Sustavni pregledi: snaga i ograničenja

Trkulja, Vladimir

1917–2017
100 godina Medicinskog fakulteta Sveučilišta u Zagrebu

Vladimir Trkulja

Zavod za farmakologiju / Department of pharmacology

Sustavni pregledi: snaga i ograničenja

Systematic reviews: strengths and limitations

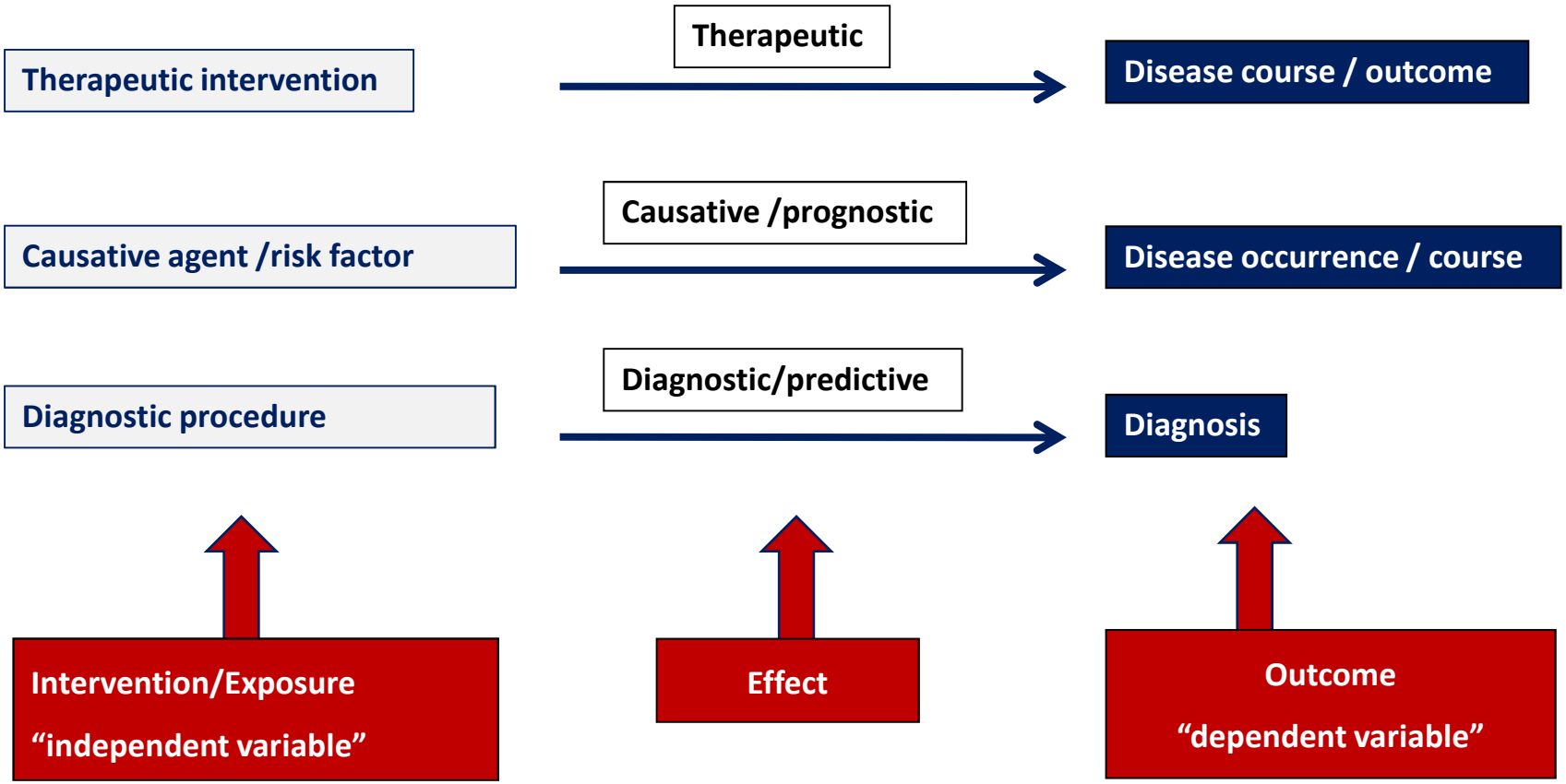
A. Basic terminology

B. The underlying logic – reasons to do it

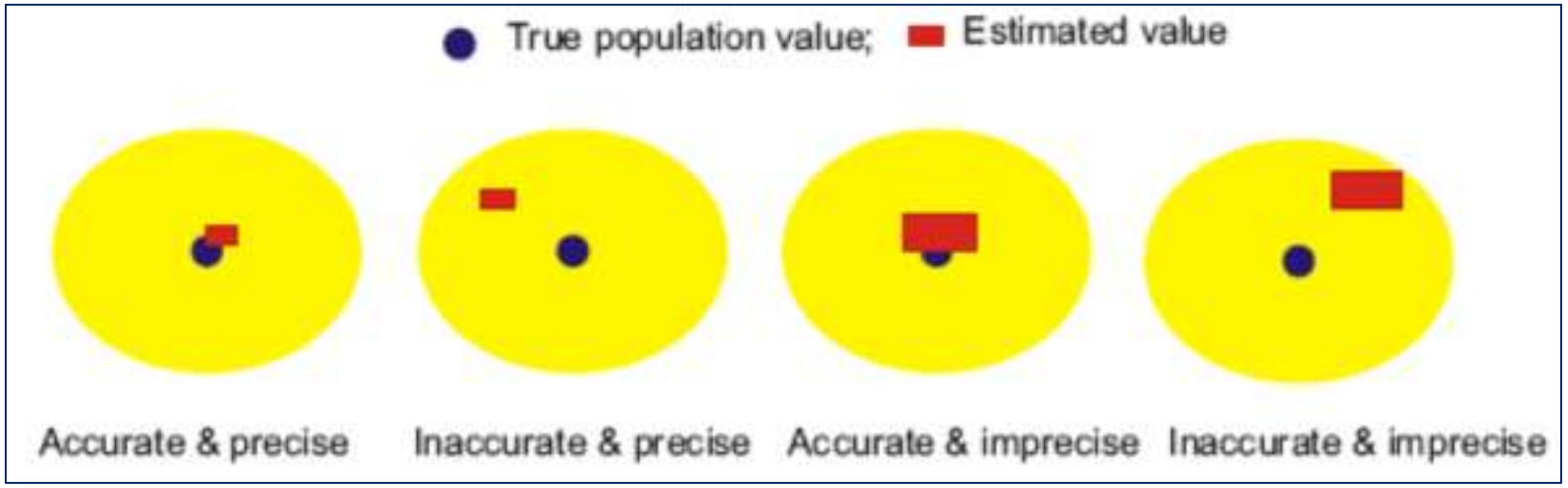
C. The potential (strengths)

D. Limitations and misconceptions

The entire (bio)medical research is about detecting/defining 3 relationships:



It is about estimation of the physical world that surrounds us, or in other words - about estimating the “population value”

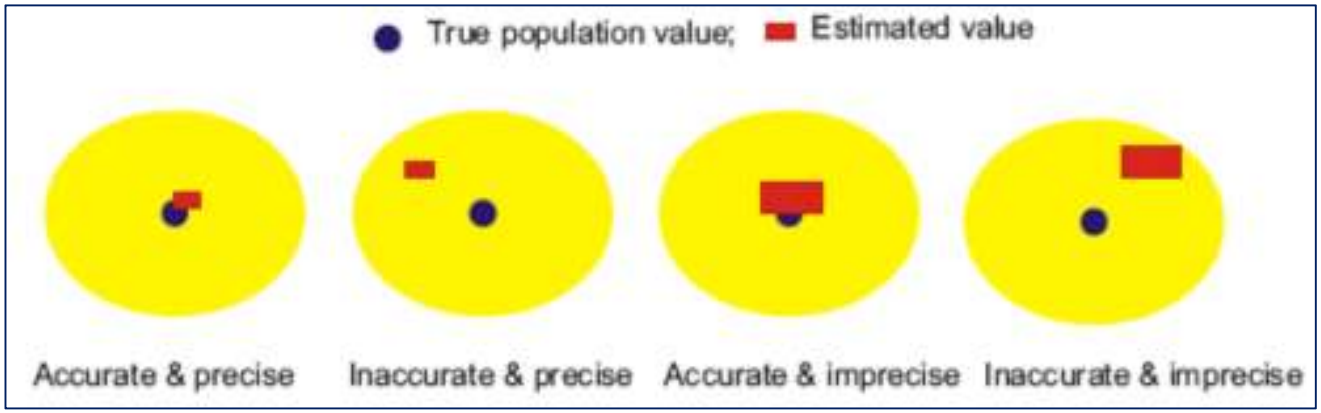


So that we could make ***population-wise*** statements/claims that govern daily medical practice:

- Treatment T cures disease D – hence, we should use it to treat the patients
- These (….) factors contribute to occurrence of this specific disease – hence we should preventively intervene
- This diagnostic test is the most reliable one for this specific condition – hence, this should be our first choice during the diagnostic process

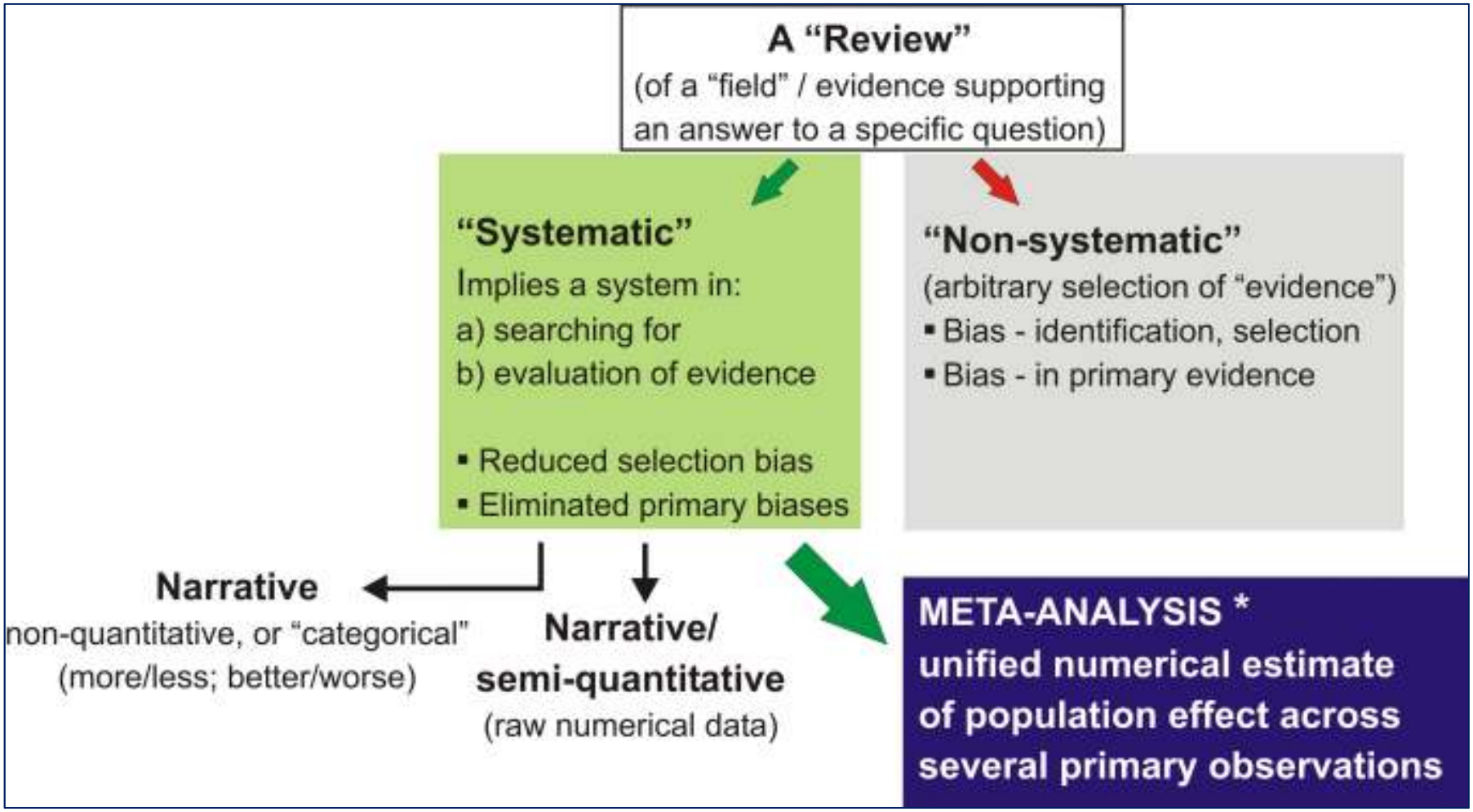
Clearly, in order to "***do more good than harm***" (Archie Cochrane -☺-):

- These estimates need to be ACCURATE – i.e., ON TARGET
- These estimates need to be (preferably) – PRECISE



- INACCURATE (off-target) or BIASED estimates – RESLUT IN **HARM**
- IMPRECISE estimates – are **inconclusive** (uninformative)

A. Basic terminology



* **Glass GV**. Primary, secondary and ***meta-analysis*** of research. *Educat Res 1976; 5:3-8.*

B. The underlying logic – reasons to do it

C. The potential (strengths)

D. Limitations and misconceptions

Lord Rayleigh (in 1884)*

"If, as it is sometimes supposed, science consisted in nothing but the laborious accumulation of facts, it would soon come to a standstill, crushed, as it were, under its own weight. The suggestion of a new idea, or the detection of a law, supersedes much that has previously been a burden on the memory, and by introducing order and coherence facilitates the retention of the remainder in an available form. ..Two processes are thus at work side by side…The work which deserves, but I am afraid does not always receive, the most credit is that in which discovery and explanation go hand in hand, in which not only are new facts presented, but their relation to old ones is pointed out."

* Systematic reviews in Health Care; meta-analysis in context. 2nd edition, Egger, Davy Smith, Altman (eds), BMJ Books 2001

Karl Pearson (in 1904)*

REPORT ON CERTAIN ENTERIC FEVER INOCULATION STATISTICS.

PROVIDED BY LIEUTENANT-COLONEL R. J. S. SIMPSON, C.M.G., R.A.M.C.

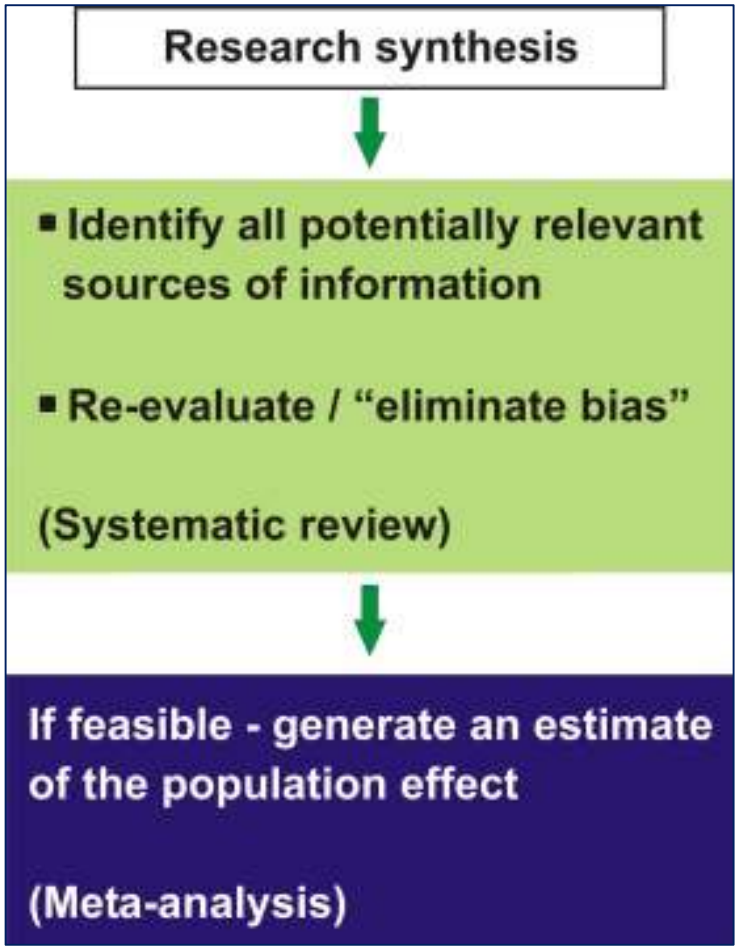
BY KARL PEARSON, F.R.S.,
Professor of Applied Mathematics, University College, London.

(why he jointly analyzed data from several studies on the preventive effect of serum inoculation against enteric fever)

*"Many of the groups….are far too small to allow of any **definite opinion** being formed at all, having regard to the size of the probable error."*

* BMJ 1904; 3:1243-1246.

B. Why do it?



Why?

- To generate more ACCURATE (unbiased) estimates
- To generate more PRECISE estimates
- To see "where we are and what to do next"

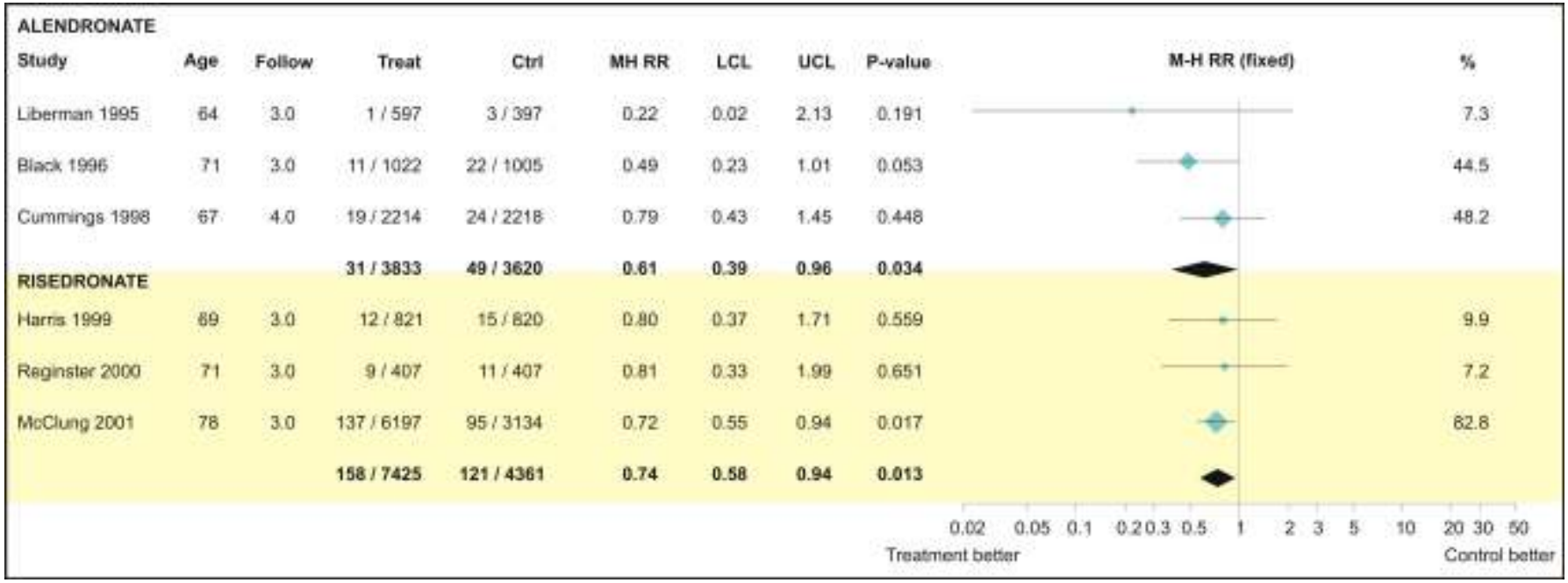
C. The potential (strengths)

D. Limitations and misconceptions

- *Legendre* (1805) and *Gauss* (1809) **combined data** from different observatories to estimate the orbit of comets and to determine meridian arcs in geodesy*
- *Birge* (1932) **combined data** from different experiments to define fundamental constants in physics*
- *Cochrane* (1937) **combined data** from different experiments in agriculture (fertilizers)*
- **Combined data** – ecological studies, market research, industry/technology*

* Hartung, Knapp, Sinha. Statistical meta-analysis with applications. Wiley 2008.

Increased precision – detect an effect not otherwise obvious (a simple example)



ALENDRONATE Study	Age	Follow	Treat	Ctrl	MH RR	LCL	UCL	P-value	M-H RR (fixed)	%
Liberman 1995	64	3.0	1 / 597	3 / 397	0.22	0.02	2.13	0.191		7.3
Black 1996	71	3.0	11 / 1022	22 / 1005	0.49	0.23	1.01	0.053		44.5
Cummings 1998	67	4.0	19 / 2214	24 / 2218	0.79	0.43	1.45	0.448		48.2
			31 / 3833	**49 / 3620**	**0.61**	**0.39**	**0.96**	**0.034**		
RISEDRONATE										
Harris 1999	69	3.0	12 / 821	15 / 820	0.80	0.37	1.71	0.559		9.9
Reginster 2000	71	3.0	9 / 407	11 / 407	0.81	0.33	1.99	0.651		7.2
McClung 2001	78	3.0	137 / 6197	95 / 3134	0.72	0.55	0.94	0.017		82.8
			158 / 7425	**121 / 4361**	**0.74**	**0.58**	**0.94**	**0.013**		

Increased precision – detect an effect not otherwise obvious*

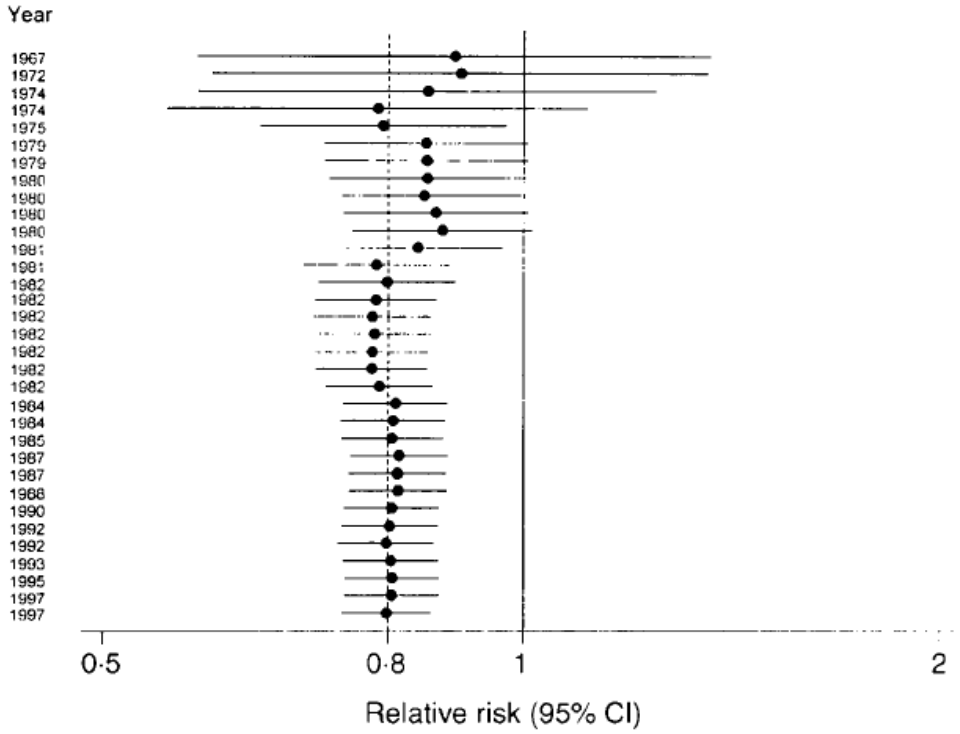
- A narrative review in BMJ in 1981 concluded about beta-blockers after myocardial infarction

> *"Thus, despite claims that they reduce arrhythmias, cardiac work, and infarct size, we still have no clear evidence that beta-blockers improve long-term survival after infarction despite almost 20 years of clinical trials"*

* Systematic reviews in Health Care; meta-analysis in context. 2nd edition, Egger, Davy Smith, Altman (eds), BMJ Books 2001

Increased precision – detect an effect not otherwise obvious*

- And actually, had anyone performed meta-analysis in 1981 – it would have been OBVIOUS that beta-blockers were effective



* Systematic reviews in Health Care; meta-analysis in context. 2nd edition, Egger, Davy Smith, Altman (eds), BMJ Books 2001

Classical examples of **detection of otherwise not obvious effects** through systematic reviews/meta-analysis that dramatically changed medical practice or health care policies

- Efficacy of ***secondary prevention of arterial thrombotic incidents*** (stroke, AMI) with antiplatelets (in 1988; RCTs, individual patient data meta-analysis)
- Efficacy of ***streptokinase in reducing AMI mortality*** (in 1992; RCTs)
- Use of ***antenatal corticosteroids to accelerate fetal lung maturation*** in women at risk of preterm birth (in 2003; RCTs)
- Passive smoking ***increases the risk of lung cancer*** (in 1990; observational studies)

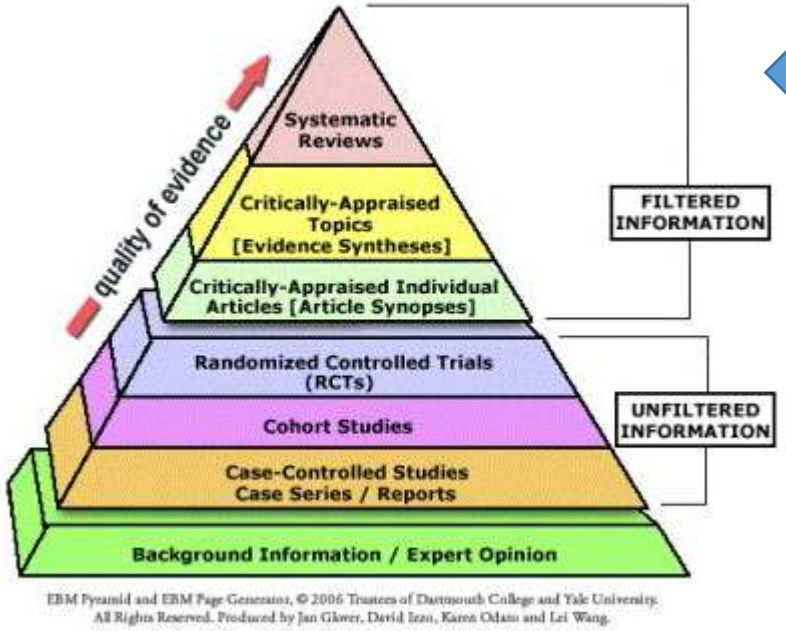
D. Limitations and misconceptions

D. Limitations & misconceptions

1. *Systematic reviews/meta-analyses are not "true research", rather a "parasitical work" by which some take the advantage from a "true work" of others*

- Critics should go back to 1880s and listen what Lord Rayleigh had to say (-☺-)
- SR/MA have all the elements of (biomedical) research:
  - A well-defined research question
  - Data acquisition
  - Data analysis
  - And interpretation
- Examples demonstrate how consideration of (filtered) evidence could be enlightening

2. *Whatever comes from a systematic review/meta-analysis "must be accurate" and is undisputable (i.e., a "top level evidence")*



Indeed, we have placed SR/MA at the top of the evidence-base pyramid

D. Limitations & misconceptions

In fact, the method HAS A POTENTIAL to be that ("top of the pyramid")

- Because all relevant data might be considered (if search adequate)
- Primary studies are re-evaluated using validated instruments ("remove bias")
- Data analysis might "correct" some errors from primary trials

HOWEVER

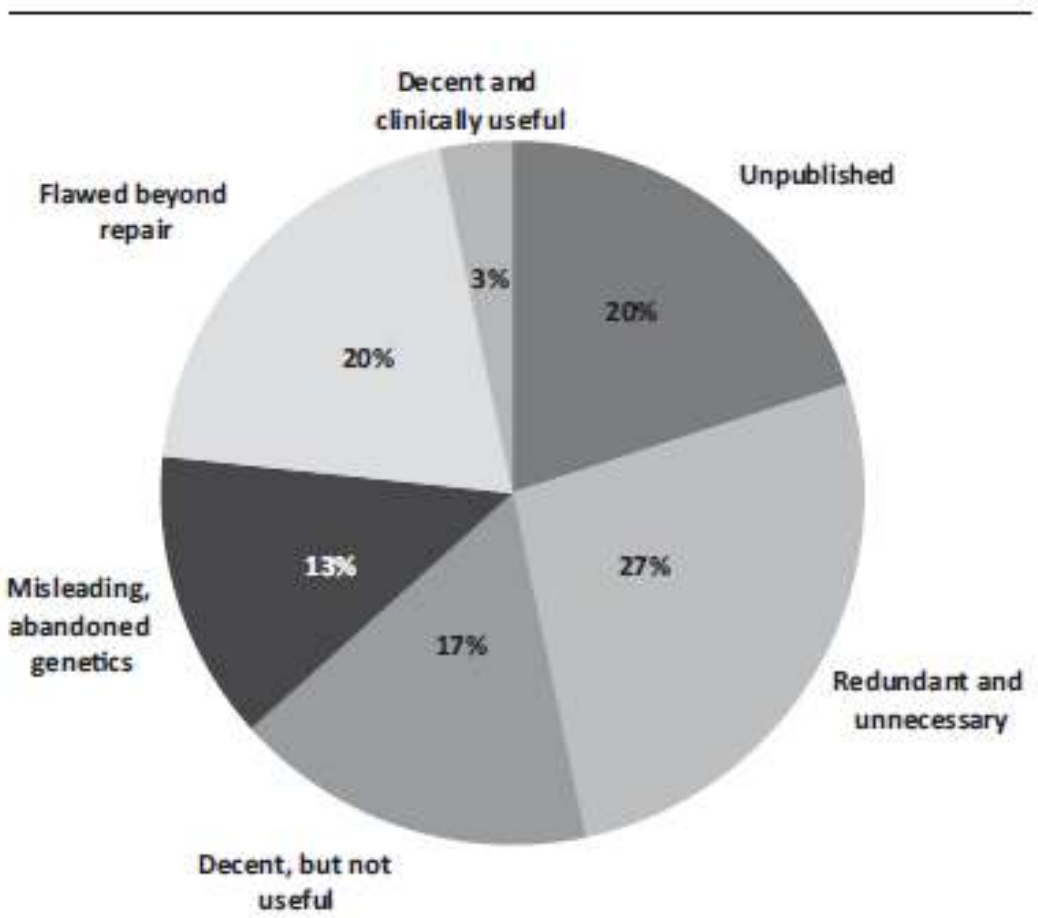
- Always *post-hoc* – cannot mend major flaws in primary studies
- Might be methodologically flawed at every level
- Overall – the results might be MISLEADING (off-target)

D. Limitations & misconceptions

1917 – 2017
100 godina Medicinskog fakulteta Sveučilišta u Zagrebu

PRIMUM NON NOCERE

- >9000 SR/MA in (bio)medicine published in 2014 (Milbank Quarterly 2016; 94:485)



D. Limitations & misconceptions

1917 – 2017
100 godina Medicinskog fakulteta Sveučilišta u Zagrebu

PRIMUM NON NOCERE

- Methodologically flawed (e.g.) (Croat Med J 2014; 55:468-480)

TABLE 2. Quality of the included reviews based on the AMSTAR (26) checklist*

	Einarson 2000 (28)	Zhang 2001 (29)	v.d Valk 2005 (30)	Li 2006 (31)	Denis 2007 (32)	Fung 2007 (33)	Aptel 2008 (34)	Cheng 2008 (35)	Hodge 2008 (36)	Loon 2008 (37)	Cheng 2009a (38)	Cheng 2009b (39)	Ejawo 2009 (40)	Hornubia 2009 (41)	v.d. Valk 2009 (42)	Orme 2010 (43)
Design *"a priori"*?	?	Y	?	Y	Y	Y	Y	Y	Y	Y	Y	?	?	Y	Y	Y
Duplicate selection/ extraction?	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y‡	?
Comprehensive search?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y‡	Y
Publication status clear?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y‡	Y
List included/ex-cluded provided?†	?	?	?	Y	?	?	?	Y	?	?	?	?	?	Y	?	Y
Study characteris-tics provided?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y‡	Y
Quality assessed?	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y‡	N
Quality accounted for conclusions?	Y	Y	Y	N	N	Y	Y	N	Y	Y	Y	Y	Y	N	N	N
Appropriate meth-od for pooling?	N	N	N	?	N	Y	?	?	Y	N	N	N	Y	?	?	Y
Publication bias assessed?	N	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	?	Y
Conflict of interest declared?	Y	N	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y
AMSTAR score	7	7	8	9	6	10	9	8	9	8	8	8	9	9	7	8

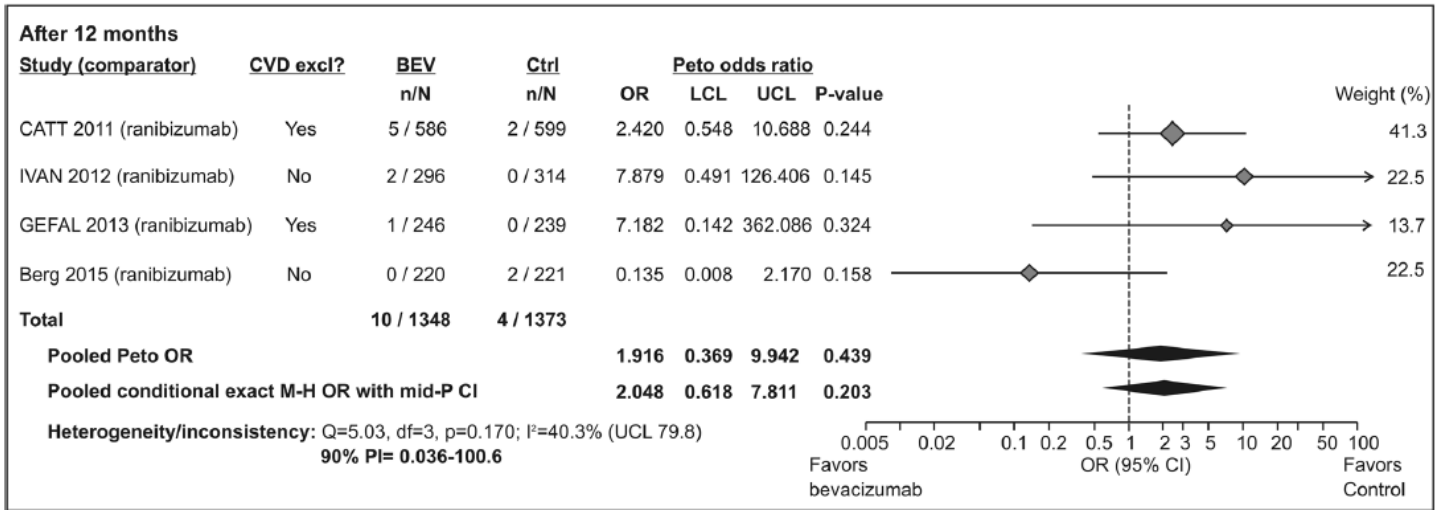
*Abbreviations: Y – yes; N – No; ? – can't tell; NA – not applicable.
†All reviews reported on included studies, but only 3 reported also on excluded studies. Hence, most reviews failed to meet this quality criterion.
‡Described in the previous publication [v.d. Valk 2005 (30)].

- Typical examples of methodological flaws
  - Inadequate/incomplete search
  - Disregarding primary study quality
  - Lack of understanding of “random-effects” and “fixed effect” concepts
  - Lack of understanding of “heterogeneity” and “inconsistency”
  - Lack of understanding of “clinical consistency”
  - Naïve data pooling (disregarding randomization)
  - Erroneous post- vs. pre- difference calculation
  - Treating observational studies as RCTs – using raw data instead of adjusted estimates
  - Inadequate methods for sparse data

  Etc.

D. Limitations & misconceptions

- A lot of limitations arises from MISINTEPRETATION (by authors or readers)



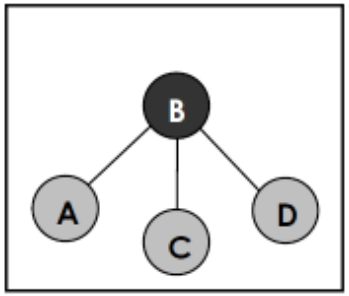
(Drug Saf 2016; doi.10.1007/s40264-016-0408-y)

- **Erroneous**: “there is trend of a higher VTE risk with bevacizumab”
- **Erroneous**: “there is no difference between bevacizumab and ranibizumab regarding VTE”
- **Appropriate**: We have no idea!

D. Limitations & misconceptions

- A lot of limitations arises from MISINTEPRETATION (by authors or readers)

Indirect: star with pair-wise contrasts



Comparison	Progression-Free Survival HR (95% CI; 95% PI)
Gefitinib vs. chemotherapy	0.44 (0.31–0.63; 0.22–0.88)
Erlotinib vs. chemotherapy	0.25 (0.15–0.42; 0.11–0.55)
Afatinib vs. chemotherapy	0.44 (0.26–0.75; 0.20–0.98)
Erlotinib vs. gefitinib	0.57 (0.30–1.08; 0.24–1.36)
Afatinib vs. gefitinib	1.01 (0.53–1.92; 0.42–2.42)
Erlotinib vs. afatinib	0.56 (0.27–1.18; 0.22–1.46)

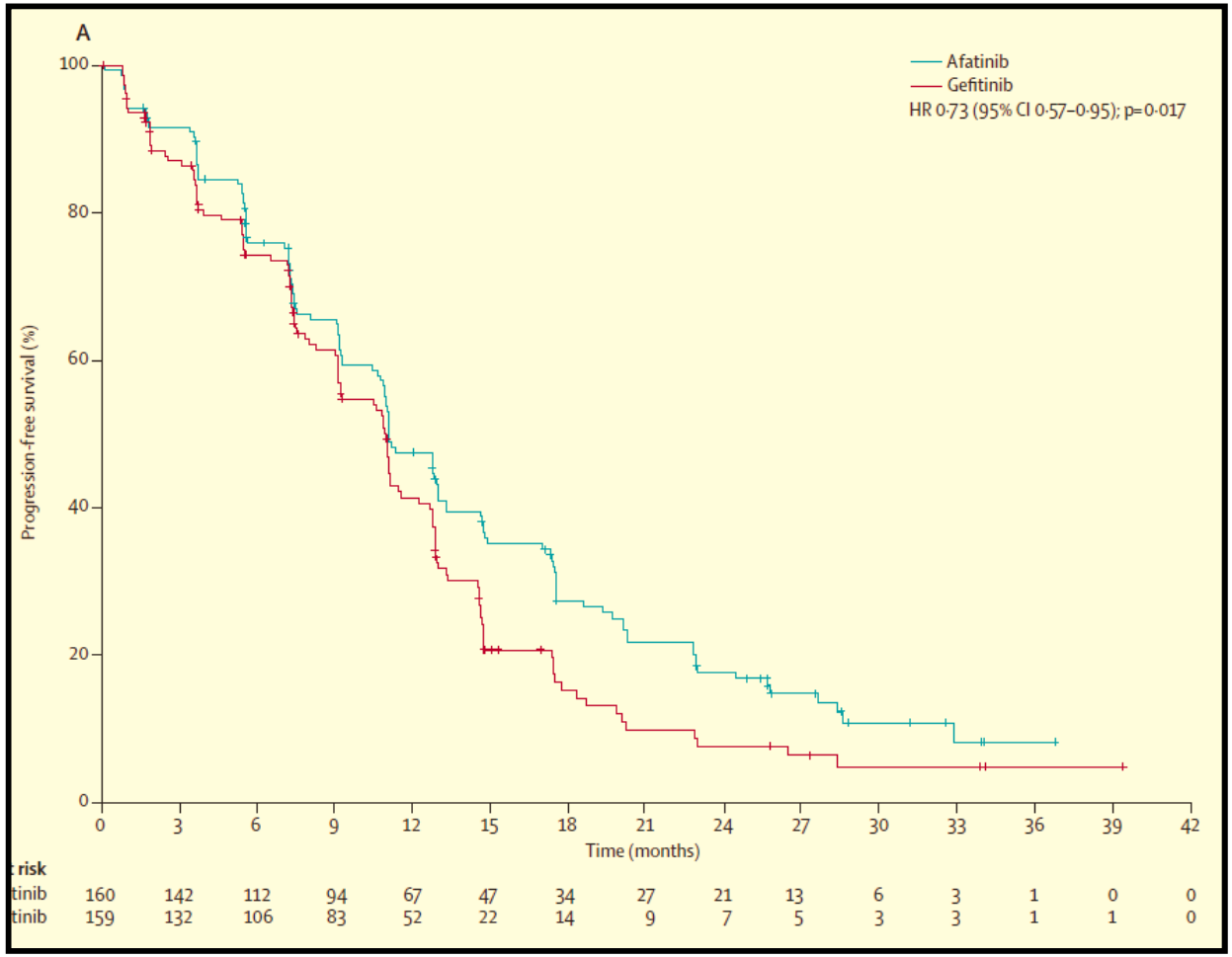
(J Thorac Oncol 2014; 9:805-811)

D. Limitations & misconceptions

1917 – 2017
100 godina Medicinskog fakulteta Sveučilišta u Zagrebu

PRIMUM NON NOCERE

- A lot of limitations arises from MISINTEPRETATION (by authors or readers)



(Lancet Oncol 2016; doi. 10.1016/S1470-2045(16)30033-x)

To CONCLUDE

- SR w/wo MA – is a **potentially** powerful tool
  - For more accurate/precise estimates of the population
  - For evaluation of accumulated evidence and definition of further goals

YET

- It is susceptible to bias and random error (as any research method)
- May be **uninformative** or **misleading**