Norwegian Institute of Public Health

Risk of bias-assessment

Norwegian Research School for Global Health

27.09.2022

Atle Fretheim, Research Director, NIPH, and Adj. Professor, OsloMet

Goal for the next hour

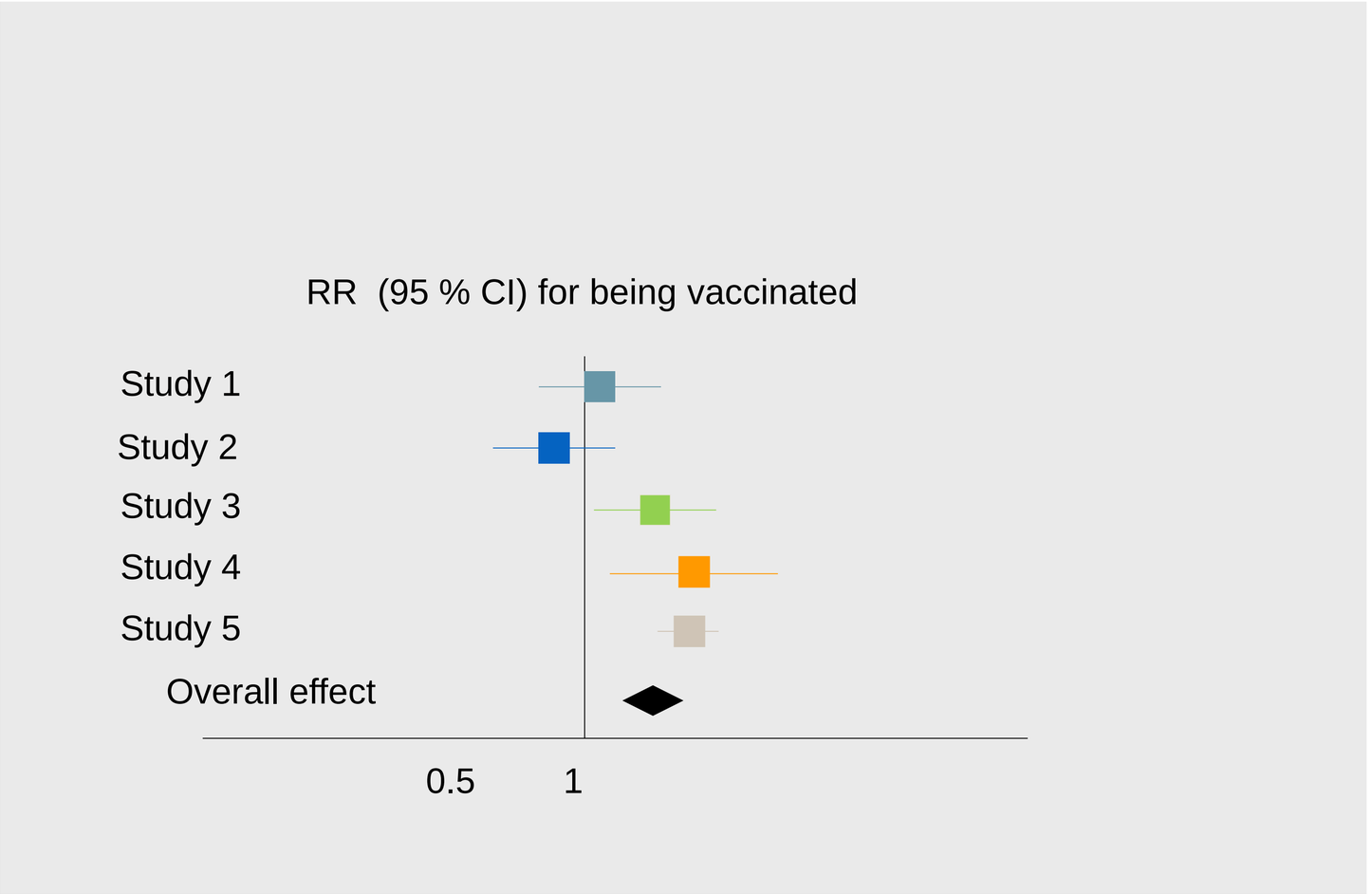
- We all have an understanding of the term «risk of bias», and are able to apply it when we assess research studies

Risk of bias assessment – a key step in doing a systematic review

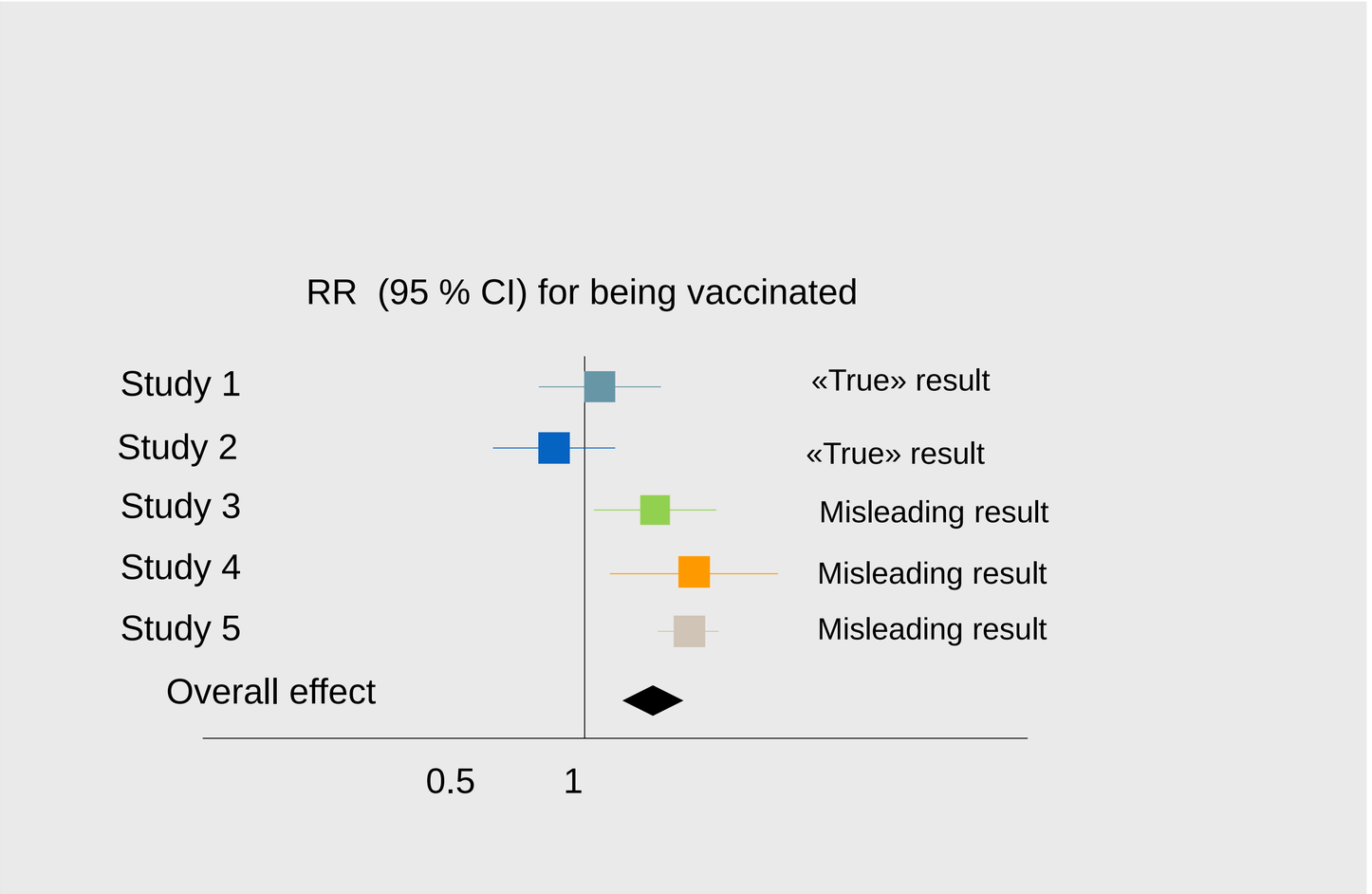
Relevant studies

Critical assessment of each of them (risk of bias-assessment)

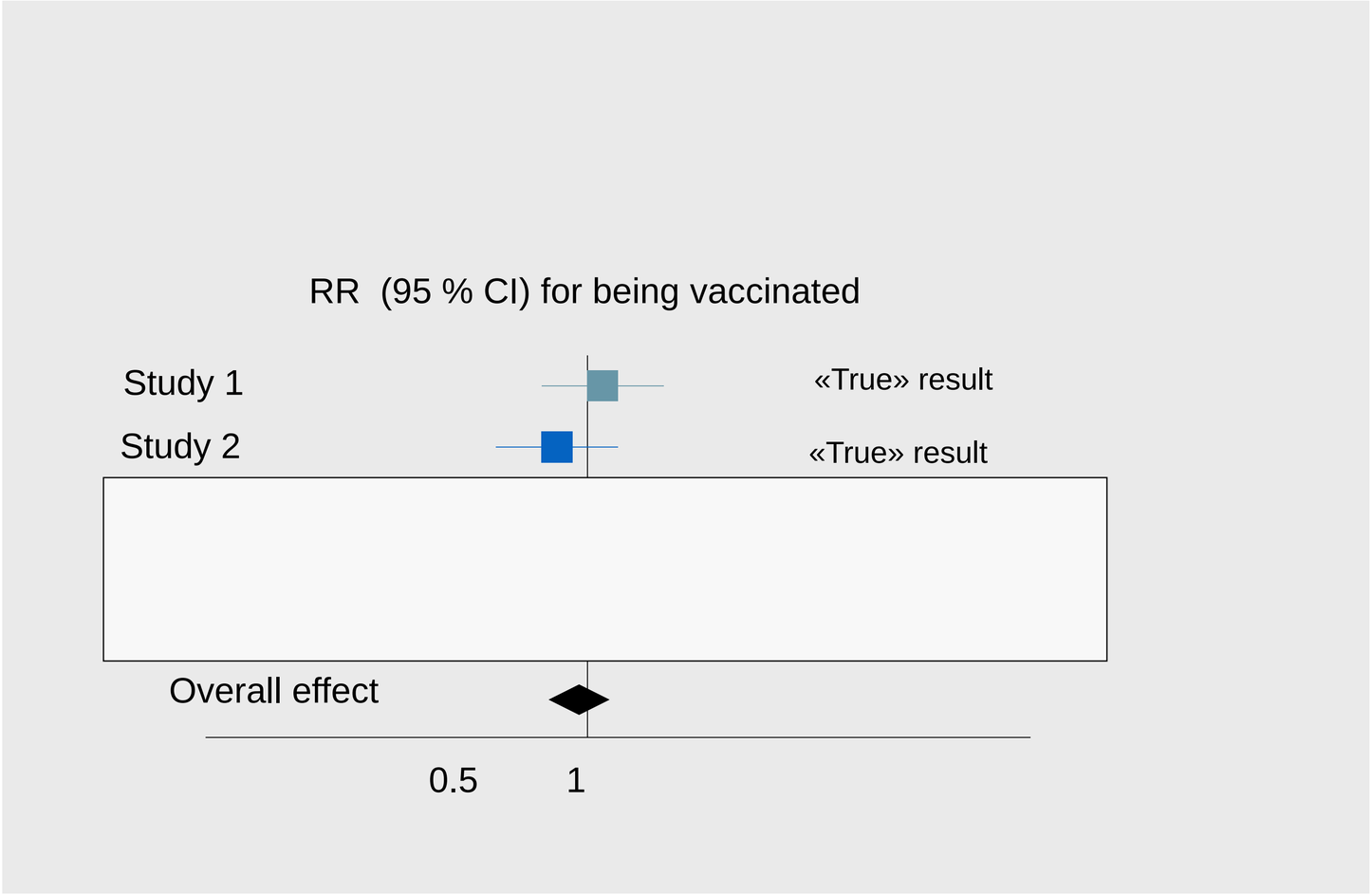
Misleading studies – misleading SR



Misleading studies – misleading SR



Misleading studies – misleading SR



«Bias»

- Systematic errors that lead to erroneous (non- «true») results

Risk of Bias (RoB)

- Directly related to internal validity – the likelihood that the results reflect «the truth»
  - Low RoB: We think the results are likely «true»
  - High RoB: We think the results may be «untrue»
- Not related to precision – an imprecise result can be «true» (is caused by random errors, not systematic errors)
- Not related to external validity (applicability, transferability) – a study finding can be «true» even if it's non-applicable to a different setting!

Risk of Bias (RoB)-assessment

- An approach to evaluating risk of systematic errors in a study, or of a study finding
- The question is: What is it about this study that gives us reason to doubt the truthfulness of the findings?
- In other words: We are looking for possible sources of «bias» in a study.
  - Several check lists out there – we recommend Cochrane's «Risk of Bias Tool»

The Cochrane RoB-tool emphasises five domains*

- bias arising from the randomization process;
- bias due to deviations from intended interventions;
- bias due to missing outcome data;
- bias in measurement of the outcome; and
- bias in selection of the reported result.

* Developed especially for RoB-assessment of RCTs, but the same approach can, in prinicple, be applied on all types of effectiveness studies.

Why these domains?

- A combination of logic/theory and empirical findings
  - E.g. it's been shown that lack of blinding can introduce substantial bias on subjective outcomes, e.g. pain (but not on objective, e.g. death)*

*Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429-38.

What about other types of studies?

- There are several check lists around
- One recommended source: https://jbi.global/critical-appraisal-tools

CRITICAL APPRAISAL TOOLS

JBI's critical appraisal tools assist in assessing the trustworthiness, relevance and results of published papers.

CRITICAL APPRAISAL TOOLS DOWNLOADS	DOWNLOAD
Checklist for Analytical Cross Sectional Studies	
Checklist for Case Control Studies	
Checklist for Case Reports	
Checklist for Case Series	

THE SERUM TREATMENT OF LOBAR PNEUMONIA

A REPORT OF THE THERAPEUTIC TRIALS COMMITTEE OF THE MEDICAL RESEARCH COUNCIL

During the last three years the Medical Research Council have assisted an inquiry at different centres in Great Britain into the therapeutic value of specific sera for lobar pneumonia, following the great development of similar work in the United States. When the Council appointed a standing Therapeutic Trials Committee in 1931 the investigation was placed under the control of that committee, and the present report summarizes the evidence obtained.

The work has been laborious; for it was little more than a critical testing for practical use of methods which were already in common knowledge, and it involved the close consideration of a very large number of cases of pneumonia. The Council wish to express their gratitude to the workers who so willingly undertook this prolonged study, and brought their results together for joint consideration.

Interest in the treatment of pneumonia has for many years been felt more keenly in Scotland, where the disease is perhaps more prevalent in the winter months, than in England. For evidence used in the present report the committee are indebted to workers in Aberdeen, Edinburgh, Glasgow, and London. The observations at Edinburgh were made by Professor D. Murray Lyon and the other physicians to the Royal Infirmary; those at Aberdeen by Professor Stanley Davidson, Dr. J. B. Ewen, and Dr. R. J. Duthie, in the City Hospital, Woodend; and those in London by Dr. R. R. Armstrong and Dr. R. Sleigh Johnson in various London County Council hospitals and also at St. Bartholomew's Hospital. The Glasgow inquiry under Dr. John Cowan, Dr. A. W. Harrington, and Dr. R. Cruickshank was developed independently with support from the Scottish branch of the British Red Cross Society, but permission has kindly been given for the use of their results. Separate reports have been, or will be, published independently by these various workers. The present summary expresses opinions agreed upon at all four centres. The practical conclusions are based directly on the evidence obtained there, but it will be evident to anyone familiar with American work that they are not widely dissimilar from those accepted in New York and Boston.

During the two winter seasons 1931–2 and 1932–3 a total of 773 cases of lobar pneumonia between the ages of 20 and 60 years were studied at Aberdeen, Edinburgh, and London. Of these, 530 belonged to either Type I or Type II, and 241 were treated by serum. The figures for Glasgow for 1930–3 were 609

agreed that a clear result under the microscope is more often obtained when rabbit sera are employed. The Glasgow observers, however, have compared the results with concentrated horse serum and with rabbit serum, and conclude that these are equally reliable, provided that the serum, whether of horse or rabbit, has an agglutinin titre of, say, 1 in 160 to 1 in 320.

In every instance in the present series, excepting a few of Dr. Armstrong's group, the final typing was recorded on the result of animal tests, Sabin's method of examining the peritoneal exudate four hours after inoculation into mice being generally used, and if that failed the examination being completed later when the mouse was dead or moribund. Whatever the method used, some special experience is required for accurate results, and it should be ascertained that the diagnostic sera are reliable. More than one specimen of sputum should be examined to confirm the type.

Serum Used

The therapeutic antisera were those made for the market either by the Lederle Antitoxin Laboratories, or Messrs. Parke, Davis and Co., or Messrs. Burroughs Wellcome and Co. The Council and the investigators are indebted to these firms for special facilities given in the supply of the sera. Almost all the observations were made with concentrated serum, the power of which to protect mice had been measured in the American Felton units.

It is clearly desirable that any further evidence as to the value of anti-pneumococcal sera shall be based on dosage expressed in stable and generally accepted units. During the progress of this investigation, action to facilitate such uniformity of notation in this country has been taken by the Standards Department of the National Institute for Medical Research. Suitable anti-pneumococcal sera of both Types I and II have been dried and are preserved under conditions ensuring permanence, the value of each having been measured in terms of the Felton unit, unofficially current in the U.S.A., by comparison with samples supplied for this purpose by Dr. Felton himself. Of these provisional British standards, that for Type I serum has for some time been on regular issue to the manufacturers of anti-pneumococcal sera for sale in this country, and a similar distribution of the Type II standard will shortly follow. Pending an international decision, and subsequent official action, this voluntary distribution should ensure that data relating to the dosage of anti-pneumococcal sera from different sources should be strictly comparable both here and in America, where a similar voluntary adoption of the Felton units is effective.

The serum has been given intravenously in all the cases under review. Intramuscular injection was found by some workers to cause painful swelling, and has the theoretical disadvantage that the antibodies may not be absorbed into the circulation as quickly as is desirable.

Ex
tre
m

THE SERUM TREATMENT OF LOBAR PNEUMONIA

A REPORT OF THE THERAPEUTIC TRIALS COMMITTEE OF THE MEDICAL RESEARCH COUNCIL

During the last three years the Medical Research Council have assisted an inquiry at different centres in Great Britain into the therapeutic value of specific sera for [illegible]

of practical medicine. The method consequently agreed upon for London, Edinburgh, and Aberdeen was that alternate cases of lobar pneumonia, taken simply in the order of their admission to hospital, should be used respectively for serum treatment and controls. So far

[illegible] who so willingly undertook this prolonged study, and brought their results together for joint consideration.

Interest in the treatment of pneumonia has for many years been felt more keenly in Scotland, where the disease is perhaps more prevalent in the winter months, than in England. For evidence used in the present report the committee are indebted to workers in Aberdeen, Edinburgh, Glasgow, and London. The observations at Edinburgh were made by Professor D. Murray Lyon and the other physicians to the Royal Infirmary; those at Aberdeen by Professor Stanley Davidson, Dr. J. B. Ewen, and Dr. R. J. Duthie, in the City Hospital, Woodend; and those in London by Dr. R. R. Armstrong and Dr. R. Sleigh Johnson in various London County Council hospitals and also at St. Bartholomew's Hospital. The Glasgow inquiry under Dr. John Cowan, Dr. A. W. Harrington, and Dr. R. Cruickshank was developed independently with support from the Scottish branch of the British Red Cross Society, but permission has kindly been given for the use of their results. Separate reports have been, or will be, published independently by these various workers. The present summary expresses opinions agreed upon at all four centres. The practical conclusions are based directly on the evidence obtained there, but it will be evident to anyone familiar with American work that they are not widely dissimilar from those accepted in New York and Boston.

During the two winter seasons 1931–2 and 1932–3 a total of 773 cases of lobar pneumonia between the ages of 20 and 60 years were studied at Aberdeen, Edinburgh, and London. Of these, 530 belonged to either Type I

agreed that a clear result under the microscope is more often obtained when rabbit sera are employed. The Glasgow observers, however, have compared the results with concentrated horse serum and with rabbit serum, and conclude that these are equally reliable, provided that the serum, whether of horse or rabbit, has an agglutinin titre of, say, 1 in 160 to 1 in 320.

In every instance in the present series, excepting a few of Dr. Armstrong's group, the final typing was recorded on the result of animal tests, Sabin's method of examining [illegible]

[illegible] Messrs. Burroughs Wellcome and Co. The Council and the investigators are indebted to these firms for special facilities given in the supply of the sera. Almost all the observations were made with concentrated serum, the power of which to protect mice had been measured in the American Felton units.

It is clearly desirable that any further evidence as to the value of anti-pneumococcal sera shall be based on dosage expressed in stable and generally accepted units. During the progress of this investigation, action to facilitate such uniformity of notation in this country has been taken by the Standards Department of the National Institute for Medical Research. Suitable anti-pneumococcal sera of both Types I and II have been dried and are preserved under conditions ensuring permanence, the value of each having been measured in terms of the Felton unit, unofficially current in the U.S.A., by comparison with samples supplied for this purpose by Dr. Felton himself. Of these provisional British standards, that for Type I serum has for some time been on regular issue to the manufacturers of anti-pneumococcal sera for sale in this country, and a similar distribution of the Type II standard will shortly follow. Pending an international decision, and subsequent official action, this voluntary distribution should ensure that data relating to the dosage of anti-pneumococcal sera from different sources should be strictly comparable both here and in America, where a similar voluntary adoption of the Felton units is effective.

The serum has been given intravenously in all the cases under review. Intramuscular injection was found by some workers to cause painful swelling, and has the theoretical disadvantage that the antibodies may [illegible]

Ex
tre
m

THE SERUM TREATMENT OF LOBAR PNEUMONIA

A REPORT OF THE THERAPEUTIC TRIALS COMMITTEE OF THE MEDICAL RESEARCH COUNCIL

During the last three years the Medical Research Council have assisted an inquiry at different centres in Great Britain into the therapeutic value of specific sera for [illegible]

agreed that a clear result under the microscope is more often obtained when rabbit sera are employed. The Glasgow observers, however, have compared the results with concentrated horse serum and with rabbit serum, and conclude that these are equally reliable, provided that the serum, whether of horse or rabbit, has an agglutinin titre of, say, 1 in 160 to 1 in 320.

In every instance in the present series, excepting a few of Dr. Armstrong's group, the final typing was recorded on the result of animal tests, Sabin's method of examining [illegible]

of practical medicine. The method consequently agreed upon for London, Edinburgh and Aberdeen was that

The researchers found that

- There were far fewer who were given than were not given the treatment (the numbers should've been equal!)
- The average age was lower among those who received the treatment (should've been equal!)

BMJ. 1999 Nov 20; 319(7221): 1372.

Ewen, and Dr. R. J. Duthie, in the City Hospital, Woodend; and those in London by Dr. R. R. Armstrong and Dr. R. Sleigh Johnson in various London County Council hospitals and also at St. Bartholomew's Hospital. The Glasgow inquiry under Dr. John Cowan, Dr. A. W. Harrington, and Dr. R. Cruickshank was developed independently with support from the Scottish branch of the British Red Cross Society, but permission has kindly been given for the use of their results. Separate reports have been, or will be, published independently by these various workers. The present summary expresses opinions agreed upon at all four centres. The practical conclusions are based directly on the evidence obtained there, but it will be evident to anyone familiar with American work that they are not widely dissimilar from those accepted in New York and Boston.

During the two winter seasons 1931–2 and 1932–3 a total of 773 cases of lobar pneumonia between the ages of 20 and 60 years were studied at Aberdeen, Edinburgh, and London. Of these, 530 belonged to either Type I

Institute for Medical Research. Suitable anti-pneumococcal sera of both Types I and II have been dried and are preserved under conditions ensuring permanence, the value of each having been measured in terms of the Felton unit, unofficially current in the U.S.A., by comparison with samples supplied for this purpose by Dr. Felton himself. Of these provisional British standards, that for Type I serum has for some time been on regular issue to the manufacturers of anti-pneumococcal sera for sale in this country, and a similar distribution of the Type II standard will shortly follow. Pending an international decision, and subsequent official action, this voluntary distribution should ensure that data relating to the dosage of anti-pneumococcal sera from different sources should be strictly comparable both here and in America, where a similar voluntary adoption of the Felton units is effective.

The serum has been given intravenously in all the cases under review. Intramuscular injection was found by some workers to cause painful swelling, and has the theoretical disadvantage that the antibodies may

The Cochrane RoB-tool emphasises five domains*

- bias arising from the randomization process;
- bias due to deviations from intended interventions;
- bias due to missing outcome data;
- bias in measurement of the outcome; and
- bias in selection of the reported result.

* Developed especially for RoB-assessment of RCTs, but the same approach can, in prinicple, be applied on all types of effectiveness studies.

The Cochrane RoB-tool emphasises five domains*

- bias arising from the randomization process;
  - Allocation sequence random?
  - Allocation adequately concealed?
- bias due to deviations from intended interventions;
  - Blinding of participants?
  - Blinding of personell?
- bias due to missing outcome data;
  - Is data lacking?
  - Differences in (reasons for) missing data between groups?
- bias in measurement of the outcome;
  - Outcome assessors blinded to allocation?
- bias in selection of the reported result.

* Developed especially for RoB-assessment of RCTs, but the same approach can, in prinicple, be applied on all types of effectiveness studies.

Trial to evaluate use of antibiotics

Randomization-process | Deviations from intended interventions | Missing data | Outcome measurement | Selective reporting?

319 children with acute otitis media

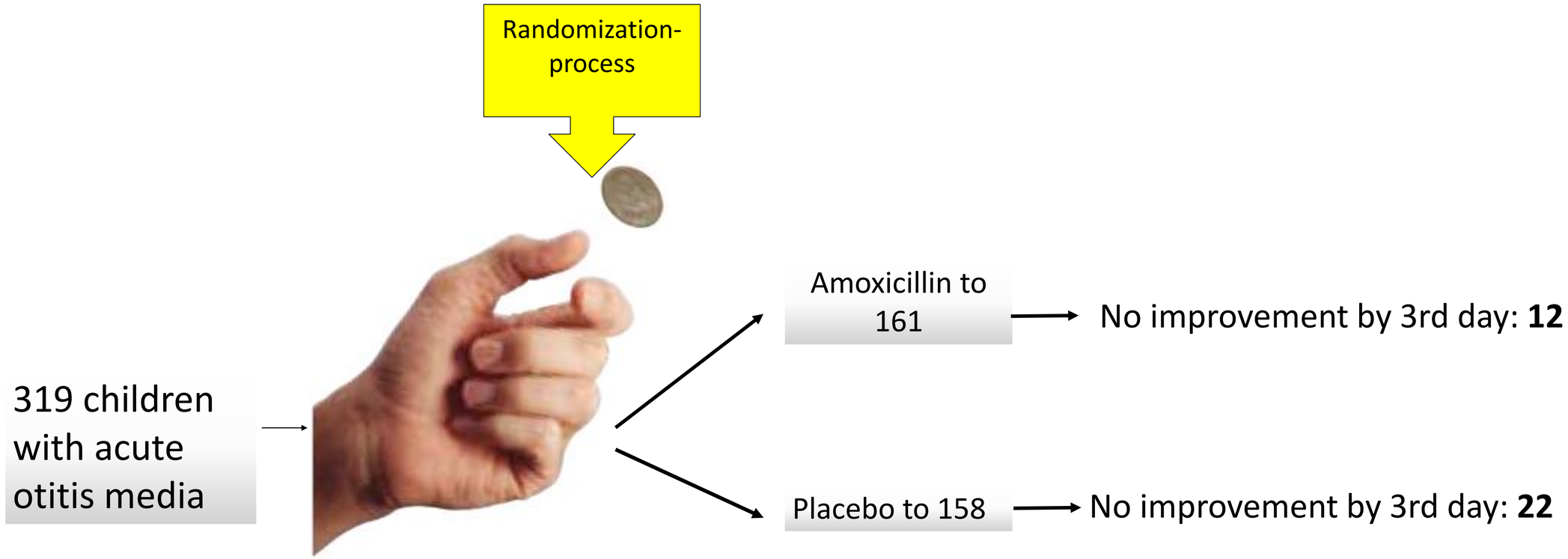
Amoxicillin to 161 → No improvement by 3rd day: **12**

Placebo to 158 → No improvement by 3rd day: **22**

Other risks of bias?

Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

Trial to evaluate use of antibiotics



Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

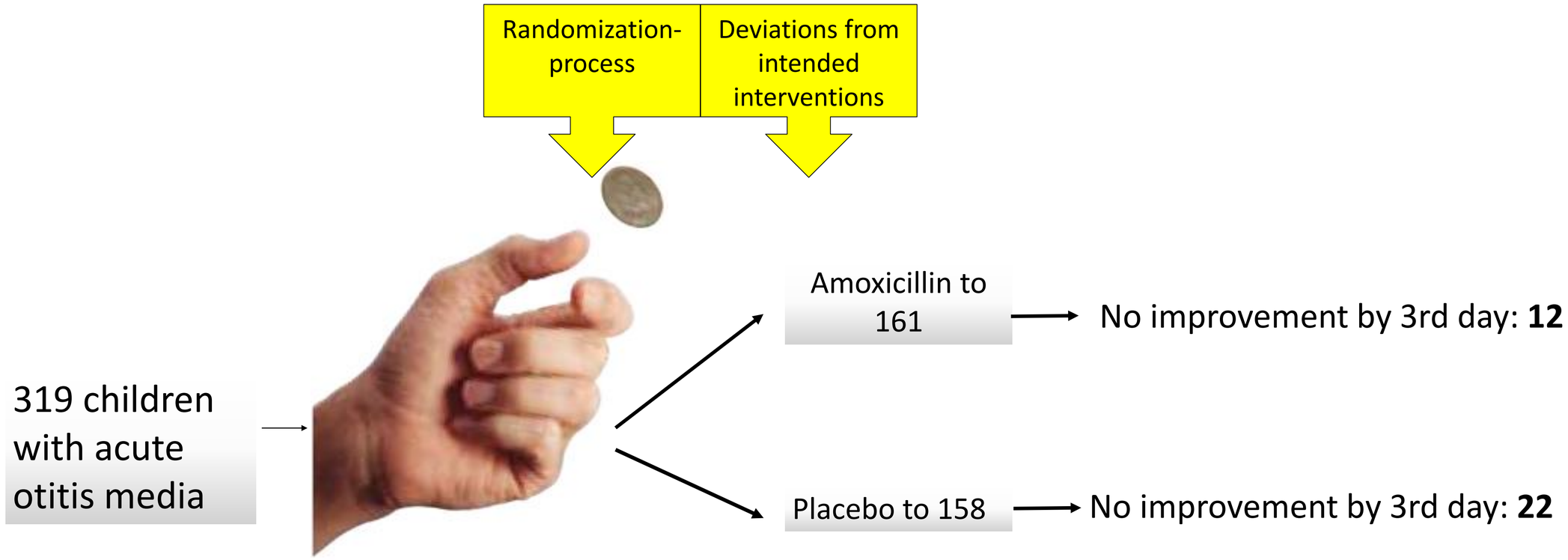
Sequence generation and concealed allocation

- *Sequence generation:* Method used to decide the order of allocation (e.g. coin tossing, pre-made randomisation liste, every other etc.)
- *Concealed allocation:* None of those involved know which group the next participant will end up in (until the participant is included in the study)

Why is the sequence generation and concealed allocation important?

- To ensure comparable groups from the start, i.e. prevent allocation bias
- Sequence generation and allocation concealment are inter-linked
  - Non-random allocation makes it difficult to achieve concealed allocation
  - Unconcealed allocation can undermine randomisation

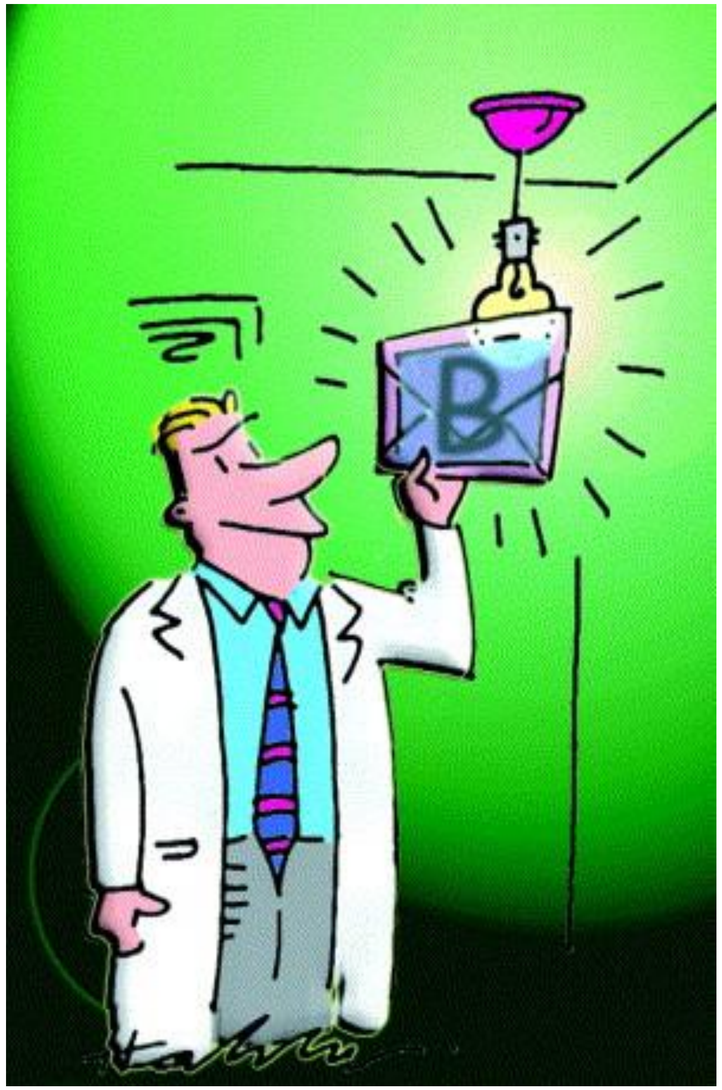
Trial to evaluate use of antibiotics



Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

Concealed allocation vs. blinding

- *Concealed allocation:* No one knows which group the next participant will be allocated to (before the participant is included in the study)
- *Blinding:* Neither personnel, participants or those who assess outcomes are aware of to which group the participants are allocated to (also after inclusion in the trial)



Blinding is important to avoid deviation from the plan

- Was the group affiliation of the participants kept secret in an adequate way during the study,
  - for the participants?
  - for the personnel?

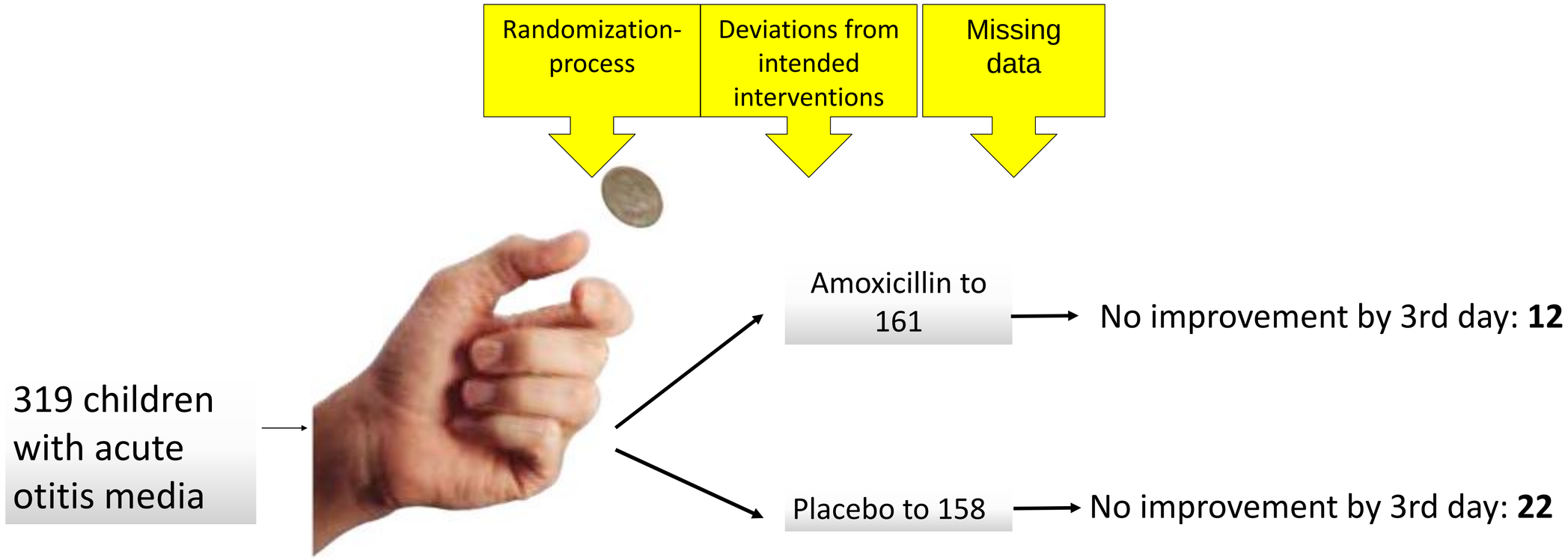
TAKE NOTE! Blinding can vary from one outcome to another within the same study (both whether blinding was done, and if lack of blinding constitutes a risk of bias).

Why is blinding important?

- If the participants know which treatment they’re receiving, that can in itself affect the result (e.g. due to expectations)

- Knowing which treatment has been given can also affect the assessment of outcomes, by the participant him-/herself, or others who assess the treatment result (will return to that)

Trial to evaluate use of antibiotics



Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

Missing data

- Were missing data managed in a satisfactory way (participants who didn’t meet or dropped out)?
- If data were not included in the analyses (exclusion of participants), was a good reason given?
- «Intention to treat» (ITT) is a key concept…
  - Optimal ITT, outcomes are assessed on all participants, and all are included in all analyses

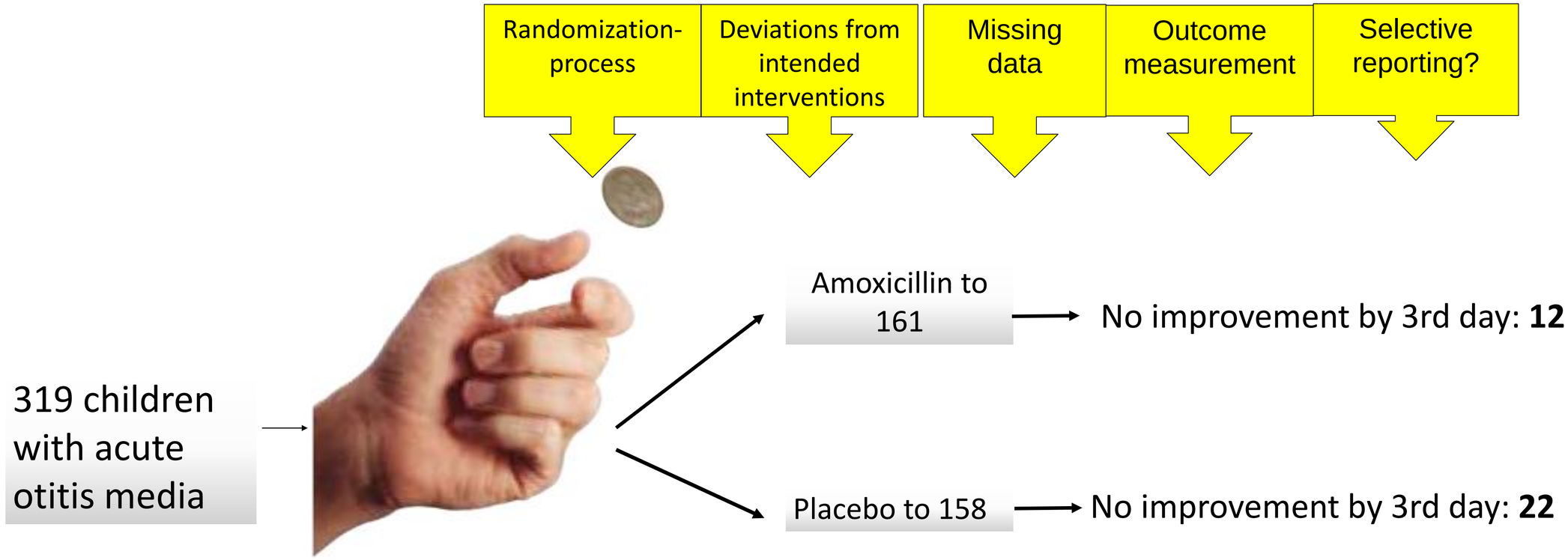
TAKE NOTE! The degree of incomplete follow up may vary from one outcome to another

Why is incomplete follow up a potential problem?

- The fate of participants who dropped out or who were excluded during the course of the study, may have an impact on the results – if they had been included in the analysis (e.g. did they drop out because they fell ill?)

	Number randomised	Risk among observed	Observed data	Hypothetical extreme risks among missing	Missing data	Complete data	Risk ratio
Intervention	500	**10%**	45/450	**80%**	40/50	85/500	**1.55**
Control	500	**10%**	45/450	**20%**	10/50	55/500	

Trial to evaluate use of antibiotics

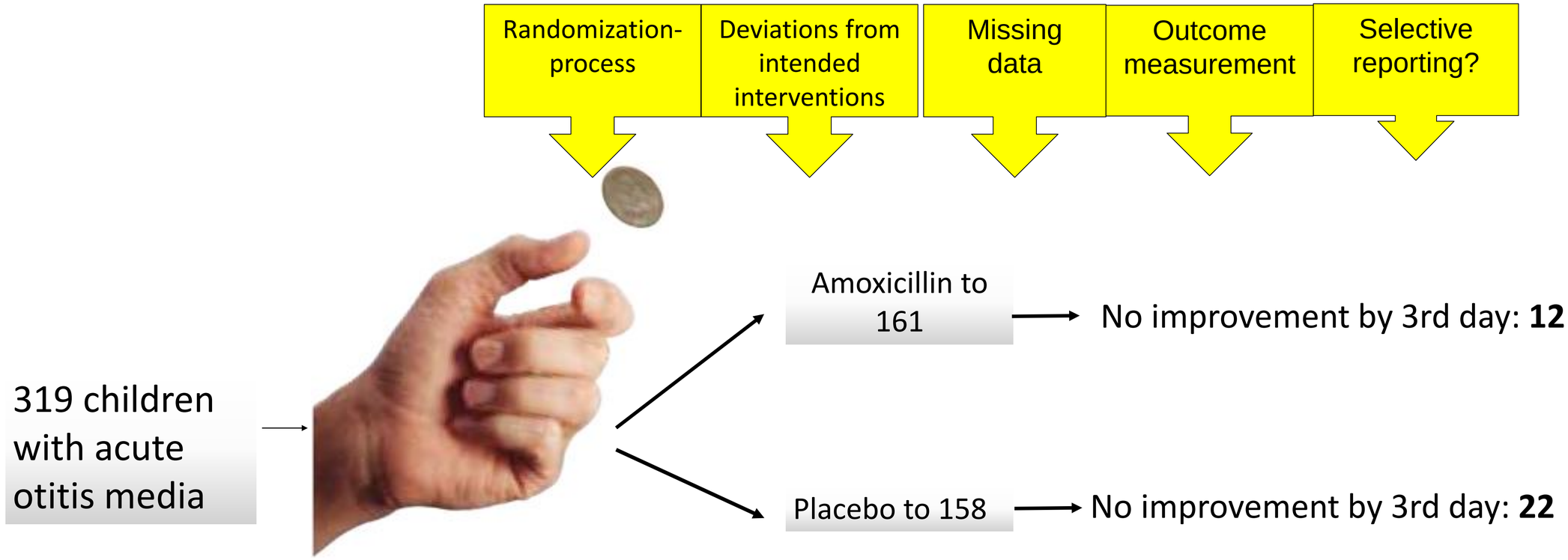


Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

Outcome measurement

- Were those who assessed outcomes unaware of allocation?

Trial to evaluate use of antibiotics



Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

Selective reporting

- Are there no signs of selective reporting of results?

Why is it important to assess the risk of selective reporting?

- It’s been shown that researchers often choose to report only *some* of their findings – typically «positive» results
- This leads to bias in systematic reviews, since usually only results that are reported are included
  - I.e. we risk being fooled if we only see the «positive» results

Trial to evaluate use of antibiotics

Randomization-process | Deviations from intended interventions | Missing data | Outcome measurement | Selective reporting?

319 children with acute otitis media

Amoxicillin to 161 → No improvement by 3rd day: **12**

Placebo to 158 → No improvement by 3rd day: **22**

Other risks of bias?

Source: Tähtinen et al, N Engl J Med 2011; 364:116-126

RoB-assessmen entails two steps

1. Figure out how the study was conducted (what did they do – what happened?)
2. Decide for yourselves whether this entails low, high or unclear risk of bias