NUFFIELD DEPARTMENT OF
PRIMARY CARE
HEALTH SCIENCES

UNIVERSITY OF
OXFORD

Critical appraisal of randomised controlled trials

Dr Kamal R. Mahtani
BSc PhD MBBS PGDip MRCGP
GP and Clinical Lecturer
Centre for Evidence Based Medicine
University of Oxford

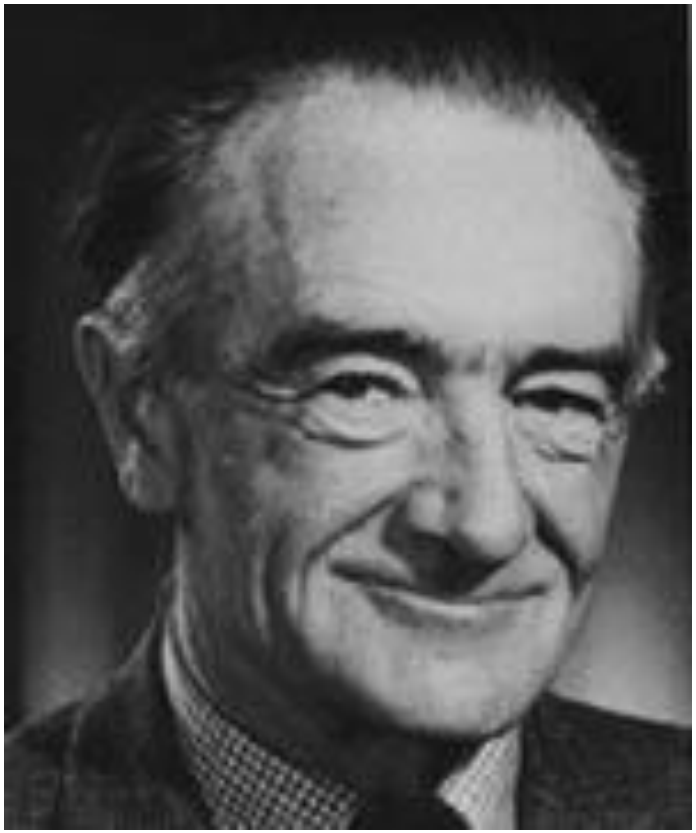
November 2014

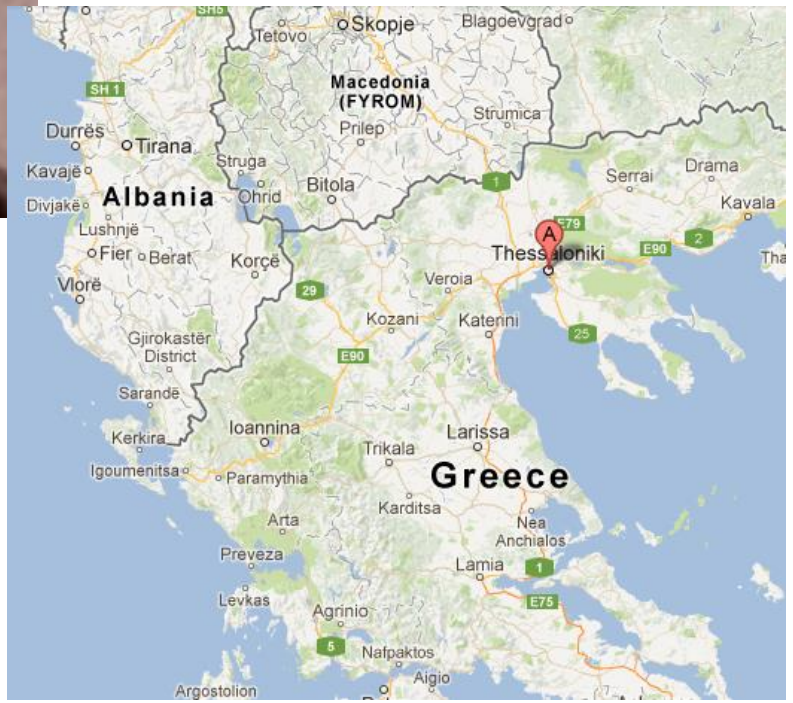
CEBM

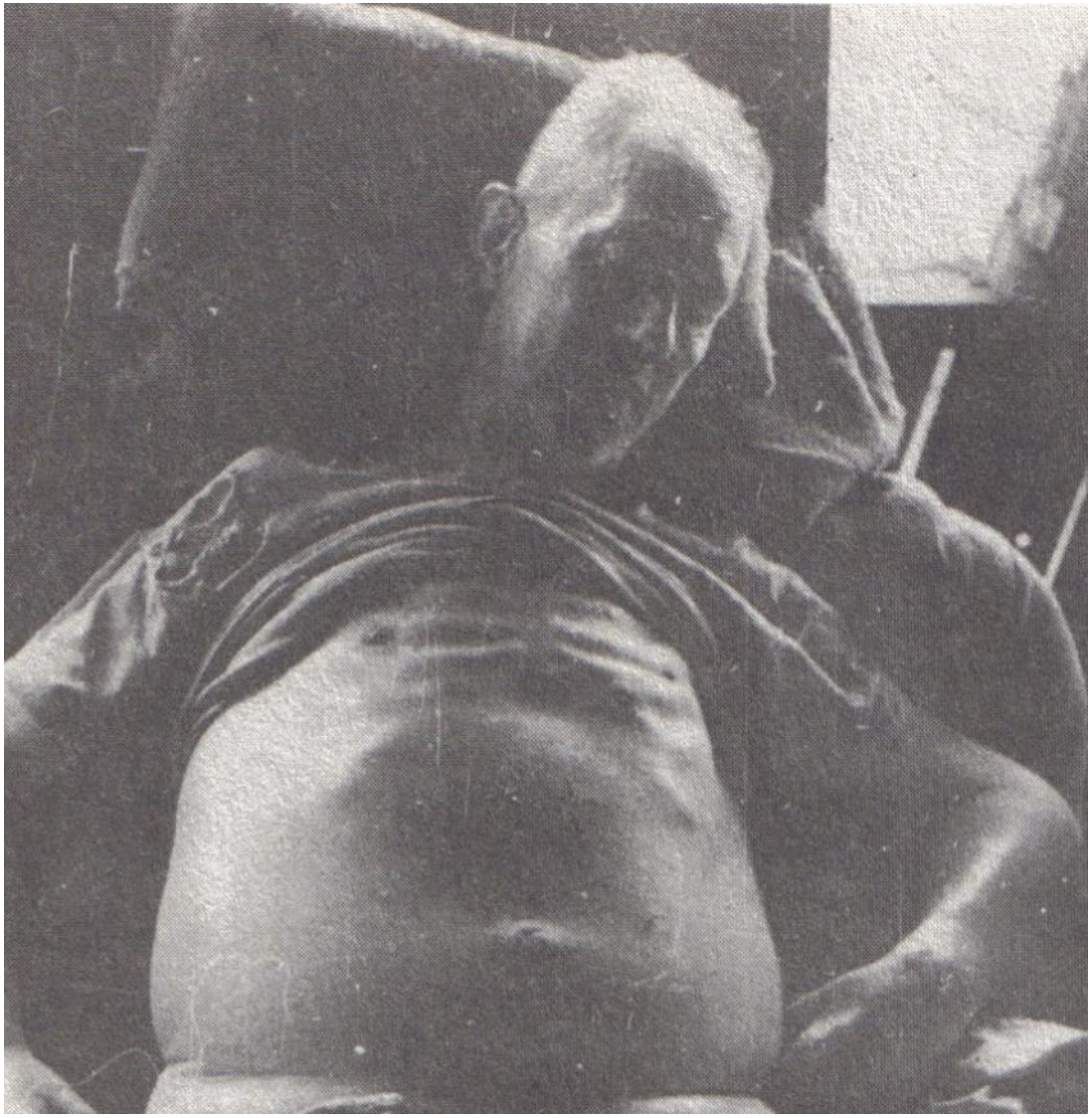
DEPARTMENT OF
PRIMARY CARE HEALTH SCIENCES

UNIVERSITY OF OXFORD

THE COCHRANE COLLABORATION®




Skopje
Tetovo
Blagoevgrad
Macedonia
(FYROM)
Strumica
Prilep
Durrës
Tirana
Kavajë
Struga
Albania
Ohrid
Bitola
Serrai
Drama
Kavala
Divjakë
Lushnjë
Fier
Berat
Korçë
Thessaloniki
Vlorë
Veroia
Kozani
Katerini
Gjirokastër
District
Sarandë
Kerkira
Ioannina
Larissa
Trikala
Igoumenitsa
Paramythia
Greece
Karditsa
Arta
Nea
Anchialos
Preveza
Lamia
Levkas
Agrinio
Nafpaktos
Aigio
Argostolion
SH 1
E79
E90
E75



Sickness in Salonica: my first, worst, and most successful clinical trial-1941.

". . . I recruited 20 young prisoners . . . I gave them a short talk about my medical hero James Lind and they agreed to co-operate in an experiment. I cleared two wards. I numbered the 20 prisoners off: odd numbers to one ward and evens to the other.

Each man in one ward received two spoonfuls of yeast daily. The others got one tablet of vitamin C from my "iron" reserve. The orderlies co-operated magnificently . . . They controlled fluid intake and measured frequency of urination.

. . . There was no difference between the wards for the first two days, but the third day was hopeful, and on the fourth the difference was conclusive . . . there was less oedema in the "yeast" ward. I made careful notes of the trial and immediately asked to see the Germans."

A. L. Cochrane (Br Med J 1984; 289: 1726-7)

“It could be argued that the trial was randomised and controlled, although this last was somewhat inadequate. In those early days, when the randomised controlled trial was little known in medicine, this was something of an achievement.”

BRITISH MEDICAL JOURNAL

LONDON SATURDAY OCTOBER 30 1948

STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS

A MEDICAL RESEARCH COUNCIL INVESTIGATION

The following gives the short-term results of a controlled investigation into the effects of streptomycin on one type of pulmonary tuberculosis. The inquiry was planned and directed by the Streptomycin in Tuberculosis Trials Committee, composed of the following members: Dr. Geoffrey Marshall (chairman), Professor J. W. S. Blacklock, Professor C. Cameron, Professor N. B. Capon, Dr. R. Cruickshank, Professor J. H. Gaddum, Dr. F. R. G. Heaf, Professor A. Bradford Hill, Dr. L. E. Houghton, Dr. J. Clifford Hoyle, Professor H. Raistrick, Dr. J. G. Scadding, Professor W. H. Tytler, Professor G. S. Wilson, and Dr. P. D'Arcy Hart (secretary). The centres at which the work was carried out and the specialists in charge of patients and pathological work were as follows:

Brompton Hospital, London.—Clinician: Dr. J. W. Crofton, Streptomycin Registrar (working under the direction of the honorary staff of Brompton Hospital); Pathologists: Dr. J. W. Clegg, Dr. D. A. Mitchison.

Colindale Hospital (L.C.C.), London.—Clinicians: Dr. J. V. Hurford, Dr. B. J. Douglas Smith, Dr. W. E. Snell; Pathologists (Central Public Health Laboratory): Dr. G. B. Forbes, Dr. H. D. Holt.

Harefield Hospital (M.C.C.), Harefield, Middlesex.—Clinicians: Dr. R. H. Brent, Dr. L. E. Houghton; Pathologist: Dr. E. Nassau.

Bangour Hospital, Bangour, West Lothian.—Clinician: Dr. I. D. Ross; Pathologist: Dr. Isabella Purdie.

Killingbeck Hospital and Sanatorium, Leeds.—Clinicians: Dr. W. Santon Gilmour, Dr. A. M. Reevie; Pathologist: Professor J. W. McLeod.

Northern Hospital (L.C.C.), Winchmore Hill, London.—Clinicians: Dr. F. A. Nash, Dr. R. Shoulman; Pathologists: Dr. J. M. Alston, Dr. A. Mohun.

Sully Hospital, Sully, Glam.—Clinicians: Dr. D. M. E. Thomas, Dr. L. R. West; Pathologist: Professor W. H. Tytler.

The clinicians of the centres met periodically as a working subcommittee under the chairmanship of Dr. Geoffrey Marshall; so also did the pathologists under the chairmanship of Dr. R. Cruickshank. Dr. Marc Daniels, of the Council's scientific staff, was responsible for the clinical co-ordination of the trials, and he also prepared the report for the Committee, with assistance from Dr. D. A. Mitchison on the analysis of laboratory results. For the purpose of final analysis the radiological findings were assessed by a panel composed of Dr. L. G. Blair, Dr. Peter Kerley, and Dr. Geoffrey S. Todd.

Introduction

When a special committee of the Medical Research Council undertook in September, 1946, to plan clinical trials of streptomycin in tuberculosis the main problem faced was that of investigating the effect of the drug in pulmonary tuberculosis. This antibiotic had been discovered two years previously by Waksman (Schatz, Bugie, and Waksman, 1944); in the intervening period its power of inhibiting tubercle bacilli in vitro and the results of treatment in

if based on adequately controlled clinical trials (Hinshaw and Feldman, 1944). The one controlled trial of gold treatment (and the only report of an adequately controlled trial in tuberculosis we have been able to find in the literature) reported negative therapeutic results (Amberson, McMahon, and Pinner, 1931). In 1946 no controlled trial of streptomycin in pulmonary tuberculosis had been undertaken in the U.S.A. The Committee of the Medical Research Council decided then that a part of the small supply of streptomycin

What's so special about RCTs?

- most rigorous way of determining:
  - a cause-effect relation exists between treatment and outcome and
  - for assessing the cost effectiveness of a treatment
- distributing the characteristics of patients that may influence the outcome randomly between the groups-no systematic differences between intervention groups

What's so special about RCTs?

- patients and trialists should remain unaware of which treatment was given until the study is completed to avoid influencing the result
- both arms treated identically except for the intervention of interest – estimating the size of the difference in predefined outcomes between intervention groups

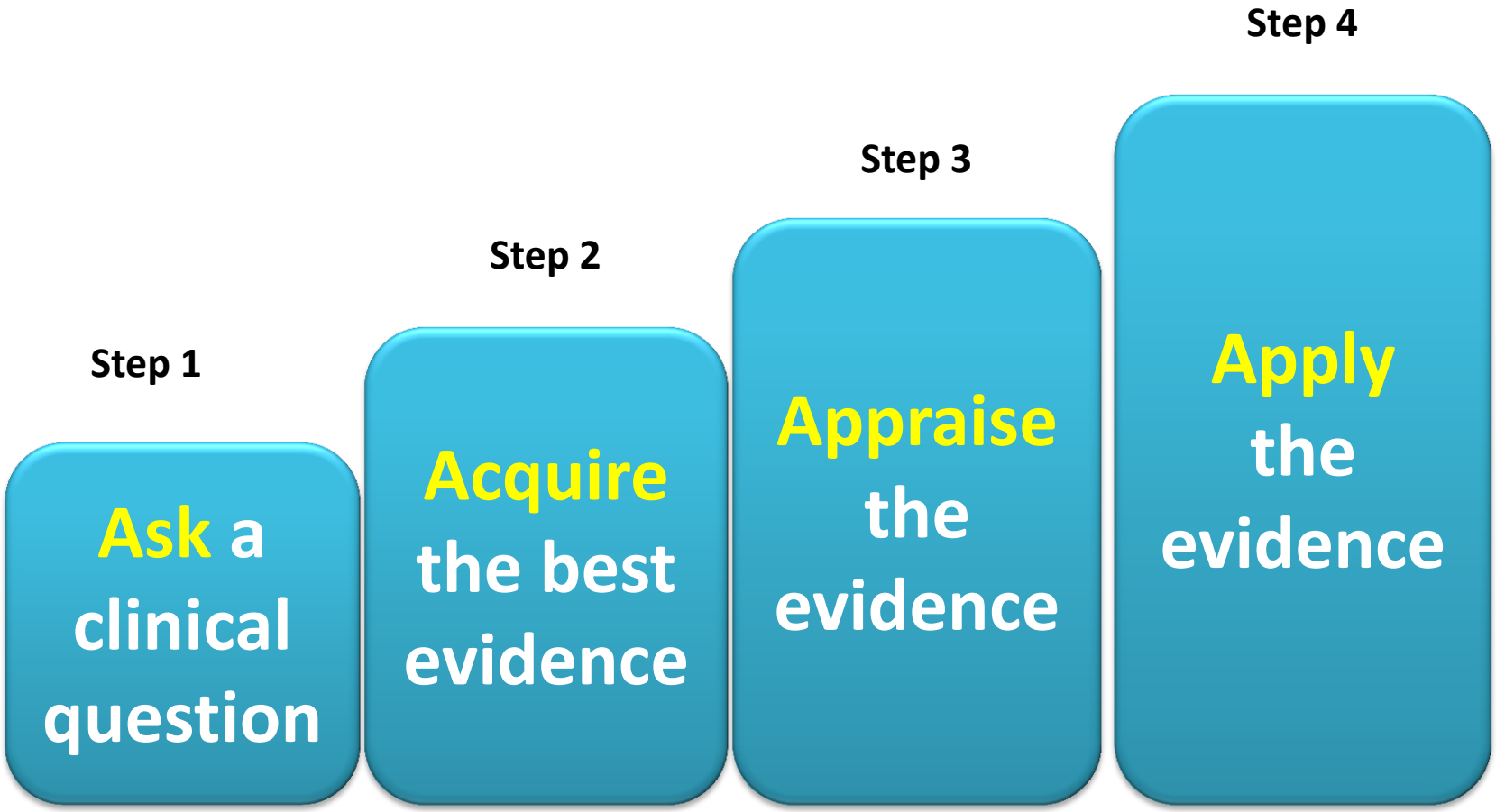
So are RCTs the gold standard for evidence?

…..depends

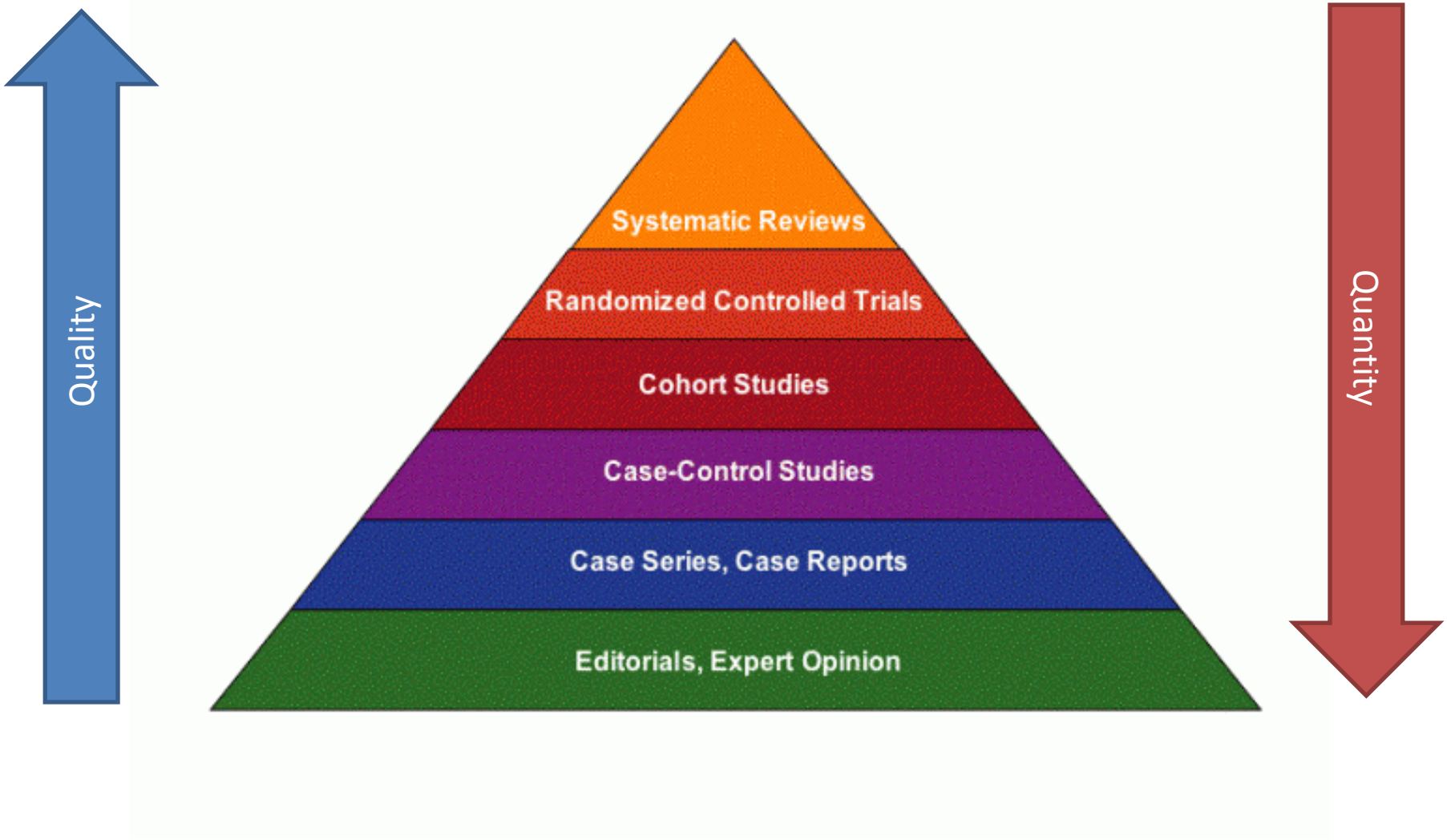
Limitations of RCTs

- Excellent vs Poor RCTs – quality varies
  - Impact on interpretation of result (external validity)?
- Expensive and time consuming
  - £250k - £millions over 2-5 years+
- May not always be the right study design to answer that question

Practicing EBM – the 4 A's

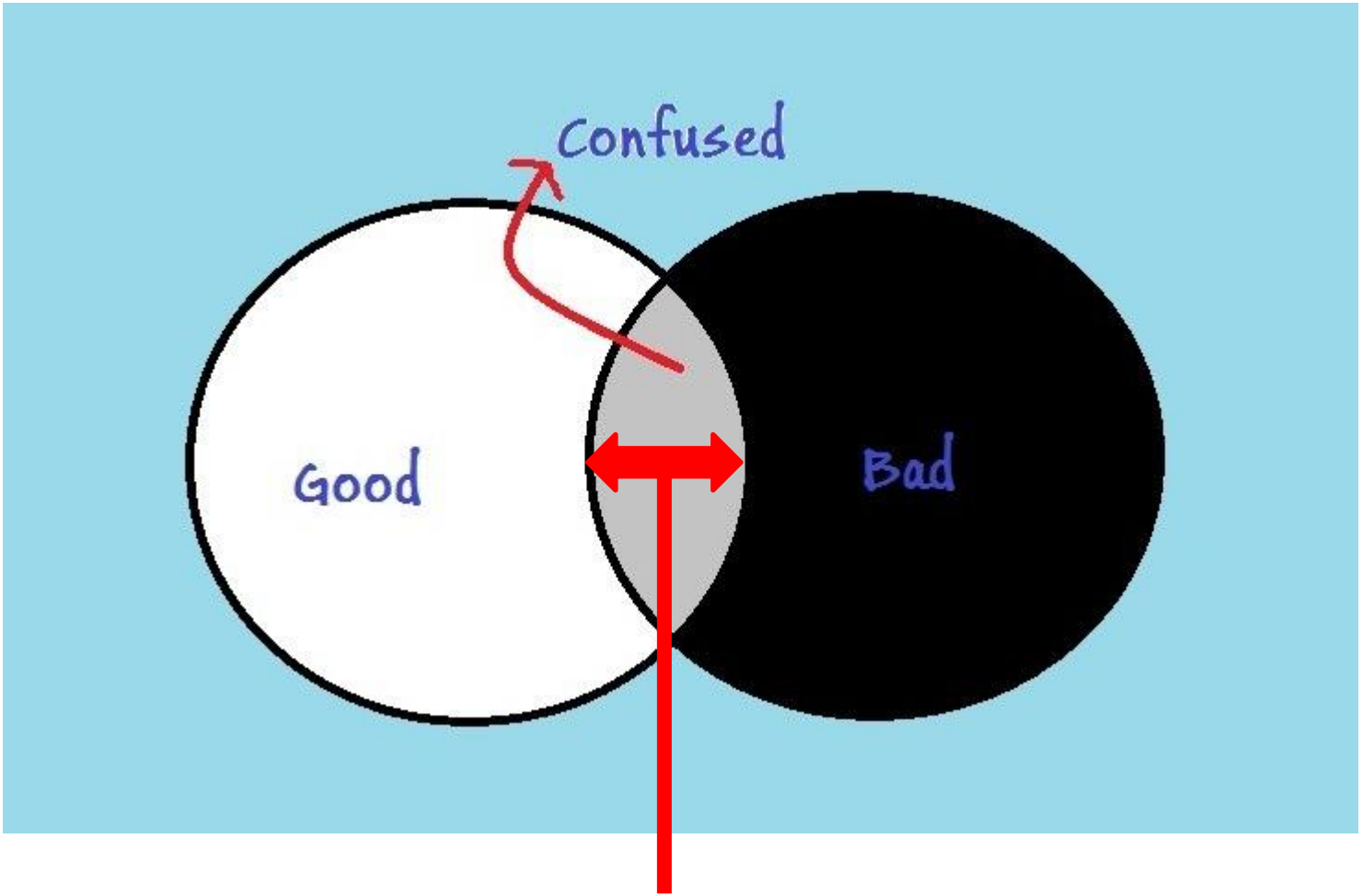


Levels of evidence



Practicing EBM – the 4 A's

Types of evidence



Risk of Bias

The degree to which the result is skewed away from the truth

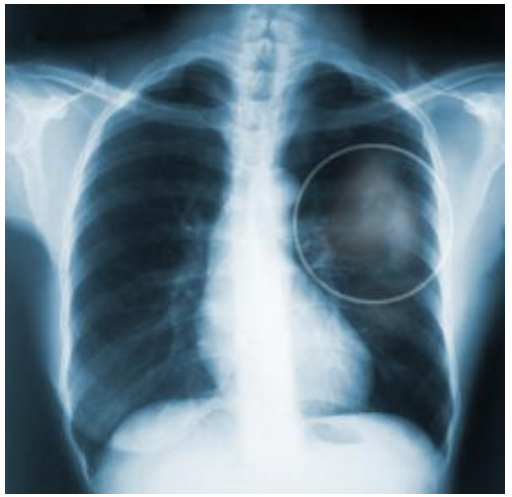
Internal validity

- extent to which observed treatment effects can be ascribed to differences in treatment and not confounding, thereby allowing the inference of causality to be ascribed to a treatment.[1]
- Systematic error (bias) could threaten the internal validity of trials, and all efforts should be made to minimise these in the design, conduct, and analysis of studies.[2]

1. http://www.bmj.com/content/344/bmj.e1004
2. http://www.ncbi.nlm.nih.gov/pubmed/18728521

Confounding factors

- Other patient features/causal factors, apart from the one being measured, that can affect the outcome of the study e.g..



External validity

- The degree to which the results of the study can be applied to other populations

Assessing risk of bias for an RCT

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

What question did the study ask?

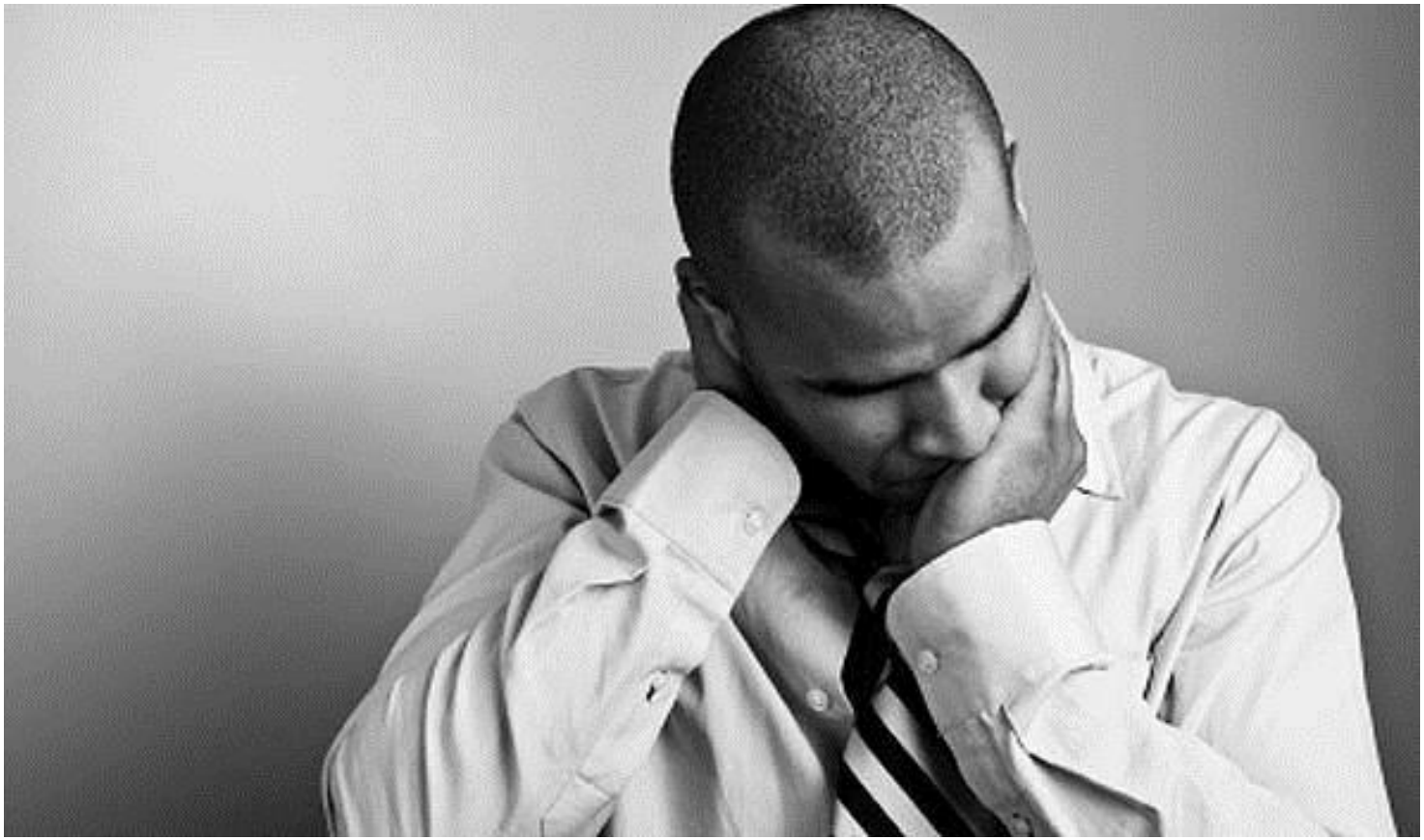
Patients –

Intervention -

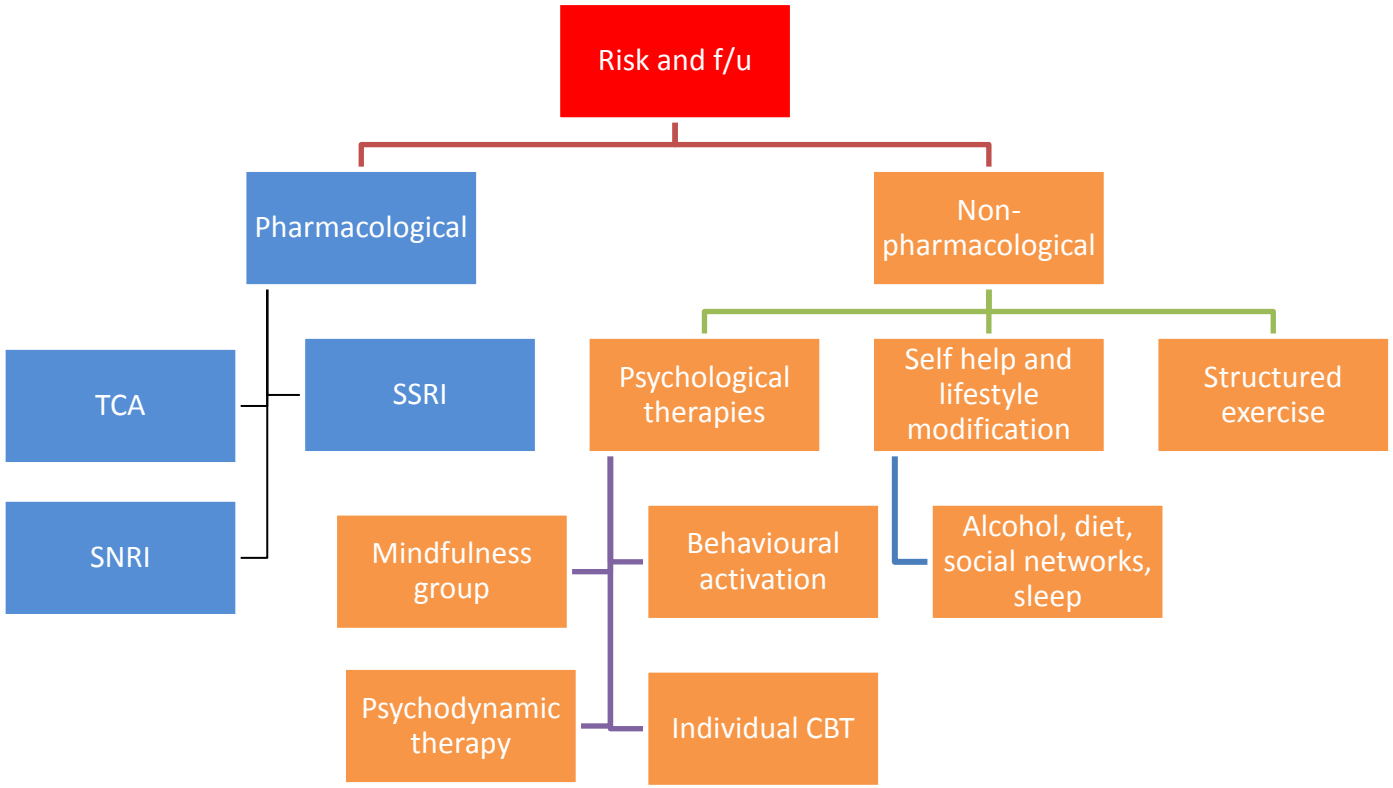
Comparison -

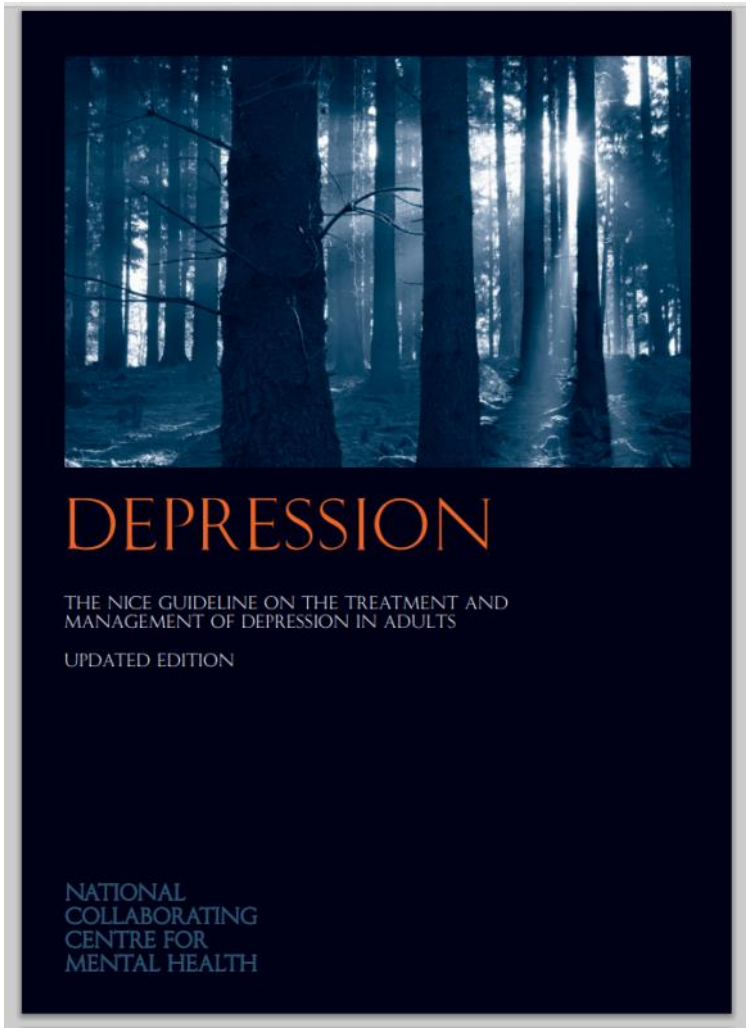
Outcome(s) -

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	The ***Results*** section should say how many patients were



Depression Management





NHS
National Institute for Health and Clinical Excellence

RECOGNISED DEPRESSION – PERSISTENT SUBTHRESHOLD DEPRESSIVE SYMPTOMS OR MILD TO MODERATE DEPRESSION

- taking regular physical exercise

Exercise does little to help the symptoms of depression, new study finds

By SUZANNAH HILLS

PUBLISHED: 08:23, 6 June 2012 | **UPDATED:** 11:01, 8 June 2012

Exercise does little to help alleviate the symptoms of depression, a new study has found.

The findings contrast with current clinical guidance which recommends exercise to help those suffering from the mental illness that affects one in six adults in Britain at any one time.

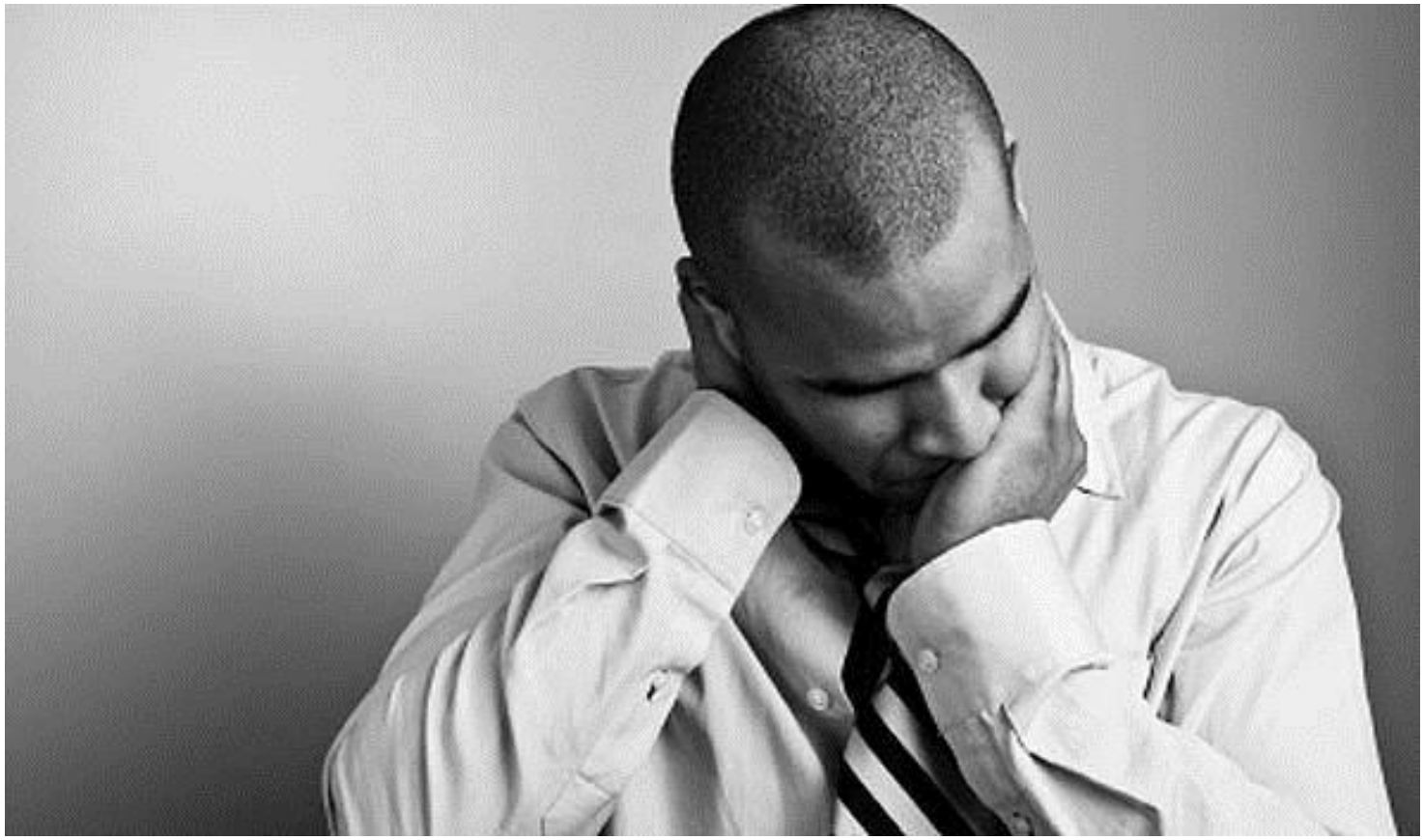
But research published in the British Medical Journal suggests that doing a physical activity combined with usual treatment did not reduce symptoms of depression more than the treatment alone.

Affecting millions: One in six adults in Britain suffer from depression at any one time

Exercise doesn't help depression, study concludes

Patients advised to get exercise fare no better than those who receive only standard care, researchers argue

Press Association
guardian.co.uk, Wednesday 6 June 2012 06.28 BST



PICO

BMJ

BMJ 2012;344:e2758 doi: 10.1136/bmj.e2758 (Published 6 June 2012)

RESEARCH

Facilitated physical activity as a treatment for depressed adults: randomised controlled trial

Melanie Chalder *research fellow*[1], Nicola J Wiles *senior lecturer*[1], John Campbell *professor*[2], Sandra P Hollinghurst *senior lecturer*[1], Anne M Haase *senior lecturer*[3], Adrian H Taylor *professor*[4], Kenneth R Fox *professor*[3], Ceire Costelloe *research associate*[1], Aidan Searle *research associate*[1], Helen Baxter *research associate*[1], Rachel Winder *associate research fellow*[2], Christine Wright *associate research fellow*[2], Katrina M Turner *lecturer*[1], Michael Calnan *professor*[5], Deborah A Lawlor *professor*[1], Tim J Peters *professor*[6], Deborah J Sharp *professor*[1], Alan A Montgomery *reader*[1], Glyn Lewis *professor*[1]

[1]School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK; [2]Primary Care Research Group, Peninsula Medical School, Exeter, UK; [3]School of Policy Studies, University of Bristol; [4]Sport and Health Sciences, University of Exeter, Exeter; [5]School of Social Policy, University of Kent, Canterbury, UK; [6]School of Clinical Sciences, University of Bristol

Abstract

Objective To investigate the effectiveness of facilitated physical activity as an adjunctive treatment for adults with depression presenting in primary care.

Design Pragmatic, multicentre, two arm parallel randomised controlled trial.

Setting General practices in Bristol and Exeter.

Participants 361 adults aged 18-69 who had recently consulted their general practitioner with symptoms of depression. All those randomised had a diagnosis of an episode of depression as assessed by the clinical interview schedule-revised and a Beck depression inventory score of 14 or more.

Interventions In addition to usual care, intervention participants were offered up to three face to face sessions and 10 telephone calls with a trained physical activity facilitator over eight months. The intervention was based on theory and aimed to provide individually tailored support and encouragement to engage in physical activity.

Main outcome measures The primary outcome was self reported symptoms of depression, assessed with the Beck depression inventory at four months post-randomisation. Secondary outcomes included use of antidepressants and physical activity at the four, eight, and 12 month

−0.54 (95% confidence interval −3.06 to 1.99; P=0.68). Similarly, there was no evidence that the intervention group reported a change in mood by the eight and 12 month follow-up points. Nor was there evidence that the intervention reduced antidepressant use compared with usual care (adjusted odds ratio 0.63, 95% confidence interval 0.19 to 2.06; P=0.44) over the duration of the trial. However, participants allocated to the intervention group reported more physical activity during the follow-up period than those allocated to the usual care group (adjusted odds ratio 2.27, 95% confidence interval 1.32 to 3.89; P=0.003).

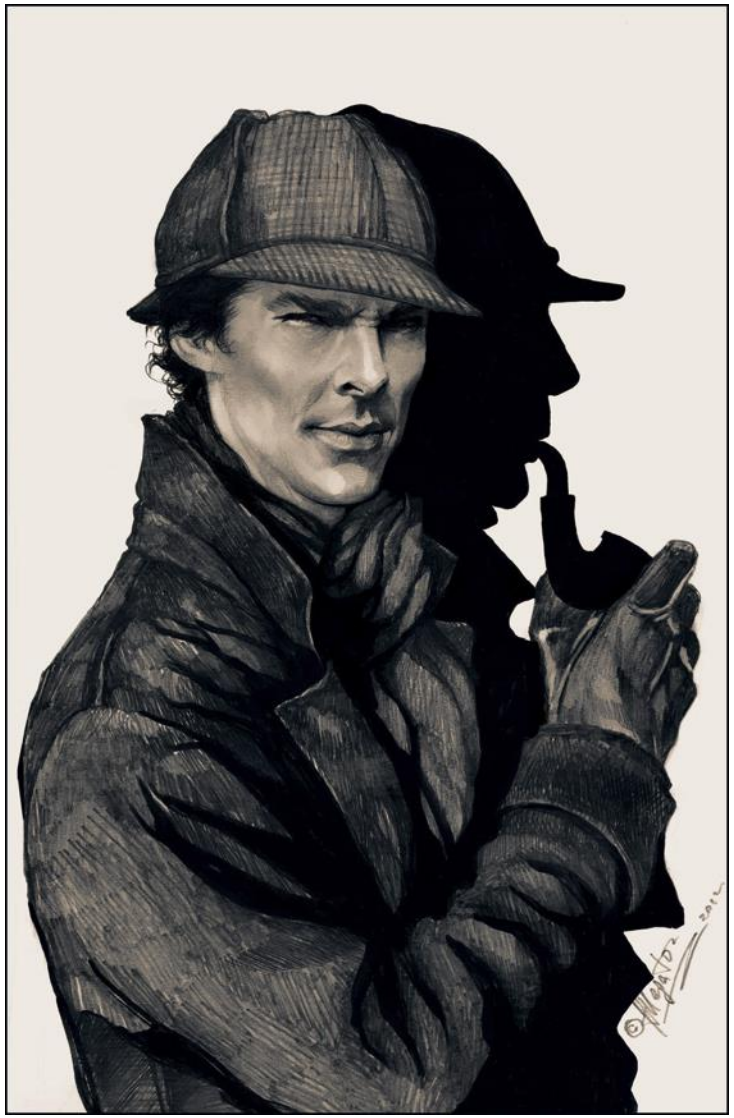
Conclusions The addition of a facilitated physical activity intervention to usual care did not improve depression outcome or reduce use of antidepressants compared with usual care alone.

Trial registration Current Controlled Trials ISRCTN16900744.

Introduction

Depression is one of the most common reasons for consulting a general practitioner within the United Kingdom, and its associated economic burden is considerable.[1] Although antidepressants are effective, many patients and healthcare professionals would like other options to be available as an alternative or adjunct to drug therapy.[2] Some evidence[3] shows

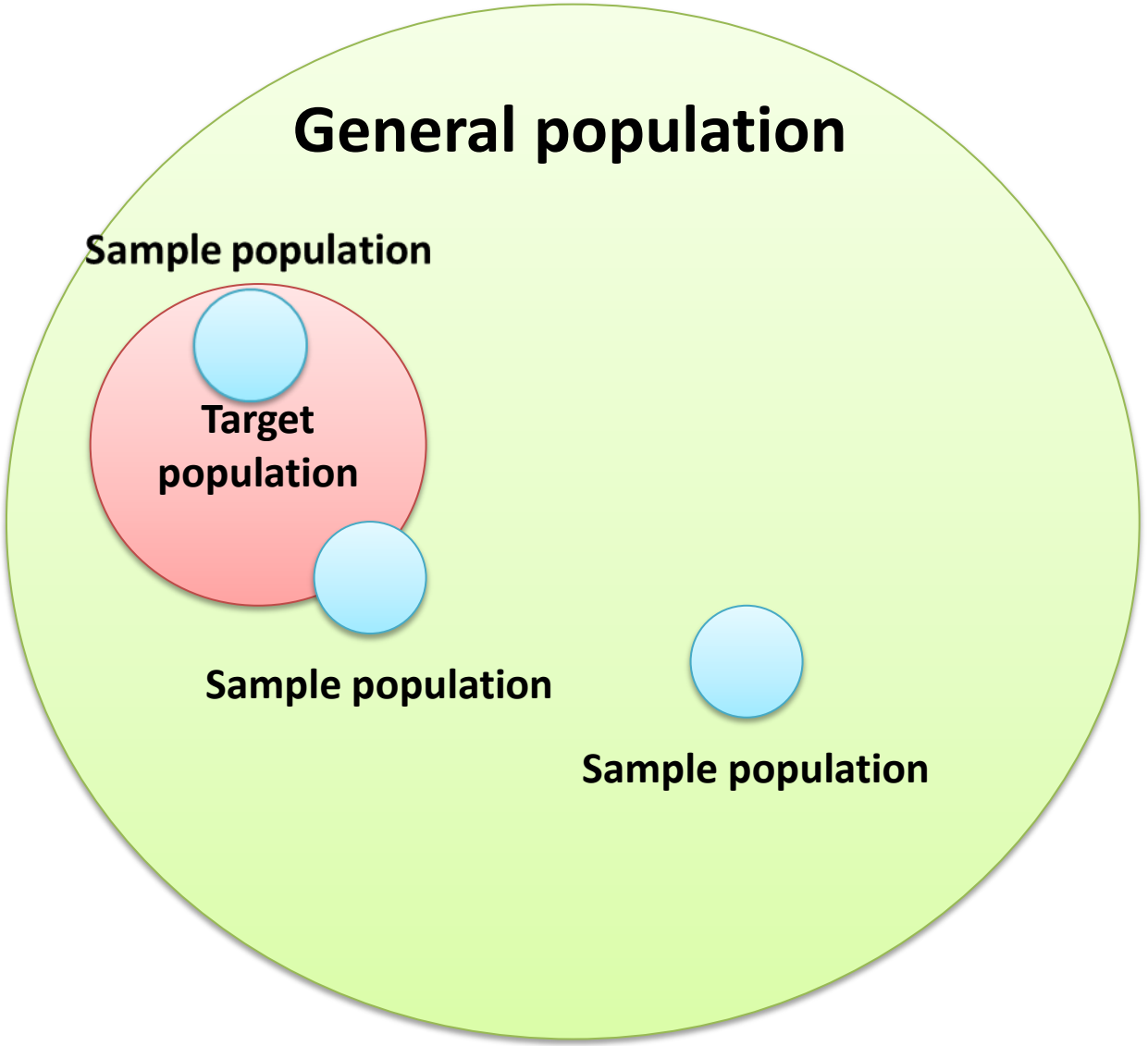
Critical appraisal….



…is like being a detective.

You need the skills to think broadly and detect the flaws that might distract you from finding the true answer.

Recruitment (selection bias)

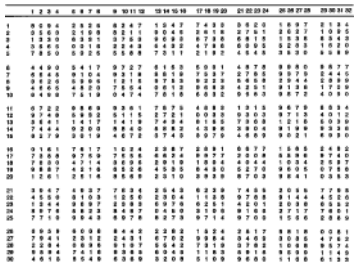


Recruitment (selection bias)

- Were the subjects representative of the target population?
  - What were the inclusion & exclusion criteria?
  - Were they appropriate?
  - How/where were they recruited from?

- Methods Recruitment of participants and baseline assessment & Results 1st para

Randomisation (selection bias)



Allocation concealment

How was the randomised sequence implemented?

BEST – most valid technique

- Central computer randomization

DOUBTFUL

- Envelopes, etc

Allocation (selection bias)

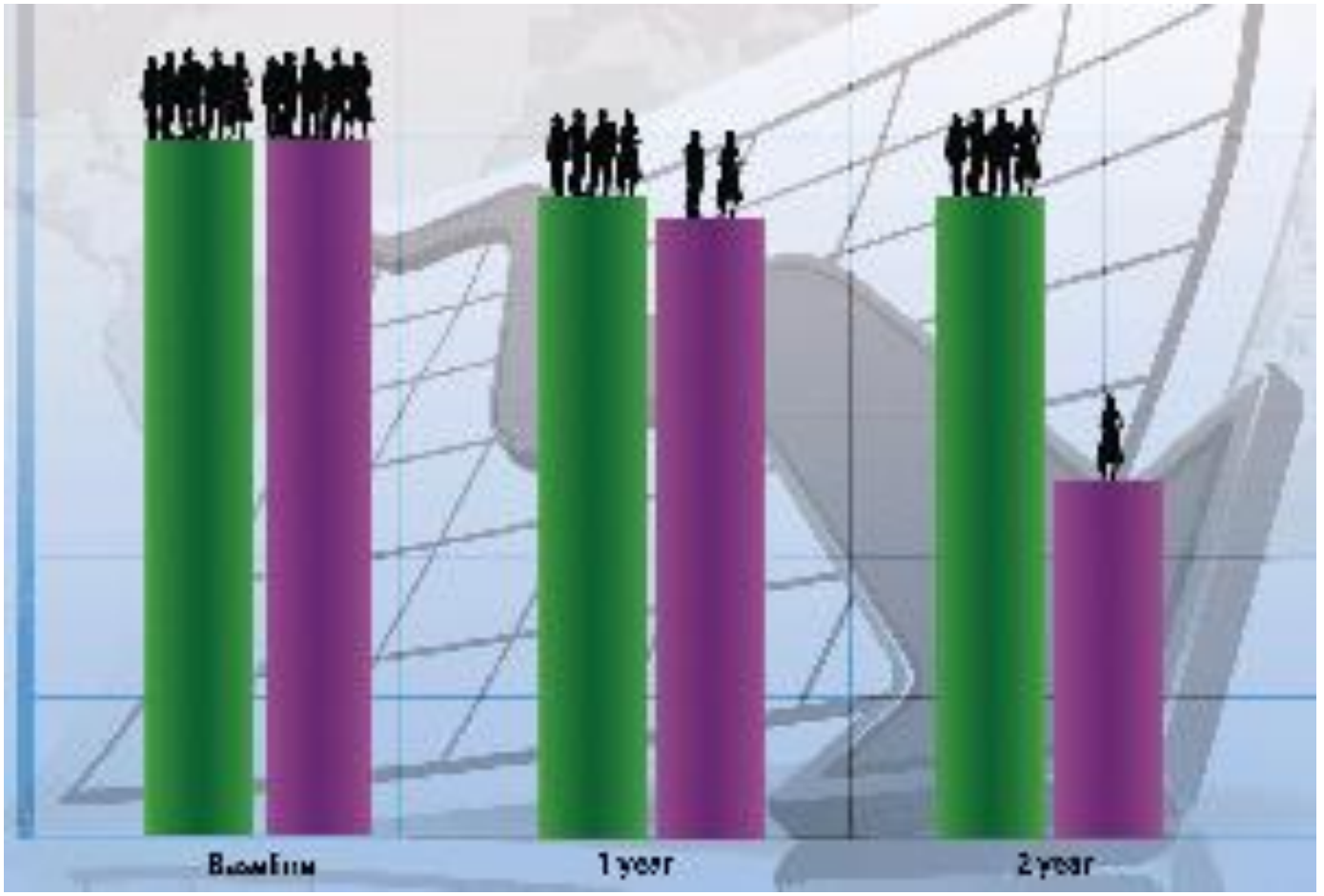
- Were the groups comparable at the start?
  - "Table 1"
- Randomised appropriately?
- Allocation to group concealed beforehand?

- Methods: Randomisation, concealment, and blinding and "Table 1"

Maintenance

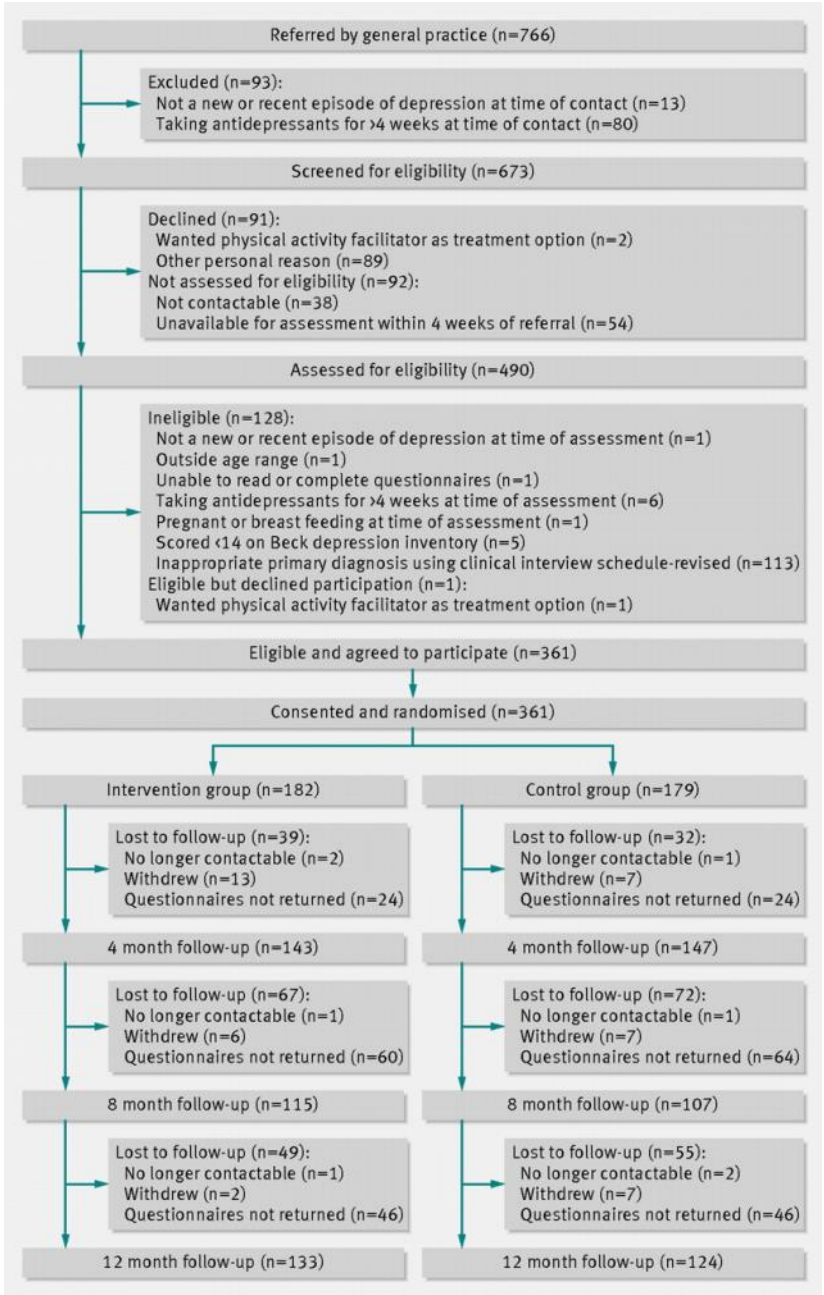
- Were both groups comparable throughout the study?
  - Managed equally bar the intervention?
    - What was the intervention?
    - What was the comparator?
- Methods: Follow up and Intervention and comparator (usual care)

Adequate follow up? (Attrition bias)

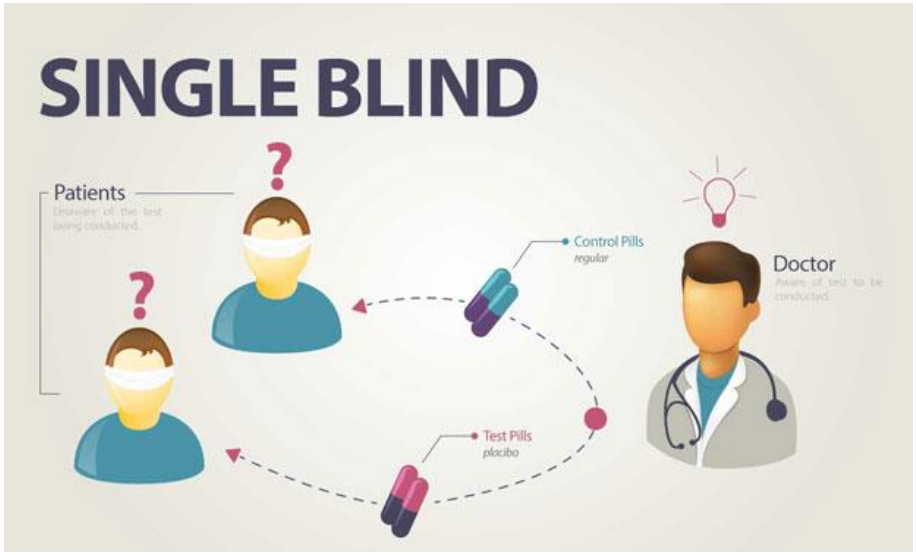


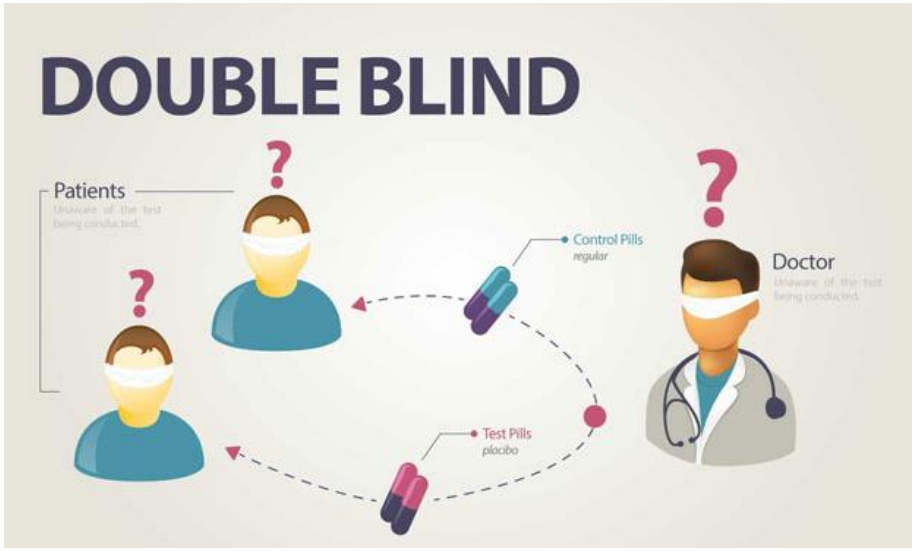
Adequate follow up? (Attrition bias)

- How many people were lost to f/u?
- Why were they lost to f/u?
- Did the researchers use an intention to treat (ITT) principle?
  - Once a participant is randomised, they should be analysed to the group they were assigned to
- Figure 1 and Statistical analysis


Referred by general practice (n=766)
Excluded (n=93):
Not a new or recent episode of depression at time of contact (n=13)
Taking antidepressants for >4 weeks at time of contact (n=80)
Screened for eligibility (n=673)
Declined (n=91):
Wanted physical activity facilitator as treatment option (n=2)
Other personal reason (n=89)
Not assessed for eligibility (n=92):
Not contactable (n=38)
Unavailable for assessment within 4 weeks of referral (n=54)
Assessed for eligibility (n=490)
Ineligible (n=128):
Not a new or recent episode of depression at time of assessment (n=1)
Outside age range (n=1)
Unable to read or complete questionnaires (n=1)
Taking antidepressants for >4 weeks at time of assessment (n=6)
Pregnant or breast feeding at time of assessment (n=1)
Scored <14 on Beck depression inventory (n=5)
Inappropriate primary diagnosis using clinical interview schedule-revised (n=113)
Eligible but declined participation (n=1):
Wanted physical activity facilitator as treatment option (n=1)
Eligible and agreed to participate (n=361)
Consented and randomised (n=361)
Intervention group (n=182)
Control group (n=179)
Lost to follow-up (n=39):
No longer contactable (n=2)
Withdrew (n=13)
Questionnaires not returned (n=24)
Lost to follow-up (n=32):
No longer contactable (n=1)
Withdrew (n=7)
Questionnaires not returned (n=24)
4 month follow-up (n=143)
4 month follow-up (n=147)
Lost to follow-up (n=67):
No longer contactable (n=1)
Withdrew (n=6)
Questionnaires not returned (n=60)
Lost to follow-up (n=72):
No longer contactable (n=1)
Withdrew (n=7)
Questionnaires not returned (n=64)
8 month follow-up (n=115)
8 month follow-up (n=107)
Lost to follow-up (n=49):
No longer contactable (n=1)
Withdrew (n=2)
Questionnaires not returned (n=46)
Lost to follow-up (n=55):
No longer contactable (n=2)
Withdrew (n=7)
Questionnaires not returned (n=46)
12 month follow-up (n=133)
12 month follow-up (n=124)

Measurement – blinding (Performance bias)





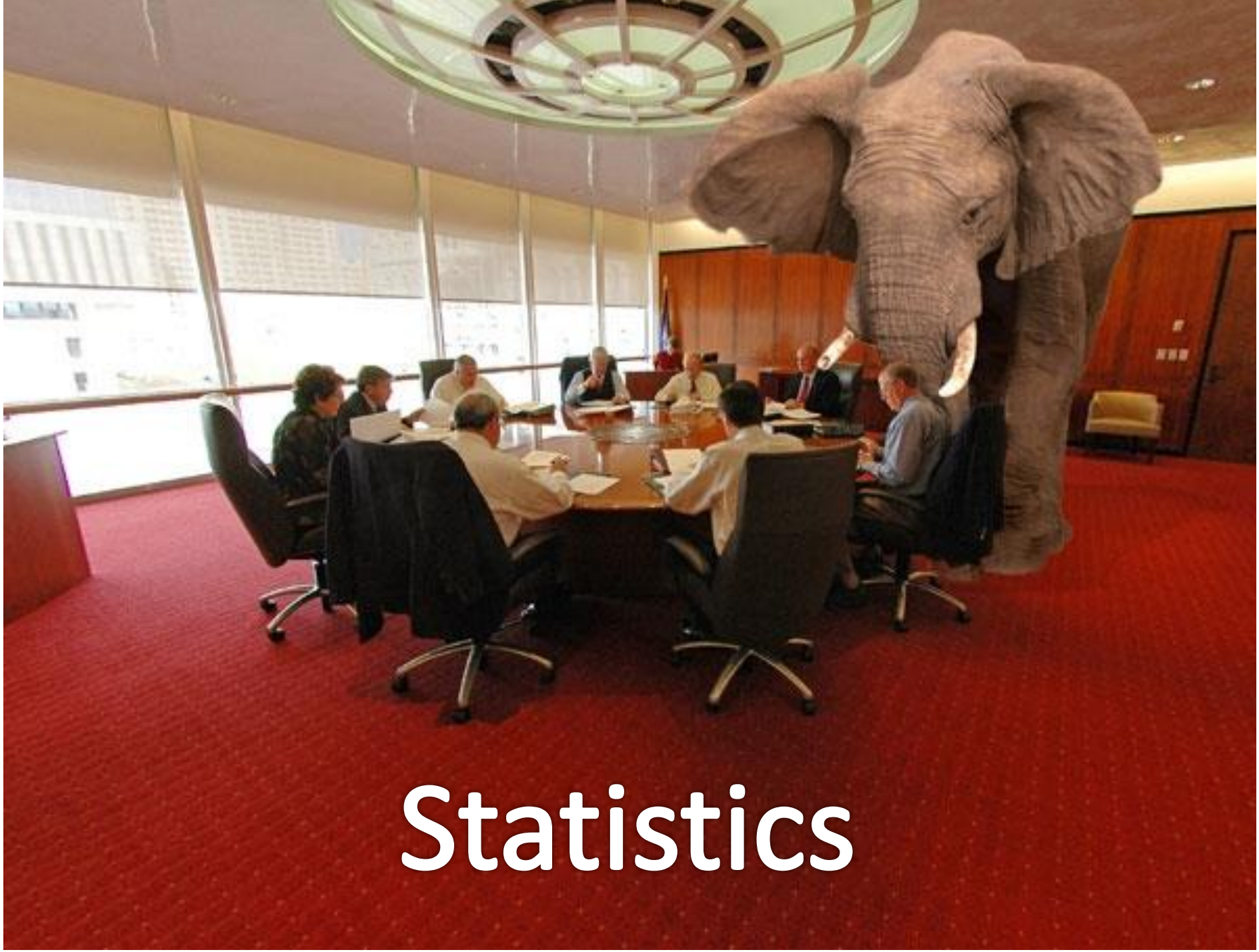
UNBLINDED

http://lc.gcumedia.com/hlt362v/the-visual-learner/the-visual-learner-v2.1.html

Measurement – blinding (Performance bias)

- Were the outcomes measured blindly by researchers and participants?

- Methods: Randomisation, concealment, and blinding


Statistics

P - values and CI

- P values
  - Measure of probability that a result is due to chance
  - The smaller the value (usually P<0.05) less likely due to chance
- Confidence intervals
  - Estimate of the range of values that are likely to include the real value
  - 95% chance of including the real value
  - Narrower the range>more reliable
  - If value does not cross 0 for a difference, or 1 for a ratio then pretty sure result is real (p<0.05)

Measurement - outcomes

- What were the outcomes?
  - Primary
  - Secondary
  - Were they appropriate?
- How were the results reported?
- Were they significant?

- Methods: Outcomes and Results

BDI-II

Date:

Name: _____ Marital Status: _____ Age: _____ Sex: _____

Occupation: _____ Education: _____

Instructions: This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the **one statement** in each group that best describes the way you have been feeling during the **past two weeks, including today**. Circle the number beside the statement you have picked. If several statements in the group seem to apply equally well, circle the highest number for that group. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite).

1. Sadness
0 I do not feel sad.
1 I feel sad much of the time.
2 I am sad all the time.
3 I am so sad or unhappy that I can't stand it.

2. Pessimism
0 I am not discouraged about my future.
1 I feel more discouraged about my future than I used to be.
2 I do not expect things to work out for me.
3 I feel my future is hopeless and will only get worse.

3. Past Failure
0 I do not feel like a failure.
1 I have failed more than I should have.
2 As I look back, I see a lot of failures.
3 I feel I am a total failure as a person.

4. Loss of Pleasure
0 I get as much pleasure as I ever did from the things I enjoy.
1 I don't enjoy things as much as I used to.
2 I get very little pleasure from the things I used to enjoy.
3 I can't get any pleasure from the things I used to enjoy.

5. Guilty Feelings
0 I don't feel particularly guilty.
1 I feel guilty over many things I have done or should have done.
2 I feel quite guilty most of the time.
3 I feel guilty all of the time.

6. Punishment Feelings
0 I don't feel I am being punished.
1 I feel I may be punished.
2 I expect to be punished.
3 I feel I am being punished.

7. Self-Dislike
0 I feel the same about myself as ever.
1 I have lost confidence in myself.
2 I am disappointed in myself.
3 I dislike myself.

8. Self-Criticalness
0 I don't criticize or blame myself more than usual.
1 I am more critical of myself than I used to be.
2 I criticize myself for all of my faults.
3 I blame myself for everything bad that happens.

9. Suicidal Thoughts or Wishes
0 I don't have any thoughts of killing myself.
1 I have thoughts of killing myself, but I would not carry them out.
2 I would like to kill myself.
3 I would kill myself if I had the chance.

10. Crying
0 I don't cry anymore than I used to.
1 I cry more than I used to.
2 I cry over every little thing.
3 I feel like crying, but I can't.

THE PSYCHOLOGICAL CORPORATION®
Harcourt Brace & Company
SAN ANTONIO
Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

_____ Subtotal Page 1

Continued on Back

0154018392
NR15645

11. Agitation
0 I am no more restless or wound up than usual.
1 I feel more restless or wound up than usual.
2 I am so restless or agitated that it's hard to stay still.
3 I am so restless or agitated that I have to keep moving or doing something.

12. Loss of Interest
0 I have not lost interest in other people or activities.
1 I am less interested in other people or things than before.
2 I have lost most of my interest in other people or things.
3 It's hard to get interested in anything.

13. Indecisiveness
0 I make decisions about as well as ever.
1 I find it more difficult to make decisions than usual.
2 I have much greater difficulty in making decisions than I used to.
3 I have trouble making any decisions.

14. Worthlessness
0 I do not feel I am worthless.
1 I don't consider myself as worthwhile and useful as I used to.
2 I feel more worthless as compared to other people.
3 I feel utterly worthless.

15. Loss of Energy
0 I have as much energy as ever.
1 I have less energy than I used to have.
2 I don't have enough energy to do very much.
3 I don't have enough energy to do anything.

16. Changes in Sleeping Pattern
0 I have not experienced any change in my sleeping pattern.
1a I sleep somewhat more than usual.
1b I sleep somewhat less than usual.
2a I sleep a lot more than usual.
2b I sleep a lot less than usual.
3a I sleep most of the day.

17. Irritability
0 I am no more irritable than usual.
1 I am more irritable than usual.
2 I am much more irritable than usual.
3 I am irritable all the time.

18. Changes in Appetite
0 I have not experienced any change in my appetite.
1a My appetite is somewhat less than usual.
1b My appetite is somewhat greater than usual.
2a My appetite is much less than before.
2b My appetite is much greater than usual.
3a I have no appetite at all.
3b I crave food all the time.

19. Concentration Difficulty
0 I can concentrate as well as ever.
1 I can't concentrate as well as usual.
2 It's hard to keep my mind on anything for very long.
3 I find I can't concentrate on anything.

20. Tiredness or Fatigue
0 I am no more tired or fatigued than usual.
1 I get more tired or fatigued more easily than usual.
2 I am too tired or fatigued to do a lot of the things I used to do.
3 I am too tired or fatigued to do most of the things I used to do.

21. Loss of Interest in Sex
0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I am much less interested in sex now.
3 I have lost interest in sex completely.

Outcomes

	Measure	Narrative	Numerical
Primary outcome: short term symptoms of depression	Beck depression inventory score	no evidence that participants in the intervention group had a better outcome at four months than those in the usual care group	difference in mean score of −0.54 (95% confidence interval −3.06 to 1.99; P=0.68)
Secondary outcomes Longer term symptoms of depression	Beck depression inventory score	no evidence of a difference between the treatment groups over the duration of the study	difference in mean Beck depression inventory score −1.20,95% confidence interval−3.42 to 1.02;P=0.29
Anti-depressant use	participants reporting use of antidepressants	no evidence to suggest any difference between the groups at either the four month follow-up point or duration of trial	adjusted odds ratio 1.20, 95% confidence interval 0.69 to 2.08; P=0.52
Physical activity	self completion seven day recall diary	there was some evidence for a difference in reported physical activity between the groups at four months post-randomisation	adjusted odds ratio 1.58, 0.94 to 2.66; P=0.08)

Conclusions of the study

BMJ 2012;344:e2758 doi: 10.1136/bmj.e2758 (Published 6 June 2012)

What is already known on this topic

Depression is a leading contributor to disability in the United Kingdom and is associated with a decrement of health greater than many other chronic diseases

Many patients and healthcare professionals would like an effective and accessible non-drug treatment for depression

Numerous studies have reported the positive effects of physical activity but most of the current evidence originates from small non-clinical samples using interventions that are not practicable in healthcare settings

What this study adds

A physical activity intervention in addition to usual care did not improve symptoms of depression or reduce the use of antidepressants compared with usual care alone

The intervention increased self reported physical activity and this effect was sustained for 12 months

Clinicians and policy makers should alert people with depression that advice to increase physical activity will not increase their chances of recovery from depression

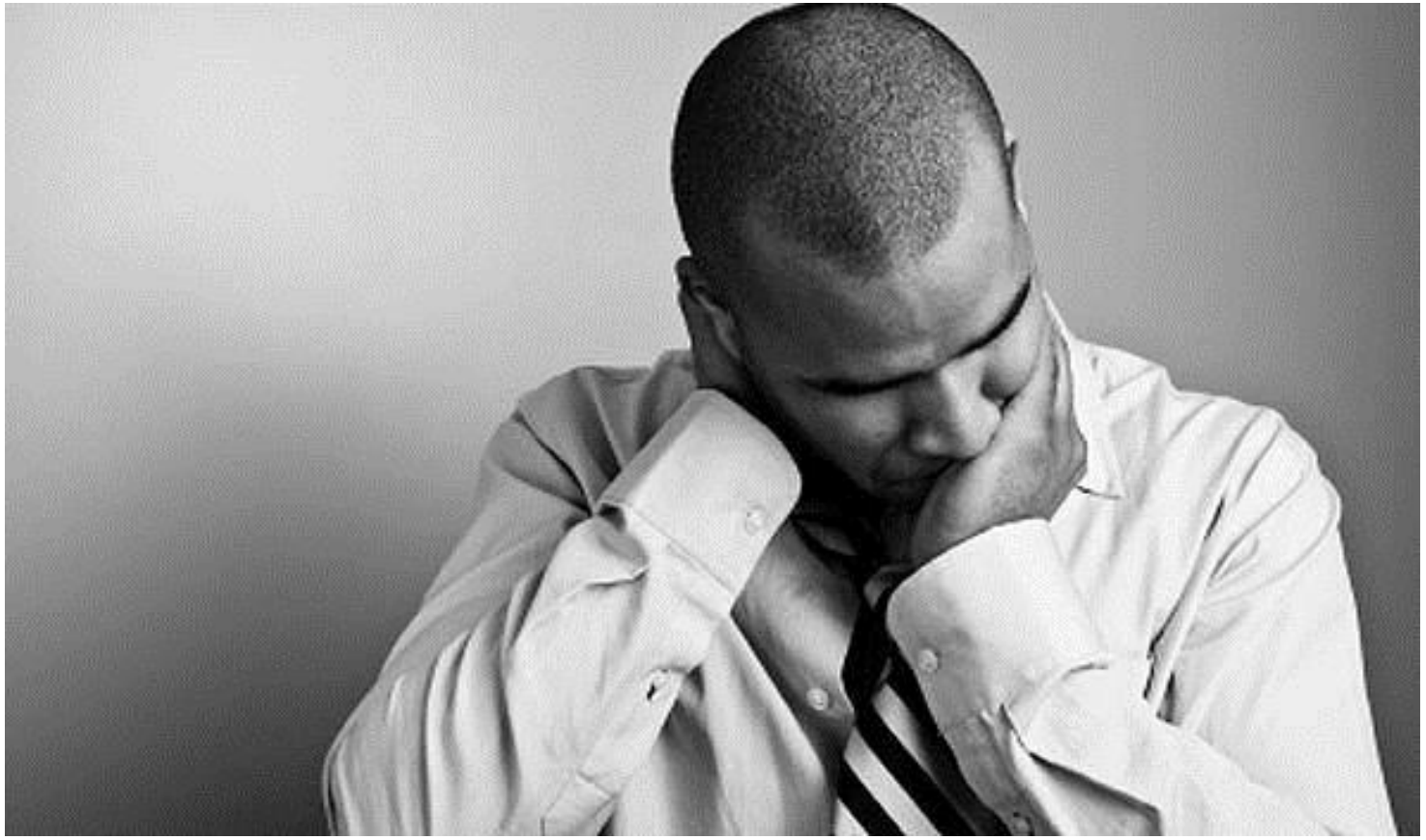
27 Pavey TG, Taylor AH, Fox KR, Hillsdon M, Anokye N, Campbell JL, et al. Effect of exercise referral schemes in primary care on physical activity and improving health outcomes: systematic review and meta-analysis. *BMJ* 2011;343:d6462.

28 Ekkekakis P, Hall EE, Petruzzello SJ. Variation and homogeneity in affective responses to physical activity of varying intensities: an alternative perspective on dose-response based on evolutionary considerations. *J Sports Sci* 2005;23:477-500.

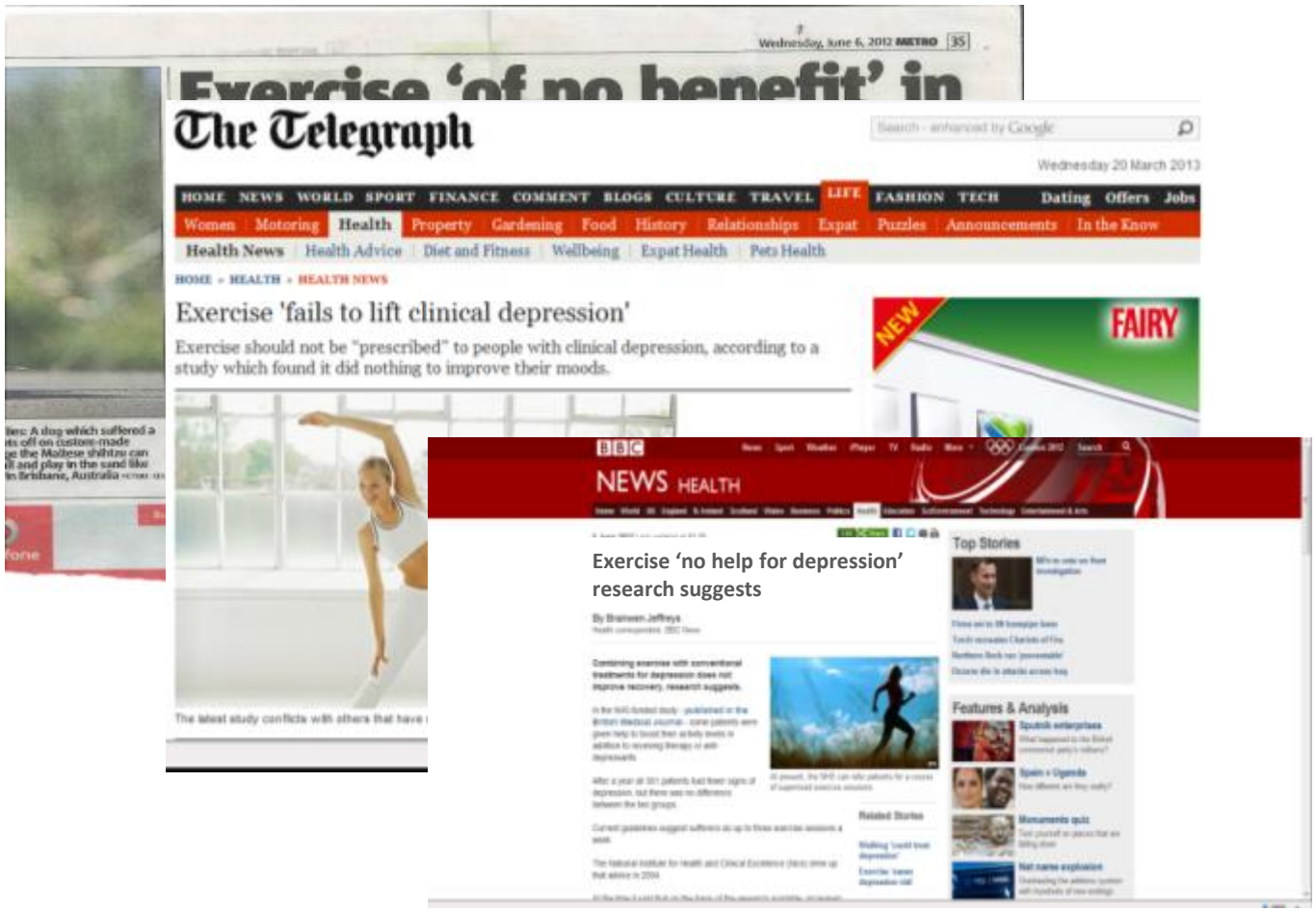
29 Searle A, Calnan M, Lewis G, Campbell J, Taylor A, Turner K. Patients' views of physical

Cite this as: *BMJ* 2012;344:e2758

External validity/applicability



Would you advocate exercise for depression based on this study?


Wednesday, June 6, 2012 METRO 35
Exercise 'of no benefit' in
The Telegraph
Wednesday 20 March 2013
HOME NEWS WORLD SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFE FASHION TECH Dating Offers Jobs
Women Motoring Health Property Gardening Food History Relationships Expat Puzzles Announcements In the Know
Health News Health Advice Diet and Fitness Wellbeing Expat Health Pets Health
HOME » HEALTH » HEALTH NEWS
Exercise 'fails to lift clinical depression'
Exercise should not be "prescribed" to people with clinical depression, according to a study which found it did nothing to improve their moods.
NEW
FAIRY
BBC
NEWS HEALTH
Exercise 'no help for depression' research suggests
Top Stories
Features & Analysis

st came the paper. Its abstract concludes:

Research Paper: *"The addition of a facilitated*
activity intervention to usual care did not impro
outcome or reduce use of antidepressants com
usual care alone."

same conclusions as they appear in the press release

Press Release: *"*
the BMJ, suggest
to usual care did
than usual care a

Press Release -
designed researc
appear to be effe

Re: Facilitated physical activity as a treatment for depressed adults: randomised controlled trial

6 June 2012

Great - another depressing headline

Those of you out there working tirelessly to get people 'enjoying' a more active lifestyle are, like me, unlikely to be overjoyed at today's headlines on the BBC website's health page.1 A nice example of bad PR in my opinion, particularly when you actually read the conclusions made by the authors of the headline source paper.2 I went a bit mad on twitter for an hour or so after I read the study but after calming down decided that a better way to get answers to my many questions would be via this letter. So, for me it would be useful to know the following:

1. How does the conclusion "advice and encouragement to increase physical activity is not an effective strategy for reducing symptoms of depression" translate to "Exercise 'no help for depression' research suggests'?1

2. Why does OR 1.58, 0.94 to 2.66 P = 0.08 translate to "some evidence" when OR of 0.66, 0.4 to 1.11 P = 0.12 translates to "no

BBC News Sport Weather iPlayer TV Radio More London 2012 Search

NEWS HEALTH

Home World UK England N. Ireland Scotland Wales Business Politics Health Education Sci/Environment Technology Entertainment & Arts

6 June 2012 Last updated at 01:29

Exercise 'no help for depression' research suggests

By Branwen Jeffreys
Health correspondent, BBC News

Combining exercise with conventional treatments for depression does not improve recovery, research suggests.

In the NHS-funded study - published in the British Medical Journal - some patients were given help to boost their activity levels in addition to receiving therapy or anti-depressants.

At present, the NHS can refer patients for a course of supervised exercise sessions

After a year all 361 patients had fewer signs of depression, but there was no difference between the two groups.

Current guidelines suggest sufferers do up to three exercise sessions a week.

The National Institute for Health and Clinical Excellence (Nice) drew up that advice in 2004.

Related Stories

Walking 'could beat depression'

Exercise 'eases depression risk'

Top Stories

MPs to vote on Hunt investigation

Films set to lift hosepipe bans

Torch recreates Chariots of Fire

Northern Rock run 'preventable'

Dozens die in attacks across Iraq

Features & Analysis

Sputnik enterprises
What happened to the British communist party's millions?

Spain v Uganda
How different are they really?

Monuments quiz
Test yourself on places that are falling down

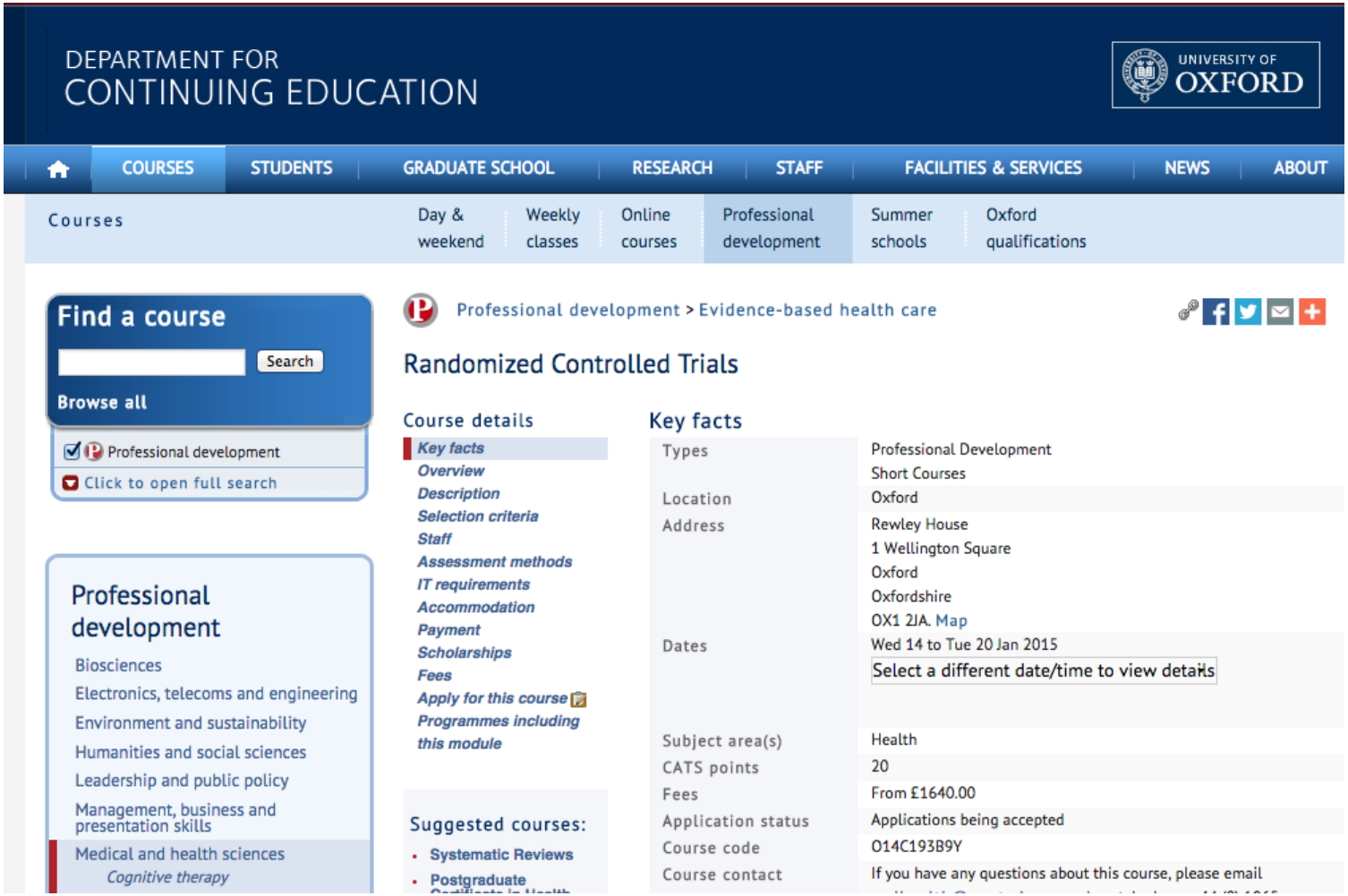
Net name explosion
Overhauling the address system with hundreds of new endings

Summary

- Lots of “evidence” in healthcare
- RCTs provide an opportunity to deliver answers to the effects if interventions
- But dependent upon minimising risk of bias
- Critical appraisal assess this
- Lots of tools to assess risk of bias
- Application (external validity) based on your interpretation of results

Want more?

RCT course



https://www.conted.ox.ac.uk/

kamal.mahtani@phc.ox.ac.uk

@krmahtani

Group work

Exercise for depression (Review)

Cooney GM, Dwan K, Greig CA, Lawlor DA, Rimer J, Waugh FR, McMurdo M, Mead GE

THE COCHRANE COLLABORATION®

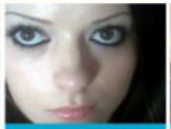
This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2013, Issue 9

http://www.thecochranelibrary.com

WILEY

Exercise for depression: critical appraisal

- 2-3 groups
- 2-3 different RCTs from same SR
- In groups:
  - Read paper – DON'T REFER BACK TO COCHRANE RV!
  - PICO
  - Critical appraisal – internal validity
  - External validity
  - Each group present their paper (PICO, appraisal)
  - Comment on the validity for 10 mins


Haunting final moments of heavily


Jordan's Queen Rania in powerful


Father burned to death inside 280C


BREAKING NEWS: Three men charged


Oscar-winning director of The


Tony Blair handed a gong for spending

For those with mental health problems, exercise can be more effective than antidepressants

- **Exercise raises self-esteem and boosts levels of 'feel good' hormones**
- **For those with mild depression, physical activity improves symptoms**

By CARA LEE FOR THE DAILY MAIL

PUBLISHED: 00:15, 23 July 2013 | **UPDATED:** 09:43, 23 July 2013

As a Zumba instructor, Karen Bedford might be expected to rave about what she teaches. But Karen has a more personal motive, for she believes exercise helped her recover from a decade-long battle with depression.

Where antidepressants failed, Zumba succeeded. 'Since Zumba became part of my life my mental state's been more stable and I haven't needed the pills,' says Karen, 49, who has two grown-up children.

'I get the odd down day, but because I've regained my confidence I face my problems much better now.'

It's a far cry from when Karen's depression emerged in 1998. Her long-term relationship had just ended and she was fighting to keep the family home in Potters Bar, Hertfordshire, when she was made redundant from her job at a publishing house.

Hemat-Far 2012

Hemat-Far 2012

Methods	RCT
Participants	University students aged 18 - 25 with depression 100% women
Interventions	1. 40 - 60 minutes of running, 3 times a week, supervised. (n = 10) 2. Control group with no active intervention (n = 10)
Outcomes	Beck Depression Inventory score
Notes	Small sample size (10 participants in each arm); specific population under study

Hemat-Far 2012

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Clinician judgement used at recruitment. After reviewing questionnaires psychiatrists "selected" 20 women
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) participants	Unclear risk	Participants not blinded to intervention; unclear effect on bias
Blinding (performance bias and detection bias) those delivering intervention	Unclear risk	No information given
Blinding (performance bias and detection bias) outcome assessors	High risk	Self report BDI
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No discussion on attrition rate
Selective reporting (reporting bias)	Low risk	BDI specified at outset and completed in results

Sims 2009

Sims 2009

Methods	RCT
Participants	Recruited from hospital databases of stroke patients discharged in last year, general practitioners and newspaper articles. Had to be over 6 months post-stroke and have depression confirmed by a psychiatrist Mean age 67.13 (range 21 to 93) 40% women N = 45
Interventions	1. Group-based moderate-intensity strengthening exercises twice a week for 10 weeks. The PRT programme included 2 high-intensity sessions/week for 10 weeks at a community-based gymnasium. (n = 23) 2. Usual care (n = 22)
Outcomes	Centre for Epidemiologic Studies for Depression scale
Notes	Intention-to-treat analysis Outcome was self-rated symptoms of depression by CES-D scale

Sims 2009

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Block randomised list
Allocation concealment (selection bias)	Low risk	Randomisation was conducted centrally by an independent person
Blinding (performance bias and detection bias) participants	Unclear risk	Participants not blind, unclear risk of bias
Blinding (performance bias and detection bias) those delivering intervention	Unclear risk	Those delivering intervention were not blind, unclear risk of bias
Blinding (performance bias and detection bias) outcome assessors	High risk	Self report outcome (depressive symptoms by CES-D scale)
Incomplete outcome data (attrition bias) All outcomes	High risk	Baseline assessment was performed in 45 people; complete data were available for 43 people at 6 months (23/23 in intervention group and 20/22 in the control)
Selective reporting (reporting bias)	Unclear risk	Reported all prespecified outcome (though we do not have access to the protocol)
Other bias	Unclear risk	Unclear

Singh 2005

Methods	RCT
Participants	People responding to a postal questionnaire who had DSM-IV depression or dysthymia Mean age 69 55% women N = 60
Interventions	1. Progressive resistance training at 80% of 1 repetition max (n = 20) 2. Resistance training at 20% of 1 repetition max (n = 20) 3. Usual care (n = 20) Each intervention group held 3 times a week for 8 weeks
Outcomes	1. Hamilton Rating Scale for depression 2. Geriatric Depression score
Notes	Not intention-to-treat (50/60 completed the study and were available for assessment) Outcome assessment blind

Singh 2005

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Adequate. Sealed opaque envelopes open after baseline assessment
Blinding (performance bias and detection bias) participants	Unclear risk	Participants not blind, unclear effect on bias
Blinding (performance bias and detection bias) those delivering intervention	Unclear risk	Those delivering the intervention were not blind to treatment allocation, unclear effect on bias
Blinding (performance bias and detection bias) outcome assessors	Low risk	HRSD performed by blinded outcome assessors
Incomplete outcome data (attrition bias) All outcomes	High risk	6/60 dropped out (2 from the high-dose, 3 from the low-dose and 1 from the usual care group)
Selective reporting (reporting bias)	Unclear risk	Prespecified outcomes in paper were reported, but no protocol
Other bias	Unclear risk	Unclear

Krogh 2009

Krogh 2009

Methods	RCT (parallel group)
Participants	Referred from general practitioners, private psychiatrists, psychologists and psychiatric wards institutions. Included if met criteria for major depression Mean age 38.9 73.9% women N = 165
Interventions	1. Strength circuit training (n = 55) 2. Aerobic (machine-based) training (n = 55) 3. Relaxation control (n = 55 Twice-weekly intervention for 32 sessions delivered over a 4-month period
Outcomes	Hamilton Rating Scale for Depression
Notes	Intention-to-treat analysis Significant drop-outs in each group Changed sample size calculation after first 50 participants on basis of observed standard deviation

Krogh 2009

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computerised restricted randomisation with a block size of 8
Allocation concealment (selection bias)	Low risk	The block size and allocation sequence were unknown to the DEMO trial staff
Blinding (performance bias and detection bias) participants	Unclear risk	Participants not blind, but unclear what influence this had on bias
Blinding (performance bias and detection bias) those delivering intervention	Unclear risk	Physiotherapists delivering the intervention were not blind. Unclear how this influenced risk of bias
Blinding (performance bias and detection bias) outcome assessors	Low risk	The assessor was blind to intervention group. The investigators asked the outcome assessors to guess intervention group. The kappa values for agreement between the right allocation and the guessed allocation were 0.15 and 0.05 for the assessments at 4 and 12 months respectively
Incomplete outcome data (attrition bias) All outcomes	Low risk	137/165 were available for follow-up at the end of the intervention. Eighteen were lost to follow-up and 10 refused to participate (8/55 in strength group, 7/55 in aerobic group and 13/55 in the relaxation group). The authors used a likelihood-based mixed-effect model with an unstructured variance matrix available in SPSS, which is able to handle missing data with higher precision and power than last observation carried forward. The authors reported no significant difference between missing participants and participants included in the analyses at either 4 or 12 months, and concluded that it was reasonable to assume that the missing data were 'missing at random'
Selective reporting (reporting bias)	Low risk	All prespecified outcomes seem to have been reported. Protocol was published in advance of the trial
Other bias	Unclear risk	The authors repeated power calculations part-way through the trial, and reduced the sample size as the standard deviation was lower than anticipated

Chu 2008

Chu 2008	
Methods	RCT
Participants	Volunteers aged 18 to 45 recruited via flyers and word of mouth from University and local physician referral. Depression severity mild to moderate, if severe required written permission from physician Mean age 26.4 (18 to 43). 100% women.
Interventions	For 10 weeks: 1. Up to 5 high-intensity aerobic exercise sessions per week (1 supervised) to expend 1000 Kcal per week (n = 15) 2. Up to 5 low-intensity aerobic exercise sessions per week (1 supervised) to expend 1000 Kcal per week (n = 11) 3. Met with investigator once per week for 30 minutes of group stretching exercises (n = 12)
Outcomes	Beck Depression Inventory-II
Notes	Analysis not intention-to-treat BDI-II self-rated depression score

Chu 2008

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not described
Allocation concealment (selection bias)	Unclear risk	Unclear; not described
Blinding (performance bias and detection bias) participants	Unclear risk	Participants not blind to treatment allocation
Blinding (performance bias and detection bias) those delivering intervention	Unclear risk	Those delivering the intervention were not blind
Blinding (performance bias and detection bias) outcome assessors	High risk	BDI-II self report was used as the primary outcome
Incomplete outcome data (attrition bias) All outcomes	High risk	16/54 dropped out (3/18 in high dose, 7/18 in low dose and 6/18 in control)
Selective reporting (reporting bias)	Unclear risk	It appears that all prespecified outcomes are reported, but no protocol
Other bias	Unclear risk	Unclear


KEEP
CALM
IT'S
LUNCH
TIME

Odds ratio

- odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure
- Interpreting OR
  - OR=1 Exposure does not affect odds of outcome
  - OR>1 Exposure associated with higher odds of outcome
  - OR<1 Exposure associated with lower odds of outcome
- E.g.… OR = 1.46
  - Odds of having the outcome are 1.46 higher in the exposed group vs control group

Odds ratio

Outcome of interest

Exposure of interest

	+	-
+	a	b
-	c	d

$$OR = \frac{a/c}{b/d}$$

Relative Risk or Risk Ratio

- the risk of the event in one group divided by the risk of the event in the other group
- Interpreting RR
    - RR =1 Exposure does not affect risk of outcome
  - Is the treatment intended to prevent an undesirable outcome?
    - RR < 1Exposure reduces the risk of the event
    - RR > 1 Exposure increases the risk of the event (possible treatment harm, adverse events)
  - Is the treatment intended to promote an outcome? (e.g. disease remission)
    - RR < 1Exposure reduces the risk of the event (disease remission)
    - RR > 1 Exposure increases the risk of the event (disease remission)

E.g.… RR = 0.46

  - Risk of getting the outcome with the exposure was 0.46 of that in the control group

RR v OR

- Often similar when event rate is low (<10%) or treatment effect is small (close to 1)
- As event rate increases (>10%)

Relative Risk or Risk Ratio

Outcome of interest

Exposure of interest

	+	-
+	a	b
-	c	d

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

Odds ratio

- odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure
- Interpreting OR
  - OR=1 Exposure does not affect odds of outcome
  - OR>1 Exposure associated with higher odds of outcome
  - OR<1 Exposure associated with lower odds of outcome
- E.g.… OR = 1.46
  - Odds of having the outcome are 1.46 higher in the exposed group vs control group

Odds ratio

Outcome of interest

Exposure of interest

	+	-
+	a	b
-	c	d

$$OR = \frac{a/c}{b/d}$$

Relative Risk or Risk Ratio

- the risk of the event in one group divided by the risk of the event in the other group
- Interpreting RR
    - RR =1 Exposure does not affect risk of outcome
  - Is the treatment intended to prevent an undesirable outcome?
    - RR < 1Exposure reduces the risk of the event
    - RR > 1 Exposure increases the risk of the event (possible treatment harm, adverse events)
  - Is the treatment intended to promote an outcome? (e.g. disease remission)
    - RR < 1Exposure reduces the risk of the event (disease remission)
    - RR > 1 Exposure increases the risk of the event (disease remission)

E.g.… RR = 0.46

  - Risk of getting the outcome with the exposure was 0.46 of that in the control group

RR v OR

- Often similar when event rate is low (<10%) or treatment effect is small (close to 1)
- As event rate increases (>10%)

Relative Risk or Risk Ratio

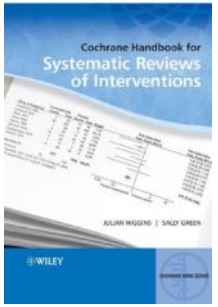
Outcome of interest

Exposure of interest

	+	-
+	a	b
-	c	d

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

Selection bias



- systematic differences between baseline characteristics of the groups

- Adequate randomisation
  - 1) Sequence generation
  - 2) Allocation concealment

Sequence generation (selection bias)

Low risk of bias

- random number table
- Using a computer random number generator
- Coin tossing
- Shuffling cards or envelopes
- Throwing dice
- Drawing of lots

High risk of bias

- Sequence generated by a a non-random component e.g
  - odd or even date of
  - birth date (or day) of admission
  - hospital or clinic record number
- judgement of the clinician
- preference of the participant
- availability of the intervention

Allocation concealment (selection bias)

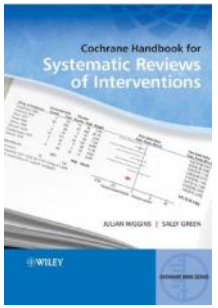
Low risk

- Central allocation (including telephone, web-based and pharmacy-controlled randomization
- Sequentially numbered drug containers of identical appearance
- Sequentially numbered, opaque, sealed envelopes.

High risk

- Alternation or rotation
- open random allocation schedule (e.g. a list of random numbers)
- envelopes were unsealed or non-opaque

Performance bias



- Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.

- Blinding of participants, personnel and outcome assessors

Blinding (Performance bias)

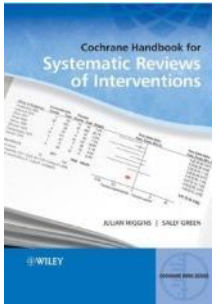
Low risk of bias

- No blinding, but outcome and the outcome measurement are not likely to be influenced
- Blinding of participants *and* personnel
- blinding of participants *or* personnel but outcome assesment unlikely to have been affected

High risk of bias

- No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding
- Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken
- No blinding

Attrition bias



- Systematic differences between groups in withdrawals from a study.
- **Attrition** refers to situations in which outcome data are not available
- **Exclusions** refer to situations in which some participants are omitted from reports of analyses, despite outcome data being available to the trialists.

Incomplete reporting (Attrition bias)

Low risk of bias

- No missing outcome data
- Reasons for missing outcome data unlikely to be related to true outcome
- Methodology ITT

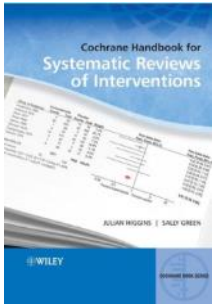
High risk of bias

- Reason for missing outcome data likely to be related to true outcome,
- “As-treated’ analysis done with substantial departure of the intervention received from that assigned at randomization

Intention to treat (ITT)

- participants in trials should be analysed in the groups to which they were randomized, regardless of whether they received or adhered to the allocated intervention.
- 2 issues:
  - estimate the effects in practice
    - Not a subgroup who adhere to the intervention
    - “Per protocol” can overestimate effects
  - Loss to follow up
    - ITT ensures the outcome is still measured on these patients

Reporting bias



- systematic differences between reported and unreported findings.

- E.g publication bias, more likely to report significant differences between intervention groups than non-significant differences.

Selective outcome reporting (Reporting bias)

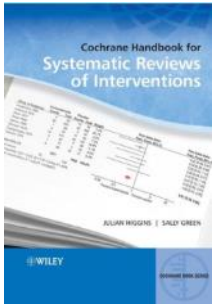
Low risk of bias

- The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way
- The study protocol is not available but it is clear that the published reports include all expected outcomes

High risk of bias

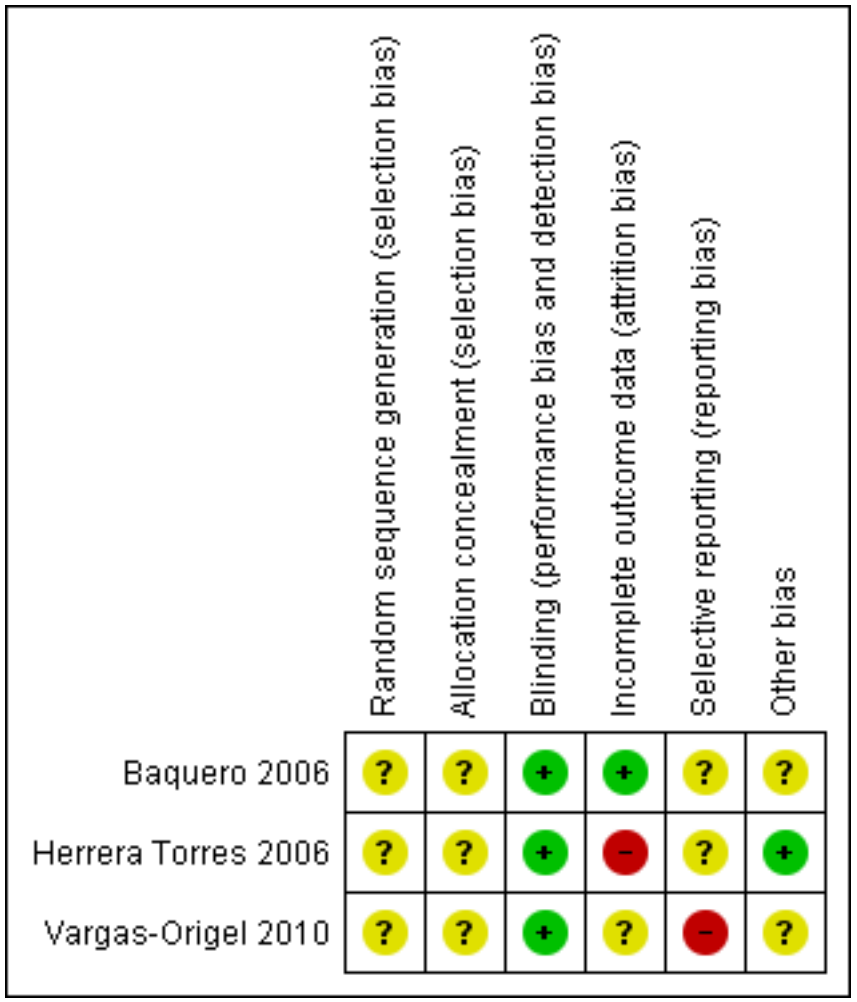
- Not all of the study's pre-specified primary outcomes have been reported
- One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified
- One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect);
- outcomes of interest in the review are reported

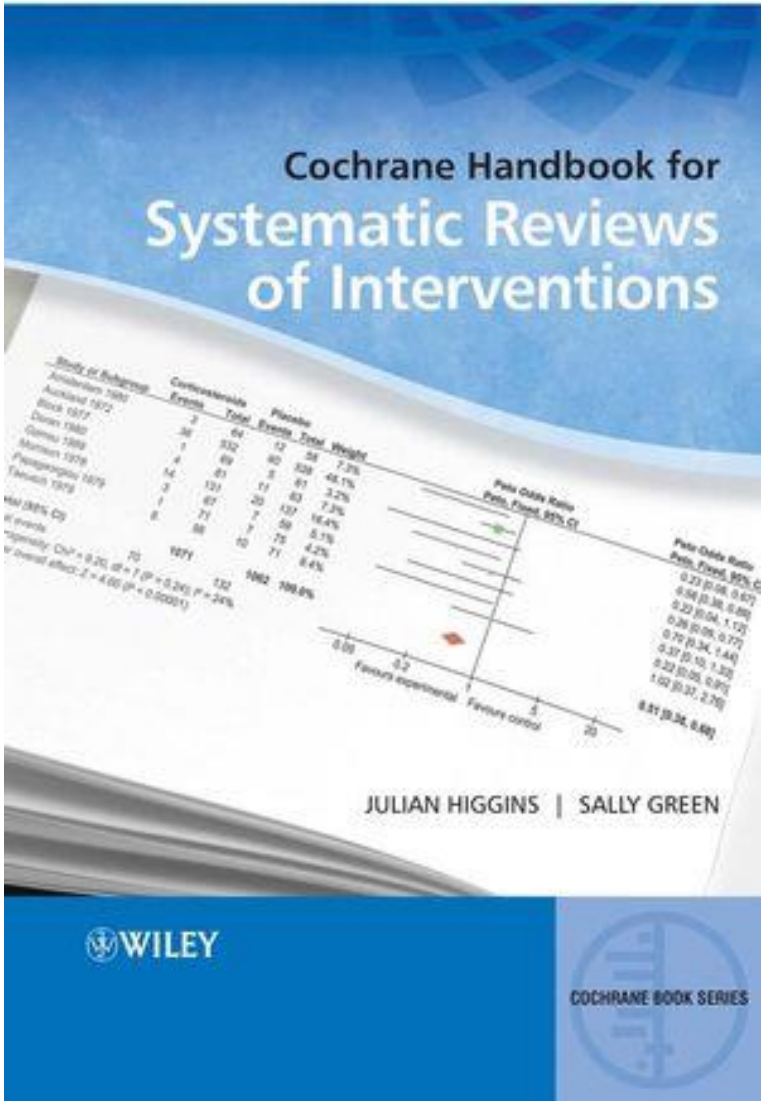
Other biases



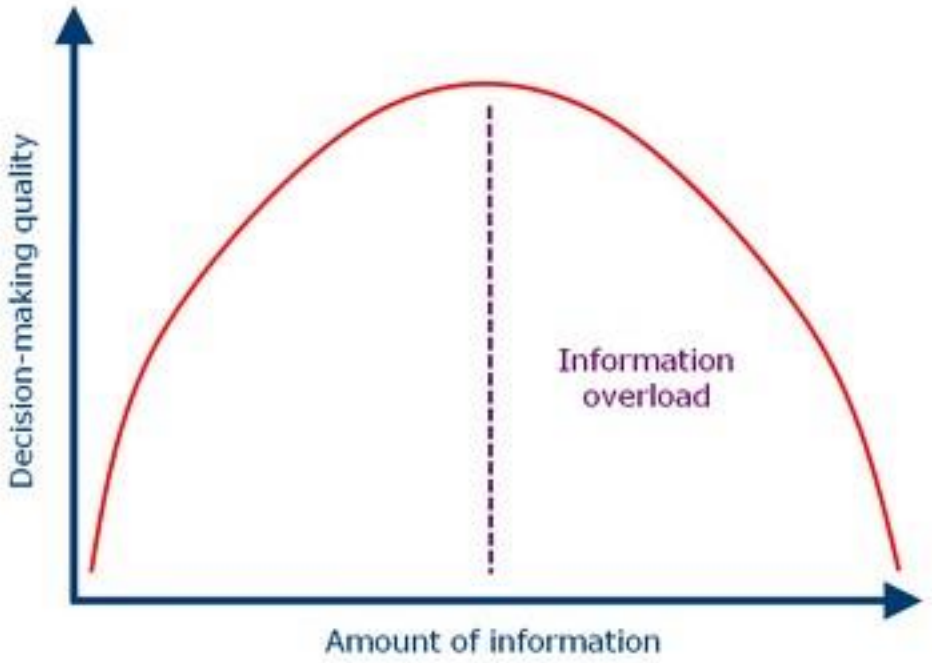
- Trial designs
  - carry-over in cross-over trials
  - recruitment bias in cluster-randomized trials
    - E.g participants may know already which group they have been allocated to because everyone in that “cluster” gets the same intervention.

Cochrane risk of bias table





http://handbook.cochrane.org/front_page.htm


Decision-making quality
Information overload
Amount of information

RRAMMbo tool map to Cochrane RoB

		Type of bias	Cochrane RoB domains
Recruitment	Were the subjects representative of the target population?	Selection bias Other sources of bias	Other sources of bias
Randomisation **A**llocation	How was randomisation carried out? Was allocation concealed?	Selection bias	Sequence generation Allocation concealment
Maintenance	Were the groups equal at the start? And maintained through equal management and f/u?	Performance bias Attrition bias	Incomplete outcome data Blinding of participants, personnel and outcome assessors
Measurement- **B**linding	Were the outcomes measured with blinded assessors/participants	Performance bias	Blinding of participants, personnel and outcome assessors
Objective outcomes (Measurement)	Were there differences in how outcomes were determined	Detection bias	Blinding of participants, personnel and outcome assessors. Other potential threats to validity

Types of bias

Type of bias	Description
Selection bias	Systematic differences between baseline characteristics of the groups that are compared.
Performance bias	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest
Attrition bias	Systematic differences between groups in withdrawals from a study
Detection bias	Systematic differences between groups in how outcomes are determined
Reporting bias	Systematic differences between reported and unreported findings