W16: How to Build an Evidence-Based Guideline – Important Epidemiological Principles

Workshop Chair: Marco Blanker, Netherlands
14 September 2016 08:35 - 10:05

Start	End	Topic	Speakers
08:35	08:40	Introduction	Marco Blanker
08:40	09:05	How to grade quality of evidence	Rufus Cartwright Kari Tikkinen
09:05	09:20	What's a risk factor?	Marco Blanker
09:20	09:35	The interpretations of odds ratios for common conditions	Ilse Hofmeester
09:35	09:55	Statistically significance vs. patient-importance	Rufus Cartwright Kari Tikkinen
09:55	10:05	Discussion	All

Aims of course/workshop

Despite the growing evidence in the field of lower urinary tract symptoms, the development and interpretation of guidelines remains difficult. This workshop aims to provide ICS members (both guideline-developers and users) with important background knowledge to enhance the quality of future guidelines.

Within the allotted time, we will focus on the following aspects:

- GRADE methodology and systematic reviews & meta-analyses
- What is a risk factor?
- Interpretation of odds ratios for common conditions.
- Statistical significance vs. clinical relevance for treatment outcomes?
- The impact of the setting from which evidence arises

Learning Objectives

After this workshop participants should be able to:

1. To know how to interpret odds ratios for common conditions.
2. To know the difference between statistical significance and clinically relevant outcomes.
3. To know about the background of the GRADE methodology and how this is applied to modern guidelines.

Learning Outcomes

After the course, the student will be able to:

- Know the difference between associated factors and true risk factors;
- Interpret odds ratios for common conditions;
- Compare odds ratios to relative risks (or rate ratios);
- Make the difference between statistical significance and clinical relevance of outcomes;
- Estimate the absolute risk difference based on relative risk reductions and prevalence rates;
- Interpret findings that result from the GRADE methodology.

Target Audience

All delegates

Advanced/Basic

Basic

Conditions for learning

This will be an interactive workshop in which participants are encouraged to have an active role. Speakers will invite participants to ask questions and respond to the presentations.

Suggested Learning before workshop attendance

http://www.gradeworkinggroup.org/#pub
Website with synopsis for:
- Explanation about The GRADE working group;
- Why rate the certainty in the evidence and strength of recommendations;
- Criteria for applying or using GRADE

Suggested Reading

- Johnston BC et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. CMAJ. 2016;188(1):25-32 (abstract and introduction)

- Blanker MH et al. No evidence (yet) to support the statement "LUTS - an independent risk factor for cardiovascular disease". BJU Int. 2016 Feb 25. doi: 10.1111/bju.13456.

- Hofmeester I et al. The association between nocturia and nocturnal polyuria in clinical and epidemiological studies: a systematic review and meta-analyses. J Urol. 2014;191(4):1028-33

Marco Blanker

Will discuss the qualifications of risk factors. Many patient characteristics are mentioned as risk factors, even from studies in which no causal associations can be distinguished. What are the requisites for a characteristic to become a "true" risk factor? The association between lower urinary tract symptoms and cardiovascular disease will illustrate this topic, by means of discussion of the (in)ability to define risk factors based on cross sectional studies.

Take home message: A risk factor is any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury (WHO definition). Therefore, longitudinal data are required to find risk factors for diseases; from cross sectional studies, at most characteristics can be defined as 'associated to' some disease.

Kari Tikkinen & Rufus Cartwright

Will compare statistical considerations and patient-importance. What do p-values tell us about the clinical relevance of a described risk difference, or risk reduction? Relative risk reductions can result in large differences in absolute risk reductions, depending on the baseline risk of patients. Ultimately, patients are interested in absolute risk (reductions), and physicians should also be. The topic is illustrated with clinical scenarios, including examples from cancer screening and pharmacological prophylaxis. Epidemiological aspects covered in this part include the interpretation of a p-value, relative risk reduction, absolute risk reduction, risk difference, number needed to treat (NNT).

Take home message: When considering treatment, patients are interested in their absolute risk reduction, which depend on their baseline risk; for a proper estimation of an absolute risk reduction, both baseline risk and relative risk reduction are needed.

Ilse Hofmeester

Will elaborate on the interpretation of odds ratios for common conditions. Often, results from epidemiological studies present large odds ratios (ORs), or at least large ORs get much attention. Many physicians regard such high ORs as relevant for their patients. As a consequence, advises may enter guidelines, but is that always relevant? From what kind of study were the ORs derived? How should ORs be interpreted for different conditions with different prevalence? Ilse Hofmeester will take the association between nocturia and nocturnal polyuria as an example.

Take home message: for the sound interpretation of odds ratios, information about the prevalence of the disease/outcome is needed; only for conditions with low prevalence, odds ratios may be interpreted as relative risks.

Rufus Cartwright & Kari Tikkinen

Many systematic reviews fail to adequately assess the quality of the evidence they synthesise, and many clinical guidelines lack transparency about their methods for deriving recommendations from that evidence. This talk will apply basic principles of clinical epidemiology to assessment of the quality of evidence, and explain the main tenets of the GRADE methodology, as the cornerstone of modern guideline development.

Take home message: GRADE provides a systematic way to assess both the quality of evidence (that is, certainty in estimates), and interpret the size of a pooled effect based on that evidence. The GRADE approach separately considers the impact of bias from design factors, inconsistency in results, indirectness, imprecision, and publication bias. GRADE allows guideline authors to reach "strong" or "weak" recommendations, reflecting the extent to which we can be confident that desirable effects of an intervention outweigh the undesirable effects, and the extent to which that balance will apply for most patients, or vary with patients' own values and preferences.

Marco H. Blanker

Affiliations to disclose†:

University of Groningen, University Medical Center Groningen, Department of General practice, Groningen, The Netherlands

† All financial ties (over the last year) that you may have with any business organisation with respect to the subjects mentioned during your presentation

Funding for speaker to attend:

- [X] Self-funded
- [ ] Institution (non-industry) funded
- [ ] Sponsored by:

How to build an evidence-based guideline

important epidemiological principles

ICS Annual meeting 2016
Kari Tikkinen
Ilse Hofmeester
Rufus Cartwright
Marco Blanker

Schedule

08:35 08:40	Introduction	Marco Blanker, Speaker, Disclosure
	The interpretations of odds ratios for common conditions	Ilse Hofmeester, Speaker, Disclosure
	Statistically significance vs. patient-importance	Rufus Cartwright, Speaker, Disclosure; Kari Tikkinen, Speaker, Disclosure
	What's a risk factor?	Marco Blanker, Speaker, Disclosure
	How to grade quality of evidence	Rufus Cartwright, Speaker, Disclosure; Kari Tikkinen, Speaker, Disclosure

General introduction

Most physicians have difficulties in interpreting effect sizes [1]

This may hamper

- sound interpretation of literature
- sound interpretation of guidelines
- sound development of guidelines

1. Johnston et al. CMAJ 2015

General introduction

Guidelines intended for patients with symptom / disease, e.g. incontinence

Guideline developers AND users need to be aware of pitfalls when interpreting guidelines

We will address some (certainly not all) pitfalls

At 10:00 you will be able to:

Interpret and distinguish different outcome measures for associations, especially Odds Ratios for common conditions

Discuss the differences between statistical significance and clinical relevance of treatment outcomes

Discuss different aspects of risk factors

Tell others about the GRADE methodology

Faculty

Kari Tikkinen, MD PhD, adjunct professor of clinical epidemiology & urology resident

Ilse Hofmeester, MD, epidemiologist & urology resident

Rufus Cartwright, MD PhD, urogynaecologist

Marco Blanker, MD PhD, general practitioner & epidemiologist

Who are you?

Personal introduction impossible, but please rise if you are a:

nurse

resident

urologist

GP

researcher

(pelvic) physiotherapist

(uro)gynaecologist

other:…

Who are you?

How do you rate your epidemiological knowledge/skills?

(please provide honest answer….)

Less than average

Average

Better than average

(What's average?)

Who are you?

Your input is more than welcome in this workshop

so feel free to interrupt,
ask questions, or even correct us

The interpretation of odds ratios for common conditions

Ilse Hofmeester

ICS Annual meeting 2016 – workshop

How to build an evidence-based guideline important epidemiological principles

What's a risk factor?

Marco H. Blanker

ICS Annual meeting 2016 – workshop

How to build an evidence-based guideline important epidemiological principles

What's a risk factor?

True or false?

Smoking is a risk factor for lung cancer

Vaginal delivery is a risk factor for Pelvic Organ Prolaps

Smoking is a risk factor for bladder cancer

Smoking is a risk factor for cardiovascular disease (CVD)

Lower urinary tract symptoms are a risk factor for CVD ???

What's a risk factor?

Lower urinary tract symptoms (LUTS) – an independent risk factor for cardiovascular disease (CVD)

G. Jackson, M.G. Kirby, R. Rosen, BJU Int 2015

Editorial comment on

BJUI Functional Urology

Increase of Framingham cardiovascular disease risk score is associated with severity of lower urinary tract symptoms

Giorgio I. Russo, Tommaso Castelli, Salvatore Privitera, Eugenia Fragalà, Vincenzo Favilla, Giulio Reale, Daniele Urzì, Sandro La Vignera*, Rosita A. Condorelli*, Aldo E. Calogero*, Sebastiano Cimino and Giuseppe Morgia

*Department of Urology, and *Department of Medical and Paediatric Sciences, Section of Endocrinology, Andrology and Internal Medicine, University of Catania, Catania, Italy*

What's a risk factor?

Lower urinary tract symptoms (LUTS) – an independent risk factor for cardiovascular disease (CVD)

G. Jackson, M.G. Kirby, R. Rosen, BJU Int 2015

What is your interpretation of this statement?

What's a risk factor?

Lower urinary tract symptoms (LUTS) – an independent risk factor for cardiovascular disease (CVD)

G. Jackson, M.G. Kirby, R. Rosen, BJU Int 2015

What is needed for this statement to be true?

What is in fact a risk factor?

World Health Organization:
A risk factor is any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury.

What's a risk factor?

Developing disease (in the future)

Causal association between risk factor & disease

True association (not explained by other variables)

Ask yourself "why would LUTS cause CVD?"

World Health Organization:
A risk factor is any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury.

What's a risk factor?

BJUI Functional Urology

Increase of Framingham cardiovascular disease risk score is associated with severity of lower urinary tract symptoms

Crosssectional study

336 Consecutive patients with BPH-related LUTS

Assessment of Framingham Heart Risk score

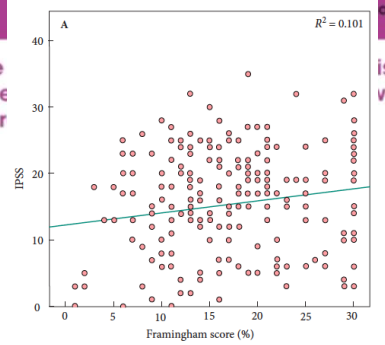
(based on age, HDL, total cholesterol level, systolic blood pressure, anti-hypertensive medication use, diabetes and current smoking status)

What's a risk factor?

BJUI

Increase
risk score
urinary tr

isease
ver



What's a risk factor?

BJUI Functional Urology

Increase of Framingham cardiovascular disease risk score is associated with severity of lower urinary tract symptoms

Risk of having moderate/severe LUTS for high CVD-risk group: OR 5.9 (age-adjusted)

Comments:

Crosssectional study
No CVD but 'risk-for CVD score' } No firm conclusion can be drawn

What's a risk factor

Rosso-study no evidence of LUTS as risk factor for CVD

More information is needed

EURURO-6911; No. of Pages 9
ARTICLE IN PRESS
EUROPEAN UROLOGY XXX (2016) XXX–XXX

available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Platinum Priority – Collaborative Review – Benign Prostatic Enlargement
Editorial by XXX on pp. x–y of this issue

Male Lower Urinary Tract Symptoms and Cardiovascular Events: A Systematic Review and Meta-analysis

What's a risk factor

Male Lower Urinary Tract Symptoms and Cardiovascular Events: A Systematic Review and Meta-analysis

Objective: To evaluate whether LUTS severity can be considered as a significant risk factor of major adverse cardiac events (MACE) in the male population.

Authors included all cross-sectional & longitudinal trials enrolling men, comparing prevalence/incidence of MACE in men with moderate to severe LUTS and those without LUTS or with mild LUTS.

What's a risk factor

Male Lower Urinary Tract Symptoms and Cardiovascular Events: A Systematic Review and Meta-analysis

Objective: To evaluate whether LUTS severity can be considered as a significant risk factor of major adverse cardiac events (MACE) in the male population

Authors included all cross-sectional & **longitudinal trials** enrolling men, comparing prevalence/incidence of MACE in men with moderate to severe LUTS and those without LUTS or with mild LUTS

What's a risk factor

Male Lower Urinary Tract Symptoms and Cardiovascular Events: A Systematic Review and Meta-analysis

5 studies with 25,494 patients and 2,291 MACE.

Authors included all cross-sectional and **longitudinal trials** enrolling men, comparing prevalence/incidence of MACE in men with moderate to severe LUTS and those without LUTS or with mild LUTS

What's a risk factor

Male Lower Urinary Tract Symptoms and Cardiovascular Events: A Systematic Review and Meta-analysis

5 studies with 25,494 patients and 2,291 MACE.

Presence of moderate to severe LUTS associated with increased incidence of MACE compared with the rest of the sample (OR: 1.68; 1.13–2.50)

BUT:

No adjustment for confounders

No exclusion of patients with MACE/CVD at baseline

What's a risk factor

Do lower urinary tract symptoms predict cardiovascular diseases in older men? A systematic review and meta-analysis

Iris I. Bouwman[1] · Maarten J. H. Voskamp[2] · Boudewijn J. Kollen[1] · Rien J. M. Nijman[2] · Wouter K. van der Heide[1] · Marco H. Blanker[1]

World J Urol 2015;33:1911–20

5 studies with 6,027 (LUTS) & 18,993 (no LUTS) men

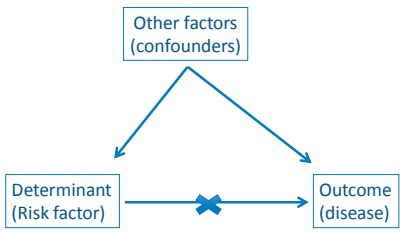
All without CVD at baseline

Follow-up period 5 - 17 years

2,780 CVD events

No clear association between CVD and LUTS [pooled effect size: hazard ratio 1.09 (95 % CI 0.90–1.31)].

What's a risk factor?



What's a risk factor?

Term might lead to confusion, as definitions differ

Most often used in epidemiology:

- particular outcome will occur after particular exposure
- an exposure that is statistically related to an outcome

Risk factors may be immutable or modifiable

Uncertainty about what strength of association is needed

What's a risk factor?

Related terms:

Risk marker: attribute/exposure associated with increased probability of outcome, but not necessarily a causal factor

Determinant: attribute/exposure that increases probability of outcome

Modifiable risk factor: a determinant that can be modified by intervention, thereby reducing the probability of disease

What's a risk factor?

In case of LUTS & CVD

- In those with CVD: LUTS seems to be associated
  BUT: CVD history itself is major predictor of new CVD
- In those without CVD: no association

Most probably: LUTS and CVD share common risk factors

If so, LUTS might be a risk marker

What's a risk factor?

Thank you for your attention

ICS Annual meeting 2016 – workshop

How to build an evidence-based guideline important epidemiological principles

How to grade the quality of evidence?

Kari Tikkinen & Rufus Cartwright

ICS Annual meeting 2016 – workshop

How to build an evidence-based guideline important epidemiological principles

Affiliations to disclose†:

No disclosures

† All financial ties (over the last year) that you may have with any business organisation with respect to the subjects mentioned during your presentation

Funding for speaker to attend:

- [X] Self-funded
- [ ] Institution (non-industry) funded
- [ ] Sponsored by:

The interpretation of odds ratios for common conditions

Ilse Hofmeester
Urologist in training – epidemiologist
The Netherlands

Risk

Which risk estimates do you know?

Absolute risk

Relative risk

Odds ratio

Risk ratio

Hazard ratio

Risk

What's the most frequently used risk estimate?

- Relative risk estimates

What's the most important risk estimate?

- (depends on aim)
- Absolute risk estimates

Both are used & misused

Example if (mis)use

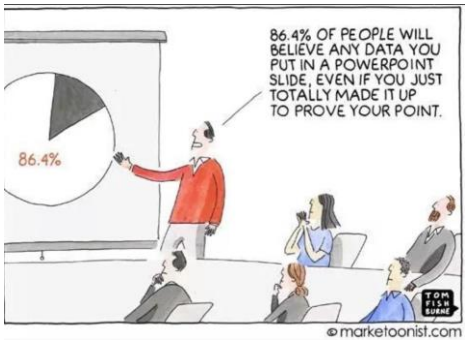
72% of alpha-blocker users experience improvement of symptoms

61% of placebo users experience improvement of symptoms

Use of 5-alpha reductase inhibitors reduces the risk of acute urinary retention (AUR) by 50%

Absolute risk reduction of AUR after 5 years: 2,5%

Examples if (mis)use

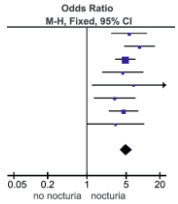


Nocturia & nocturnal polyuria

In-depth example of interpretation of OR

Risk of having nocturnal polyuria based on nocturia status

Results of meta-analyses
(Hofmeester, J Urol 2014)

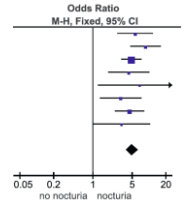


Nocturia & nocturnal polyuria

What's your interpretation of this OR?

People with nocturia have nocturnal polyuria 5 times more often than those without nocturia

Don't know



Interpretation of odds ratio's

Back to basics!

Relative risk estimates are based on absolute risk estimates in 2 or more groups

Absolute risk estimates important for interpretation

Interpretation of odds ratios

Prevalence NP 5%	NP +	NP -	Total
Nocturia +	20	230	250
Nocturia -	30	720	750
Total	50	750	1000

Interpretation of odds ratios

Prevalence NP 25%	NP +	NP -	Total
Nocturia +	100	150	250
Nocturia -	150	600	750
Total	250	750	1000

Prevalence of disease – influence on OR

Prevalence NP 5.0%	Nocturnal Polyuria	No Nocturnal Polyuria	Total
Nocturia +	20	230	250
Nocturia -	30	720	750
Total	50	950	1000

Prevalence NP 25.0%	Nocturnal Polyuria	No Nocturnal Polyuria	Total
Nocturia +	100	150	250
Nocturia -	150	600	750
Total	250	750	1000

Prevalence NP 60.0%	Nocturnal Polyuria	No Nocturnal Polyuria	Total
Nocturia +	240	10	250
Nocturia -	360	390	750
Total	600	400	1000

Prevalence	5%	25%	60%
Odds ratio	2.09	2.67	26.00
Relative risk	2.00	2.00	2.00

Interpretation of odds ratio's

Association between OR and RR
depends on prevalence of condition

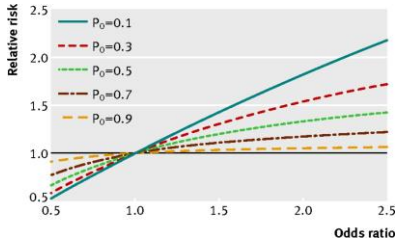
Odds ratio's look like relative risks,

but only if prevalence of condition is small
ORs may be interpreted as RR
Rare disease assumption

Interpretation of odds ratio's

Association between OR and RR
depends on prevalence of condition

$$Relative\ risk = \frac{OR}{(1-p)+(p*OR)}$$



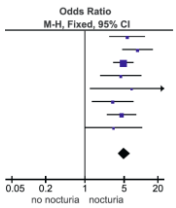
p_0 = prevalence of condition

Nocturia & nocturnal polyuria

What's your interpretation of this OR?

People with nocturia have nocturnal polyuria
5 times more often than
those without nocturia

Don't know

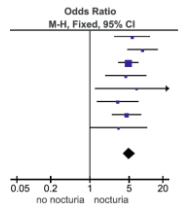


Nocturia & nocturnal polyuria

What's your interpretation of this OR?

People with nocturia have nocturnal polyuria
5 times more often than
those without nocturia

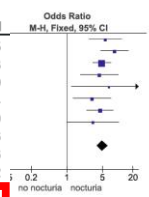
Don't know = correct
Important info was lacking



Nocturia & nocturnal polyuria

What's your interpretation of this OR?

Study or Subgroup	nocturia Events	Total	no nocturia Events	Total
NPi 0.33: Bing [18]	44	75	15	75
NPi 0.33: Rembratt [15]	97	116	40	108
NPi 0.33: van Doorn [10]	340	370	483	689
NPi 0.33:Swithinbank [11]	25	33	81	194
NPi 0.35: Johnson [19]	22	35	2	10
NPi 0.35: Ku [16]	27	38	29	66
NUP/daytimeUP 1: Udo [17]	69	84	185	366
NUV 10ml/kgBW: Homma [14]	19	39	5	22
	643	790	840	1530



Prevalence = (643+840)/(790+1530) = **63.9%**

Nocturia & nocturnal polyuria

What's your interpretation of this OR?

Prevalence of nocturnal polyuria 63.9%
(well above 10%)

$$Relative\ risk = \frac{OR}{(1-p)+(p*OR)}$$

Relative risk: 1.41

In summary

Relative risk estimates most often used

Absolute risk estimates are important for interpretation

For proper interpretation of odds ratio's, information on prevalence of condition is vital

Clinical Urology and Epidemiology Working Group

How to grade quality of evidence

Rufus Cartwright (@roofus)
Department of Epidemiology and Biostatistics
Imperial College London, UK

Kari Tikkinen (@KariTikkinen)

Departments of Urology and Public Health,
Helsinki University Hospital, Academy of Finland and University of Helsinki, Finland

W16: How to Build an Evidence-Based Guideline – Important Epidemiological Principles
46th Annual Scientific Meeting of the International Continence Society
September 14, 2016 – Tokyo, Japan

Guidelines and clinicians

- increasingly, clinicians rely on formal guidelines
- strong recommendations
  - strong methods
  - large precise effect
  - few down sides of therapy
- weak recommendations
  - weak methods
  - imprecise estimate
  - small effect
  - substantial down sides

Proliferation of systems ☹

Common international grading ☺

- GRADE (*G*rades of *r*ecommendation, *a*ssessment, *d*evelopment and *e*valuation)
- international group
  - Australian NMRC, SIGN, USPSTF, WHO, NICE, Oxford CEBM, CDC, CC
- ~ 35 meetings over last 14 years
    - (~10 – 80 attendants – now 300 contributors)

THE COCHRANE COLLABORATION®

World Health Organization

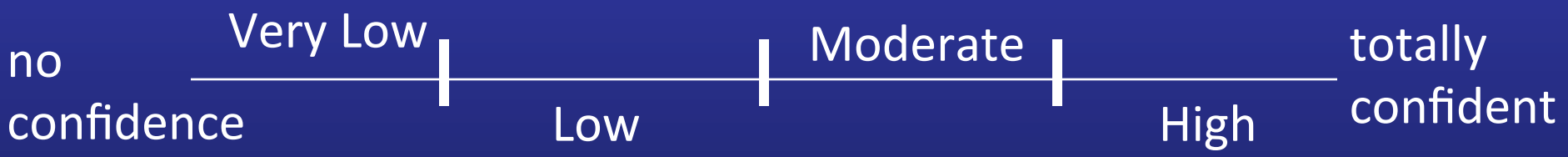
UpToDate

ACP American College of Physicians
INTERNAL MEDICINE | *Doctors for Adults*℠

80+ Organizations

What are we grading?

two components



strength of recommendation:
strong and weak

Grading system – for what?

- interventions
  - management strategy 1 versus 2

- what grade is **not** about
  - individual studies (body of evidence)

What GRADE is not primarily about

- diagnostic accuracy questions
  - in patients with a sore leg, what is the accuracy of a blood test (D-Dimer) in sorting out whether a deep venous thrombosis is the cause of the pain
- prognosis
- what it is about: diagnostic impact
  - are patients better off (improved outcomes) when doctors use the d-dimer test

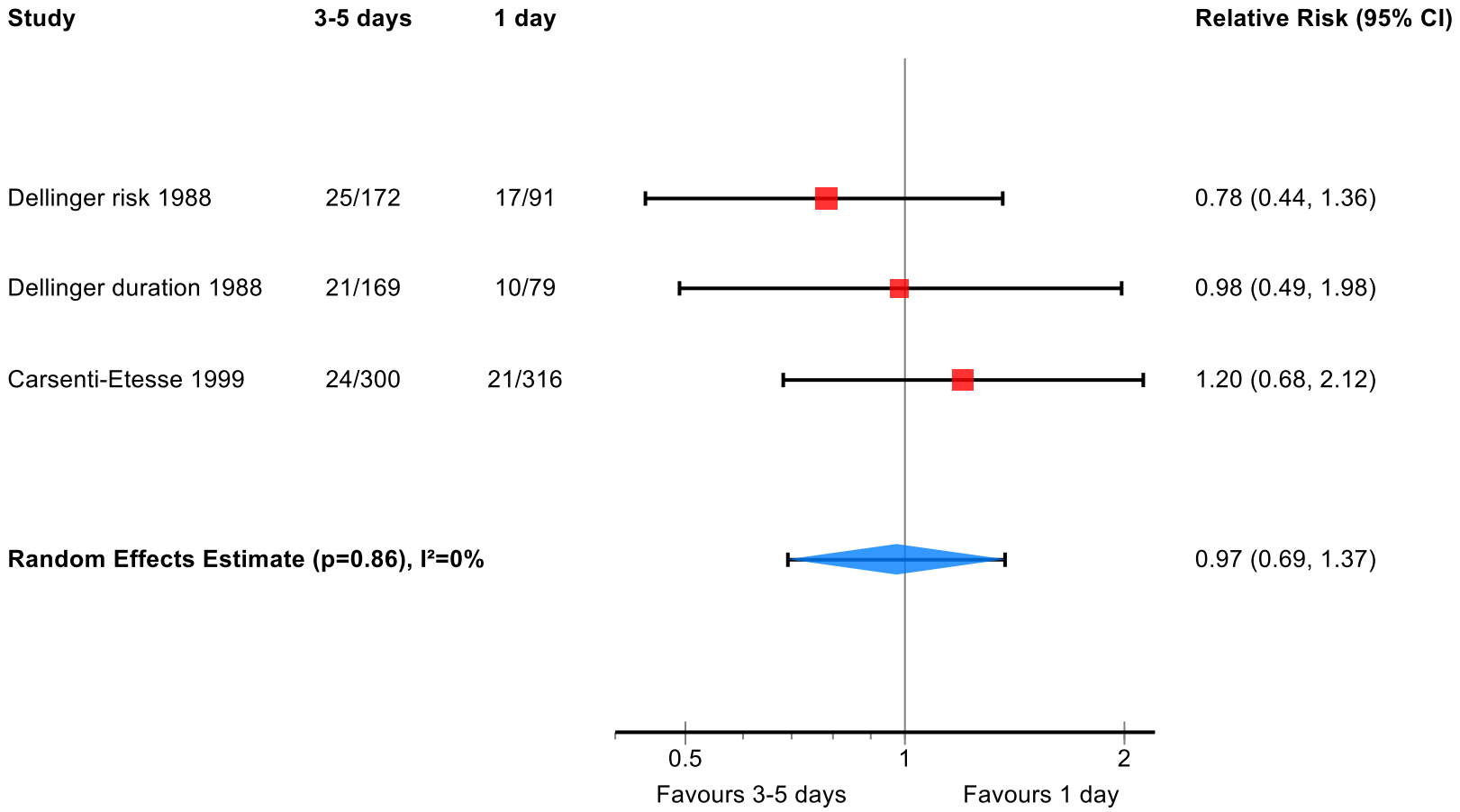
Determinants of quality

- RCTs start high
- observational studies start low
- what can lower confidence?

What can lower confidence?

- clue 1
  - lack of blinding in an RCT
- clue 2
  - RCT loses ½ patients to follow-up
- high risk of bias in RCTs lowers confidence

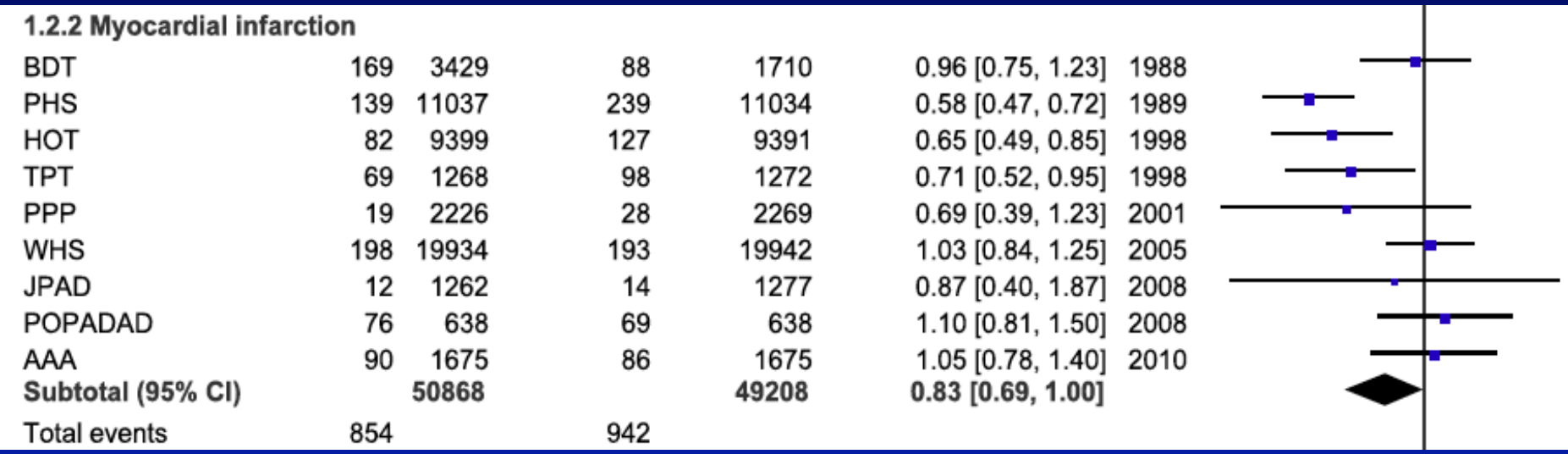
Clue: Have a look at the forest plot below – Infections with short and long term antibiotics after open fractures



Any concerns?

Another reason for rating down: imprecision

Clue: Have a look at the forest plot below
Aspirin in primary prophylaxis



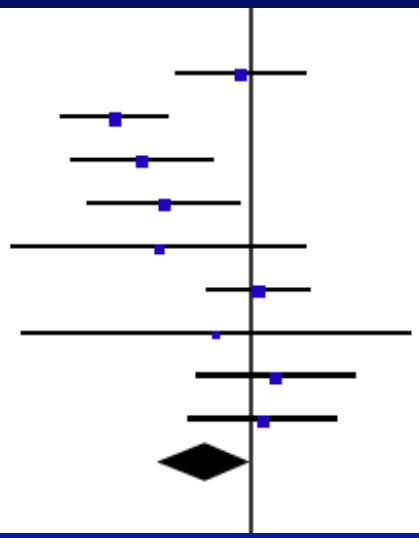
1.2.2 Myocardial infarction						
BDT	169	3429	88	1710	0.96 [0.75, 1.23]	1988
PHS	139	11037	239	11034	0.58 [0.47, 0.72]	1989
HOT	82	9399	127	9391	0.65 [0.49, 0.85]	1998
TPT	69	1268	98	1272	0.71 [0.52, 0.95]	1998
PPP	19	2226	28	2269	0.69 [0.39, 1.23]	2001
WHS	198	19934	193	19942	1.03 [0.84, 1.25]	2005
JPAD	12	1262	14	1277	0.87 [0.40, 1.87]	2008
POPADAD	76	638	69	638	1.10 [0.81, 1.50]	2008
AAA	90	1675	86	1675	1.05 [0.78, 1.40]	2010
Subtotal (95% CI)		**50868**		**49208**	**0.83 [0.69, 1.00]**	
Total events	854		942			

Any concerns?

Another reason for rating down: inconsistency

1.2.2 Myocardial infarction

BDT	169	3429	88	1710	0.96 [0.75, 1.23]	1988
PHS	139	11037	239	11034	0.58 [0.47, 0.72]	1989
HOT	82	9399	127	9391	0.65 [0.49, 0.85]	1998
TPT	69	1268	98	1272	0.71 [0.52, 0.95]	1998
PPP	19	2226	28	2269	0.69 [0.39, 1.23]	2001
WHS	198	19934	193	19942	1.03 [0.84, 1.25]	2005
JPAD	12	1262	14	1277	0.87 [0.40, 1.87]	2008
POPADAD	76	638	69	638	1.10 [0.81, 1.50]	2008
AAA	90	1675	86	1675	1.05 [0.78, 1.40]	2010
Subtotal (95% CI)		**50868**		**49208**	**0.83 [0.69, 1.00]**	
Total events	854		942			



Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 27.51$, df = 8 (P = 0.0006); $I^2 = 71\%$

Test for overall effect: Z = 1.99 (P = 0.05)

More reasons to lose confidence

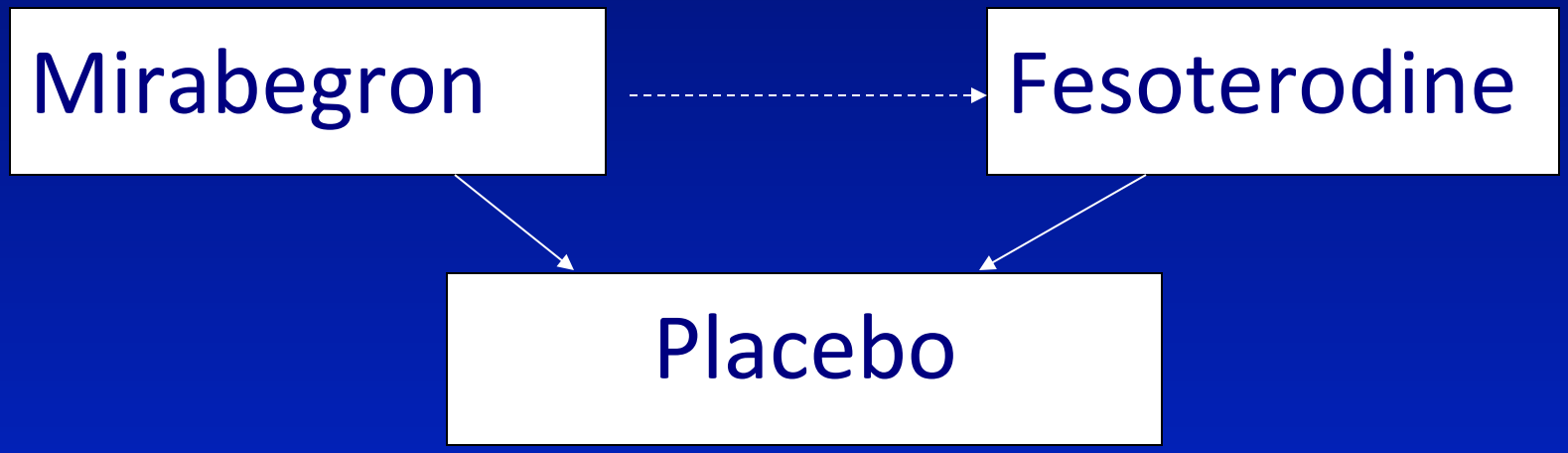
- RCTs show less UI after new intervention
  - patients in RCTs 40 to 70
  - your patient 90
- are you confident?
- indirectness of population
  - older, sicker or more co-morbidity

More reasons to lose confidence

- operation for lap mesh prolapse repair
- technically challenging
  - frequent complications
- RCTs: lap surgery decreases recurrence
  - only top surgeons participate in the RCTs
- are you confident?
- indirectness of intervention

Directness

interested in A versus B
available data A vs C, B vs C



Another reason to lose confidence

- some trials never get published
- "negative" studies more likely
- biased sample of studies
  - overestimates of treatment effect

Positive results more likely to get published

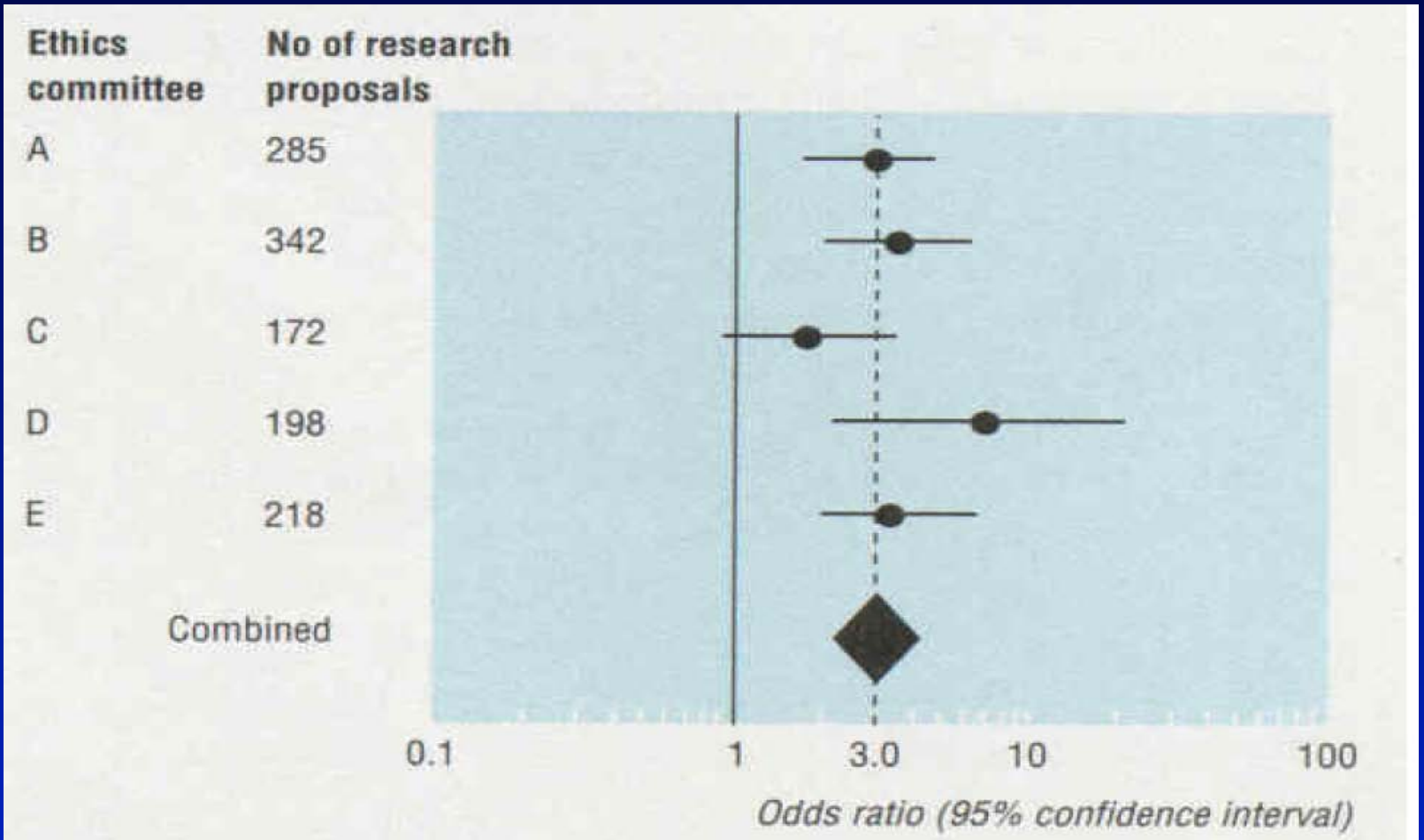
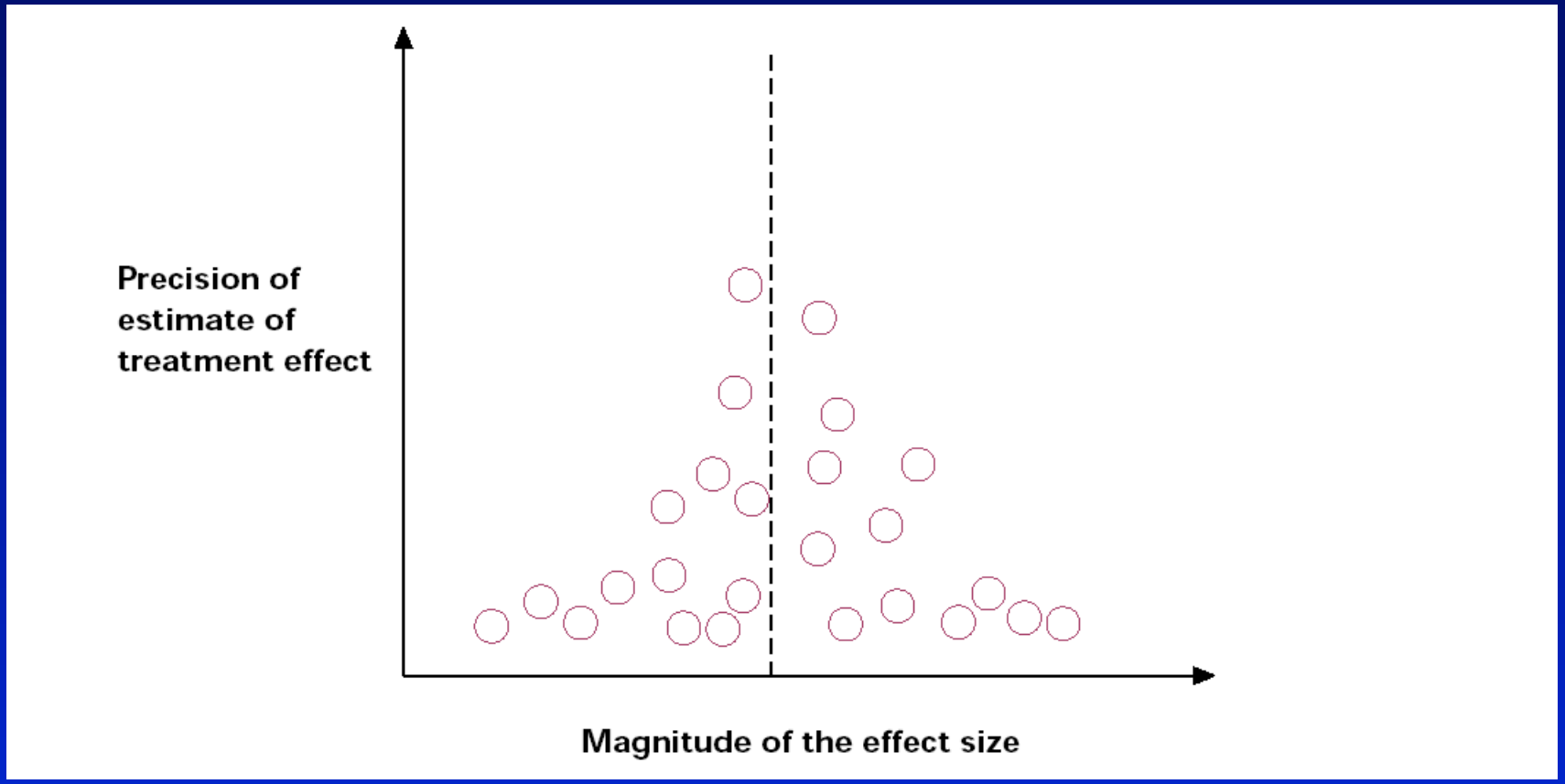


Fig 1 Meta-analysis of five studies examining association of significant results and publication among research proposals submitted to ethics committees. The unadjusted odds ratios were combined by using a fixed effects model

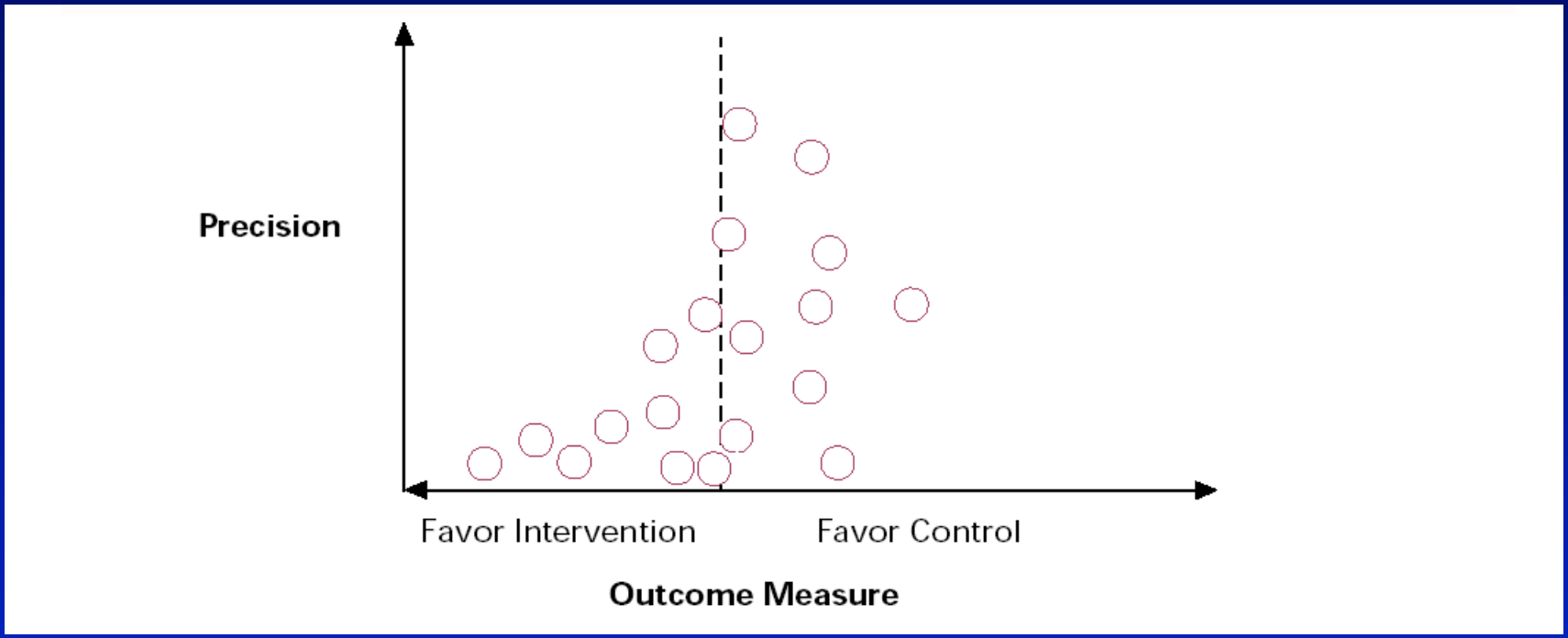
How to demonstrate?

Funnel plot



How to demonstrate?

Publication bias



Determinants of quality

- RCTs start high
- observational studies start low
- what can lower quality?

- risk of bias
- inconsistency
- indirectness
- imprecision
- publication bias

What can raise quality?

- large magnitude can rate up one level
  - very large two levels
- common criteria
  - everyone used to do badly
  - almost everyone does well
  - quick action
- hip replacement for hip osteoarthritis
- mechanical ventilation in respiratory failure

Confidence assessment criteria

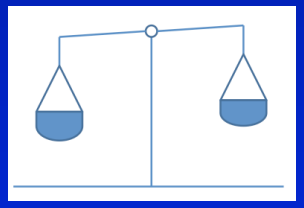
Study Design	Confidence in estimates	Lower if	Higher if
Randomised trial →	High	Risk of bias - 1 Serious -2 Very serious Inconsistency -1 Serious -2 Very serious Indirectness -1 Serious -2 Very serious Imprecision -1 Serious -2 Very serious Publication bias -1 Likely -2 Very likely	Large effect +1 Large +2 Very large Dose response +1 Evidence of a gradient All plausible confounding +1 Would reduce a demonstrated effect or +1 Would suggest a spurious effect when results show no effect
	Moderate		
Observational study →	Low		
	Very low		

Strength of Recommendation

- strong recommendation
  - benefits clearly outweigh risks/hassle/cost
  - risk/hassle/cost clearly outweighs benefit

- what can downgrade strength?

- low confidence in estimates
- close balance between up and downsides



Risk/Benefit tradeoff

- aspirin after myocardial infarction
  - 25% reduction in relative risk
  - side effects minimal, cost minimal
  - benefit obviously much greater than risk/cost
- warfarin in low risk atrial fibrillation
  - warfarin reduces stroke vs ASA by 50%
  - but if risk only 1% per year, ARR 0.5%
  - increased bleeds by 1% per year

Conclusion

- clinicians, policy makers need summaries
  - quality of evidence
  - strength of recommendations
- explicit rules
  - transparent, informative
- GRADE
  - simple, transparent, systematic
  - increasing wide adoption
  - great opportunity for teaching EBHC

Clinical Urology and Epidemiology Working Group

Statistical considerations versus patient-importance

Rufus Cartwright (@roofus)
Department of Epidemiology and Biostatistics
Imperial College London, UK

Kari Tikkinen (@KariTikkinen)

Departments of Urology and Public Health,
Helsinki University Hospital, Academy of Finland and University of Helsinki, Finland

W16: How to Build an Evidence-Based Guideline – Important Epidemiological Principles
46th Annual Scientific Meeting of the International Continence Society
September 14, 2016 – Tokyo, Japan

Calibrating Your Enthusiasm

Your flight is cancelled due to bad weather

Your flight will arrive earlier than scheduled due to very good weather and nice tailwind

Interpreting the Evidence

Willingness to fund mammography screening

- *program A* reduces the rate of dying from breast cancer by 33% (p=0.001)
- *program B* increases the rate of patients not dying from breast cancer from 99.82% to 99.88% (p=0.001)
- *program C* means that 1,667 women needed to be screened yearly for 7 years to prevent one death from breast cancer (p=0.001)

Breast Cancer Screening

Breast cancer death rates (p=0.001)

- unscreened 0.18% (18 out of 10,000)
- screened 0.12% (12 out of 10,000)

Relative risk reduction: (0.18% - 0.12%) / 0.18% = 33%

Breast cancer death rates

- unscreened 0.18% means 99.82% don't die
- screened 0.12% means 99.88% don't die

Absolute risk reduction: 0.18% - 0.12% = 0.06%

Number needed to screen: 100/0.06 = 1,667

P-value same, tells nothing about magnitude

Example: VA hypertension study

Mortality after 5 years of treatment

	Controls	Treated	RRR
DBP (90 – 104)	0.074	0.059	$\frac{0.074 - 0.059}{0.074}$
			20%

DBP, diastolic blood pressure

Relative risk reduction (RRR)

	Control	Treat-ment	RRR
TOD+	0.20	0.16	20%
TOD-	0.057	0.045	21%

TOD, target organ damage

Absolute risk reduction (ARR)

	Control	Treat-ment	RRR	ARR
TOD+	0.20	0.16	20%	4%
TOD-	0.057	0.045	21%	1.2%

TOD, target organ damage

Number needed to treat (NNT)

	Control	Treat-ment	RRR	ARR	NNT
TOD+	0.20	0.16	20%	4%	25
TOD-	0.057	0.045	21%	1.2%	83

TOD, target organ damage

Patient with DVT

Completes 6 months prophylaxis

Question: continue or not?

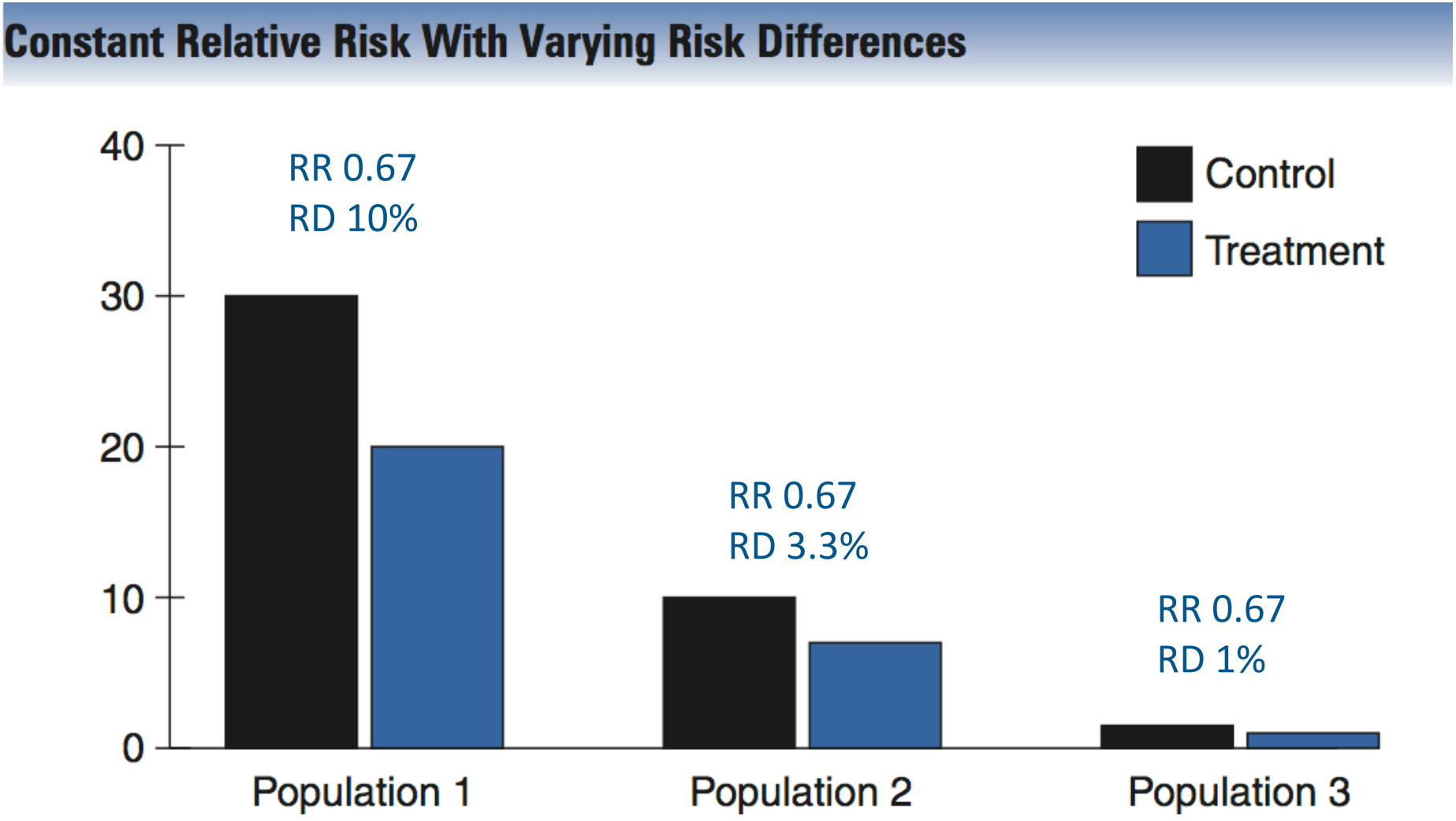
Doctor: continuing reduces risk of recurrence by 33%

- chance unlikely to explain the difference (p=0.001)

What does patient understand?

Is there something missing?

Patient with DVT



Patients with atrial fibrillation

$CHADS_2$: congestive heart failure; hypertension; age >75; diabetes; prior stroke

Risk of stroke varies

- $CHADS_2$ 0: 8 per 1,000 per year
- $CHADS_2$ 1: 22 per 1,000 per year
- $CHADS_2$ 2: 45 per 1,000 per year
- $CHADS_2$ 3: 96 per 1,000 per year

Warfarin constant 2/3 relative risk reduction

- $CHADS_2$ 0: 5 per 1,000 per year
- $CHADS_2$ 1: 14 per 1,000 per year
- $CHADS_2$ 2: 40 per 1,000 per year
- $CHADS_2$ 3: 64 per 1,000 per year

Measures of Relative Effect

- Relative risk
- Relative risk reduction
- Odds ratio
- Relative odds reduction
- Hazard ratio

Small, medium or large?

VTE prophylaxis in 65 year old man, COPD exacerbation, anticipated walking in hall day 3, hospitalization

RRR 50%

Baseline risk 4/1,000

Risk difference 2/1,000 so, NNT 500

Balance in favour of treatment?

VTE, venous thromboembolism

Small, medium or large?

VTE prophylaxis in 65 year old man, disseminated cancer, severe pneumonia, likely bed-bound for at least 3 days

RRR 50%

Baseline risk 100/1,000

Risk difference 50/1,000 so, NNT 20

Balance in favour of treatment?

Summary

Relative estimates: RR, OR, HR

Absolute estimates: RD (ARR), NNT

Ultimately patients interested in absolute risk (reductions)

Patients not interested in p-values or relative estimates

Relative risk reductions constant across patients, absolute risk reductions not

So, to get absolute risk reductions, need baseline risk and relative risk reductions