 6.7 ± 5.5 days haloperidol, 4.8 ± 3.4 days risperidone, $P = .23$
	Risperidone	Final: haloperidol 1.67 ± 1.32, risperidone			

Comparison: 03 Treatment versus Placebo
Outcome: 01 Effect of treatment on mortality

Study	Treatment n/N	Control n/N	OR (95%CI Fixed)	Weight %	OR (95%CI Fixed)
Brown 1998	24 / 472	35 / 499		9.6	0.71[0.42,1.21]
Geoffrey 1997	120 / 2850	182 / 2838		51.8	0.64[0.51,0.81]
Mason 1996	56 / 2051	84 / 2030		24.4	0.65[0.46,0.92]
Peters 2000	5 / 81	4 / 78		1.1	1.22[0.31,4.71]
Scott 1998	31 / 788	46 / 792		13.1	0.66[0.42,1.06]
Total(95%CI)	236 / 6242	351 / 6237		100.0	0.66[0.56,0.78]

Test for heterogeneity chi-square=0.92 df=4 p=0.92
Test for overall effect z=-4.82 p<0.00001

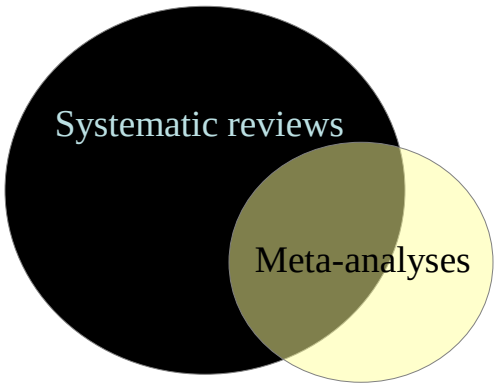
.1 .2 1 5 10
Favours treatment Favours control

2012/9/14

10

What is a meta-analysis?

Optional part of a systematic review



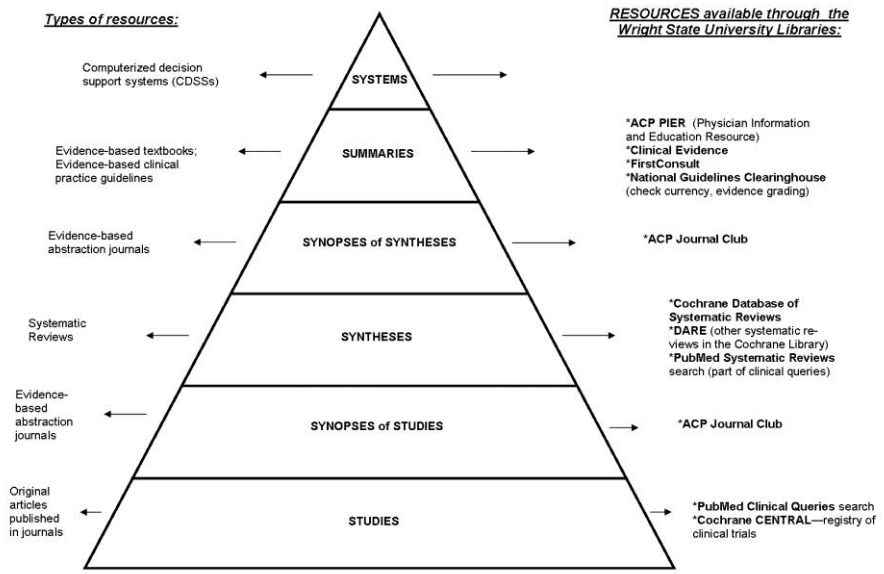
2012/9/14

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

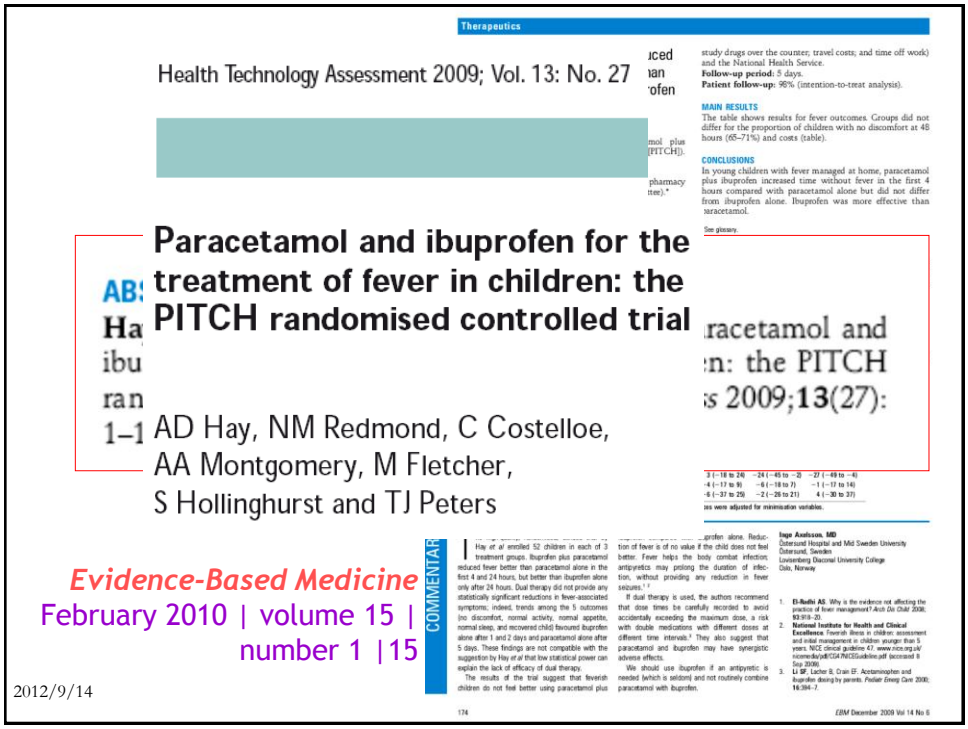
Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard**	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or *n*-of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, *n*-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or *n*-of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non -randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

2012/9/14 OCEBM Levels of Evidence Working Group*. "The Oxford 2011 Levels of Evidence".

"6S" Hierarchy of Preappraised Evidence
From: Haynes et al: ACP J Club, 2009 Sep 15; 151:JC3-2



BS, 1/10



Therapeutics

Health Technology Assessment 2009; Vol. 13: No. 27

Paracetamol and ibuprofen for the treatment of fever in children: the PITCH randomised controlled trial

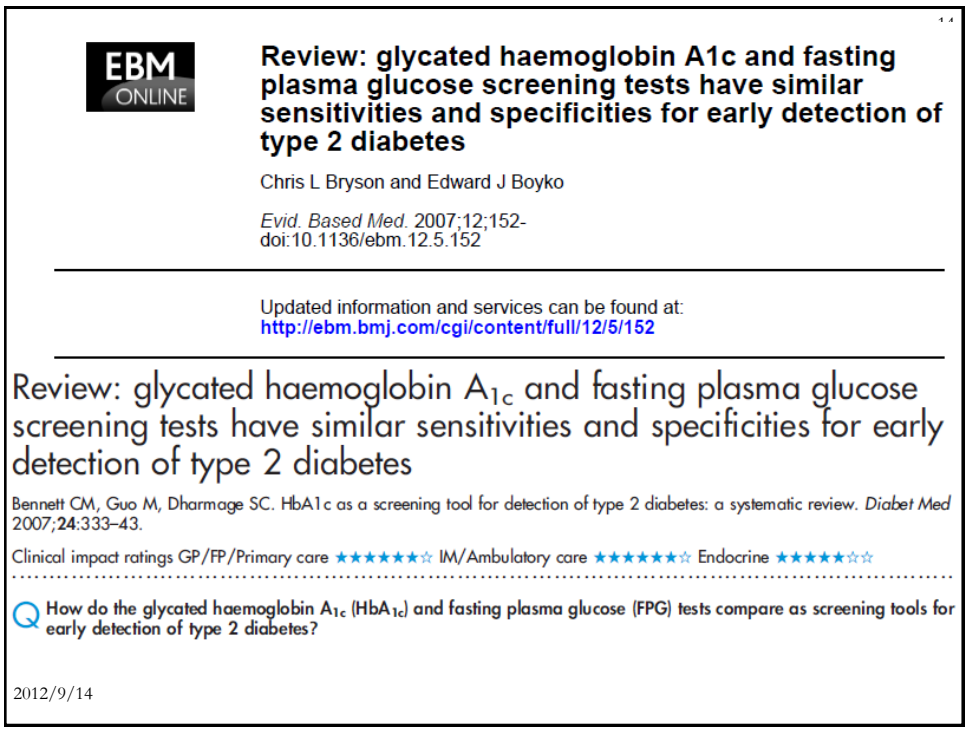
AD Hay, NM Redmond, C Costelloe, AA Montgomery, M Fletcher, S Hollinghurst and TJ Peters

Evidence-Based Medicine
February 2010 | volume 15 | number 1 |15

COMMENTARY

Inge Axelsson, MD
Östersund Hospital and Mid Sweden University
Östersund, Sweden
Lovisenberg Diaconal University College
Oslo, Norway

2012/9/14



EBM ONLINE

Review: glycated haemoglobin A1c and fasting plasma glucose screening tests have similar sensitivities and specificities for early detection of type 2 diabetes

Chris L Bryson and Edward J Boyko

Evid. Based Med. 2007;12;152-
doi:10.1136/ebm.12.5.152

Updated information and services can be found at:
http://ebm.bmj.com/cgi/content/full/12/5/152

Review: glycated haemoglobin A_{1c} and fasting plasma glucose screening tests have similar sensitivities and specificities for early detection of type 2 diabetes

Bennett CM, Guo M, Dharmage SC. HbA1c as a screening tool for detection of type 2 diabetes: a systematic review. *Diabet Med* 2007;**24**:333–43.

Clinical impact ratings GP/FP/Primary care ★★★★★★☆ IM/Ambulatory care ★★★★★★☆ Endocrine ★★★★★☆☆

Q How do the glycated haemoglobin A_{1c} (HbA_{1c}) and fasting plasma glucose (FPG) tests compare as screening tools for early detection of type 2 diabetes?

2012/9/14

METHODS

Data sources: Medline, EMBASE/Excerpta Medica (1994 to September 2004), and bibliographies of relevant articles.

Study selection and assessment: English language cross-sectional studies that compared the HbA_{1c} test with the FPG test as screening tools for detecting type 2 diabetes, reported sensitivities and specificities using the 75 g oral glucose tolerance test (OGTT) as the reference standard, and reported HbA_{1c} test results in a format compatible with that of the Diabetes Control and Complications Trial. 9 studies (n = 19 500, 13–92 y) published in 1998–2004 met the selection criteria; 4 studies were community-based, and 5 were hospital-based. Quality assessment of individual studies was based on the sampling frame and size, measurements of HbA_{1c} and FPG, adequacy of test descriptions, and 80% verification with OGTT.

Outcomes: sensitivity, specificity, and likelihood ratios (LRs).

MAIN RESULTS

In ≥2 studies, receiver operating characteristic curves identified 5.9%, 6.1%, and 6.2% as optimal HbA_{1c} cut-points and 5.6 mmol/l and 6.1 mmol/l as FPG cut-points for diagnosing diabetes. The sensitivities, specificities, and LRs for the HbA_{1c} and FPG tests were similar for detecting diabetes (table). At equivalent cut-points, both the HbA_{1c} and FGP tests had generally lower sensitivities (about 50%) for detecting impaired glucose tolerance.

CONCLUSIONS

The glycated haemoglobin A_{1c} and fasting plasma glucose screening tests have similar sensitivities and specificities for early detection of type 2 diabetes. The HbA_{1c} and FPG tests have lower sensitivities for detecting impaired glucose tolerance.

For correspondence: Dr C M Bennett, University of Melbourne, Victoria, Australia. c.bennett@unimelb.edu.au

Commentary

The systematic review by Bennett *et al* compared HbA_{1c}, a measure commonly used to assess diabetes control, with FPG, using OGTT as the gold standard in screening for type 2 diabetes. Currently, the American Diabetes Association (ADA) recommends 2 tests to screen for diabetes among asymptomatic individuals: plasma glucose obtained after 8 hours of fasting or an OGTT given according to the World Health Organization protocol.[1] Both screening tests require patient preparation, which may promote missed opportunities for screening, and because of burden, OGTT is not usually done except in pregnant women. A third test, random plasma glucose, is recommended for diagnostic testing in symptomatic patients. Unless there is clear evidence of hyperglycaemia, it is recommended that each of these tests be repeated.

HbA_{1c} testing is an attractive alternative as it reflects average plasma glucose concentrations over 60–90 days and does not require additional burden from patients other than phlebotomy. However, although the results of the OGTT, FPG, and HbA_{1c} tests are correlated, the latter 2 identify (or miss) different patients, and a combined screening strategy may prove more effective for minimising the false negative rate.[2] Concerns about HbA_{1c} standardisation have also been raised by the ADA and International Diabetes Federation, neither of which currently recommends HbA_{1c} as a screening test for type 2 diabetes.

Any single HbA_{1c} cut-point to rule in or rule out diabetes would lead to significant misdiagnosis. Hence, separate HbA_{1c} values may be needed: one that clearly rules out diabetes mellitus and a higher HbA_{1c} value that clearly rules it in. Patients with values between these thresholds would need to proceed to an OGTT. At this time, clinicians should continue to use the current approach of screening high risk patients with an FPG test and recommending an OGTT if the value is ≥5.6 mmol/l.[3]

Chris L Bryson, MD, MS
Edward J Boyko, MD, MPH
Veterans Affairs Puget Sound Health Care System
Seattle, Washington, USA

1 American Diabetes Association. *Diabetes Care* 2007;**30**:S42–7.
2 Perry RC, Shankar RR, Fineberg N, *et al. Diabetes Care* 2001;**24**:465–71.
3 American Diabetes Association. *Diabetes Care* 2007;**30**:S4–41.

When we concerned about EBM - in real world

- Updated Guideline
- Find a systematic review (meta-analysis)
  - Appraise SR
- If a SR could not be fined
  - Find relevant RCTs
  - Appraise RCTs
  - -------

2012/9/14

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Hearing Impairment in Childhood Bacterial Meningitis Is Little Relieved by Dexamethasone or Glycerol

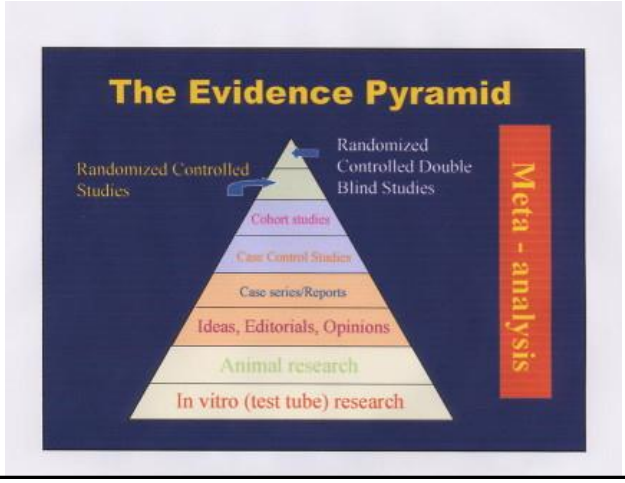
Heikki Peltola, Irmeli Roine, Josefina Fernández, Antonio González Mata, Inés Zavala, Silvia Gonzalez Ayala, Antonio Arbo, Rosa Bologna, José Goyo, Eduardo López, Greta Miño, Solange Dourado de Andrade, Seppo Sarna and Tapani Jauhiainen

Pediatrics 2010;125;e1-e8; originally published online Dec 14, 2009;
DOI: 10.1542/peds.2009-0395

METHODS. Children aged 2 months to 16 years with meningitis were treated with ceftriaxone but were double-blindly randomly assigned to receive adjuvant dexamethasone intravenously, glycerol orally, both agents, or neither agent. We used the Glasgow coma scale to grade the presenting status. The end points were the better ear's ability to detect sounds of >40 dB, ≥60 dB, and ≥80 dB, with these thresholds indicating any, moderate-to-severe, or severe impairment, respectively. All

3. Critical appraisal – example as intervention

1. Meta-analysis 2. Systemic review



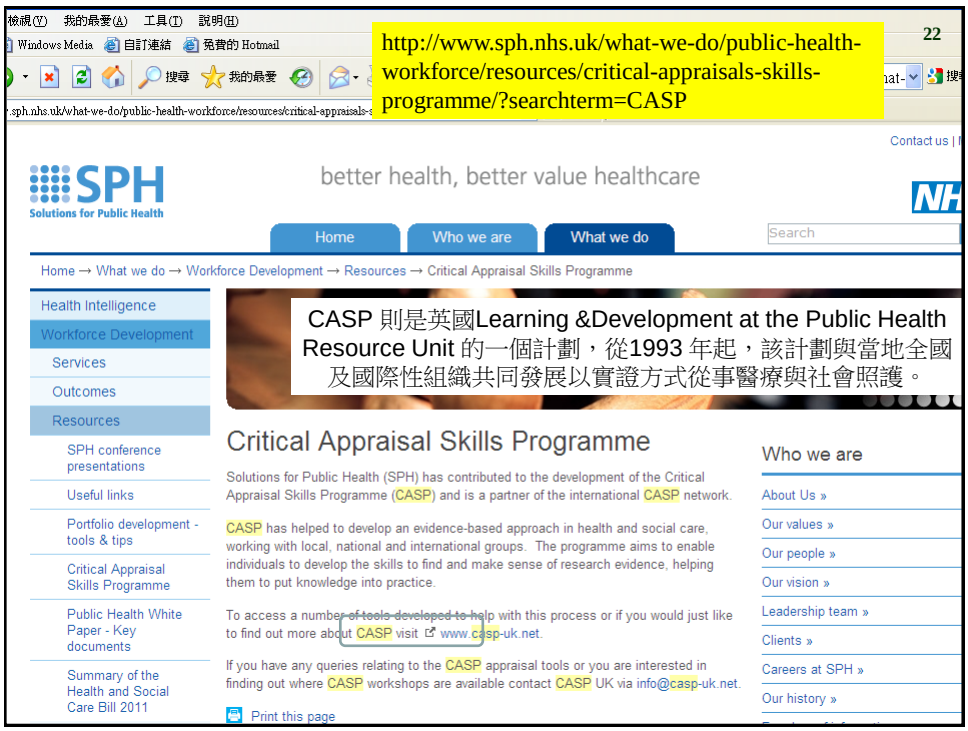
Grade of Recommendation	Level of Evidence	Therapy
[A]	1a	Systemic review of RCTs
	1b	Single RCT
	1c	'All-or-none'
[B]	2a	Systemic review of cohort studies
	2b	Cohort study or poor RCT
	2c	'Outcomes' research
	3a	Systemic review of case-control studies
	3b	Case-control study
[C]	4	Case series
[D]	5	Expert opinion, physiology, bench research

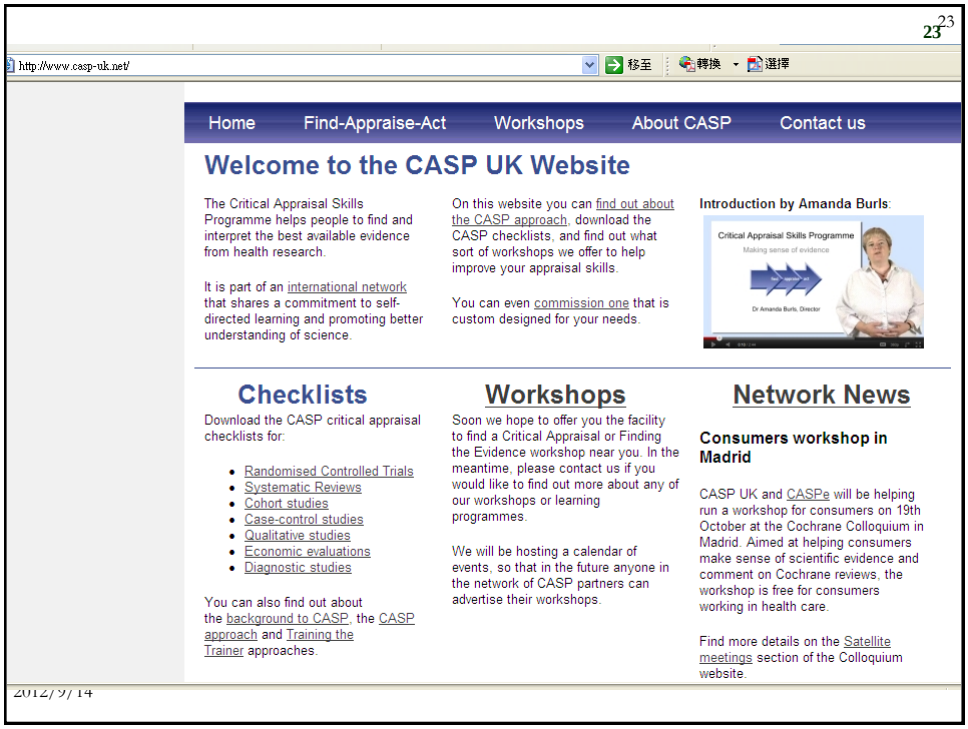
3. Critical Appraise

- Not journal reading
- 清楚描述挑選文獻的原因(publication year, sample size, study quality, race or ethnicity ---)
- 正確使用評讀工具
- Validity and importance
- 正確評定Level of evidence (比賽時更須要)

Critical Appraisal Tools

- Systemic reviews
- Randomized controlled trials
- Cohort studies
- Case control studies
- Qualitative research studies
- Economic evaluation studies
- Diagnosed test studies




23
http://www.casp-uk.net/
移至
轉換
選擇
Home
Find-Appraise-Act
Workshops
About CASP
Contact us
Welcome to the CASP UK Website
The Critical Appraisal Skills Programme helps people to find and interpret the best available evidence from health research.
It is part of an international network that shares a commitment to self-directed learning and promoting better understanding of science.
On this website you can find out about the CASP approach, download the CASP checklists, and find out what sort of workshops we offer to help improve your appraisal skills.
You can even commission one that is custom designed for your needs.
Introduction by Amanda Burls:
Critical Appraisal Skills Programme
Making sense of evidence
Dr Amanda Burls, Director
Checklists
Download the CASP critical appraisal checklists for:
Randomised Controlled Trials
Systematic Reviews
Cohort studies
Case-control studies
Qualitative studies
Economic evaluations
Diagnostic studies
You can also find out about the background to CASP, the CASP approach and Training the Trainer approaches.
Workshops
Soon we hope to offer you the facility to find a Critical Appraisal or Finding the Evidence workshop near you. In the meantime, please contact us if you would like to find out more about any of our workshops or learning programmes.
We will be hosting a calendar of events, so that in the future anyone in the network of CASP partners can advertise their workshops.
Network News
Consumers workshop in Madrid
CASP UK and CASPe will be helping run a workshop for consumers on 19th October at the Cochrane Colloquium in Madrid. Aimed at helping consumers make sense of scientific evidence and comment on Cochrane reviews, the workshop is free for consumers working in health care.
Find more details on the Satellite meetings section of the Colloquium website.
2012/9/14


http://www.cebm.net/index.aspx?o=1157
Google
檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)
趨勢科技工具列
我的最愛
CEBM > EBM Tools > Critical Appraisal > Overview
首頁(M)
閱讀郵件
Home
About the CEBM
Publications
Training & Education
EBM Tools
Resource Centre
Blog
Contact
CEBM
CENTRE FOR EVIDENCE BASED MEDICINE
UNIVERSITY OF OXFORD
Quickfinder
CEBM > EBM Tools > Critical Appraisal > Overview
Asking Focused Questions
Finding the Evidence
Critical Appraisal
Study Designs
Explanations and Examples
Critical Appraisal Sheets
EBM Calculators
CATmaker
Making a Decision
Evaluating Performance
Designing Research
Critical Appraisal
This section contains useful tools and downloads for the critical appraisal of medical evidence. Example appraisal sheets are provided together with several helpful examples. Below, you can download our calculators, as well as our PC-based software tool CATmaker:
Critical Appraisal Sheets
Systematic Review Critical Appraisal Sheet
Diagnostic Critical Appraisal Sheet
Prognosis Critical Appraisal Sheet
RCT Critical Appraisal Sheet
PICO Critical Appraisal Sheet (PDF)
PICO Critical Appraisal Sheet (MS-Word)
Educational Prescription Critical Appraisal Sheet (PDF)
Explanations & Examples
Pre-test probability
SpPin and SnNout
Likelihood Ratios
NNTs
Calculators
All-purpose 2x2 Table
The "CATmakers Scratching Post"
Download (Adobe Flash format)
Interactive Nomogram
CATmaker
CATmaker is a computer-assisted critical appraisal tool, which helps you
What's New
Levels of Evidence #2
Be the first to review our updated 'Levels of Evidence' table.
More Information & FREE Downlo
Workshop on Evidence-Based Practice
(1 day) 24th June 2011
More Information
iPhone App
...critical evidence-based decision making
Doctor's Toolbo
Clinical Prediction Rules, evidenc summaries & excerpts from Best Practice.
Download from iTunes
完成
網際網路

VIP

- **V** (Valid/Reliability)： Are the results of the article valid? 研究方法的探討;”我們能相信這篇文獻嗎?”
- **I** (Important/Impact)： What are the results? 結論的分析;”我們相信它，但這個結論重要嗎?”
- **P** (Practice/Applicability)： Will the results help me in caring for my patients? 如何運用來照顧病人;”如果我們相信它，這個結論可以應用在我們所有的病患嗎?”

Critical Appraisal for Therapy Articles

THERAPY STUDY: **Are the results of the trial valid? (Internal Validity)**

What question did the study ask?

Patients -

Intervention -

Comparison -

Outcome(s) -

1a. R- Was the assignment of patients to treatments randomised?	
What is best?	Where do I find the information?
Centralised computer randomisation is ideal and often used in multi-centred trials. Smaller trials may use an *independent* person (e.g. the hospital pharmacy) to "police" the randomization.	The ***Methods*** should tell you how patients were allocated to groups and whether or not randomisation was concealed.
This paper: Yes ☐ No ☐ Unclear ☐	
Comment:	
1b. R- Were the groups similar at the start of the trial?	
What is best?	Where do I find the information?
If the randomisation process worked (that is, achieved comparable groups) the groups should be similar. The more similar the groups the better it is. There should be some indication of whether differences between groups are statistically significant (ie. p values)	The ***Results*** should have a table of "Baseline Characteristics" comparing the randomized groups on a number of variables that could affect the outcome (ie. age, risk factors etc). If not, there may be a description of group similarity in the first paragraphs of the ***Results*** section.
This paper: Yes ☐ No ☐ Unclear ☐	
Comment:	
2a. A - Aside from the allocated treatment, were groups treated equally?	
What is best?	Where do I find the information?
Apart from the intervention the patients in the different groups should be treated the same, eg., additional treatments or tests.	Look in the ***Methods*** section for the follow-up schedule, and permitted additional treatments, etc and in ***Results*** for actual use.
This paper: Yes ☐ No ☐ Unclear ☐	
Comment:	
2b. A - Were all patients who entered the trial accounted for? - and were they analysed in the groups to which they were randomised?	
What is best?	Where do I find the information?
Losses to follow-up should be minimal – preferably less than 20%. However, if few patients have the outcome of interest, then even small losses to follow-up can bias the results. Patients should also be analysed in the groups to which they were randomised – *'intention-to-treat analysis'*.	The ***Results*** section should say how many patients were randomised (eg., Baseline Characteristics table) and how many patients were actually included in the analysis. You will need to read the results section to clarify the number and reason for losses to follow-up.
This paper: Yes ☐ No ☐ Unclear ☐	
Comment:	
3. M - Were measures objective or were the patients and clinicians kept "blind" to which treatment was being received?	
What is best?	Where do I find the information?
It is ideal if the study is 'double-blinded' – that is, both patients and investigators are unaware of treatment allocation. If the outcome is *objective* (eg., death) then blinding is less critical. If the outcome is *subjective* (eg., symptoms or function) then blinding of the outcome assessor is critical.	First, look in the ***Methods*** section to see if there is some mention of masking of treatments, eg., placebos with the same appearance or sham therapy. Second, the ***Methods*** section should describe how the outcome was assessed and whether the assessor/s were aware of the patients' treatment.
This paper: Yes ☐ No ☐ Unclear ☐	
Comment:	

What were the results?

1. How large was the treatment effect?

Most often results are presented as dichotomous outcomes (yes or not outcomes that happen or don't happen) and can include such outcomes as cancer recurrence, myocardial infarction and death. Consider a study in which 15% (0.15) of the control group died and 10% (0.10) of the treatment group died after 2 years of treatment. The results can be expressed in many ways as shown below.

What is the measure?	**What does it mean?**
Relative Risk (RR) = risk of the outcome in the treatment group / risk of the outcome in the control group.	The relative risk tells us **how many times more likely** it is that an event will occur in the treatment group relative to the control group. An **RR** of **1** means that there is no difference between the two groups thus, the treatment had **no effect**. An RR < 1 means that the treatment decreases the risk of the outcome. An RR > 1 means that the treatment increased the risk of the outcome.
In our example, the RR = 0.10/0.15 = 0.67	Since the RR < 1, the treatment decreases the risk of death.
Absolute Risk Reduction (ARR) = risk of the outcome in the control group - risk of the outcome in the treatment group. This is also known as the **absolute risk difference.**	The absolute risk reduction tells us the absolute difference in the rates of events between the two groups and gives an indication of the baseline risk and treatment effect. An **ARR** of **0** means that there is no difference between the two groups thus, the treatment had **no effect**.
In our example, the ARR = 0.15 - 0.10 = 0.05 or 5%	The absolute benefit of treatment is a 5% reduction in the death rate.
Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. An alternative way to calculate the RRR is to subtract the RR from 1 (eg. RRR = 1 - RR)	The relative risk reduction is the complement of the RR and is probably the most commonly reported measure of treatment effects. It tells us the reduction in the rate of the outcome in the treatment group relative to that in the control group.
In our example, the RRR = 0.05/0.15 = 0.33 or 33% Or RRR = 1 - 0.67 = 0.33 or 33%	The treatment reduced the risk of death by 33% relative to that occurring in the control group.
Number Needed to Treat (NNT) = inverse of the ARR and is calculated as 1 / ARR.	The number needed to treat represents the number of patients we need to treat with the experimental therapy in order to prevent 1 bad outcome and incorporates the duration of treatment. Clinical significance can be determined to some extent by looking at the NNTs, but also by weighing the NNTs against any harms or adverse effects (NNHs) of therapy.
In our example, the NNT = 1/ 0.05 = 20	We would need to treat 20 people for 2 years in order to prevent 1 death.

2. How precise was the estimate of the treatment effect?

The true risk of the outcome in the population is not known and the best we can do is estimate the true risk based on the sample of patients in the trial. This estimate is called the **point estimate**. We can gauge how close this estimate is to the true value by looking at the confidence intervals (CI) for each estimate. If the confidence interval is fairly narrow then we can be confident that our point estimate is a precise reflection of the population value. The confidence interval also provides us with information about the statistical significance of the result. If the value corresponding to **no effect** falls outside the 95% confidence interval then the result is statistically significant at the 0.05 level. If the confidence interval includes the value corresponding to **no effect** then the results are not statistically significant.

Will the results help me in caring for my patient? (External Validity/Applicability)

The questions that you should ask before you decide to apply the results of the study to your patient are:

- Is my patient so different to those in the study that the results cannot apply?
- Is the treatment feasible in my setting?
- Will the potential benefits of treatment outweigh the potential harms of treatment for my patient?

JAMA®

Online article and related content current as of November 18, 2010.

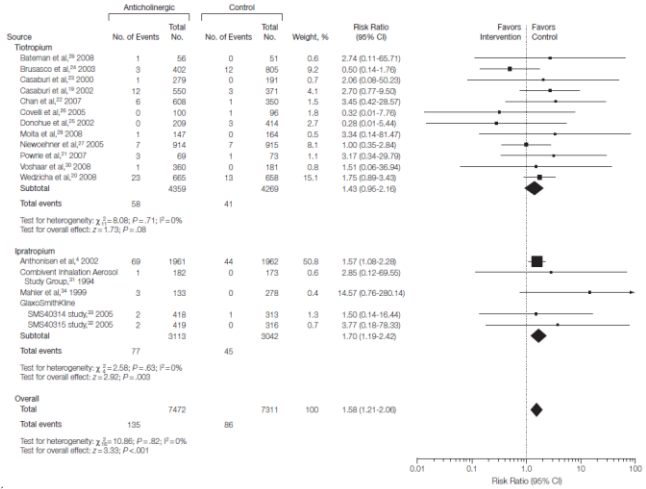
Inhaled Anticholinergics and Risk of Major Adverse Cardiovascular Events in Patients With Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-analysis

Sonal Singh; Yoon K. Loke; Curt D. Furberg

JAMA. 2008;300(12):1439-1450 (doi:10.1001/jama.300.12.1439)

27

Figure 2. Meta-analysis of Randomized Controlled Trials of Inhaled Anticholinergics vs Control for Major Adverse Cardiovascular Outcomes Composite



20

Cardiovascular outcomes composite indicates cardiovascular death, myocardial infarction, and stroke. Size of the data markers indicates weight of the study. CI indicates confidence interval.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 OCTOBER 9, 2008 VOL. 359 NO. 15

A 4-Year Trial of Tiotropium in Chronic Obstructive Pulmonary Disease

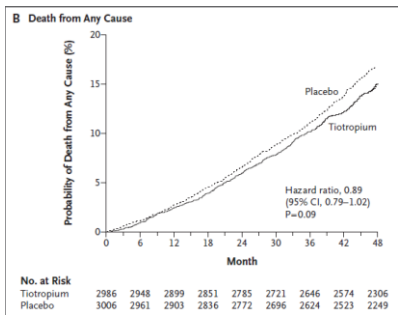


Table 4. Incidence Rate of Serious Adverse Events per 100 Patient-Years.*

Adverse Event	Tiotropium (N=2986)	Placebo (N=3006)	Relative Risk for Tiotropium vs. Placebo (95% CI)
Cardiac	3.56	4.21	0.84 (0.73–0.98)†
Angina	0.51	0.36	1.44 (0.91–2.26)
Atrial fibrillation	0.74	0.77	0.95 (0.68–1.33)
Cardiac failure	0.61	0.48	1.25 (0.84–1.87)
Congestive heart failure	0.29	0.48	0.59 (0.37–0.96)†
Coronary artery disease	0.21	0.37	0.58 (0.33–1.01)
Myocardial infarction	0.69	0.97	0.71 (0.52–0.99)†
Lower respiratory	11.32	13.47	0.84 (0.77–0.92)†
Bronchitis	0.37	0.31	1.20 (0.73–1.98)
COPD exacerbation	8.19	9.70	0.84 (0.76–0.94)†
Dyspnea	0.38	0.62	0.61 (0.40–0.94)†
Pneumonia	3.28	3.46	0.95 (0.81–1.11)
Respiratory failure	0.90	1.31	0.69 (0.52–0.92)†

* Listed are the incidence rates of serious adverse events (excluding lung cancer) that were reported by more than 1% of patients in either study group, according to organ class during the study period (from the first day of administration of a study drug until the last day plus 30 days).
† P<0.05.

2012/9/14

N Engl J Med 2008;359:1543-54.

29

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

The Safety of Tiotropium — The FDA's Conclusions

Theresa M. Michele, M.D., Simone Pinheiro, Sc.D., and Solomon Iyasu, M.D., M.P.H.

N Engl J Med 2010; 363:1097-1099 | September 16, 2010

- Because of the strength of the UPLIFT data, the absence of a strong signal related to stroke or cardiovascular events with tiotropium, and the potential methodologic limitations of the Singh meta-analysis, the FDA concluded that current data do not support the conclusion that there is an increased risk of stroke, heart attack, or death associated with tiotropium HandiHaler

Safety Data from Pooled Analysis of Tiotropium Trials and UPLIFT.*

Attribute	29 Pooled Trials (N=13,544)	UPLIFT (N=5992)
Study duration	1–12 mo	48 mo
Patient-years (placebo group)	3065	8499
Patient-years (tiotropium group)	4571	9222
Relative risk (95% CI)		
Stroke	1.37 (0.73–15.6)	0.95 (0.70–1.29)
Myocardial infarction		0.71 (0.51–0.99)
Death from cardiovascular causes†	0.97 (0.54–1.75)	0.73 (0.56–0.95)
Death from any cause		0.85 (0.74–0.98)

* Data from UPLIFT (Understanding Potential Long-Term Impacts on Function with Tiotropium) are for the treatment period plus 30 days of follow-up, not including vital status for patients who withdrew from the trial. Data may be found at www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Pulmonary-AllergyDrugsAdvisoryCommittee/ucm190461.htm. CI denotes confidence interval.

† Deaths include those categorized as resulting from an adverse event in the cardiac system organ class or the vascular system organ class, myocardial infarction, stroke, sudden death, cardiac death, or sudden cardiac death.

2012/9/14

30

VIP

- V (Valid/Reliability)： Are the results of the article valid? 研究方法的探討;”我們能相信這篇文獻嗎?”
- I (Important/Impact)： What are the results? 結論的分析;”我們相信它，但這個結論重要嗎?”
- P (Practice/Applicability)： Will the results help me in caring for my patients? 如何運用來照顧病人;”如果我們相信它，這個結論可以應用在我們所有的病患嗎?”

2012/9/14

What were the results?

1. How large was the treatment effect?

Most often results are presented as dichotomous outcomes (yes or not outcomes that happen or don't happen) and can include such outcomes as cancer recurrence, myocardial infarction and death. Consider a study in which 15% (0.15) of the control group died and 10% (0.10) of the treatment group died after 2 years of treatment. The results can be expressed in many ways as shown below.

What is the measure?	What does it mean?
Relative Risk (RR) = risk of the outcome in the treatment group / risk of the outcome in the control group.	The relative risk tells us **how many times more likely** it is that an event will occur in the treatment group relative to the control group. An **RR** of 1 means that there is no difference between the two groups thus, the treatment had **no effect**. An RR < 1 means that the treatment decreases the risk of the outcome. An RR > 1 means that the treatment increased the risk of the outcome.
In our example, the RR = 0.10/0.15 = 0.67	Since the RR < 1, the treatment decreases the risk of death.
Absolute Risk Reduction (ARR) = risk of the outcome in the control group - risk of the outcome in the treatment group. This is also known as the **absolute risk difference.**	The absolute risk reduction tells us the absolute difference in the rates of events between the two groups and gives an indication of the baseline risk and treatment effect. An **ARR** of **0** means that there is no difference between the two groups thus, the treatment had **no effect**.
In our example, the ARR = 0.15 - 0.10 = 0.05 or 5%	The absolute benefit of treatment is a 5% reduction in the death rate.
Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. An alternative way to calculate the RRR is to subtract the RR from 1 (eg. RRR = 1 - RR)	The relative risk reduction is the complement of the RR and is probably the most commonly reported measure of treatment effects. It tells us the reduction in the rate of the outcome in the treatment group relative to that in the control group.
In our example, the RRR = 0.05/0.15 = 0.33 or 33% Or RRR = 1 - 0.67 = 0.33 or 33%	The treatment reduced the risk of death by 33% relative to that occurring in the control group.
Number Needed to Treat (NNT) = inverse of the ARR and is calculated as 1 / ARR.	The number needed to treat represents the number of patients we need to treat with the experimental therapy in order to prevent 1 bad outcome and incorporates the duration of treatment. Clinical significance can be determined to some extent by looking at the NNTs, but also by weighing the NNTs against any harms or adverse effects (NNHs) of therapy.
In our example, the NNT = 1/ 0.05 = 20	We would need to treat 20 people for 2 years in order to prevent 1 death.

2. How precise was the estimate of the treatment effect?

The true risk of the outcome in the population is not known and the best we can do is estimate the true risk based on the sample of patients in the trial. This estimate is called the **point estimate**. We can gauge how close this estimate is to the true value by looking at the confidence intervals (CI) for each estimate. If the confidence interval is fairly narrow then we can be confident that our point estimate is a precise reflection of the population value. The confidence interval also provides us with information about the statistical significance of the result. If the value corresponding to **no effect** falls outside the 95% confidence interval then the result is statistically significant at the 0.05 level. If the confidence interval includes the value corresponding to **no effect** then the results are not statistically significant.

Will the results help me in caring for my patient? (External Validity/Applicability)

The questions that you should ask before you decide to apply the results of the study to your patient are:

- Is my patient so different to those in the study that the results cannot apply?
- Is the treatment feasible in my setting?
- Will the potential benefits of treatment outweigh the potential harms of treatment for my patient?

2012

- The magnitude of the treatment effect (effect size)

British Journal of Dermatology (1979) **100**, 113.

Clinical and Laboratory Investigations

Comparison of chlorambucil, azathioprine or cyclophosphamide combined with corticosteroids in the treatment of lupus nephritis

Dichotomous outcome

TABLE 3. Range of urinary abnormalities seen in patients

Morphological type	No. of patients	Hypertension	Proteinuria	Miscroscopic haematuria	Nephrotic syndrome
Lupus glomerulo-nephritis					
Mild	23	5	18	21	—
Moderate	61	16	48	51	27
Severe	62	41	62	62	43
Mixed membrano-proliferative lesion	17	7	7	3	5

Continuous outcome

TABLE 6. Serum urea (mg/100 ml) in patients on corticosteroids only and in patients on steroids + azathioprine

	Corticosteroids only	Corticosteroids + azathioprine
Before treatment	85 (±23·1)	81 (±11·7)
After 6 months	121 (±19·2)	117 (±9·3)
After 12 months	163 (±17·3)	154 (±5·7)
After 24 months	237 (±12)	219 (only one survivor)

2012/9/14

Calculation of OR/RR 計算RR

Treatment	Event	
	Positive	Negative
Exposed (experimental)	a	b
Not exposed	c	d

- RR= Relative Risk = Risk ratio
  - 相對危險、相對風險
  - 治療組風險與對照組風險的比值
  - RR= EER/CER= (a/a+b)/(c/c+d)

Calculation of OR/RR 計算OR

Treatment	Event	
	Positive	Negative
Exposed (experimental)	a	b
Not exposed	c	d

- OR = Odds Ratio = Relative Odds = 勝算比
  - 實驗組中發生目標疾病的勝算與控制組中發生目標疾病的勝算比值
  - OR= EEO/CEO = (a/b)/(c/d)= ad/bc
- Odds (勝算)
  - a ratio of events to non-events
  - 發生某事件的人數與未發生該事件的人數的比值

35

NNT vs NNH

- **NNT** : number needed to treat
  - **益一需治數;使用試驗性治療多少病人，才能預防一個額外的壞結果**
  - **(治療多少人才有1人有效)**
- **越小越好;依照慣例，將帶有小數點的NNT直接進位至整數**
- NNH: number needed to harm
  造成一個病人受傷害需治療的病人數（治療的不良作用）

參考網站: www.cebm.utoronto.ca

36

Important

- EER: experimental event rate
- CER: control event rate
- ARI: absolute risk increase （絕對風險增加度）
  - ARI= |EER – CER|
- RRI: relative risk increase（相對風險增加度）
  - RRI= |EER-CER| /CER
- ARR: absolute risk reduction
- NNT= 1/ARR
- NNH= 1/ARI

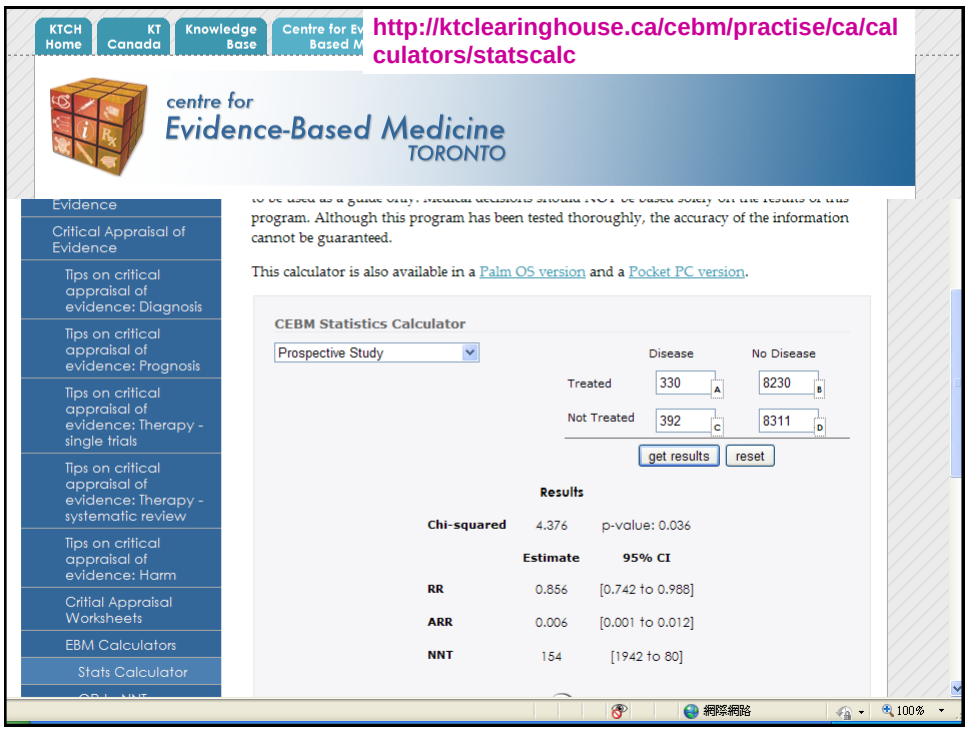
37

Operating characteristics of the HbA_{1c} and FPG tests for early detection of type 2 diabetes*

Diagnostic tests	Number of trials (n)	Cut-point	Sensitivity	Specificity	+LR	-LR
Hb1Ac	1 (866)	≥5.9%	76%	86%	5.4	0.28
	1 (111)		95%	67%	2.9	0.07
	1 (2877)	≥6.1%	78%	79%	3.7	0.29
	1 (111)		81%	84%	5.1	0.23
	1 (505)	≥6.2%	43%	99%	107	0.58
	1 (111)		81%	88%	6.8	0.22
FPG	1 (2877)	≥5.6 mmol/l	88%	79%	4.3	0.15
	1 (505)		80%	86%	5.6	0.24
	1 (111)	≥6.1 mmol/l	48%	98%	24	0.53
	1 (10 447)		64%	94%	10	0.39

*HbA_{1c} = glycated haemoglobin A_{1c}; FPG = fasting plasma glucose; LR = likelihood ratio. Diagnostic terms defined in glossary.

2012/9/14



2. How precise was the estimate of the treatment effect?
The true risk of the outcome in the population is not known and the best we can do is estimate the true risk based on the sample of patients in the trial. This estimate is called the **point estimate**. We can gauge how close this estimate is to the true value by looking at the confidence intervals (CI) for each estimate. If the confidence interval is fairly narrow then we can be confident that our point estimate is a precise reflection of the population value. The confidence interval also provides us with information about the statistical significance of the result. If the value corresponding to **no effect** falls outside the 95% confidence interval then the result is statistically significant at the 0.05 level. If the confidence interval includes the value corresponding to **no effect** then the results are not statistically significant.

- How precise are these results? (CASP)
  - If the result is precise enough to make a decision
  - If a confidence interval were reported. Would your decision about whether or not to use this intervention be the same at the upper confidence limit as a the lower confidence limit?
  - If a p-value is reported where confidence intervals are unavailable

VIP

- V (Valid/Reliability)： Are the results of the article valid? 研究方法的探討;”我們能相信這篇文獻嗎?”
- I (Important/Impact)： What are the results? 結論的分析;”我們相信它，但這個結論重要嗎?”
- P (Practice/Applicability)： Will the results help me in caring for my patients? 如何運用來照顧病人;”如果我們相信它，這個結論可以應用在我們所有的病患嗎?”

41

Will the results help me in caring for my patient? (ExternalValidity/Applicability)

The questions that you should ask before you decide to apply the results of the study to your patient are:
- Is my patient so different to those in the study that the results cannot apply?
- Is the treatment feasible in my setting?
- Will the potential benefits of treatment outweigh the potential harms of treatment for my patient?

- 病患的差異
- 可運用的資源
- 病患的偏好
- 病患的生物因素（biologic issues）
  - ”同樣的治療應用在不同的病患族群是否有不同的反應？”
  - “我們的病人與研究中的病人是否非常不同，以致無法應用在研究結果？”
- 社會經濟因素（social and economic issues）的考量
  - 評估這個結果的可行性
  - ”這個治療適用於我們的診療環境嗎？病患的配合度如何？醫療提供者的配合度及能力如何？”
- 流行病學因素（epidemiological issues）
  - 我們的病人是否有其他共病狀況，可能改變治療的結果？影響有多大？病人可能從治療中得到什麼好處或壞處？經由治療而減少的不良後果是否比不治療有明顯的差別？”

2012/9/14

陳杰峰 王慈蜂：醫學文獻評讀概念、方法與等級介紹

42

實證醫學五步驟

- 恰當運用 (Evidence Application) – application: apply the evidence to your clinical practice
- 衡量結果 (Outcome Evaluation) – auditing: evaluating your effectiveness and efficiency in executing step 1-4 and seeking ways to improve them both for next time

2012/9/14

Auditing-evaluate performance

- Self-evaluation
- Evaluation by expert or peer
- Audit by organization
- Audit by third party (NHI, Insurance)
- Audit by computer

Evaluating evidence-based performance

- Does a trainee perform the 5 steps in the course of patient care activities?
- Does this clinician perform evidence-based clinical manoeuvres and affect desirable patient outcomes?

Evidence-based medicine, 2006, 11, 99-101.

45

實證醫學之臨床應用

2012/9/14

46

從vancomycin trough level for MRSA infection發現藥物相關問題

2012/9/14

47

提昇vancomycin藥物血中濃度之臨床應用-從A(ask)到A(audit) - Background information

- From pharmacodynamics/pharmacokinetics
  - AUC/MIC > 400
  - MIC <= 1 ug/ml
  - Trough level increased to 15-20 ug/ml in severe or complicated MRSA infection
- 高雄榮總二線以上抗生素管制-感染專科醫師
- Vancomycin TDM-臨床藥師介入提供藥物動力學服務

2012/9/14

48

提昇Vancomycin藥物血中濃度之臨床應用-從A(ask)到A(audit) 專案 2009-2010

- Asking:
  - 嚴重或複雜MRSA感染須使用vancomycin之成人，維持在較高之trough level(15-20ug/ml)是否能比維持在傳統之trough level(5-15ug/ml)有較好的預後?
- Accessing:
  - PICO
  - Studies included or excluded?

2012/9/14

49

提昇Vancomycin藥物血中濃度之臨床應用-從A(ask)到A(audit) 專案 2009-2010

- Auditing:
  - Evaluating our effectiveness and efficiency in executing step 1-4 and seeking ways to improve them both for next time
- Methods
  - Self-evaluation
  - Evaluation by expert or peer
  - Audit by organization
  - Audit by third party (NHI, Insurance)
  - Audit by computer

2012/9/14

50

提昇Vancomycin藥物血中濃度之臨床應用-從A(ask)到A(audit) 專案 2009-2010

	Journal	Title	Study design	Inclusion criteria
Hidayat LK	Archives of Internal Medicine 2006.166:2138-44	High dose vancomycin therapy for MRSA infections	Prospective, cohort	1. >=18y/o **2. nosocomial MRSA infection** 3. vancomycin therapy >=72hrs
Jeffres MN	Chest 2006. 130:947-55	Predictors of **mortality** for MRSA health-care-associated pneumonia	Retrospective, cohort	1. MRSA health-care-associated **pneumonia** 2. vancomycin >=72hrs
Jeffres MN	Clinical Therapeutics 2007. 29:1107-15	A retrospective analysis of possible **renal toxicity** associated with vancomycin in patients with health care-associated MRSA pneumonia	Retrospective, cohort	1. MRSA health-care-associated **pneumonia** 2. vancomycin >=72hrs

2012/9/14

51

提昇vancomycin藥物血中濃度之臨床應用-從A(ask)到A(audit) 專案

表一、相關文獻或準則

	Study design	Setting	Patients	F/U	Interventions	numbers	Trough concentration (μg/mL)	mortality	Resolution of fever after 72hrs of vancomycin
Jeffres 2006	Retrospective, cohort	Single center	health-care-associated pneumonia	6.5yrs	≥15μg/mL	34	20.4 ± 3.2	35.3%	87.5%
					<15μg/mL	68	9.4 ± 3.2	29.4%	69.7%
Ben Lomaestro 2009	Guideline	Therapeutic monitoring of vancomycin in adult patients: A consensus review of the American Society of Health-System Pharmacists, the Infectious Diseases Societh of America, and the Societh of Infectious Diseases Pharmacists. Am J Health-Syst Pharm. 2009; 66:82-98							
American Thoracic Societh Documents	Guidelines	Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. Am J Respir Crit Care Med. 2005;171(4):388-416.							

2012/9/14

52

Thanks for your attention

2012/9/14