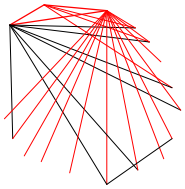
ISPE 2013 Mid-Year Meeting & Symposium

Hilton Munich City | Munich, Germany | April 11-13, 2013



Mixed Treatment Comparisons (MTC) – Concepts and Problems –

Ralf Bender, Sibylle Sturtz

Institute for Quality and Efficiency in Health Care (IQWiG), Cologne

Disclosures

No relationships to disclose

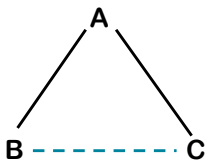
Agenda

- Introduction
- Concepts and methods
  - Simple adjusted indirect comparison
  - Mixed treatment comparison (MTC)
  - Basic assumptions
- Problems
  - Network size
  - Inconsistency
- Importance of MTC
- Conclusions

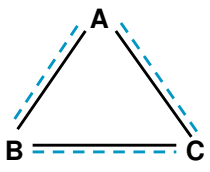
Idea

- Indirect comparison: IC
- Effect of intervention C relative to B: $d_{BC} = d_{AC} - d_{AB}$
- Mixed treatment comparison (MTC) meta-analysis
  (Also called: *Multiple treatment meta-analysis, Network meta-analysis*)

Lu & Ades , JASA 2006



IC



MTC

Introduction

Reasons for MTC

- For many clinical indications there are often several possible interventions
- Combined analysis of all relevant data is to be prefered for health care decisions

Problems solved by MTC

- Direct comparisons between active interventions A and B may not always be available
- Even if direct evidence is available, there may be only a few studies
- Direct evidence from separate pairwise comparisons cannot determine which of several interventions is most effective

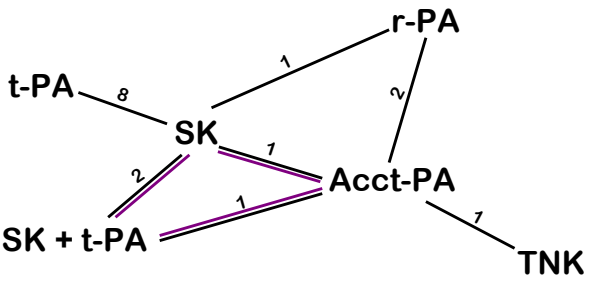
Introduction

Example: Thrombolysis

Boland et al., HTA 2003

6 treatments for acute myocardial infarction:

(1) Streptokinase (SK)
(2) Tissue plasminogen activator (t-PA)
(3) Accelarated alteplase (Acct-PA)
(4) Tenecteplase (TNK)
(5) Reteplase (r-PA)
(6) SK+t-PA



14 studies, 15 possible pairwise comparisons

Introduction

Thrombolysis: Results from Boland et al. (*HTA*, 2003)

- Pairwise comparisons performed by applying usual meta-analyses
- " … SK is as effective as t-PA …"
  " … TNK is as effective as Acct-PA …"
  " … r-PA is at least as effective as SK …"
- "… SK is as effective as, or inferior to Acct-PA … "
  "… r-PA is as effective as Acct-PA or not …"
- " … two further questions on indirect comparisons arise, whether TNK is superior to SK or not and whether r-PA is as effective as TNK or not …"

Introduction

Results Thrombolysis (MTC FE model)

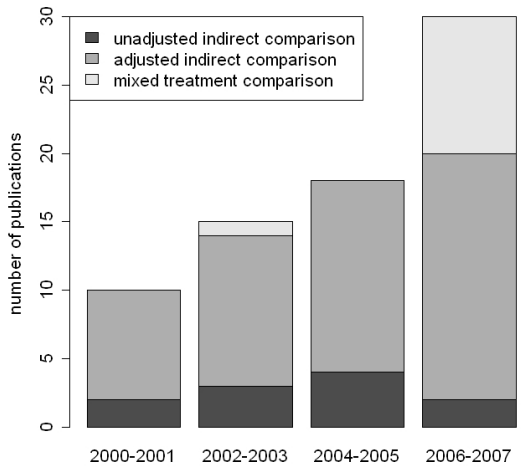
direct / MTC	SK	t-PA	Acct-PA	SK+t-PA	r-PA	TNK
SK	**	1,00 [0,94; 1,06]	0,86	0,96	0,95 [0,79; 1,12]	
t-PA	1,00 [0,94; 1,06]					
Acct-PA		0,87	**	1,12	1,02	1,01
SK + t-PA	0,96	0,97	1,11	**		
r-PA	0,90	0,91	1,04	0,94	**	
TNK	0,87 [0,74; 1,00]	0,88	1,01	0,91	0,97	**

- Consistent
- Possibly with narrower credibility intervals
- Additional information (indirect comparisons)

Probability of treatment *x* being best

	MTC FE model	
Treatment *x*	35 day mortality (%)	P(best)
(1) SK	6,5	0,0 %
(2) t-PA	6,4	0,0 %
(3) Acct-PA	5,6	40,0 %
(4) SK + t-PA	6,2	1,0 %
(5) r-PA	5,8	15,0 %
(6) TNK	5,6	43,0 %

Indirect comparisons are of increasing popularity



Schöttker et al., DIMDI 2009

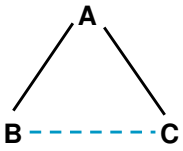
Naive Approaches

Unadjusted indirect comparisons

- Meta-analysis using data of single arms:
  MA using all data for intervention A,
  MA using all data for intervention B,
  MA using all data for intervention C, …
- Breaks randomisation
- Should never be used!
- A correct analysis has to be based on the estimated effects of each RCT

Bucher Approach

Simple adjusted indirect comparison



- No direct evidence available
- One mutual comparator
- On log-odds scale:

$$\hat{d}_{BC}^{indirect} = \hat{d}_{AC}^{direct} - \hat{d}_{AB}^{direct}$$

$$Var(\hat{d}_{BC}^{indirect}) = Var(\hat{d}_{AC}^{direct}) + Var(d_{AB}^{direct})$$

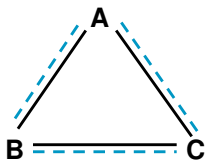
- Assumes independence of pairwise comparisons
- Extensions available for several direct comparisons linked by common comparators ('ladder' design)
- Not applicable for more complex networks

Bucher et al., JCE 1997
Wells et al., CADTH 2009

Frequentist Network Approaches

Lumley, Stat. Med. 2002
White et al., RSM 2012

Frequentist network meta-analysis



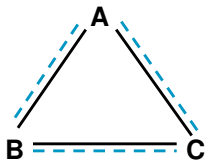
- Combination of direct and indirect evidence in a complete network
- At least one closed loop required
- Original approach by Lumley (2002) only for 2-arm trials
- Extended by White et al. (2012) to situation of multi-arm trials

- Computations:
  Any software for linear mixed models can be used (SAS, R, Stata etc.)

Bayesian Network Approaches

Lu & Ades, JASA 2006

Mixed treatment comparison (MTC) meta-analysis



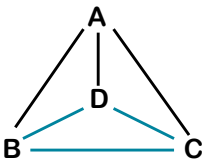
- Also called: Multiple treatment meta-analysis, network meta-analysis
- Combination of direct and indirect evidence in a complete network
- Applicable in all kinds of (connected) networks
- Can be applied to multi-armed trials
- Study level covariates can be incorporated
- Most flexible approach

- Bayesian approach requires specification of prior distributions

Bayesian Network Approaches

Lu & Ades, JASA 2006

Models for MTC: Basic & Functional Parameters



Take A as reference treatment

Treatment effects of B,C,D relative to A:
Basic parameters with priors

$$d_{Ak} \sim N(0,10000) \qquad k=B,C,D$$

Remaining contrasts: Functional parameters:

$$d_{BC} = d_{AC} - d_{AB}$$
$$d_{BD} = d_{AD} - d_{AB}$$
$$d_{CD} = d_{AD} - d_{AC}$$

Bayesian Network Approaches

Lu & Ades, JASA 2006

Fixed effects (FE) model

For intervention k in study j:

$$r_{jk} \sim Bin(p_{jk}, n_{jk})$$

$$logit(p_{jk}) = \begin{cases} \mu_{jb} & b = A,B,C \text{ if } k = b \\ \mu_{jb} + d_{bk} & \text{if k 'after' b} \end{cases}$$

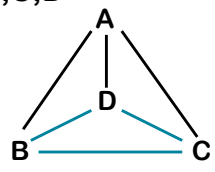
μ_{jb}: study specific effect of b

Functional p. $d_{bk} = d_{Ak} - d_{Ab}$

Basic p. $d_{Ak} \sim N(0,10000)$

$\mu_{jb} \sim N(0,10000)$

k=B,C,D



Bayesian Network Approaches

Lu & Ades, JASA 2006

Random effects (RE) model

For intervention k in study j:

$$r_{jk} \sim Bin(p_{jk}, n_{jk})$$

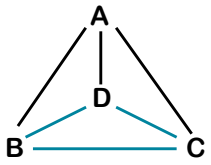
$$logit(p_{jk}) = \begin{cases} \mu_{jb} & b = A,B,C \text{ if } k = b \\ \mu_{jb} + \delta_{bk} & \text{if k 'after' b} \end{cases}$$

μ_{jb}: study specific effect of b

Functional p. $\delta_{jbk} \sim N(d_{bk}, \sigma^2) \sim N(d_{Ak} - d_{Ab}, \sigma^2)$

Basic p. $d_{Ak} \sim N(0,10000)$ k=B,C,D

$\mu_{jb} \sim N(0,10000)$



Bayesian Network Approaches

Computations: WinBUGS

- **NICE Decision Support Unit:**
  http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm
- **Multi-Parameter Evidence Synthesis Research Group**
  http://www.bris.ac.uk/social-community-medicine/projects/mpes/mtc/
  WinBUGS Code for MTC meta-analyses:
  - FE model
  - RE model for 2- and 3-arm trials
  - RE model for multi-arm trials

Basic assumptions for IC and MTC

Song et al., BMJ 2009

- **Similarity assumption:**
  Trials are similar concerning moderators of the relevant treatment effect
- **Homogeneity assumption:**
  Trials are sufficiently homogeneous to be quantitatively combined

⇨ **Same assumptions as for usual pairwise MA**

✚ **Consistency assumption:**
Direct and indirect evidence estimate the same effect

- **Definition of relevant interventions**
  - **Primary interventions**
  - **Comparators**
  - **Connecting interventions**
- **Information retrieval**
  - **Systematic literature search**
  - **Network may be never "complete"**
- **Assessment of assumptions**
  - **Similarity: PICO for the whole network**
  - **Homogeneity: Pairwise meta-analyses (forest plot, I^2, Q)**
  - **Consistency: In the framework of MTC meta-analysis**

Assessing similarity

- **Similarity assumption:**

Comparability of studies regarding possible effect modifiers across all interventions

- **PICO for the whole network**
- **Subjective evaluation of study characteristics**
- **Subgroup analyses**
- **Meta-regression**

Song et al., BMJ 2009

Assessing homogeneity

- **Homogeneity assumption:**

Sufficient homogeneity of effect estimates across all studies comparing interventions

- **Assessment of homogeneity in each pairwise meta-analysis**
- **Forest plots**
- **Tests for heterogeneity: Cochrans' Q**
- **Measures for heterogeneity: τ^2, I^2**

Assessing consistency

- **Consistency assumption:**

Comparability of effect estimates from direct and indirect evidence

- **Assessment of consistency within MTC meta-analysis**
- **Tests for inconsistency**
- **Models containing inconsistency parameters**
- **Graphical approaches**
- **No clear standard yet!**

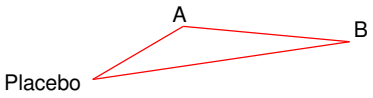
Direct comparison

Possible networks

Example

Possible networks

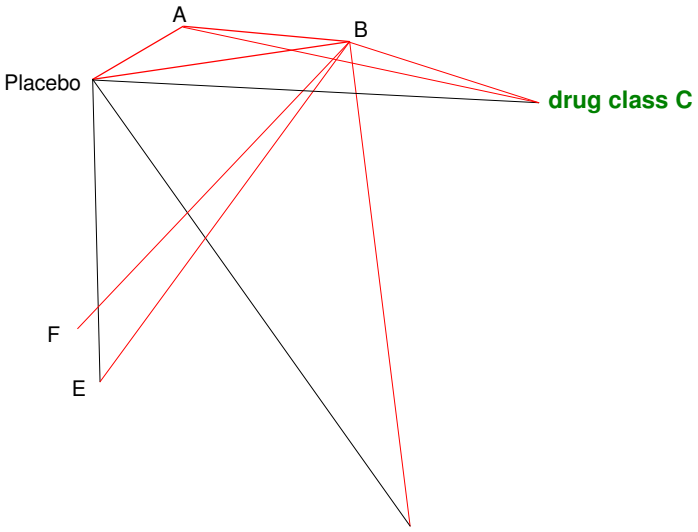
MTC incl. placebo arms



Example

Possible networks

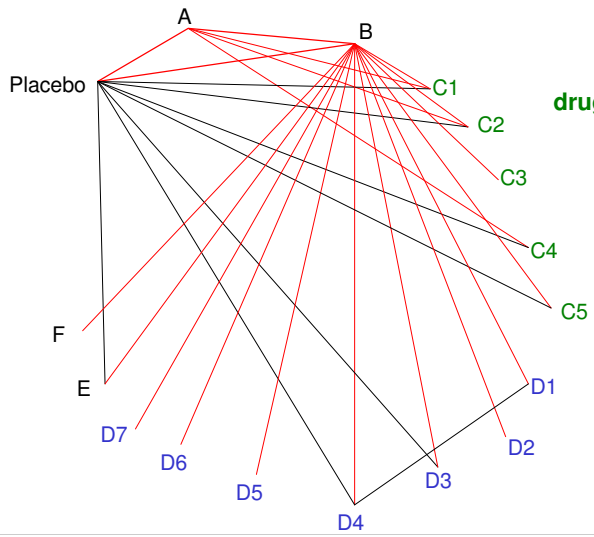
MTC incl. drug classes



Example

Possible networks

MTC incl. all drugs sep.



Example 1

Possible networks

	OR [95% CI]
Direct comparison	1.33 [0.93; 1.91]
MTC incl. placebo arms	1.10 [0.89; 1.35]
MTC incl. drug classes	1.14 [0.96; 1.35]
MTC incl. all drugs sep.	1.12 [0.93; 1.33]

Random effects MTC

Example 2

Possible networks

		OR [95% CI]
Direct comparison		0.77 [0.46; 1.28]
MTC incl. placebo arms		1.02 [0.74; 1.37]
MTC incl. drug classes		1.07 [0.90; 1.28]
MTC incl. all drugs sep.		1.09 [0.89; 1.32]

Random effects MTC

Definition of relevant interventions

Original Article

Research Synthesis Methods

Received 28 June 2011, Revised 10 July 2012, Accepted 19 July 2012 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1057

Unsolved issues of mixed treatment comparison meta-analysis: network size and inconsistency

Sibylle Sturtz[a*†] and Ralf Bender[a,b]

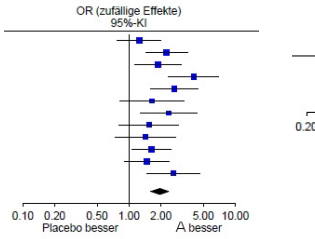
⇨ **Impact of network size:**
Larger networks are based upon more evidence but have more potential for heterogeneity and inconsistency

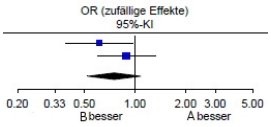
Network size

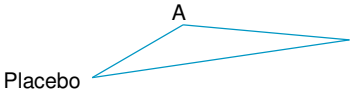
Issues regarding network size

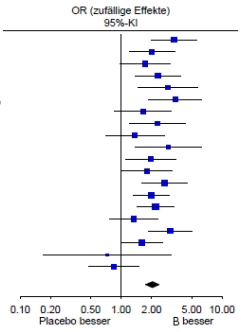
- **Results of different networks may or may not differ**
- **CI width in general smaller in large networks – but not always!**
- **Claim of inclusions of "all relevant evidence" represents an unworkably vague phrase (*Cooper et al., Value Health 2011*)**
- **Literature search for a "complete" evidence base may be a never ending story**
- **Network size connected with consistency**

Homogeneity









For 6 interventions:
15 possible pairs

For 12 interventions:
66 possible pairs!

Methods used to investigate inconsistency

- **Test for inconsistency**
  (*Bucher et al., JCE 1997; Caldwell et al., JCE 2010*)
- **Graphical approach:**
  Leverage vs. Bayesian deviance residuals
  (*Dias et al., Stat. Med. 2010*)

Test for consistency based on Bucher's approach:

$$z_{BC} = \frac{\hat{\omega}_{BC}}{\sqrt{Var(\hat{\omega}_{BC})}} \sim N(0,1)$$

$$\text{with } \hat{\omega}_{BC} = \hat{d}_{BC}^{direct} - \hat{d}_{BC}^{indirect}$$

$$\text{and } Var(\hat{\omega}_{BC}) = Var_{BC}^{direct} + Var_{BC}^{indirect}$$

$$= Var_{BC}^{direct} + Var_{AB}^{direct} + Var_{AC}^{direct}$$

Direct	Placebo	A
A	1.99 [1.65; 2.39]	p=0.466
B	2.04 [1.74; 2.38]	1.33 [0.93; 1.91]

Bucher et al., JCE 1997
Caldwell et al., JCE 2010

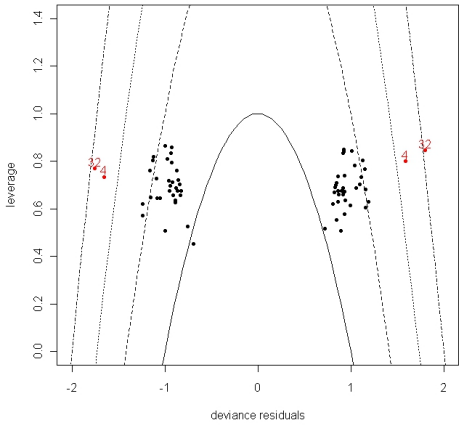
Example

Inconsistency

All studies (MTC)	Placebo	A
A	1.93 [1.61; 2.29]	
B	2.11 [1.82; 2.44]	1.10 [0.89; 1.35]

- Leverage Plot
- Curves of form $x^2+y=c$, c=1,2,3,4

⇨ Inconsistency



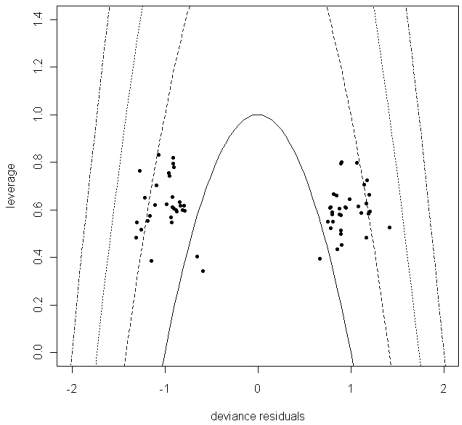
Direct	Placebo	A
A	1.99 [1.65; 2.39]	p=0.466
B	2.04 [1.74; 2.38]	1.33 [0.93; 1.91]

Example

Inconsistency

Without No. 4+32	Placebo	A
A	1.84 [1.58; 2.13]	
B	2.19 [1.93; 2.50]	1.19 [1.00; 1.42]

- Deviance: no evidence of inconsistency
- Test for consistency n.s.
- Results for MTC and MA comparable
- A vs B: MTC smaller CI



Direct	Placebo	A
A	1.86 [1.59; 2.19]	p=0.653
B	2.13 [1.87; 2.44]	1.33 [0.93; 1.91]

Important Questions:

- **Which network is the most appropriate?**
- **Is lumping of drug classes reasonable?**
- **Which amount of inconsistency is relevant?**
- **Which method is appropriate to evaluate inconsistency?**
  - **Graphical approach seems to be useful**
  - **Test for inconsistency is not reliable** (for α=0.05)

- **Many things are still unclear!**
- **More experience is required!**

Problematic Issue:

- **Only consistent networks should be used in practice**
- **All relevant evidence should be used**

Challenge

Network meta-analyses should only be performed if

- **Similarity assumption is sufficiently plausible**
- **Pairwise meta-analyses do not show relevant heterogeneity**
- **Network does not show relevant inconsistency**

Song et al., BMJ 2011

Challenge

Network meta-analyses should only be performed if

- **Similarity assumption is sufficiently plausible**
- **Pairwise meta-analyses do not show relevant heterogeneity**
- **Network does not show relevant inconsistency**

OPEN ACCESS Freely available online

PLoS one

Indirect Comparisons: A Review of Reporting and Methodological Quality

Sarah Donegan*, Paula Williamson, Carrol Gamble, Catrin Tudur-Smith

Centre for Medical Statistics and Health Evaluation, University of Liverpool, Liverpool, United Kingdom

Donegan et al., PLoS One 2010

⇨ ***"This review shows that the underlying assumptions are not routinely explored or reported when undertaking indirect comparisons."***

Example: ACEI/ARB in Diabetes

Diabetologia (2012) 55:566–578
DOI 10.1007/s00125-011-2398-8

META-ANALYSIS

Reno-protective effects of renin–angiotensin system blockade in type 2 diabetic patients: a systematic review and network meta-analysis

P. Vejakama · A. Thakkinstian · D. Lertrattananon · A. Ingsathit · C. Ngarmukos · J. Attia

A network meta-analysis was performed to compare indirectly all treatment effects.

Typical problems:

- **Heterogeneity in meta-analyses explored – but without consequence**
- **No assessment of consistency in network meta-analysis**

Example: 3 types of stents

BMJ

RESEARCH

Drug eluting and bare metal stents in people with and without diabetes: collaborative network meta-analysis

Christoph Stettler, senior research fellow,[1,2,3] Sabin Allemann, research fellow,[1,2] Simon Wandel, research fellow,[1] Adnan Kastrati, professor of cardiology,[4] Marie Claude Morice, professor of cardiology,[5] Albert Schömig, professor of medicine,[4] Matthias E Pfisterer, professor of cardiology,[6] Gregg W Stone, professor of medicine,[7] Martin B Leon, professor of medicine ,[7] José Suárez de Lezo, professor of cardiology,[8] Jean-Jacques Goy, professor of interventional cardiology,[9] Seung-Jung Park, professor of cardiology,[10] Manel Sabaté, associate professor of cardiology,[11] Maarten J Suttorp, head of department ,[12] Henning Kelbaek, associate professor of cardiology ,[13] Christian Spaulding, professor of cardiology ,[14] Maurizio Menichelli, interventional cardiologist,[15] Paul Vermeersch, interventional cardiologist,[16] Maurits T Dirksen, training fellow in cardiology,[17] Pavel Cervinka, cardiologist,[18] Marco De Carlo, vice director,[19] Andrejs Erglis, associate professor of cardiology,[20] Tania Chechi, interventional cardiologist,[21] Paolo Ortolani, interventional cardiologist,[22] Martin J Schalij, professor of cardiology,[23] Peter Diem, head of division ,[2] Bernhard Meier, professor of cardiology,[24] Stephan Windecker, head of invasive cardiology ,[24,25] Peter Jüni, head of division [1,25]

Stettler et al., BMJ 2008

3 web appendixes (17 pages in total):

- **Comprehensive description of applied models**
- **Methods to assess goodness-of-fit, heterogeneity, inconsistency**
- **Description of intermediate data**

INTERNATIONAL BIOMETRIC SOCIETY

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

gmds Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.

Stellenwert von Ergebnissen aus indirekten Vergleichen
Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR
Autoren: Ralf Bender, Carsten Schwenke, Claudia Schmoor, Dieter Hauschke

GMDS Geschäftsstelle

Beatrix Behrendt
Industriestraße 154
D-50996 Köln

<u>Joint statement of IQWiG, GMDS and IBS-DR (07.03.2012):</u>

Network meta-analyses lead to lower certainty of results compared to meta-analyses of direct head-to-head studies

http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf

(In-)direct comparisons

„Direct randomized comparisons of treatments are usually more trustworthy than indirect comparisons …"

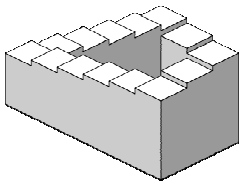
Ioannidis, CMAJ 2009

Always?

Reviews and Overviews

Why Olanzapine Beats Risperidone, Risperidone Beats Quetiapine, and Quetiapine Beats Olanzapine: An Exploratory Analysis of Head-to-Head Comparison Studies of Second-Generation Antipsychotics

Heres et al., Am. J. Psychiatry 2006



Sometimes an infinite stair
(Impossible figure, by M.C. Escher)

Conclusions

- **MTC represents an important and promising method for health technology assessment**
- **In practice, choice of an appropriate network required**
- **Only consistent networks should be used**
- **A network may never be "complete"**
- **Exclusion of a few studies may be reasonable**
- **Clear standards for identification of inconsistency and dealing with inconsistency are currently lacking**