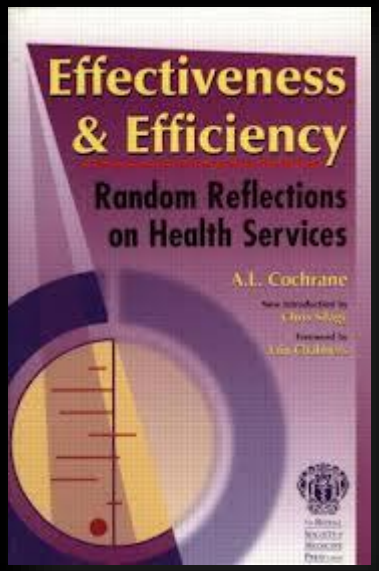
Real World Critical Appraisal and EBM

Carl Heneghan
Director CEBM & Professor of EBM
Nuffield Dept of Primary Care Health Sciences,
University of Oxford

CEBM

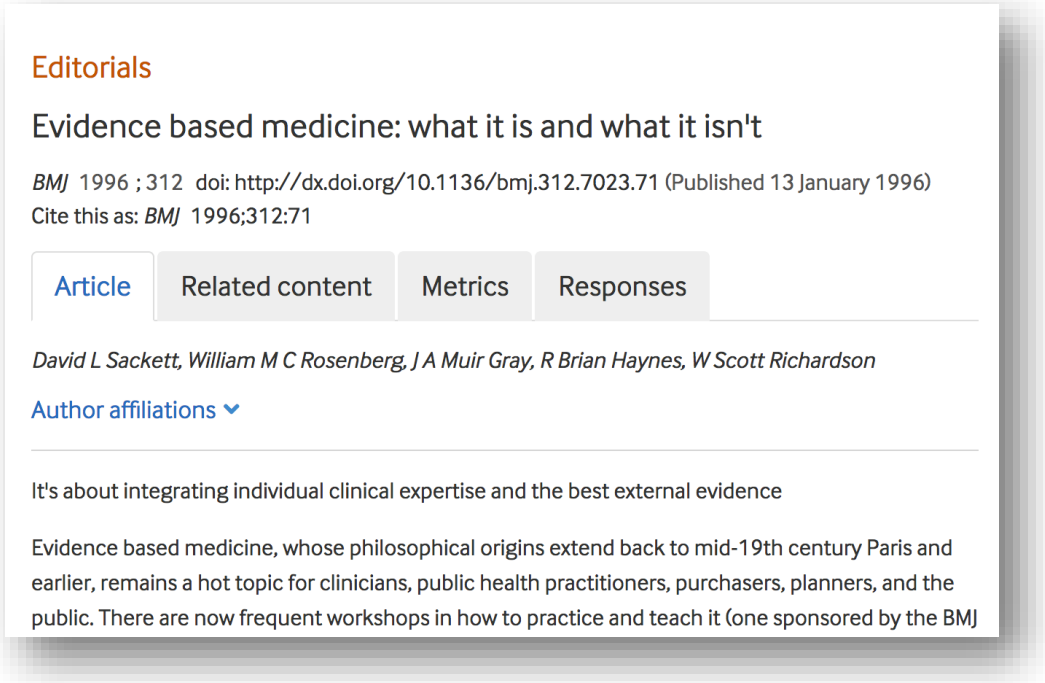
UNIVERSITY OF OXFORD



- Archibald (Archie) Cochrane's most influential mark on healthcare was his 1971 publication, “Effectiveness and Efficiency.”
-
- This book strongly criticized the lack of reliable evidence behind many of the commonly accepted healthcare interventions at the time.

- His criticisms spurred rigorous evaluations of healthcare interventions and highlighted the need for evidence in medicine

What is Evidence-Based Medicine?



Editorials

Evidence based medicine: what it is and what it isn't

BMJ 1996 ; 312 doi: http://dx.doi.org/10.1136/bmj.312.7023.71 (Published 13 January 1996)
Cite this as: *BMJ* 1996;312:71

Article | Related content | Metrics | Responses

David L Sackett, William M C Rosenberg, J A Muir Gray, R Brian Haynes, W Scott Richardson

Author affiliations

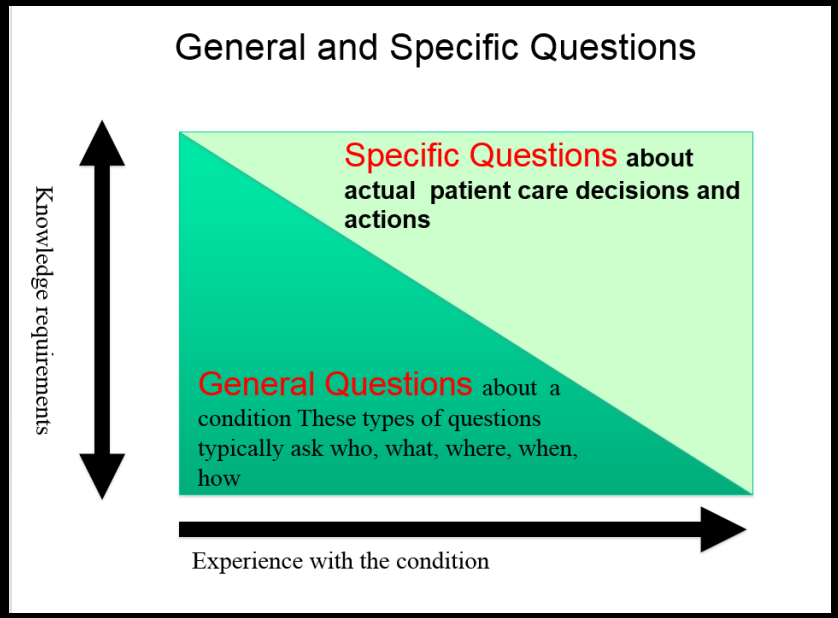
It's about integrating individual clinical expertise and the best external evidence

Evidence based medicine, whose philosophical origins extend back to mid-19th century Paris and earlier, remains a hot topic for clinicians, public health practitioners, purchasers, planners, and the public. There are now frequent workshops in how to practice and teach it (one sponsored by the BMJ

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

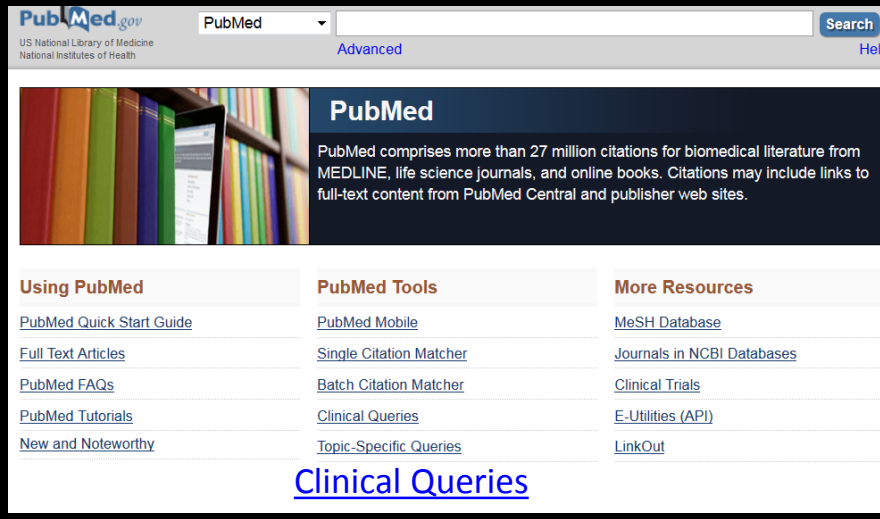
E in EBM

1 Learn to ask Questions efficiently



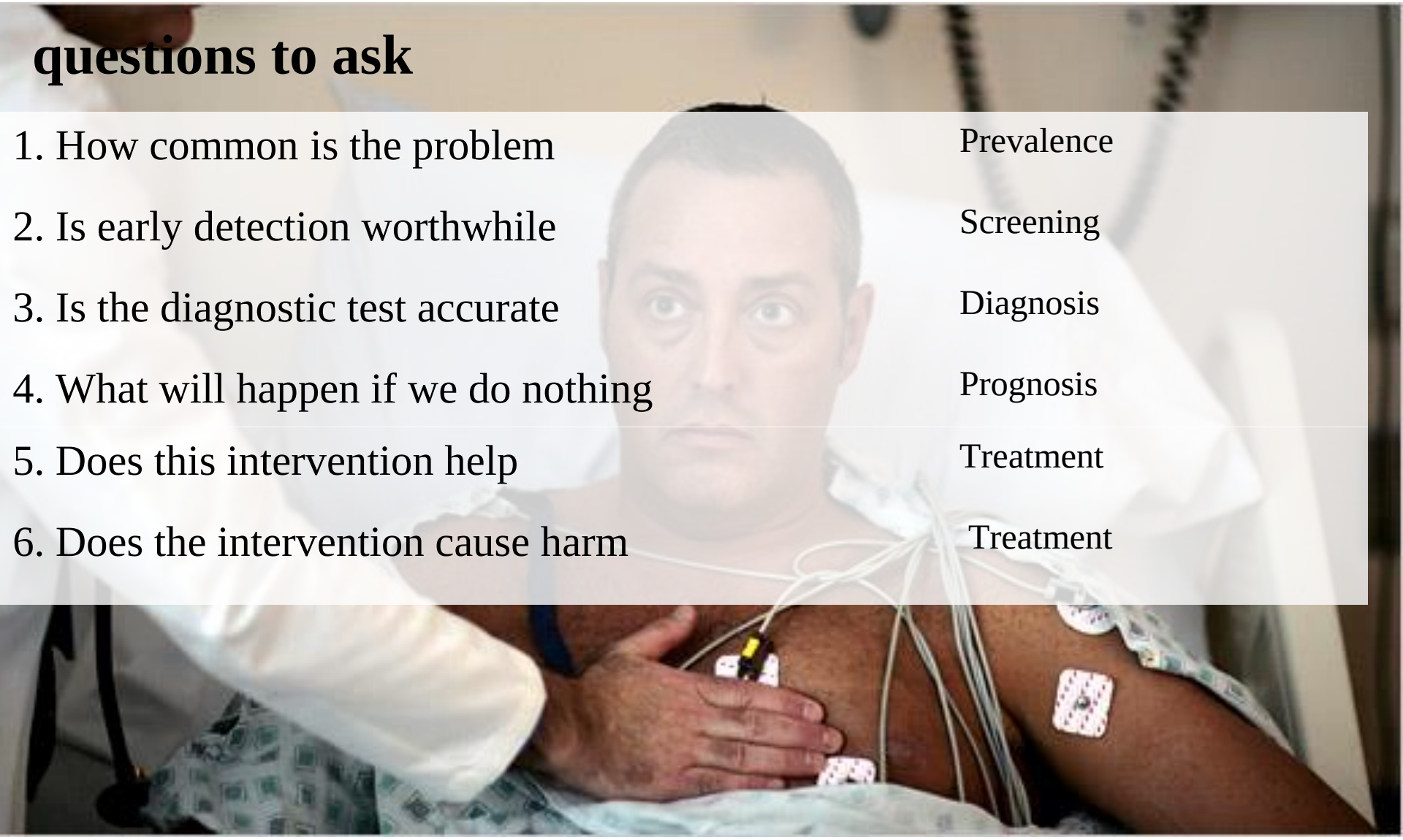
Population Intervention Comparator Outcome

2. Search for evidence effectively



Clinical Queries

questions to ask

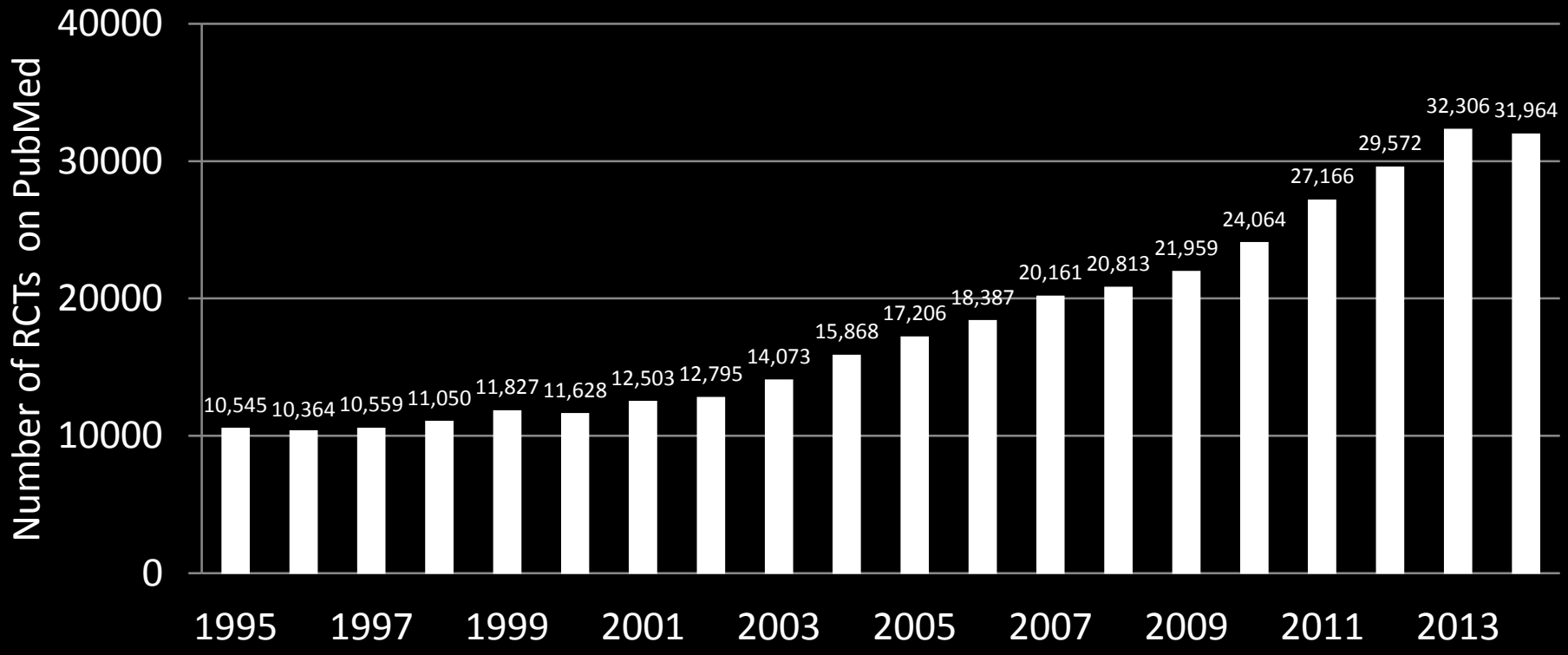


1. How common is the problem	Prevalence
2. Is early detection worthwhile	Screening
3. Is the diagnostic test accurate	Diagnosis
4. What will happen if we do nothing	Prognosis
5. Does this intervention help	Treatment
6. Does the intervention cause harm	Treatment

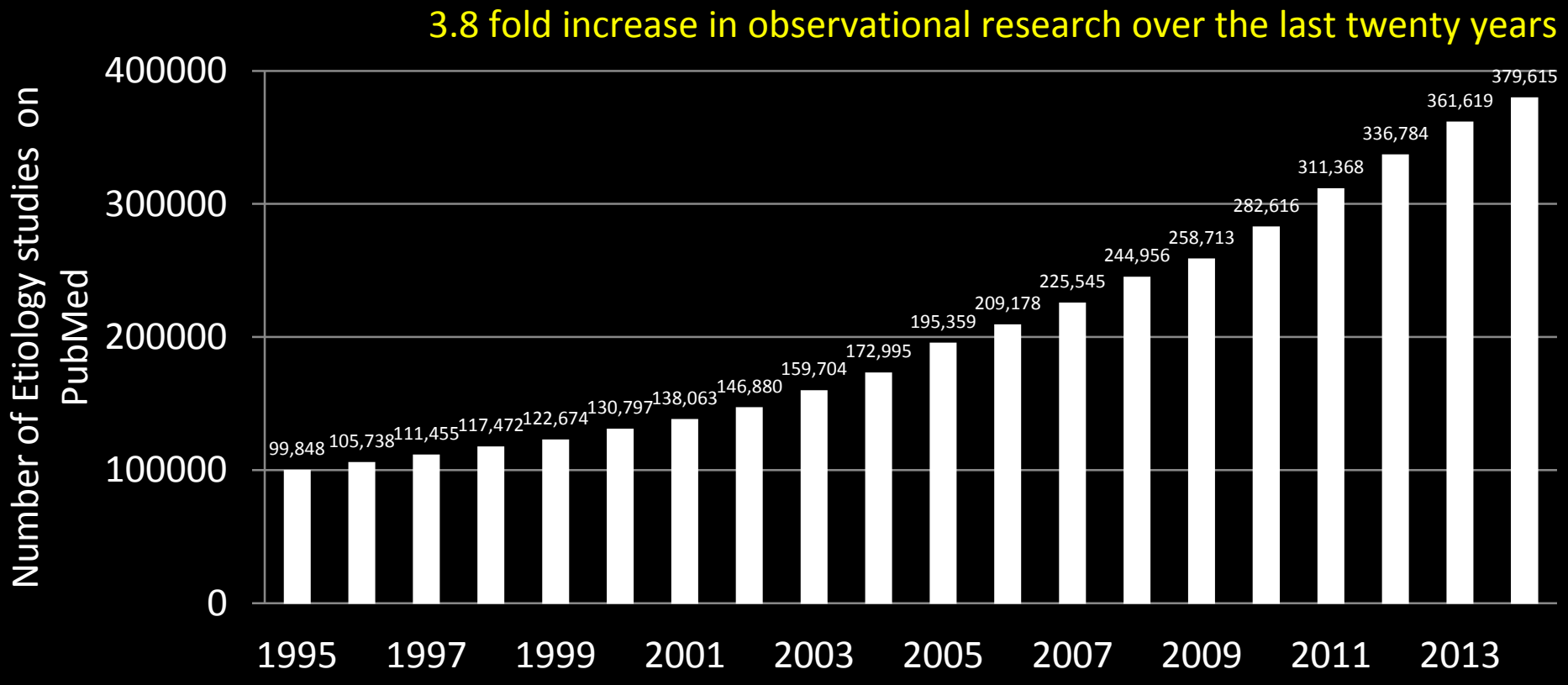
There are huge shortcomings in the way that evidence based medicine operates today: bad quality research, evidence that is withheld, piecemeal dissemination, a failure to respect patients' priorities, and more.

We need less research, better research, and research done for the right reasons.
Doug Altman. BMJ 1994 The Scandal of Poor Medical Research

(3 fold increase in RCTs over 20 years)

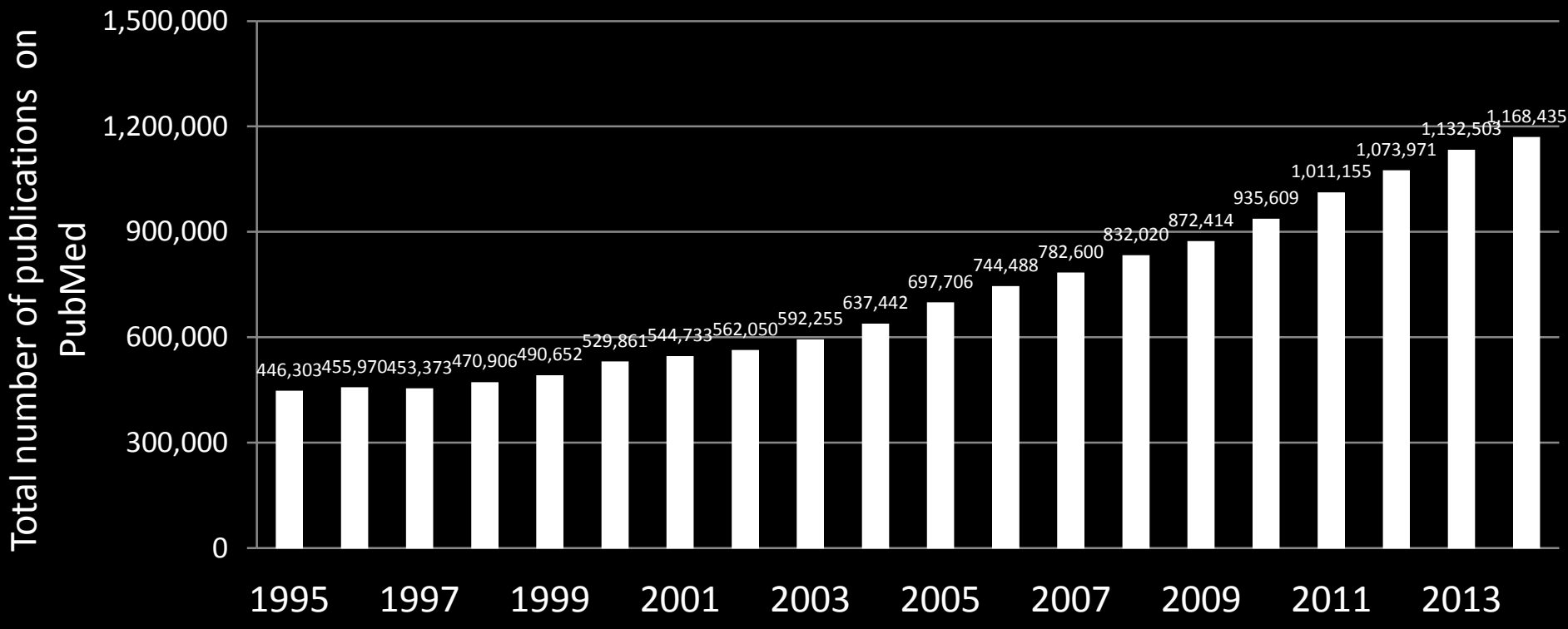


Observational research: etiology filter - risk*[Title/Abstract] OR risk*[MeSH:noexp] OR risk *[MeSH:noexp] OR cohort studies[MeSH Terms] OR group[Text Word] OR groups[Text Word] OR grouped [Text Word])

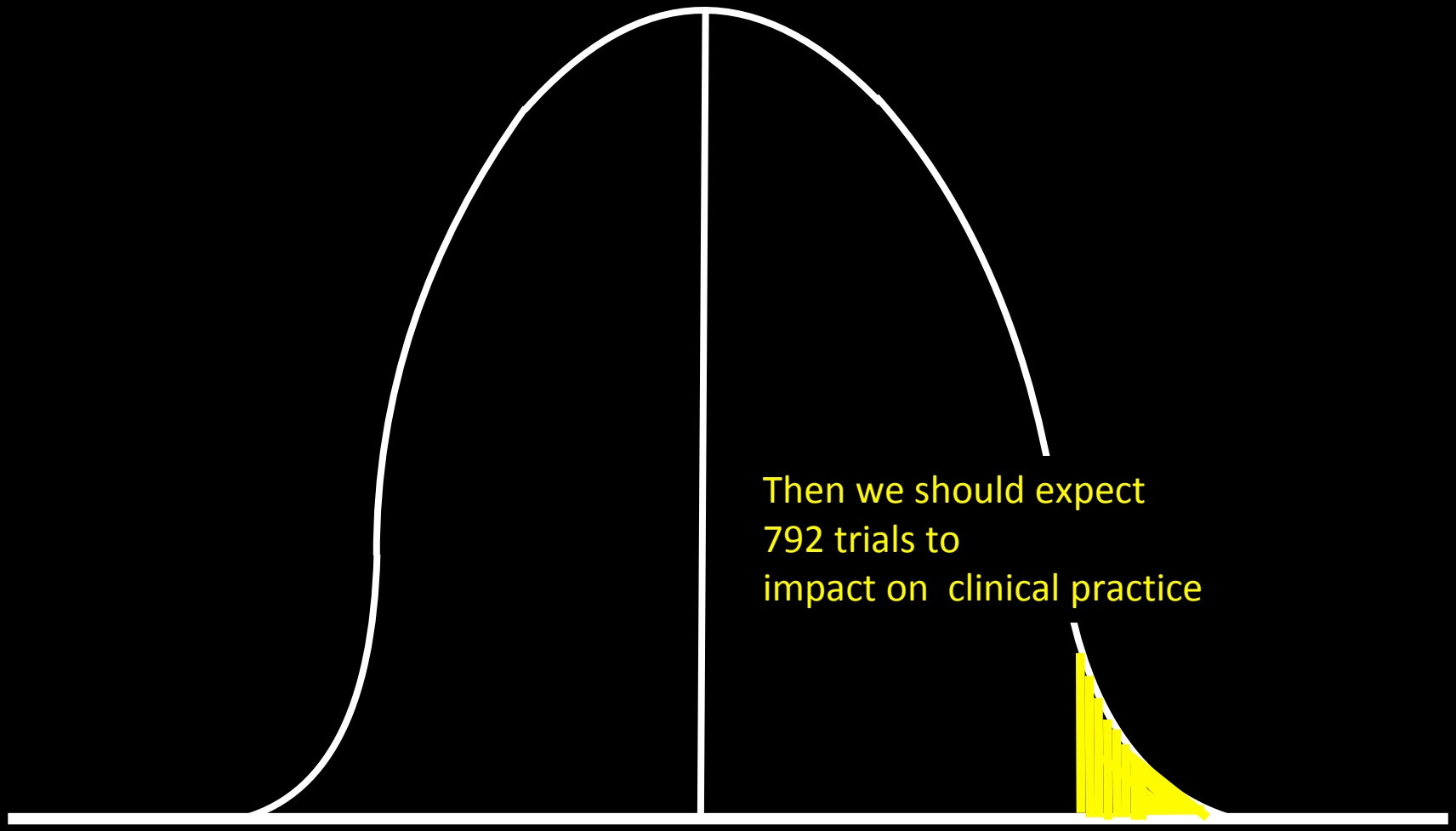


Growth in all Research on Pubmed -

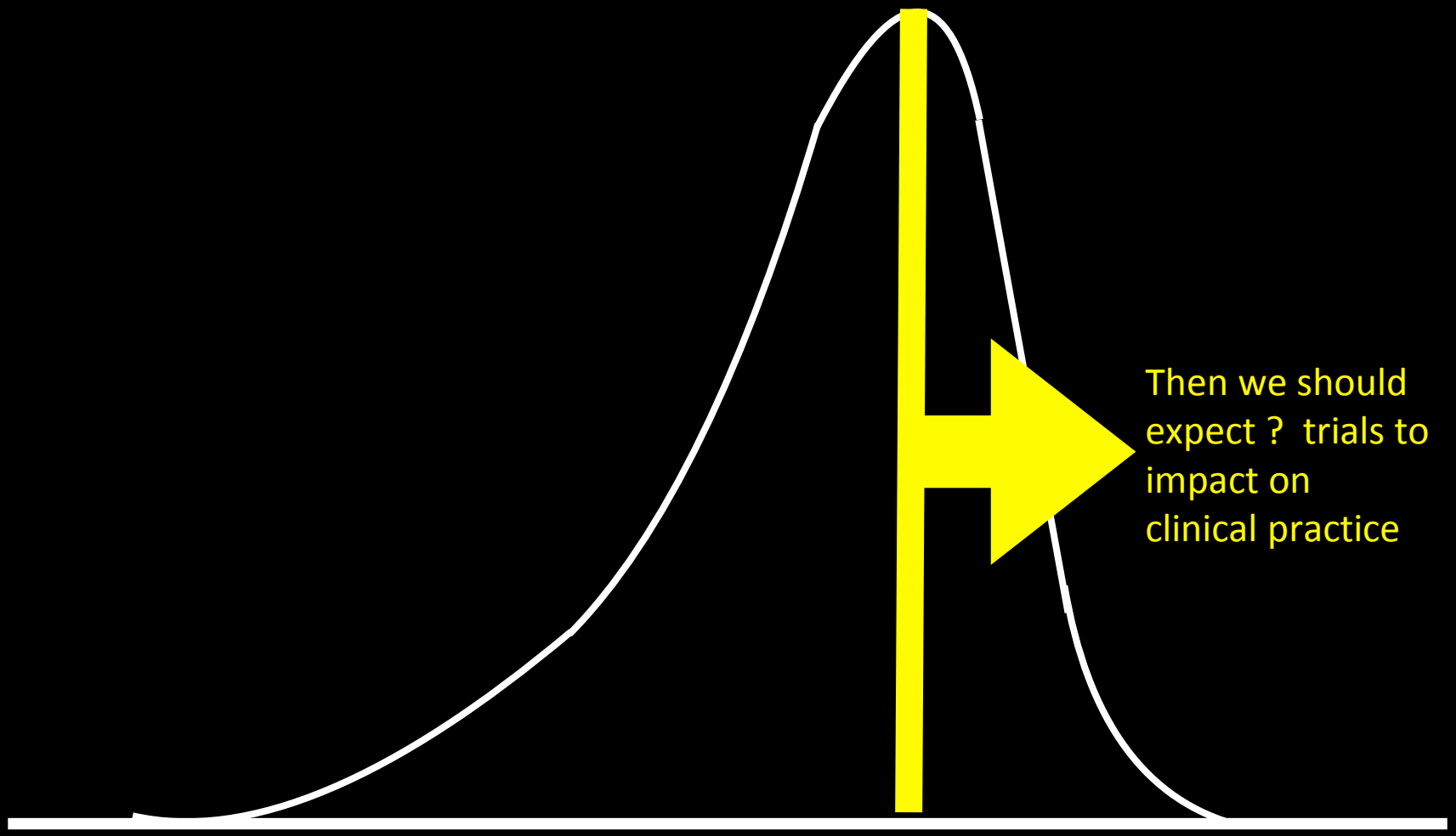
(2.6 fold increase in all research on PubMed in the last 20 years)



If 31,694 trials on average have no effect on clinical practice



But improvements in the design and conduct of research should mean t
net effect is skewed towards benefit from the outset.

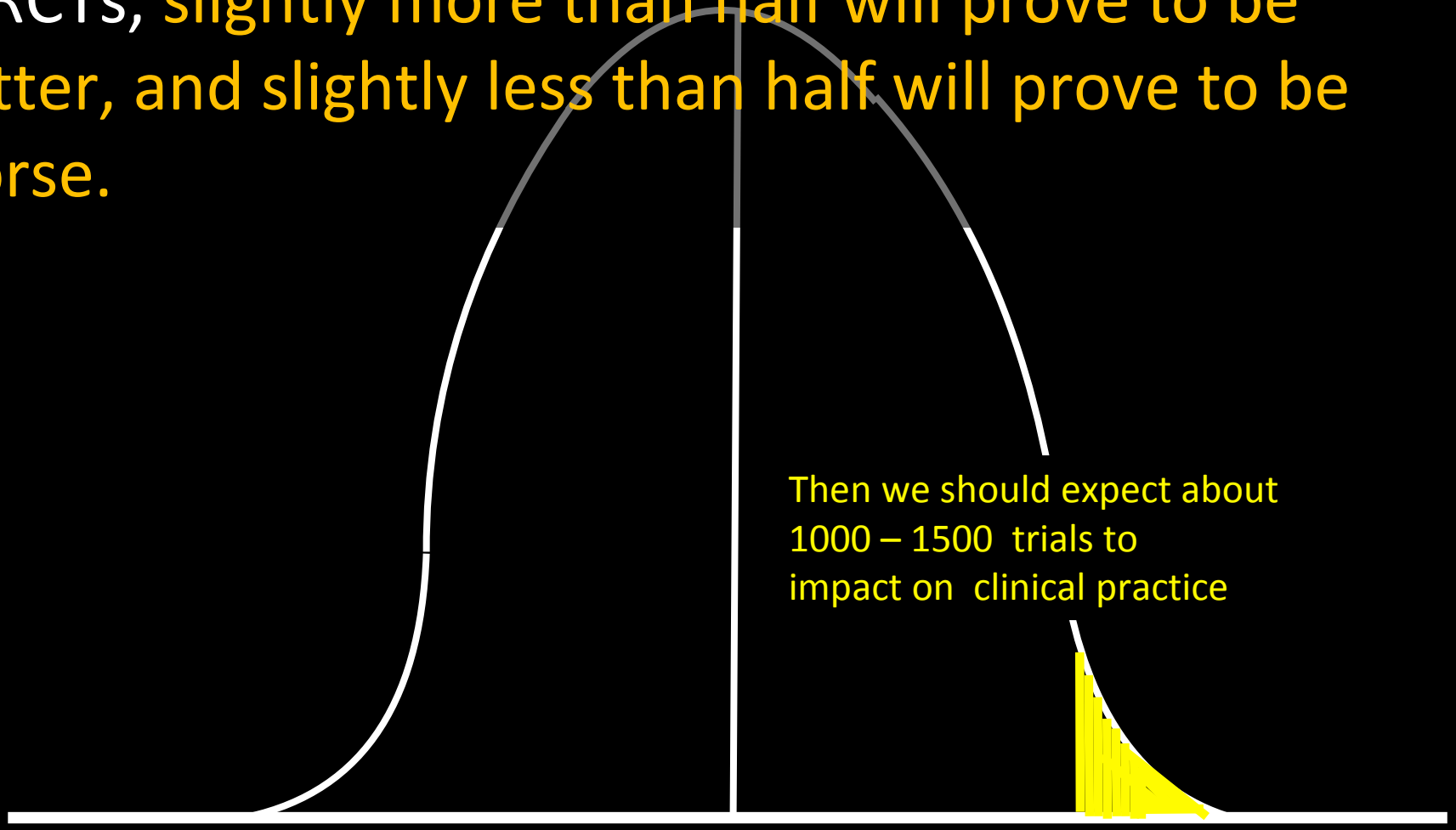


The proportion of proposed new treatments that are ’successful’ is of ethical, scientific, and public importance. We investigated how often new, experimental treatments evaluated in RCTs are superior to established treatments

New treatments are only slightly superior to established treatments when tested in RCTs.

Additionally, results have remained stable over time and that the success rate of new treatments has not changed over the last half century of clinical trials.

Society can expect that when new experimental treatments are tested against established treatments in RCTs, slightly more than half will prove to be better, and slightly less than half will prove to be worse.



If 31,694 trials on average have no effect on clinical practice

Significant problems exist with the E in EBM

Why does so little research translate into practice?

Why does so little research translate into practice?

Three main problems:

1. **External validity** - The results of the trials should apply to the populations we see in practice.

2. **Internal validity** – A trial's validity should be based on the study design, the quality of the data acquisition, adherence to the protocol, the quality of the reporting and the impact of conflicts of interest have on the interpretation of the results.

3. **Clinical Significance** – the significance should be based on the relevance of the outcomes and the trade off between benefits and harms

Sally Davis: Chief Medical Officer

In 2015, Sally Davies, the UK's Chief Medical Officer, requested a review to restore public trust in the safety and effectiveness of medicines, because patients increasingly see *"doctors as over medicating and clinical scientists afflicted by conflicts of interest, and therefore it is difficult for the public to trust either."*

Review

Synthèse

Dr. Sackett is Director of the Trout Research & Education Centre at Irish Lake, Markdale, Ont.

This article has been peer reviewed.

CMAJ 2001;165(9):1226-37

Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only formula a clinician-trialist is ever likely to need (or understand!)

David L. Sackett

> Because statistics has too often been presented as a bag of specialized computational tools, with morbid emphasis on calculation, it is no wonder that survivors of such courses regard their statistical tools as instruments of torture [rather] than as diagnostic aids in the art and science of data analysis.
> — George W. Cobb[1]

Review

Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only formula a clinician-trialist is ever

The "only formula" of physiological statistics

Go to:

The formula is ridiculously simple, and looks like this (Equation 1):

Confidence in the balance of benefits to harms

$$= \frac{\text{Effect Size}}{\text{Bias}} \times \sqrt{\text{Sample size}}$$

Optimize

Achieve

Patient benefit $=$ $\frac{\text{Outcome}}{\text{Bias}}$ $\times$ Optimal Information size

Minimize

Optimize

Confidence in the results Patient Benefit $= \frac{\text{Outcome}}{\text{Bias}} \times$ Optimal Information size

First question to ask:

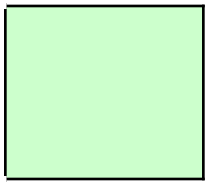
Does this make a difference to my patient?

If the answer is no, consider **stopping**

Estimate NNT

How many 60-year-old patients with mild **hypertension** would you have to treat with **diuretics** for 5 years to prevent 1 **stroke**?

NNT

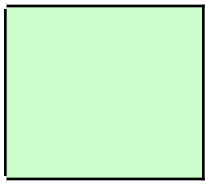


Ref: http://www.cche.net/usersguides/ebm_tips.asp

Estimate NNT

NNT

How many people with **myocardial infarction** would you have to treat with **ß-blockers** for 2 years to prevent 1 **death**?

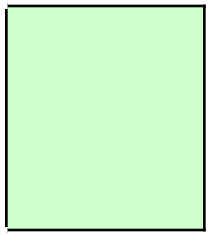


Ref: http://www.cche.net/usersguides/ebm_tips.asp

Estimate NNT

NNT

How many people with **acute myocardial infarction** would you have to treat with **streptokinase** to prevent 1 person from **dying** in the next 5 weeks?



Ref: http://www.cche.net/usersguides/ebm_tips.asp

NNTs from Controlled Trials

	Control Event %	Treatment Event %	Absolute Risk Reduction %	NNT
Population: hypertensive 60-year-olds Therapy: oral diuretics Outcome: stroke over 5 years	2.9	1.9	1	100
Population: myocardial infarction Therapy: ß-blockers Outcome: death over 2 years	9.8	7.3	2.5	40
Population: acute myocardial infarction Therapy: streptokinase (thrombolytic) Outcome: death over 5 weeks	12	9.2	2.8	36

Ref: http://www.cche.net/usersguides/ebm_tips.asp

Heneghan *et al. Trials* (2017) 18:122
DOI 10.1186/s13063-017-1870-2

Trials

COMMENTARY Open Access

Why clinical trial outcomes fail to translate into benefits for patients

Carl Heneghan*, Ben Goldacre and Kamal R. Mahtani

Abstract

Clinical research should ultimately improve patient care. For this to be possible, trials must evaluate outcomes that genuinely reflect real-world settings and concerns. However, many trials continue to measure and report outcomes that fall short of this clear requirement. We highlight problems with trial outcomes that make evidence difficult or impossible to interpret and that undermine the translation of research into practice and policy. These complex issues include the use of surrogate, composite and subjective endpoints; a failure to take account of patients' perspectives when designing research outcomes; publication and other outcome reporting biases, including the under-reporting of adverse events; the reporting of relative measures at the expense of more informative absolute outcomes; misleading reporting; multiplicity of outcomes; and a lack of core outcome sets. Trial outcomes can be developed with patients in mind, however, and can be reported completely, transparently and competently. Clinicians, patients, researchers and those who pay for health services are entitled to demand reliable evidence demonstrating whether interventions improve patient-relevant clinical outcomes.

Keywords: Clinical outcomes, Surrogate outcomes, Composite outcomes, Publication bias, Reporting bias, Core outcome sets

thebmj

BMJ 2014;349:g6501 doi: 10.1136/bmj.g6501 (Published 21 November 2014)

RESEARCH

Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews

OPEN ACCESS

Pooja Saini *research associate*[1], Yoon K Loke *professor*[2], Carrol Gamble *professor*[3], Douglas G Altman *professor*[4], Paula R Williamson *professor*[3], Jamie J Kirkham *lecturer*[3]

[1]Department of Public Health and Policy, University of Liverpool, Liverpool, UK; [2]Norwich Medical School, University of East Anglia, Norwich, UK; [3]Department of Biostatistics, University of Liverpool, Liverpool, L69 3GA, UK; [4]Centre for Statistics in Medicine, University of Oxford, Oxford, UK

Abstract

Objective To determine the extent and nature of selective non-reporting of harm outcomes in clinical studies that were eligible for inclusion in a cohort of systematic reviews.

Design Cohort study of systematic reviews from two databases.

Setting Outcome reporting bias in trials for harm outcomes (ORBIT II) in systematic reviews from the Cochrane Library and a separate cohort of systematic reviews of adverse events.

Participants 92 systematic reviews of randomised controlled trials and non-randomised studies published in the Cochrane Library between issue 9, 2012 and issue 2, 2013 (Cochrane cohort) and 230 systematic reviews published between 1 January 2007 and 31 December 2011 in other publications, synthesising data on harm outcomes (adverse event cohort).

Methods A 13 point classification system for missing outcome data on harm was developed and applied to the studies.

Results 86% (79/92) of reviews in the Cochrane cohort did not include full data from the main harm outcome of interest of each review for all of the eligible studies included within that review; 76% (173/230) for the adverse event cohort. Overall, the single primary harm outcome was inadequately reported in 76% (705/931) of the studies included in the 92 reviews from the Cochrane cohort and not reported in 47% (4159/8837) of the 230 reviews in the adverse event cohort. In a sample of primary studies not reporting on the single primary harm outcome in the review, scrutiny of the study publication revealed that outcome reporting bias was suspected in nearly two thirds (63%, 248/393).

Conclusions The number of reviews suspected of outcome reporting bias as a result of missing or partially reported harm related outcomes from at least one eligible study is high. The declaration of important harms and the quality of the reporting of harm outcomes must be improved in both primary studies and systematic reviews.

Correspondence to: J J Kirkham jjk@liv.ac.uk

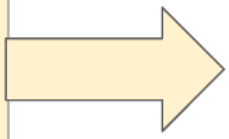
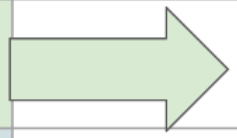
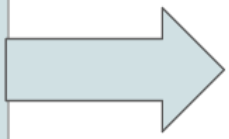
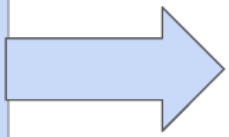
Introduction

> "When we looked at that data, it actually showed an increase in harm amongst those who got the active treatment, and we ditched it because we weren't expecting it and we were concerned that the presentation of these data would have an impact on people's understanding of the study findings."[1]

Health technology assessment is a form of policy research that examines evidence of safety, efficacy, and cost effectiveness of a healthcare technology to provide guidance and recommendations to support decisions about treatment. Health technology assessment systematic reviews of clinical studies aim to include all relevant studies conducted on a particular topic and to provide an unbiased summary of their results, thus producing the best evidence on the benefits and harms of medical treatments. However, research has shown that the validity of systematic reviews can be affected by outcome reporting bias in the primary studies, which has been defined as selection (on the basis of the results) of a subset of the original variables recorded for inclusion in a study publication.[2]

The prevalence and impact of outcome reporting bias has recently been investigated in a large unselected cohort of Cochrane systematic reviews (ORBIT (Outcome Reporting Bias In Trials) I study).[3] This study, which focused on a single primary outcome for each review, found that over half the reviews (157/283, 55%) could not include data for the review primary outcome from all eligible studies. Additionally, interviews were conducted with trialists to understand the reasons for discrepancies between outcomes specified in the study protocol and those reported in the study publication.[1] The

Why clinical trial outcomes fail to translate into benefits for patients

Design	Badly Chosen		Surrogate Composite Subjective Complex Scales Lack of relevance to patients and decision makers
Methods	Badly Collected		Missing data Poorly Specified
Publication	Selectively reported		Publication Bias Reporting Bias Underreporting of Adverse events Switched Outcomes
Interpretation	Inappropriately interpreted		Relative measures Spin Multiplicity Core outcome sets

COPD: Journal of Chronic Obstructive Pulmonary Disease, 4:191–194
ISSN: 1541-2555 print / 1541-2563 online

DOI: 10.1080/15412550701471231

informa
healthcare

INTRODUCTION

How Can We Assess Outcomes of Clinical Trials: The MCID Approach

Barry Make (makeb@njc.org)

Division of Pulmonary Sciences and Critical Care Medicine, National Jewish Medical and Research Center, and University of Colorado School of Medicine, Denver, Colorado

ABSTRACT

Interpreting changes in outcomes of clinical trials in chronic obstructive pulmonary disease should be viewed from a broader perspective than only the statistical significance of the findings. The minimal clinical difference in outcome measures provides a conceptual framework to assist in clinical trial interpretation and a methodology to assess the clinical relevance of study results. Use of distribution-based techniques, comparison with other external measures, and opinions from experts, clinicians and patients can assist in minimal clinically important difference development. Although the minimal clinically important difference has been suggested for a wide range of outcomes of importance in chronic obstructive pulmonary disease, many have not been subjected to rigorous analysis. For newer tools such as activity monitors and questionnaires and measures not widely employed such as laboratory-based exercise tests, minimal clinically important differences remain to be determined.

INTRODUCTION

The focus of this article is on the assessment of three important outcomes of new therapies for patients with chronic obstructive pulmonary disease (COPD): exercise, health-related quality of life, and activity. It is tempting and seemingly logical to suggest that all 3 constructs are closely related. Other articles in this supplement indicate reasonable correlations between quality of life and activity, and exercise and quality of life (1, 2). However, the evidence indicates relations are less clear between exercise and activity, and exercise and quality of life (3). If all 3

The concepts of exercise, quality of life and activity are not well understood, the measures used to assess these parameters are not routinely employed in clinical practice, there is only a fair relationship of these measures to patient symptoms, and the outcome tools and their scoring are not readily apparent to patients and physicians. In addition, the clinical significance of changes in these outcomes may not be readily apparent. This article reviews the use of MCIDs as a potential tool to interpret the results of clinical trials using exercise, activity and quality-of-life outcomes. Certainly the MCID is not the only possible

Expanding the Definition of Clinical Differences: From Minimally Clinically Important Differences to Really Important Differences. Analyses in 8931 Patients with Rheumatoid Arthritis

FREDERICK WOLFE, KALEB MICHAUD, and VIBEKE STRAND

ABSTRACT. ***Objective.*** Minimally clinically important differences (MCID) have become an important way to interpret data of randomized clinical trials (RCT), but do not reflect the degree of improvement consistent with a "really important difference" (RID). To define RID, we compared mean and/or least desirable clinical states with best and/or most desirable states.

Methods. In total, 8931 patients with rheumatoid arthritis (RA) < 65 years of age completed the Health Assessment Questionnaire (HAQ) and Medical Outcomes Survey Short Form 36 Physical Component Score (PCS). Definitions of RID were based on values for HAQ and PCS corresponding with the best and worst category of the following conditions: disabled vs not disabled: joint replacement vs no joint replacement; ≤ poverty level vs > poverty level; very satisfied with health vs not; and independent in participation activities vs not.

Results. In contrast to published MCID values for the HAQ of ~0.22, RID was as high as 0.76 using objective reference conditions and 0.87 using the subjective measure of dependence vs independence. The HAQ score of independent RA patients was 0.38 (SD 0.45), and was 0.42 (SD 0.53) for those very satisfied with their health. The difference in HAQ scores between disabled and working patients was ~0.75. PCS differences were similarly increased.

Conclusion. RID values are 3 to 4 times greater than MCID values. Although MCID are meaningful statistics for RCT, the RID percentage achieved [(actual improvement/RID) × 100%] is a simple way to put the results of RCT in a broader perspective that gives an idea of how much additional treatment effect is needed. (J Rheumatol 2005;32:583–9)

Optimize

Patient benefit $=$ $\frac{\text{Outcome}}{\text{Bias}} \times$ Optimal Information size

First questions to ask:

1. Does this make a difference to my patient?
2. Does this apply to my patient?
3. Is the treatment feasible in my setting?

If the answer is no, then consider **Stopping**

Reply Reply All Forward

Tue 06/06/2017 11:22

Fiona Lethbridge <lethbridge@sciencemediacentre.org>

RE: EMBARGOED: 'Even moderate drinking linked to a decline in brain health, finds study' - QUOTES NEEDED

To Carl Heneghan

From: Fiona Lethbridge [mailto:lethbridge@sciencemediacentre.org]
Sent: 05 June 2017 16:06
To: Fiona Lethbridge <lethbridge@sciencemediacentre.org>
Subject: RE: EMBARGOED: 'Even moderate drinking linked to a decline in brain health, finds study' - QUOTES NEEDED
Importance: High

Hi all,

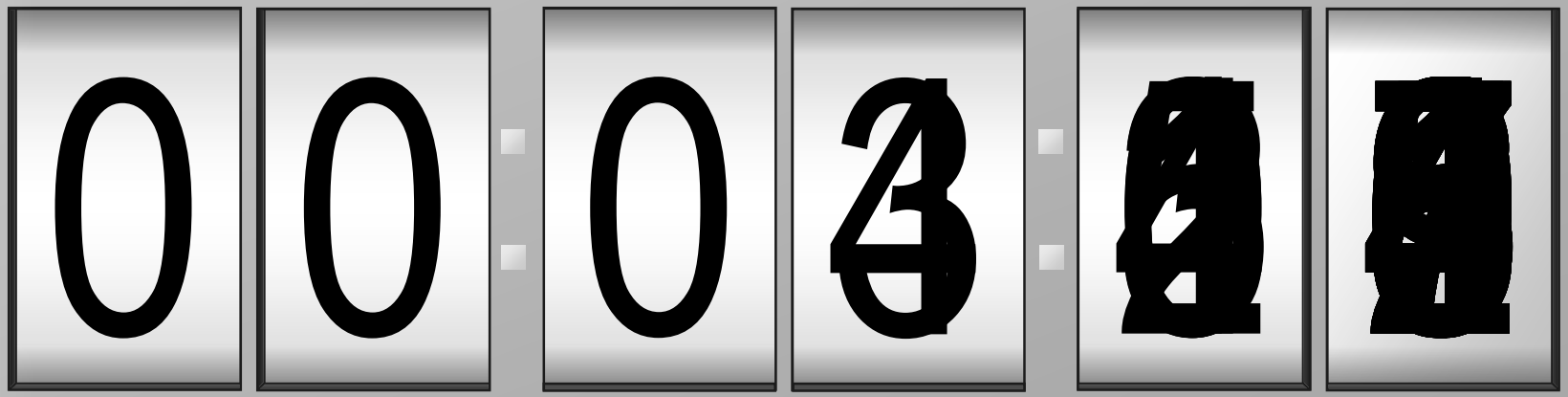
Hope you're well. Just a reminder I'd like expert quotes on this by tomorrow morning (Tuesday 6 June) if this is your area of expertise and you're able to help – let me know if you can and I'll send you the paper and editorial.

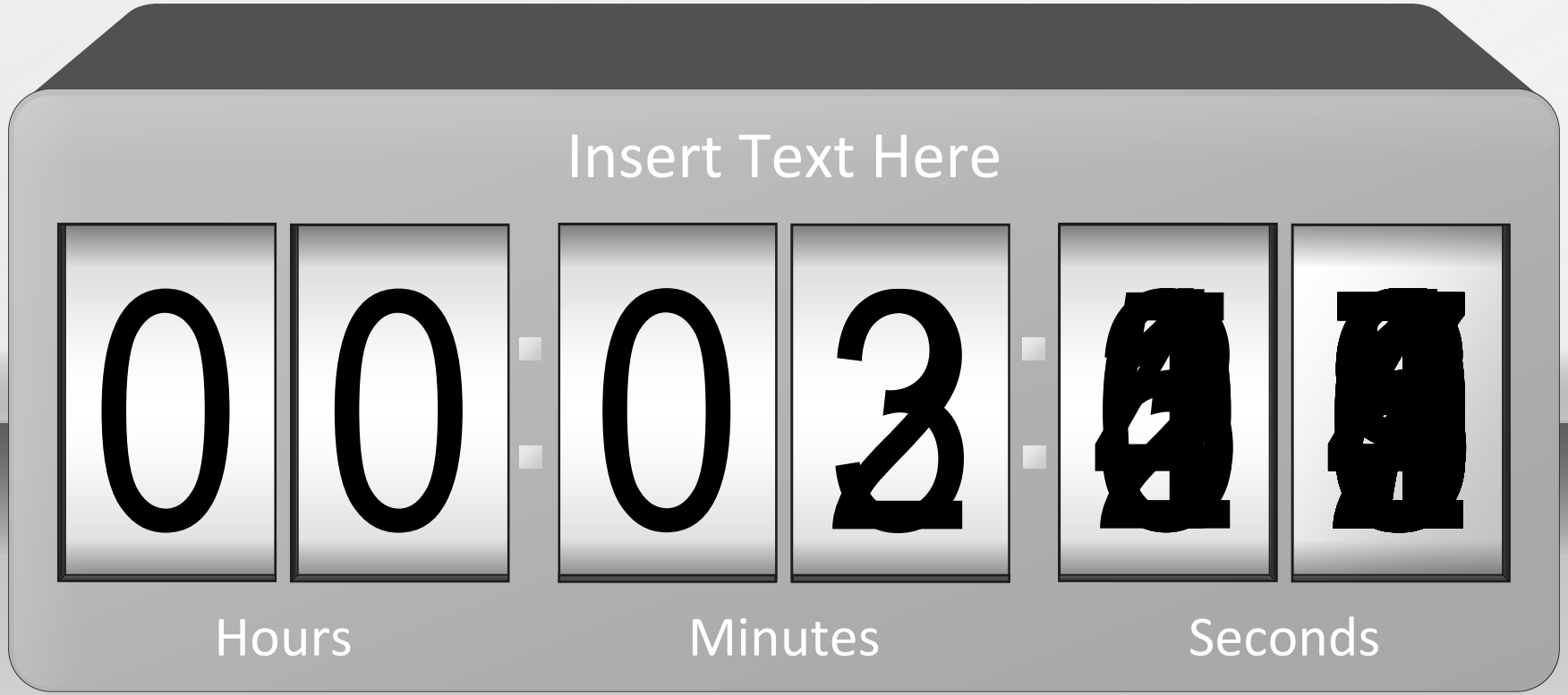
Thanks in advance!

Cheers,

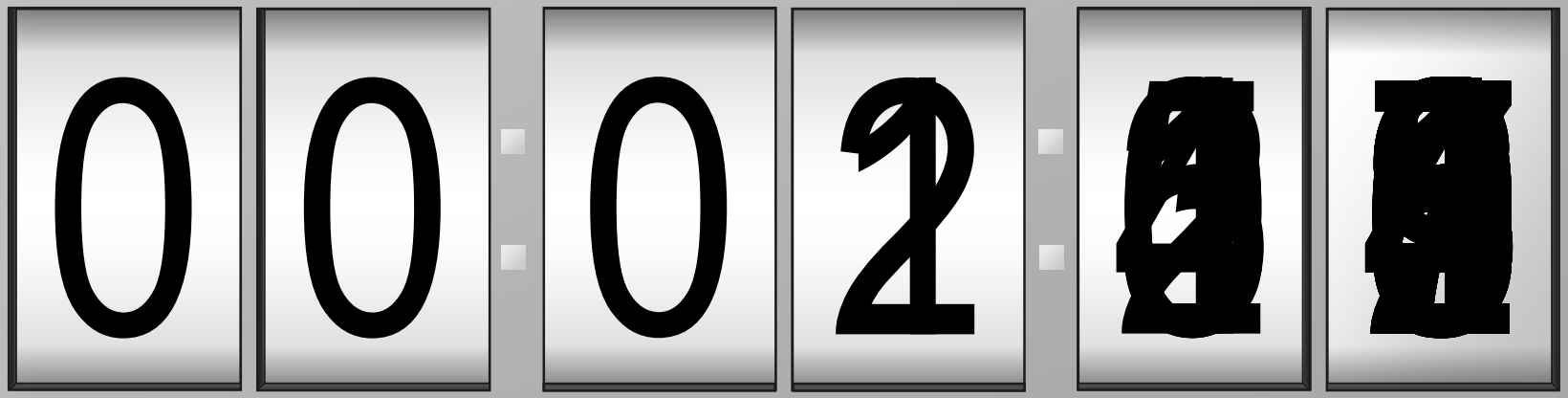
Fiona

Insert Text Here

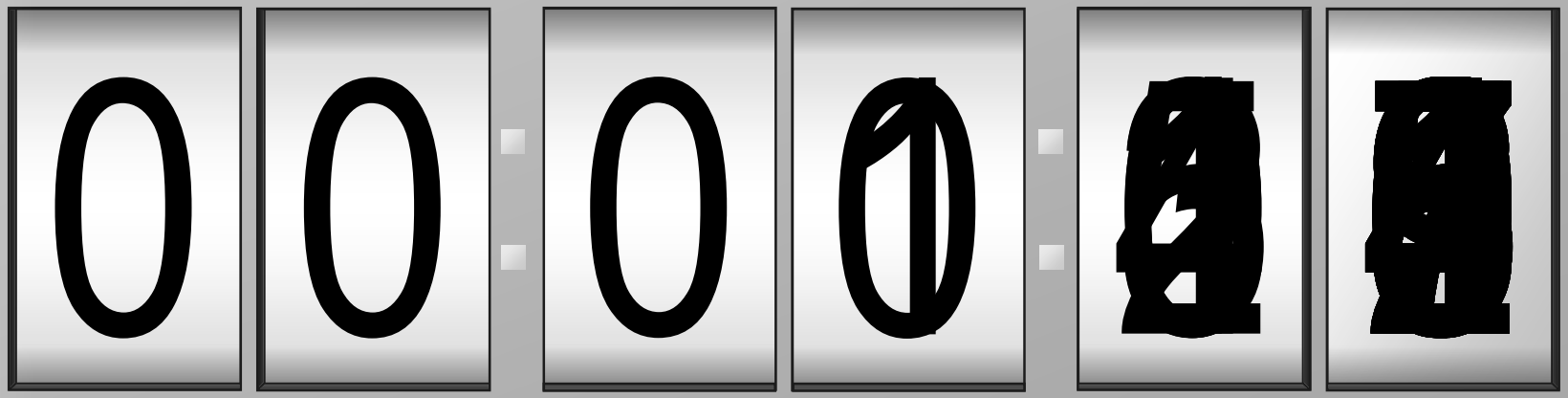



Insert Text Here
00:03:
Hours
Minutes
Seconds

Insert Text Here



Insert Text Here



Tue 06/06/2017 13:30
Fiona Lethbridge <lethbridge@sciencemediacentre.org>
SMC Roundup V2: Drinking alcohol and brain health, BMJ [EMBARGO 23:30 Tues 6 June]

To lethbridge@sciencemed

Science M

EMBARGOED UNTIL

Expert reacti

NEW COMMENT Pr

"Does moderate alc
significant methodol

"For example, the ris
discrepancy.

"Also, over the 30 ye
alcohol consumptior
https://bmcmedicine

"I also want to make
atrophy (see table 2)
times higher – 24/31

"These type of studi

"Using all the availat
coronary death, hea

Prof Carl Heneghan, Director of the Centre for Evidence-Based Medicine, University of Oxford, said:

"Does moderate alcohol consumption lead to adverse brain outcomes and cognitive decline? The BMJ cohort study that tries to address this question has significant problems that prevent firm conclusions being drawn – the **sample size is small**, there are **significant methodological problems** and the results are **sometimes at odds with each other.**

"For example, the **risk of right-sided hippocampal atrophy was significantly greater at >14 alcohol units** a week compared with abstinence, but for left-sided atrophy this was only the case at >30 units a week. Yet, there is no reason given for this discrepancy.

"Also, over the 30 year period, weekly alcohol intake did not increase in the study participants. This suggests there were significant problems with self-reported alcohol intake in the cohort, as it does change over time. Analysis of life-course trajectories of alcohol consumption in the UK reports a rapid increase in 'alcohol intake during adolescence leading to a peak in early adulthood, followed by a plateau in mid-life, and then a decline into older ages', see: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-015-0273-z

"I also want to make an important point for interpreting results. **The use of odds ratios** in the paper are misleading – they suggest the effect is much greater than it actually is. In the no alcohol group approximately half of the people had hippocampal atrophy (see table 2). The results in the abstract suggest that if you are a heavy drinker then your risk is five times higher. This is not the case – as relatively your risk as heavy drinker can be only twice as high if the baseline risk is 50% (it's actually about 1.8 times higher – 24/31 (74%) versus 9/22 (41%)). "These type of studies also cannot account for all the confounders, and therefore they cannot, and should not, conclude causation.

"**Using all the available evidence provides a much more balanced approach for the public on deciding how much to drink. As an example, a recent BMJ observational study pointed to** non-drinking being associated with increased risk of heart attack, coronary death, heart failure and stroke when compared with moderate drinking, see: http://www.bmj.com/content/356/bmj.j909."

Because the new study was observational and not experimental, no firm conclusions could be drawn about cause and effect. The authors also acknowledged that the sample size was small.

Outside experts gave the study mixed reviews.

"It shows evidence for 'hidden' damage to the brain," commented Paul Matthews of Imperial College London, who highlighted the value of the advanced imaging techniques used.

Jennifer Wild, a senior researcher in clinical psychology at the University of Oxford, said the results showed a "robust link" between what most people would consider casual drinking and brain degeneration later in life.

Other scientists expressed skepticism about some of the methodology, and the self-reporting of alcohol consumption.

"Over the 30-year period, weekly intake did not increase in the study participants," noted Carl Heneghan, director of the Centre for Evidence-Based Medicine at the University of Oxford.

This flies in the face of a known pattern of increasing consumption through early adulthood, and a gradual tapering off into older age, he pointed out.

Patient benefit = Outcome / Bias × Optimal Information size

Assess

First questions to ask:

1. Does this make a difference to my patient?
2. Does this apply to my patient?
3. Is the treatment feasible in my setting?

If yes then do you believe the results assess bias

How to decide on what makes an outcome important?

WHO Position Paper on Mammography Screening Scoping Document

"Breast cancer is the leading cancer in women worldwide both in the developed and developing world. So far the only screening method proving to be effective is mammography screening. However, there is currently great controversy regarding the benefits and harms of mammography screening. The same evidence is interpreted in different ways by different groups of experts and methodologists who tend to have extreme positions in favor or against this type of screening…There is also great controversy on the age groups in which screening should be recommended. Consequently, there is an urgent need for WHO to provide policy makers, patients and health care providers with a clear, objective and independent guidance on the benefits and harms of mammography screening in different age groups."

http://apps.who.int/iris/bitstream/10665/137339/1/9789241507936_eng.pdf?ua=1

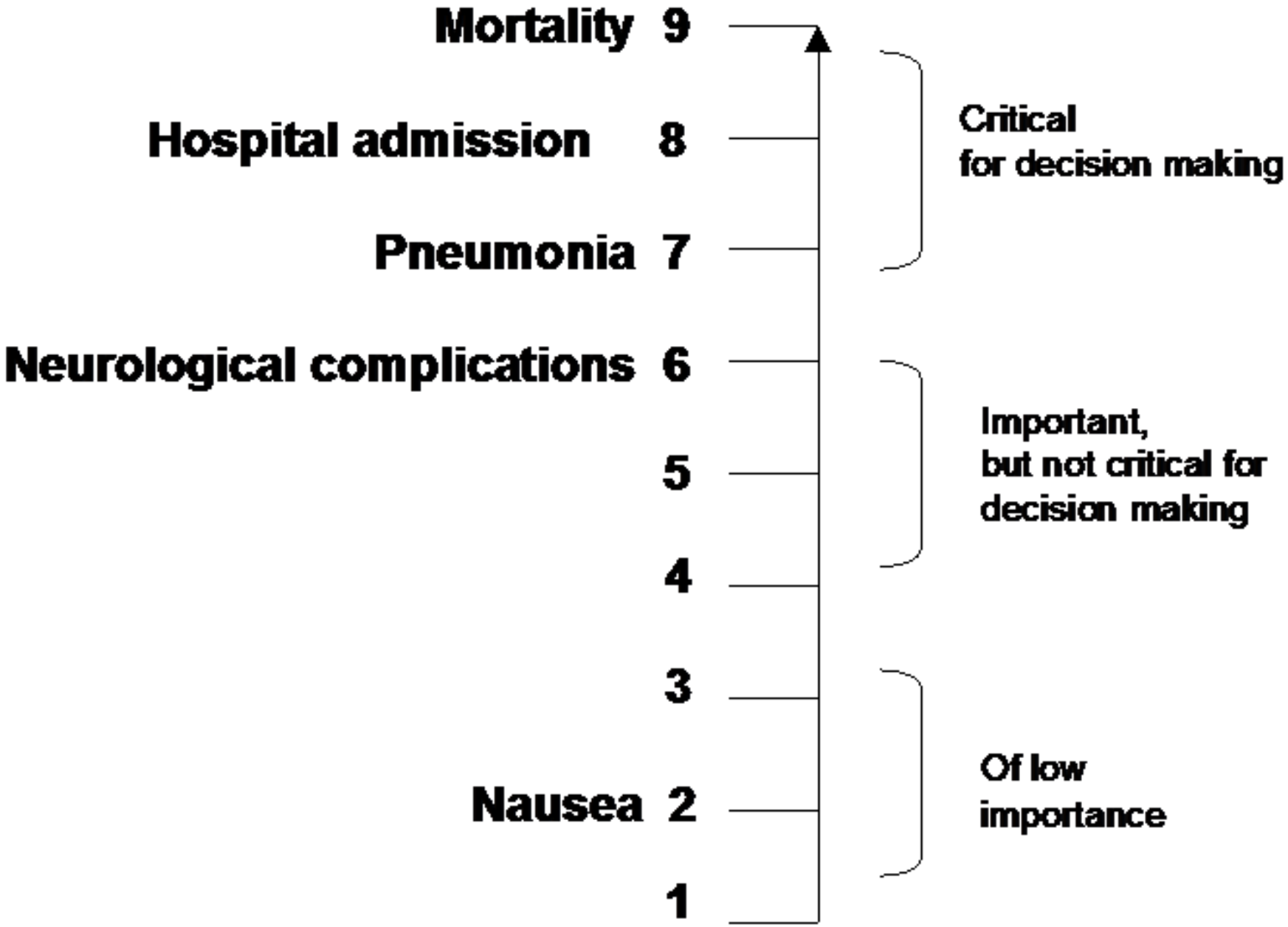
2. How to decide outcomes

- Patient-centred
- Consider input from clinicians, patients, public
- Consider harms
- Should be extensive
- Think “outside the box”

Why should we prioritize outcomes?

- Helps to focus research and evidence syntheses
- Helps to define the domains of the evidence profile and summary of findings tables
- The quality of the evidence for critical outcomes gives a measure of the overall evidence base for the PICO-related research question

3. How to Prioritise outcomes



How should we decide which outcomes to prioritise?

- Should be based on **“importance and not evidence”**
- Include both critical and important outcomes
- Can use a panel, clinical experts, patient groups, literature
- Outcomes should be ranked by importance anonymously
- Use a measure (mean or median) to rank
- Outcomes with wide variability should be discussed
- Obtain a consensus for all outcomes

Outcomes of interest in mammography screening

(What would be the critical outcomes and what would be the important outcomes in terms of mammography screening?)

Critical outcomes

1.
2.
3.

Important outcomes

WHO Position Paper on Mammography Screening Scoping Document

"Breast cancer is the leading cancer in women worldwide both in the developed and developing world. So far the only screening method proving to be effective is mammography screening. However, there is currently great controversy regarding the benefits and harms of mammography screening. The same evidence is interpreted in different ways by different groups of experts and methodologists who tend to have extreme positions in favor or against this type of screening…There is also great controversy on the age groups in which screening should be recommended. Consequently, there is an urgent need for WHO to provide policy makers, patients and health care providers with a clear, objective and independent guidance on the benefits and harms of mammography screening in different age groups."

http://apps.who.int/iris/bitstream/10665/137339/1/9789241507936_eng.pdf?ua=1

Outcomes of interest in mammography screening

(What would be the critical outcomes and what would be the important outcomes in terms of mammography screening?)

Critical outcomes

- Breast cancer-specific mortality
- Disability-adjusted life-years (DALYs) gained
- Health-related quality of life

Important outcomes

- All-cause mortality
- Overtreatment
- Reduction in mastectomies
- Overdiagnosis
- Cumulative false-positives

Patient benefit = Outcome / Bias × Optimal Information size

Assess

First questions to ask:

1. Does this make a difference to my patient?
2. Does this apply to my patient?
3. Is the treatment feasible in my setting?

If yes then do you believe the results assess bias

What gives rise to the highest amount of bias when assessing evidence for treatment effects?

Types of study evidence affects the quality

2005

THE LANCET
Volume 366, Issue 9493, 8–14 October 2005, Pages 1267–1278

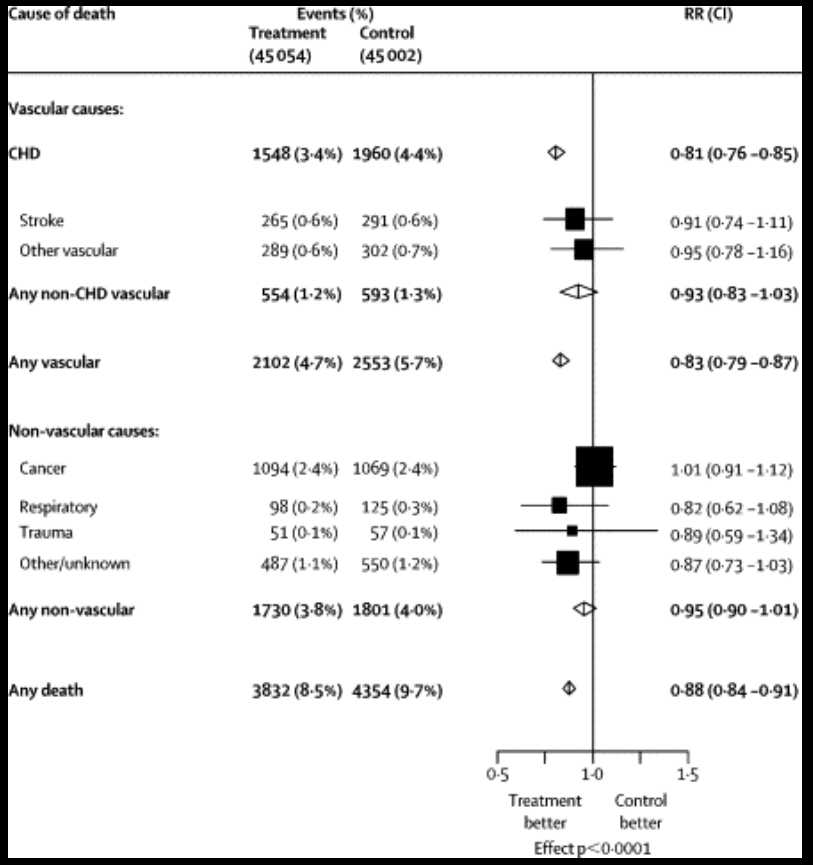
Fast track — Articles

Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins

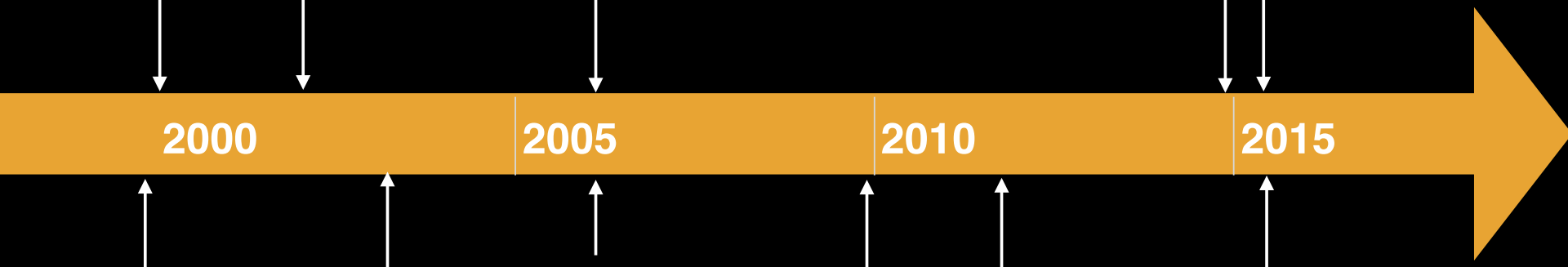
Cholesterol Treatment Trialists' (CTT) Collaborators ***‡

3832 (8·5%) deaths among the 45 054 participants allocated a statin compared with 4354 (9·7%) among the 45 002 controls.

Represents a 12% proportional reduction in all-cause mortality per mmol/L LDL cholesterol reduction (RR 0·88, 95% CI 0·84–0·91; p<0·0001)



Heart Attack evidence



1999
HOPE trial published on ramipril in the **NEJM**

2000
Lancet: publishes effect of low dose b-blockers in the elderly

2002
Lancet: HPS study shows benefit of statins independent of baseline level

2003
Lancet: primary PTCA more effective than thrombolysis

2005
Editorial in **BMJ** calls for aspirin for everyone

Lancet: 2005 COMMIT trial beta blocker should be withheld if the patient is unstable

Lancet: cholesterol triallist collaboration 90,056 participants in 14 RTCs

2005
Lancet: COMMIT trial of early metoprolol in 45,852 participants

2009
Lancet: aspirin is of uncertain net benefit for primary prevention

2010
In hospital mortality **5%**

2013
Routine use of Oxygen questioned in a **Cochrane** review

2015
American College of Physicians **guidelines** warn against screening asymptomatic individuals

2015
Paramedic 2 Trial tests effectiveness of adrenaline in cardiac arrest

Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial

*The DREAM (Diabetes REduction Assessment with ramipril and rosiglitazone Medication) Trial Investigators**

FINDINGS:

At the end of study, 59 individuals had dropped out from the rosiglitazone group and 46 from the placebo group. 306 (11.6%) individuals given rosiglitazone and 686 (26.0%) given placebo developed the composite primary outcome (hazard ratio 0.40, 95% CI 0.35-0.46; p<0.0001); 1330 (50.5%) individuals in the rosiglitazone group and 798 (30.3%) in the placebo group became normoglycaemic (1.71, 1.57-1.87; p<0.0001). Cardiovascular event rates were much the same in both groups, although 14 (0.5%) participants in the rosiglitazone group and two (0.1%) in the placebo group developed heart failure (p=0.01).

Interpretation Rosiglitazone at 8 mg daily for 3 years substantially reduces incident type 2 diabetes and increases the likelihood of regression to normoglycaemia in adults with impaired fasting glucose or impaired glucose tolerance, or both.

Prevention of diabetes

Drug trials show promising results, but have limitations

Diabetes affects one in 20 adults worldwide and 333 million cases are projected worldwide by 2025.[1] Treatment can prevent some of the microvascular and macrovascular complications, but diagnosis is often delayed until complications present,[2] so attention has focused on prevention and early screening. Two strategies currently exist for reducing the onset of diabetes—lifestyle interventions and drugs.

confidence interval 0.35 to 0.46, P<0.0001). Ramipril did not reduce the risk of diabetes.

These results are promising, but they should be interpreted with caution. The mean fasting plasma concentration of glucose in both groups at baseline was 5.8 mmol/l, whereas the two hour impaired glucose tolerance test had a value of 8.7 mmol/l. The study popula-

"Furthermore, despite the population being at low risk of heart failure (10 year risk 0.33%) a significant increase (0.4%) in heart failure was seen in the rosiglitazone group compared with placebo (7.03, 1.60 to 30.9, number needed to harm at three years 250)."

study found that lifestyle interventio

similar reduction in risk was found in

ventions are labour intensive—one s

achieve target weight reduction and

Although lifestyle interventions pro

results in research settings, they are dif

even in well funded healthcare system

Considerable interest has focused

tion of diabetes with drugs. For instan

Prevention Program Research Grou

31% reduction in the incidence o

metformin at 2.8 years.[5] Previously

shown to be effective in controlling bl

but was removed from the market be

liver toxicity.[5] In people with obesity

shown to reduce the risk of diabete

compared with placebo.[6]

More recently came the publicatio

reduction assessment with ramipril a

medication (DREAM) trial.[7,8] In this

$25m (£13m; €20m), 5269 people ov

impaired fasting glucose or impaired g

or both, and no previous cardiovascular

domised to receive either rosiglitazone 8 mg daily or pla-

onset diabetes. As in the metformin trial in 2002, this

BMJ 333:764 doi:10.1136/bmj.38996.709340.BE (Published 12 October 2006)

Editorial

Prevention of diabetes

Carl Heneghan (carl.heneghan@dphpc.ox.ac.uk), deputy director, **M Thompson**, clinical lecturer, **R Perera**, senior statistician

+ Author Affiliations

Drug trials show promising results, but have limitations

Diabetes affects one in 20 adults worldwide and 333 million cases are projected worldwide by 2025.[1] Treatment can prevent some of the microvascular and macrovascular complications, but diagnosis is often delayed until complications present,[2] so attention has focused on prevention and early screening. Two strategies currently exist for reducing the onset of diabetes—lifestyle interventions and drugs.

ROSIGLITAZONE WHAT WENT WRONG

Over 10 years after the diabetes drug rosiglitazone was approved by regulators, and despite studies on tens of thousands of people, questions remain about its cardiovascular safety. An investigation by **Deborah Cohen** looks at why this happened.

RCTs

What gives rise to bias when assessing evidence for treatment effects in RCTs?

Assessing risk of bias for an RCT

- Recruitment
  - Were the subjects representative of the target population?
- Random Allocation
  - Were the groups comparable at the start?
- Maintenance
  - Were both groups treated the same bar the intervention
- Measurement
  - Unbiased (Blinded)
  - Objective
- What was the size of the effect?

RCTs. Summary title:

What question did the study ask?
PICO:

1: What difference do the results make?

- Does the outcome make a difference to my patient?
- Does this apply to my patient?
- Is the treatment feasible in my setting?

If the results do make a difference, are feasible and do apply to your patient then assess the impact of the biases (Go to question 2).

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.
Absolute Risk Reduction (ARR) = risk of the outcome in the control group minus risk of the outcome in the treatment group.	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.
Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. (RRR = 1 - RR)	The relative risk reduction tells us the reduction in the rate of the outcome in the treatment group relative to that 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.

How precise was the estimate of the treatment effect?

If the confidence interval is fairly narrow then we can be confident our point estimate is a precise reflection of the population value. The CI also provides us with information about the statistical significance of the result.

2: What is the impact of the biases on the results of the study?

Was the assignment of patients to treatments randomised?

What is best? *Centralised computer randomisation* is ideal, Smaller trials may use an *independent* person

Yes ☐ No ☐ Unclear ☐

Were groups similar at the start of the trial?

Check whether differences between groups are significant

Yes ☐ No ☐ Unclear ☐

Aside from the allocated treatment, were groups treated equally?

Apart from the intervention the patients in the different groups should be treated the same.

Yes ☐ No ☐ Unclear ☐

CEBM Rapid Critical appraisal

Were all patients in the trial accounted for and analysed in the groups they were randomised to?

If few patients have the outcome of interest, then even small losses to follow-up can bias results. Patients should be analysed in the groups to which they were randomised – *'intention-to-treat analysis'*.

Yes ☐ No ☐ Unclear ☐

Were measures objective or were patients & clinicians kept "blind" to treatments received?

Ideal if the study is 'double-blinded' –both patients and investigators are unaware of treatment allocation. If the outcome is *objective* then blinding is less critical.

Yes ☐ No ☐ Unclear ☐

Were other biases presents such as reporting bias or conflicts of interest, funding bias, etc.

Many biases exists that can undermine the validity of the results

Yes ☐ No ☐ Unclear ☐

Summarise the major methodological concerns and biases that affect validity of the outcome:

3: Based on the likely impact of the results and the threats to validity please provide a lay summary (with a bottom line) that you can communicate to patients and/or clinicians that will directly inform care.

Carl Heneghan. www.cebm.net 2017
University of Oxford,

Systematic reviews

What gives rise to bias when assessing evidence for treatment effects in SRs?

Steps in a Systematic Review

1 Is there a clear research Question? (Q)

2 Did they find all relevant studies? (F)

3 Did they Assess Study quality? (A)

4 Step 4: Did they Summarize the evidence? (S)

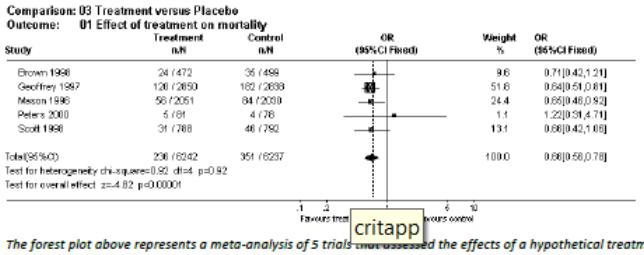
Systematic Reviews Summary title:

What question did the study ask?
PICO:

1: What difference do the results make?

- Does the outcome make a difference to my patient?
- Does this apply to my patient?
- Is the treatment feasible in my setting?

Meta-analysis is used to combine results from individual studies and an overall summary estimate is calculated. Individual results need to be expressed in a standard way, such as relative risk, odds ratio or mean difference between the groups. Results are displayed in a forest plot.



The forest plot above represents a meta-analysis of 5 trials that assessed the effects of a hypothetical treatment on mortality. Individual studies are represented by a black square (point estimate) and a horizontal line (95% CI) of the odds ratio. The size of the squares reflects the weight of the study. The solid vertical line corresponds to line of 'no effect' (odds ratio of 1.0). When the CI includes 1 it indicates the result is not significant at conventional levels (P>0.05).

The diamond at the bottom represents the combined or pooled odds ratio of all 5 trials with its 95% confidence interval. In this case, it shows that the treatment reduces mortality by 34% (OR 0.66 95% CI 0.56 to 0.78). Notice that the diamond does not overlap the 'no effect' line (the confidence interval doesn't include 1) so we can be assured that the pooled OR is statistically significant. The test for overall effect also indicates statistical significance (p<0.0001).

Exploring heterogeneity: Heterogeneity can be assessed using the "eyeball" test or with statistical tests, such as Cochran Q test. With the "eyeball" test look for overlap of the trial CIs with the summary estimate. In the example above the dotted line running vertically through the combined ORs crosses the horizontal lines of all the individual studies indicating homogenous studies.

If Cochran Q is statistically significant there is definite heterogeneity. If it is not but the ratio of Cochran Q and the degrees of freedom (Q/df) is > 1 there is possible heterogeneity. If Cochran Q is not statistically significant and Q/df is < 1 then heterogeneity is very unlikely. In the example above Q/df is <1 (0.92/4= 0.23) and the p-value is not significant (0.92) indicating no heterogeneity.

CEBM Rapid Critical appraisal

2: What is the impact of the biases on the results of the study?

Were relevant studies were missed?

What is best? The starting point is the major bibliographic databases (e.g., Medline, Cochrane, EMBASE, etc) but should include a search of reference lists from relevant studies, contact with experts to inquire about unpublished studies. The search should not be limited to English language only, and should include MESH terms and text words.

Yes ☐ No ☐ Unclear ☐

Were the criteria used to select articles for inclusion appropriate?

What is best? The inclusion or exclusion of studies should be clearly defined a priori, and should specify the patients, interventions or exposures and outcomes of interest. The study design will also be a key eligibility criteria.

Yes ☐ No ☐ Unclear ☐

Were the studies of sufficient quality to not threaten the validity of the results?

What is best? The article should describe how the quality of each study was assessed using predetermined quality criteria appropriate to the type of clinical question (e.g., randomization, blinding and completeness of follow-up)

Quality of the evidence:

Were the results similar from study to study?

Ideally, the results of the different studies should be homogeneous. If heterogeneity exists the authors may estimate whether the differences are significant (chi-square test). Possible reasons for heterogeneity should be explored.

Yes ☐ No ☐ Unclear ☐

Summarise the major methodological concerns and biases that affect validity of the outcome:

3: Based on the likely impact of the results and the threats to validity please provide a lay summary (with a bottom line) that you can communicate to patients and/or clinicians that will directly inform care.

Carl Heneghan. www.cebm.net 2017
University of Oxford,

Figure 2:

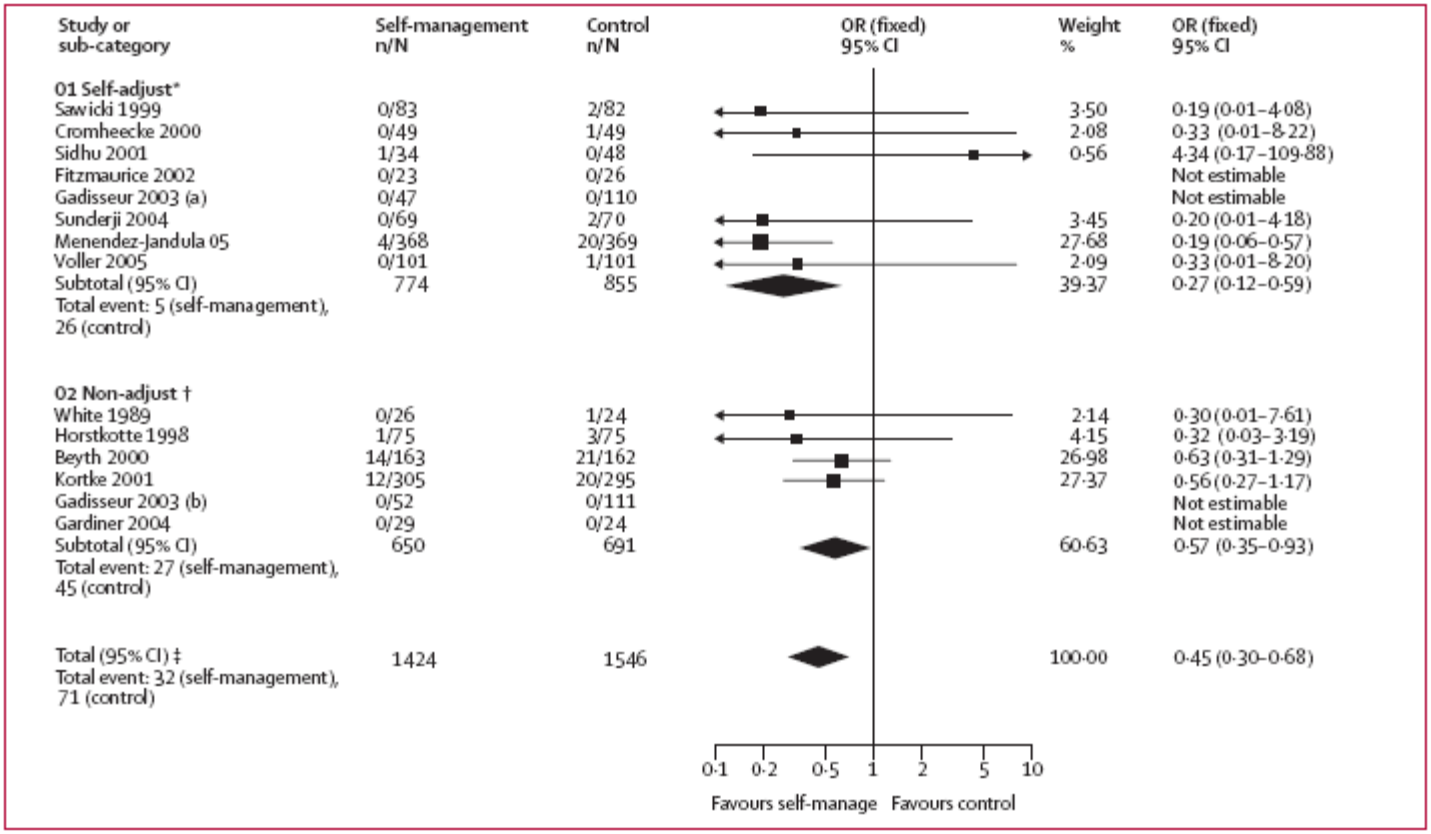


Figure 2: Self-monitoring and thromboembolic events from fixed-effects model

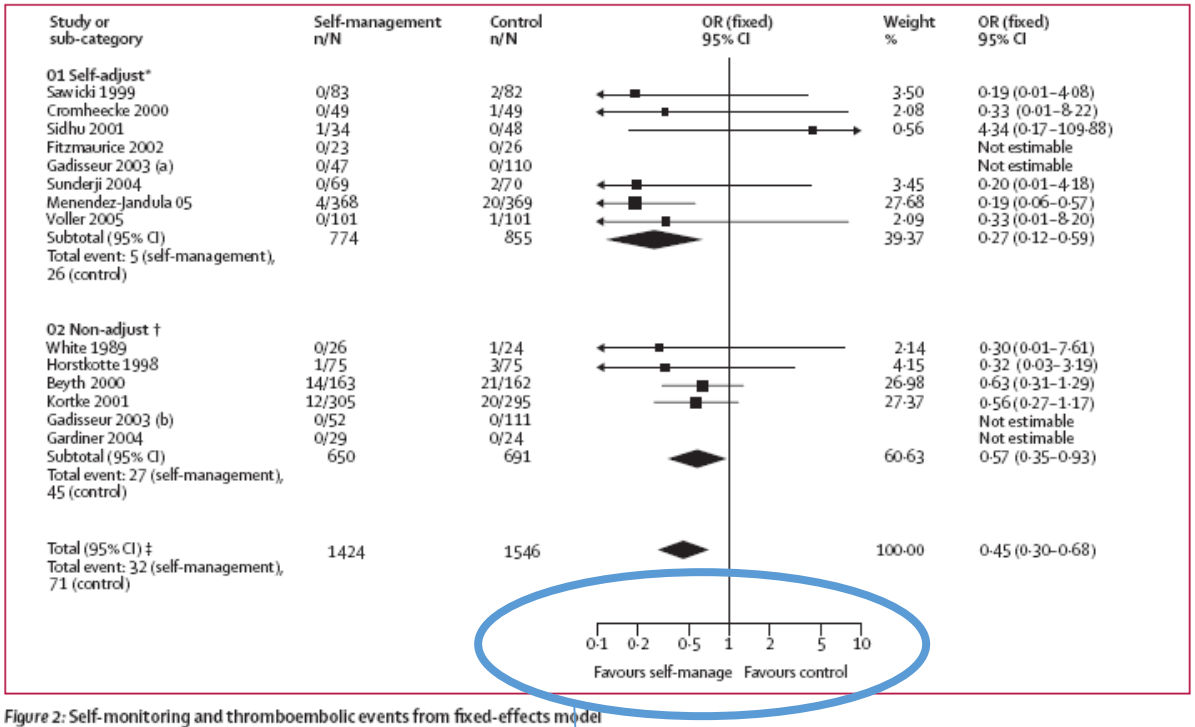
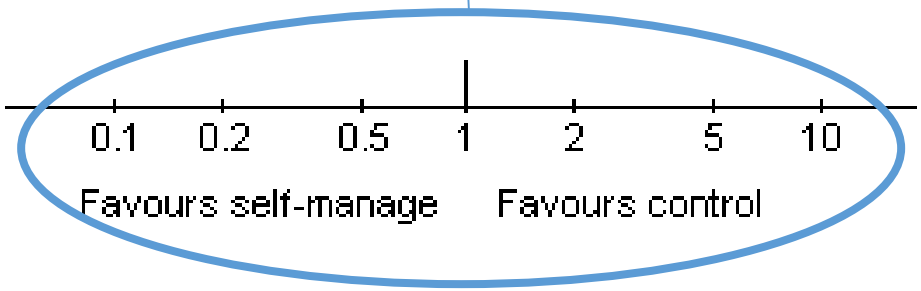
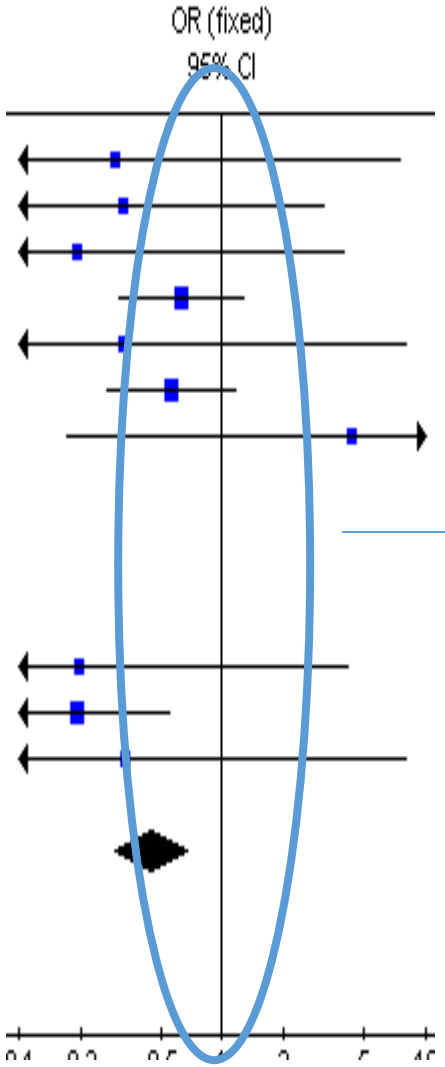


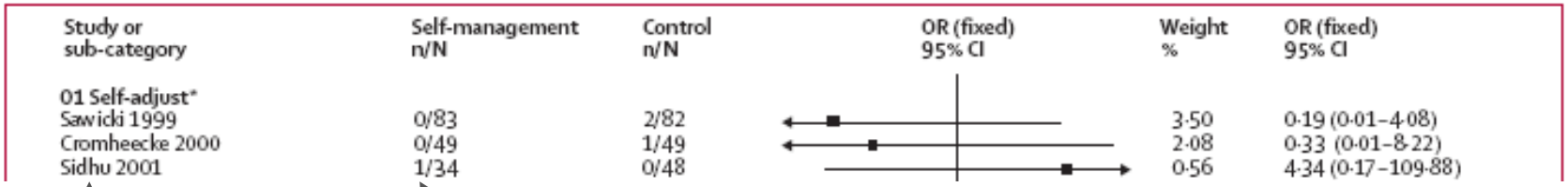
Figure 2: Self-monitoring and thromboembolic events from fixed-effects model



At the bottom there's a horizontal line. This is the scale measuring the treatment effect.



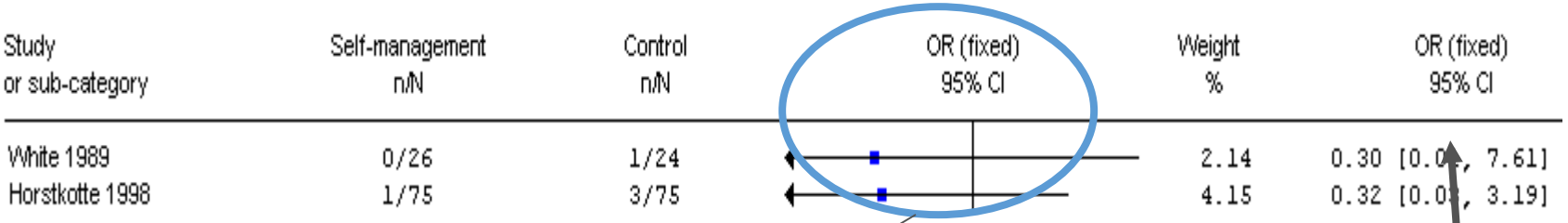
The vertical line in the middle is where the treatment and control have the same effect – there is no difference between the two



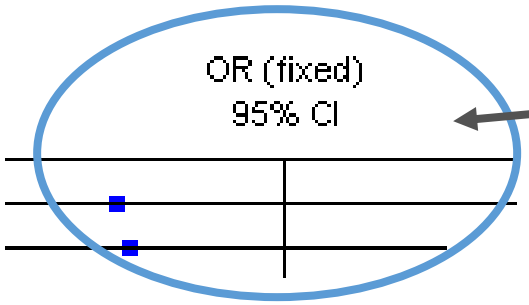
For each study there is an id

The data for each trial are here, divided into the experimental and control groups

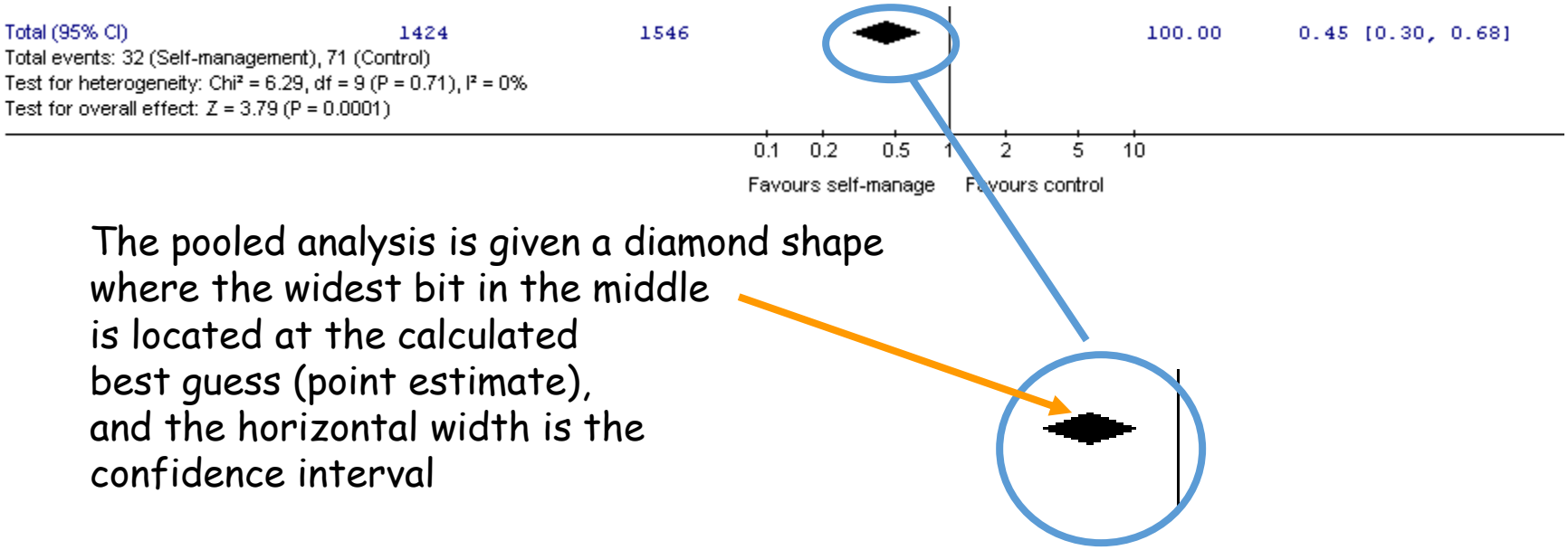
This is the % weight given to this study in the pooled analysis



The data shown in the graph are also given numerically



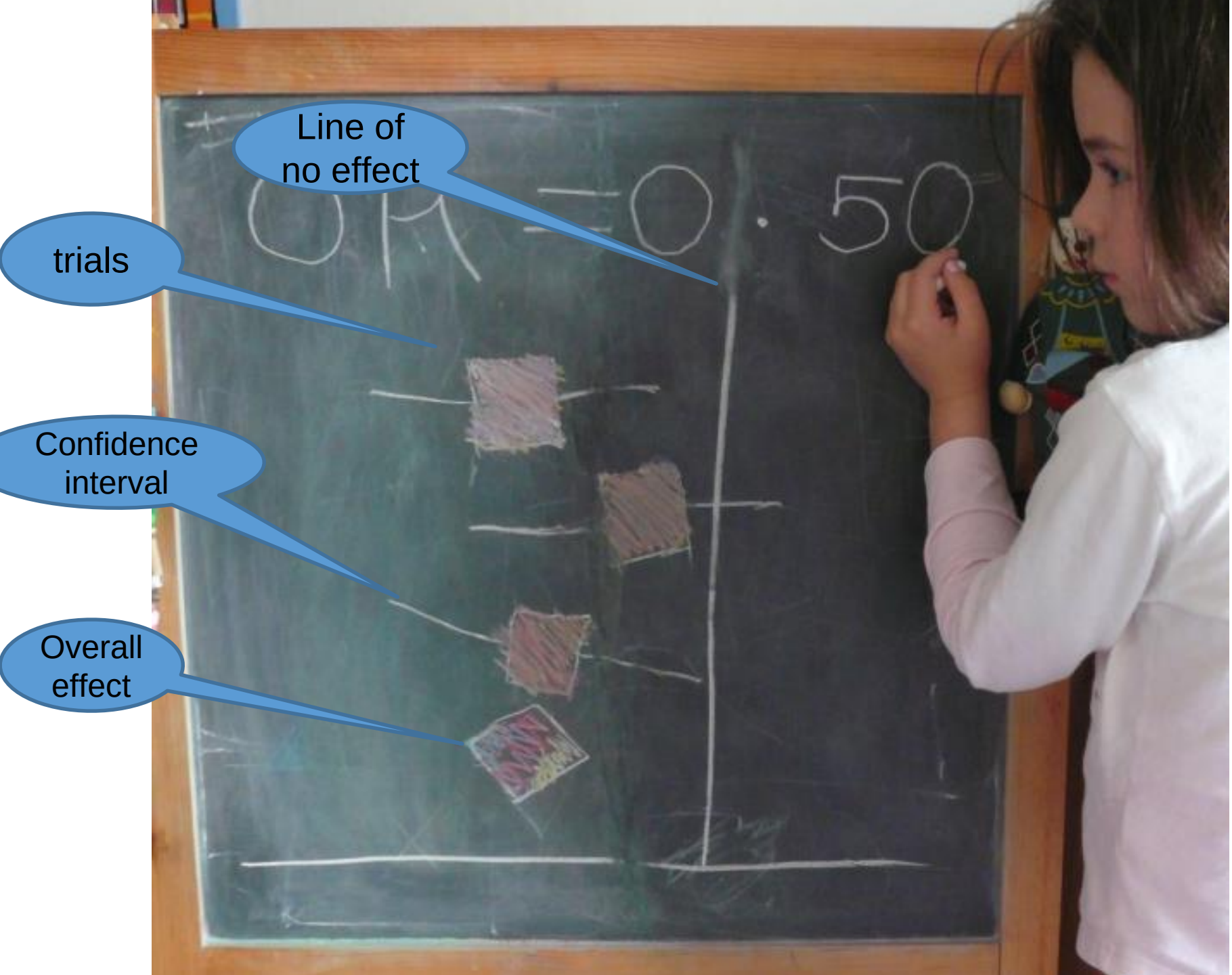
The label above the graph tells you what statistic has been used



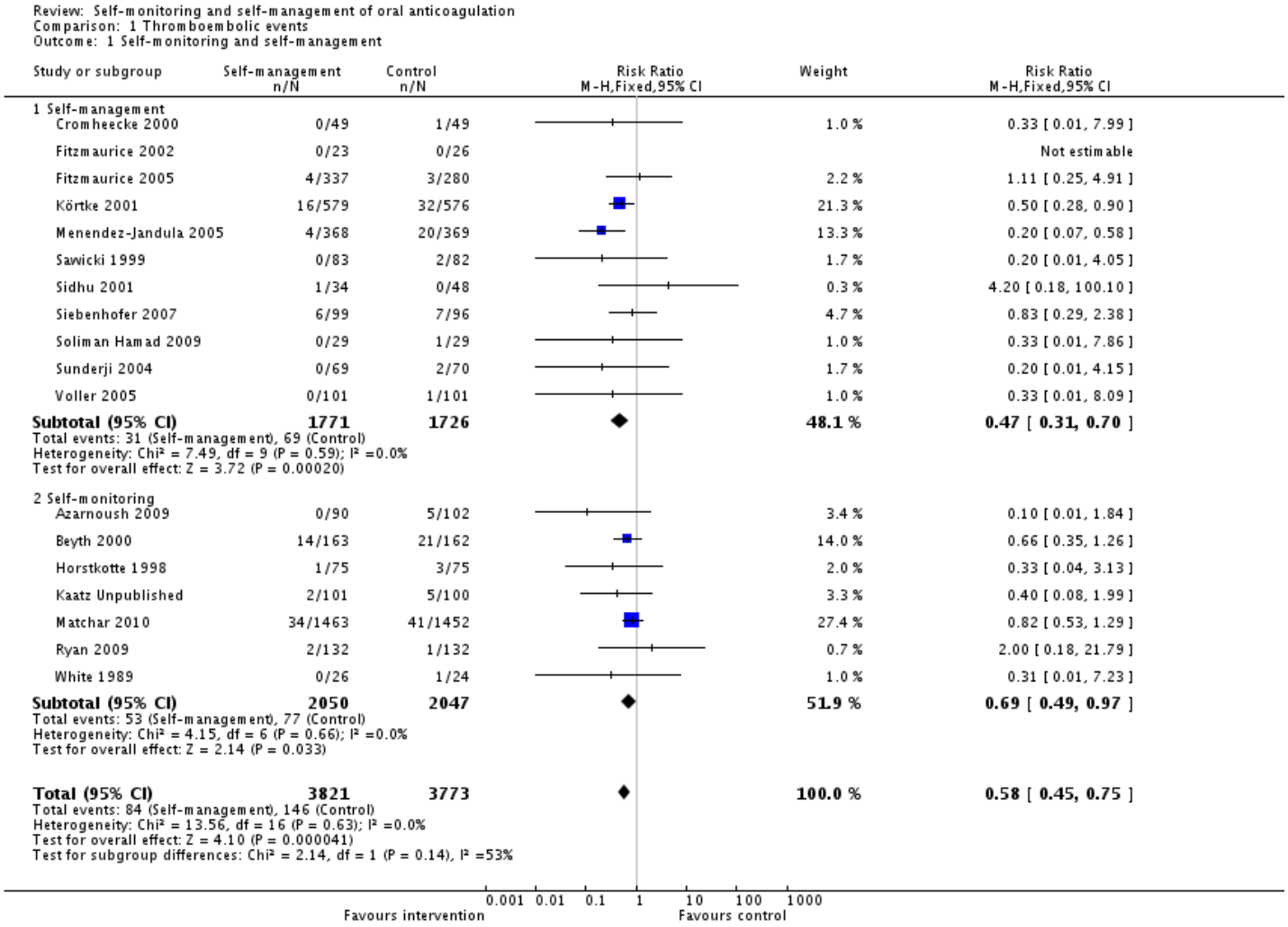
The pooled analysis is given a diamond shape where the widest bit in the middle is located at the calculated best guess (point estimate), and the horizontal width is the confidence interval

Note on interpretation

If the confidence interval crosses the line of no effect, this is equivalent to saying that we have found no statistically significant difference in the effects of the two interventions

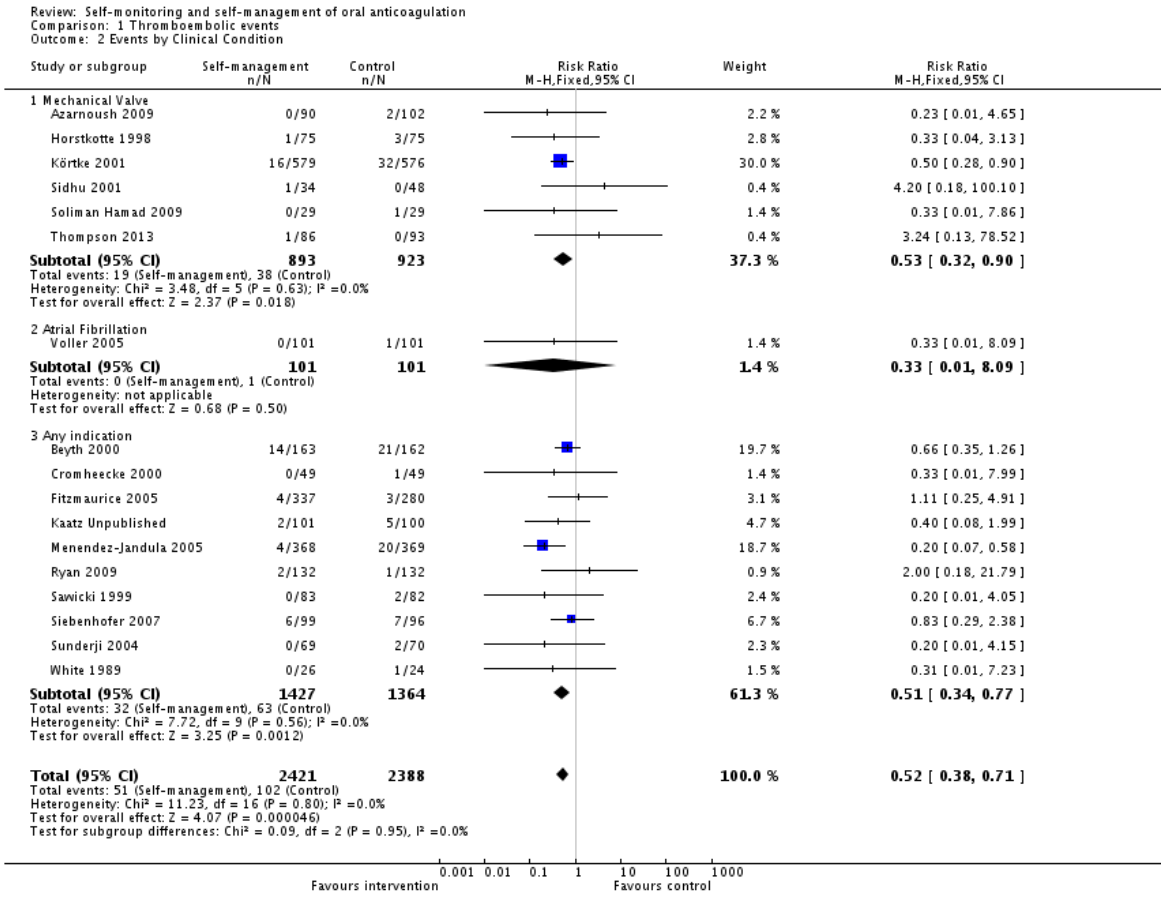

Line of no effect
trials
Confidence interval
Overall effect
OR=0.50

Analysis 1.1. What can we tell from this graph?



Cochrane Database of Systematic Reviews
5 JUL 2016 DOI: 10.1002/14651858.CD003839.pub3
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003839.pub3/full#CD003839-fig-00101

Analysis 1.2 What can we tell from this graph?



Cochrane Database of Systematic Reviews
5 JUL 2016 DOI: 10.1002/14651858.CD003839.pub3
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003839.pub3/full#CD003839-fig-00102

Thank You

@carlheneghan

CEBM

UNIVERSITY OF OXFORD