



FUNDED BY

NIHR National Institute for Health and Care Research

Developing a tool for detecting problematic RCTs in health systematic reviews: the INSPECT-SR project

Jack Wilkinson, Centre for Biostatistics, University of Manchester. 🥑 @jd_wilko

Steering Group: Calvin Heal, Georgios Antoniou, Stephanie Boughton, Lisa Bero, Jamie Kirkham.

Some of the research discussed in this presentation is funded by the NIHR Research for Patient Benefit programme (NIHR203568). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.



Disclosures

- Currently hold or have recently held grants from NIHR, Health Research Council of New Zealand, UK Government, British Skin Foundation, Wellcome Trust.
- Stats Editor roles at Cochrane Gynaecology and Fertility, Fertility and Sterility, BJOG, Reproduction and Fertility, Journal of Hypertension



For the lawyers

- I'm not accusing anyone of fraud, data fabrication/falsification, or any other form of research misconduct here.
- I will say that some trials are unlikely to be authentic or are not trustworthy. The data or results do not appear to be compatible with a genuine RCT.
- I make no claims that this is due to deliberate action on behalf of investigators/ authors (vs catastrophic errors in data management, for example).



Outline

- 1. Detecting problematic studies in the context of health systematic reviews: the INSPECT-SR project.
- 2. Some principles for investigating potentially problematic RCTs.



The University of Manchester

1. Detecting problematic studies in the context of health systematic reviews: the INSPECT-SR project.

Ivermectin for COVID-19

Ivermectin for Prevention and Treatment of COVID-19 Infection: A Systematic Review, Metaanalysis, and Trial Sequential Analysis to Inform Clinical Guidelines

b Bryant, Andrew MSc^{1,*}; Lawrie, Theresa A. MBBCh, PhD²; Dowswell, Therese PhD²; Fordham, Edmund J. PhD²; Mitchell, Scott MBChB, MRCS³; Hill, Sarah R. PhD¹; Tham, Tony C. MD, FRCP⁴

Bryant et al., 2021

Risk ratio for death:

0.38 (95% CI 0.19 to 0.73)

15 trials

Hill et al., 2021

Risk ratio for death:

0.49 (95% CI 0.28 to 0.86)

12 trials

Meta-analysis of Randomized Trials of Ivermectin to Treat SARS-CoV-2 Infection

Andrew Hill,¹ Anna Garratt,² Jacob Levi,³ Jonathan Falconer,⁴ Leah Ellis,⁵ Kaitlyn McCann,⁵ Victoria Pilkington,⁶ Ambar Qavi,⁵ Junzheng Wang,⁵ and Hannah Wentzel⁵

Ivermectin for COVID-19

- SRs widely covered in media and social media.
- Used by antivax groups

Our Systematic Review...

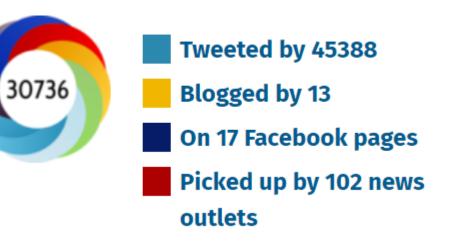
Our peer-reviewed study clearly shows that ivermectin prevents and treats Covid-19 and has the potential to save and improve countless lives.

- 2.6 million views
- Ranked <u>7th</u> of 20 million articles of a similar age.



A just-published, peer-reviewed study already clearly shows that ivermectin prevents and treats Covid-19 and has the potential to save and improve countless lives in the UK and worldwide right now.

The strength of evidence for ivermectin has this week been supercharged by publication of a gold standard review of 24 randomised trials conducted in 15 countries among more than 3400 people worldwide proving infections fall and deaths are dramatically reduced when ivermectin is administered. Published in the American Journal of Therapeutics the most rigorous statistical standards were applied by world-leading researchers biostatistician Mr Andrew Bryant and medical doctor and researcher Dr. Tess Lawrie.



The catch...

• It now appears that several of the trials were not authentic

-		10.00			14		and the second s	8.2	17. 8.25			1.00	10 DO MARK		THOOP IS	1000000	77 Prot	1.44
			1.0		-	and a second sec		14,6	- 44 J II-			1.00	- College and the second		-00000-040	104-Duk	10.000	
	e. 🕞	· · · · ·	1.4	14	1.4	dated in the second	-	4.6	1.00		87	10.000	all the second s		100 AUGUST	18,400,000	P & all	1.41
5					-	10.00 m		14	4.6- 11.6-			4.44	and a second sec	- Andrea			description of dead	A
				124	12	10.0		44	1.12 1.01			0.00	and the second se			- Property in	N Base	fact.
	-						10 million 1000					2.44	And the second sec					
		100			10.00	A-4-4+4 1		4.6					- Contract of the local division of the loca			_	- Carlos Rose Departure	
3	-	1 A 10	1.454			100.00	-	48	10, 100			1000	A REAL PROPERTY AND A REAL			000000		- Eral
n		1 1 1 1		1.00				44	A.511 A.514		45		and a second sec		indexed)-dia		(Al Anna)	
-	-	- I	- 14	14	104	1.0		11	110 0.00			in the second	1000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		minimum real real		and the second second second	Sec. 5
	1		1.4		10			4.0	141 10 81			1 Mar 10	and the second second	a second s		Carlos de la carlo de la c	in sec.	
			1.4	14	100	10.0	down the		111. 184			10.000	A REAL PROPERTY OF A REAL PROPER	Station 1			IN ADDRESS IN SUCC	10.00
		1 C T T T											and the second sec	and the				
	ALC: NO				100	1000						4.4	and the second s				autorities to have	6×11
7	1	 Internet 	1.04	114	100	that is a second se			1.			1000	A REAL PROPERTY OF A REAL PROPER				CONTRACTOR OF MALL	10.01
	4				100	and the second se			tan man	- Bil		-	and the second se	- A Distant And		A reported	34.00	
	-	- 10 March 10			1.00	anata. Cr.		44				-	and the second se				contraction of the local	
					100.0	mana distante di	100		1.01 10.01			-	and the second se	8-881-91-		1.00000	10 mm	diam'r.
									1.00. 00.00				and a local second s	distant in the second		period		4141
2	· · · · ·	1.2			100.00	1000							and the second se					
		1.0			100	- Contra - C			101 1.04			and a	- Andrewsky			- Delignar	and the second s	
,		100 Million	1.14	1.1	100	and the second s	1.4	48.	10.0 m.km		-	10.000	Construction of the second sec	and the second sec		100.00	24 Bar	-Exa
	18. B	10.00	_			tradition, dependent	ALC: NO.	11.	And Disc			4.4.	AND I SHOULD BE AND IN THE ADDRESS OF ADDRES	the second se	and the second	a fille	To Read	64.0
	10 · ·		1.14	10	1.00				10.00			10.001	Lo with	# phenode			-market about the Amort	
	14 1	1 M	- 54		1.0	- med	G		144 844			818.1	ul 10u		ideal range	BALL D	. If have	
							-	44					with the			(interface)		
1				- 04	10.0	312118			1.00, 1.00					Classification -				
	N1	1						34	11. 1474			94.91	of sail	CONSTR.			SUBJECT Paren	
•	·. ·	1	1.00	- 3.A.	100	79471	concernence of the second s	4.6 -	1.104.400	- 2-		-2.541	and a line work day.	1	-640000-000	 (0.4,4)-(1) 		
	1 N	1 II			1.44			1.4	121 1895	. 85		8.41	27 Mail	1 / Social			Builden Brit Maker	1
		- C-1	1.0	- 67	1000	10000.0	G	4.1					on table of Ro.			Los directo		1.80
		1000							111 10 10			1.0.1	LUMA .	C Minutes		- Contraction		1.1
							· · · · · · · · · · · · · · · · · · ·											
	4 1	1. 10.00	- 27				17		100 000			-0.41	149M				10000 Col. 40 Fight	100
	6	R.m.					14	14.	117 115			1.41	20.0002538000.000		10110-0	16.5.92181	There .	
•	M	 Cont 	0.0	1.0	1.00	and a second	1.0	4.6	+ + + + + + + + + + + + + + + + + +		- 40	-0.040	0.000-04			ine sinci-	- 14 Acres	
ė	12 Y	1 . A	4	1.00	24	mark backs	G	11.	1.87 1.81			1.4.	10.100		ALC: NOTE: N	to search and	1. Aug.	1.44
			- 14		1.4	and the second second	1	4.4.1			**	10.000	and the index	• 17000-100			and distances in some	
	2								10. 180	10		9.40	an and a strength	Contract of			Induction Value	
	· · · · ·	1.11.00				and down finds								Consent.				144
	· · · · ·	1000	1.17	1.10	110		100	4.0				2000	+200-0	L'ancient.		100000000		1.1.1
	A	21				and a second sec	10,000	1.5	122 0.61	8.4		141	A CANONELLINE		100.00	1.01/10/10/10/1	These and the second se	1.4
		· · · · ·		1.00		and a			4 (D)			40.041	and the second se		Market Diffe	·	1 determine	4
í	are 1	1.00		1.4	3.4	1. W.	1	1.1	1.01 1.00				A share	and the second s	1000		structures in pass	4.4
	- C						1.1		1.00 . 1.0.			2.44	- Carlos		econi-t-	(and (and (4
		1.1	- 17		12.0													
1		100 100			10.00	10-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-	Sec. 1		10. 11.00			11.00	A DECEMBER & LOP				Contract during produce	100
1				1.0	100	and the second	10.040	4.4	11, 4144			10,041	and the state of t	and the second se		insuited i	10 Page 1	
		1 1 1			in all	and the second se	14	18	1.01. 1.01			15.81	A DATE OF COMPANY OF COMPANY		100103-01	99,051	1 M Marc	114
			1.00				alla.	11.	1.00 B &-			1.44	Cont. D'il Addis 1 in		and a state of the		And and a second second	
				- 10g	100	and a second sec		11	1.00. 0.00			-	and the second se	-A.400-001		24.49		1.1
1		1.00				1000		- 22 -	100 100			2.24	and the second sec	and the second second			And State Street	
1		100			100	(and)												
1		1.0		C. 14		10.00	-		100.000			P.(P)	and the second se				TANKA TANA	- 14
1								4+	(in) - 20 and			10.00	- pole	Bulletin -		1.141.414	- Distant	4++
,				1.14	Test.	16.0	1.4		11. U.P.			10.000	- dente	and the second se		- principal of	in her	100
í			- 12			marks .		41				10.00	Area .	a contain.			owners and the second	4
1					1.00		and the second	22	1.71 0.64				and the second se	a second		-	Mass	
1		1.1				16/46 811814							and the second se			1000		100
1					100		-		1.00, 8.00			-	the second se	A CONTRACTOR OF	and the second second	in the second	Citates.	
1			1.14	1.14	1.44	10 C	1.4	. M	1.00 1.00			4.87	and the second diversion of th		manufacture of	10,000	Struker .	111.
1			_			1000	-	1.4	10.00			10 March 10	A REAL PROPERTY AND A REAL PROPERTY A REAL PROPERTY A REAL PROPERTY AND A REAL PROPERTY AND A REAL PROPERTY A REAL PROPERTY AND A REAL PROPERTY A REAL	and the second se		for stand and a	The second se	E

Analysis by Nick Brown (@sTeamTraen) at steamtraen.blogspot.com

Meta-analyses restricted to 'credible' trials

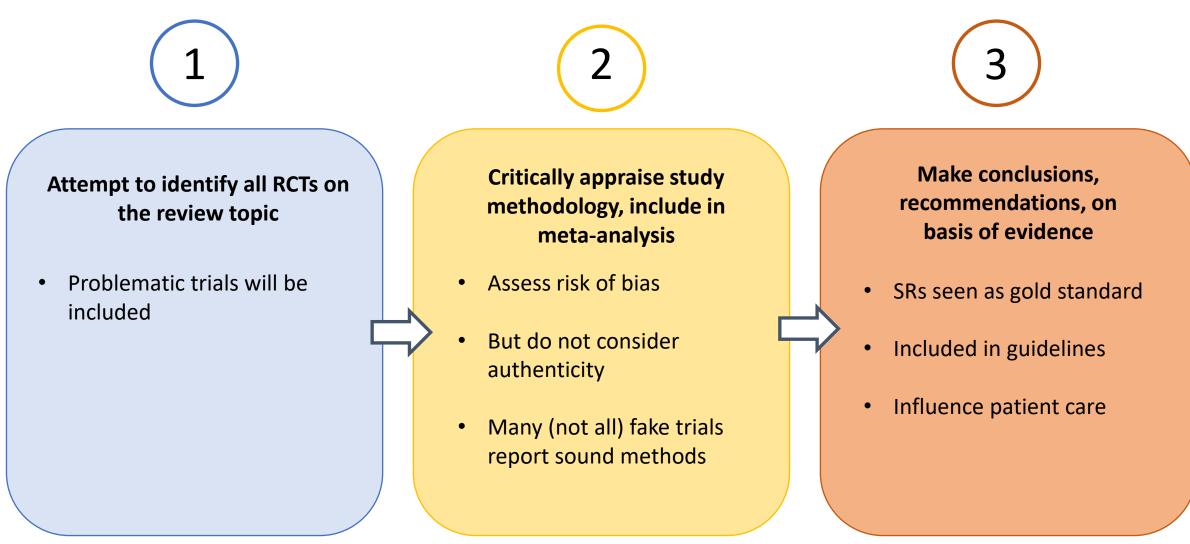
Hill et al., retracted their systematic review (1/2):

- "The significant effect of ivermectin on survival was dependent on the inclusion of studies with a high risk of bias or potential medical fraud."
- Risk ratio for death 0.96 (95% CI 0.56 to 1.66, 4 studies)

Popp et al., 2022 (Cochrane) excluded seven trials overall

- Asymptomatic or mild disease: Risk ratio for death 0.77 (95% CI 0.47 to 1.25, 6 trials)
- Moderate to severe disease: Risk ratio for death 0.60 (95% CI 0.14 to 2.51, 3 trials, 1 with no events)

Systematic reviews: Fake data to patient care pipeline



Vitamin K and the Prevention of Fractures

Systematic Review and Meta-analysis of Randomized Controlled Trials

Sarah Cockayne, MSc; Joy Adamson, PhD; Susan Lanham-New, PhD; Martin J. Shearer, PhD, MRCPath; Simon Gilbody, DPhil; David J. Torgerson, PhD

Does tranexamic acid prevent postpartum haemorrhage? A systematic review of randomised controlled trials

K Ker, H Shakur, I Roberts

Psychological therapies for the management of chronic pain (excluding headache) in adults (Review)

Williams ACDC, Fisher E, Hearn L, Eccleston C

3 out of 5 trials subsequently identified as fake.

26 trials. 8 had identical or similar text, 2 no ethical approval.

3 of 27 trials from one investigator suggested to be implausible (huge effects, no attrition).



When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F



When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

• Do not include studies until serious concerns about trustworthiness have been resolved.



When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

- Do not include studies until serious concerns about trustworthiness have been resolved.
- How do we define a 'problematic study'?



When beauty is but skin deep: dealing with problematic studies in systematic reviews

Stephanie L Boughton, Jack Wilkinson, Lisa Bero

Managing potentially problematic studies

https://bit.ly/3SsJO9F

- Do not include studies until serious concerns about trustworthiness have been resolved.
- How do we define a 'problematic study'?
- How can we detect them?



NIHR National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

- 1. Convene a panel of people with expertise and experience of investigating problematic studies.
- 2. Create an extensive list of methods for detecting problematic studies.
- 3. Apply the list to a sample of systematic reviews (feasibility, impact on review conclusions)
- 4. Enter the items into a Delphi process
- 5. Prospective testing in production and update of systematic reviews.



NIHR National Institute for Health and Care Research

INveStigating ProblEmatic Clinical Trials in Systematic Reviews

FUNDED B

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

- 1. Convene a panel of people with expertise and experience of investigating problematic studies.
- 2. Create an extensive list of methods for detecting problematic studies.



- 3. Apply the list to a sample of systematic reviews (feasibility, impact on review conclusions)
- 4. Enter the items into a Delphi process
- 5. Prospective testing in production and update of systematic reviews.



Making an extensive list of methods

Methods to assess research misconduct in healthrelated research: A scoping review

Esmee M Bordewijk • Wentao Li 2 2 • Rik van Eekelen • ... Marian Showell • Ben W Mol •

Madelon van Wely . Show all authors

Published: May 22, 2021 • DOI: https://doi.org/10.1016/j.jclinepi.2021.05.012 • 🜔 Check for updates

Experts identified warning signs of fraudulent research: a qualitative study to inform a screening tool

102 checks or tests identified

- Implemented as online survey of experts
- "Are we missing anything?"



Lisa Parker • Stephanie Boughton • Rosa Lawrence • Lisa Bero 🕺 🖂



Preliminary classifications and examples

Inspecting results in the paper

Are the results substantially divergent from others in the meta-analysis?

Inspecting conduct, governance and transparency

Is the recruitment of participants plausible within the stated time frame for the research? **Inspecting the research team**

Have other studies by the research team been retracted, or do they have expressions of concern?

Inspecting text and publication details

Is there evidence of copied work, such as duplicated or partially duplicated tables? Inspecting individual participant data

Does the dataset contain repeated sequences of baseline values?

		СВТ		Act	ive contr	ol		Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Alda 2011	36.9	8.3	56	37.1	10.5	53	3.7%	-0.02 [-0.40 , 0.35]	
Carson 2006	14	12.7	60	15	10.4	33	3.6%	-0.08 [-0.51 , 0.34]	
Ersek 2008	4.9	1.9	123	5	2.1	101	3.9%	-0.05 [-0.31 , 0.21]	-
Greco 2004	1.98	0.87	32	1.97	0.91	33	3.5%	0.01 [-0.48 , 0.50]	
Kaapa 2006	3.3	2.5	59	3.4	2.4	61	3.8%	-0.04 [-0.40 , 0.32]	
Keefe 1990	4.61	1.73	31	5.67	1.65	35	3.5%	-0.62 [-1.12 , -0.13]	
Keefe 1996	4.21	1.48	28	5.22	2.06	27	3.4%	-0.56 [-1.10 , -0.02]	
Kraaimaat 1995	14.8	4.3	24	15.4	4.6	28	3.4%	-0.13 [-0.68 , 0.41]	
Litt 2009	2.7	1.4	52	2.7	1.3	49	3.7%	0.00 [-0.39 , 0.39]	
Lumley 2014	2.7	0.7	130	2.7	1.1	134	4.0%	0.00 [-0.24 , 0.24]	-
Lumley 2017	4.7	1.7	75	5.2	1.7	76	3.8%	-0.29 [-0.61 , 0.03]	
Mangels 2009	15.9	5.3	232	16.4	5.8	131	4.0%	-0.09 [-0.31 , 0.12]	-
Monticone 2013	2.7	1	45	5	1.3	45	3.4%	-1.97 [-2.47 , -1.46]	
Monticone 2016	1.4	1.2	75	4.5	1.8	75	3.7%	-2.02 [-2.41 , -1.62]	
Monticone 2017	2.1	0.9	85	5.3	1.5	85	3.7%	-2.58 [-2.98 , -2.17]	_ —
Nicholas 2013	4.0	2.1	43	5.5	2.1		3.770	-0.33 [-0.72 , 0.00]	
Smeets 2006	42.3	25.6	55	44.6	28.9	52	3.7%	-0.08 [-0.46 , 0.30]	
Tavafian 2011	-65.8	22.6	92	-56.4	23.6	97	3.9%	-0.40 [-0.69 , -0.12]	
Thieme 2006	3.5	1	42	3.8	1.1	40	3.6%	-0.28 [-0.72 , 0.15]	
Thorn 2011	5.3	2.4	32	4.6	2.3	29	3.5%	0.29 [-0.21 , 0.80]	
Thorn 2018	5.4	2.3	83	5.7	2	80	3.9%	-0.14 [-0.45 , 0.17]	-
Thorsell 2011	7.2	2.9	52	8	2.5	38	3.6%	-0.29 [-0.71 , 0.13]	
Turner 2006	5.2	1.9	72	5.2	2.1	76	3.8%	0.00 [-0.32 , 0.32]	_
van Eijk 2013	5.5	2.1	108	5.5	2.1	95	3.9%	0.00 [-0.28 , 0.28]	-
Vitiello 2013	4.3	3.5	232	4.2	2.9	122	4.0%	0.03 [-0.19 , 0.25]	+
Vlaeyen 1996	1	1.8	42	0.4	1.8	30	3.5%	0.33 [-0.14 , 0.80]	
Zautra 2008	32.5	19.3	51	27.5	18	40	3.7%	0.26 [-0.15 , 0.68]	
Total (95% Cl)			2017			1718	100.0%	-0.33 [-0.56 , -0.10]	
Heterogeneity: Tau ² = 0.32; Chi ² = 293.71, df = 26 (P < 0.00001); l ² = 91%									•
Test for overall effect: Z = 2.82 (P = 0.005)									-2 -1 0 1 2
Test for subgroup diffe	erences: No	t applicat	ole						Favours CBT Favours active contro

Psychological therapies for chronic pain

Williams, et al. 2020 https://pubmed.ncbi.nl m.nih.gov/32794606/



Preliminary classifications and examples

Inspecting results in the paper

Are the results substantially divergent from others in the meta-analysis?

Inspecting conduct, governance and transparency

Is the recruitment of participants plausible within the stated time frame for the research? **Inspecting the research team**

Have other studies by the research team been retracted, or do they have expressions of concern?

Inspecting text and publication details

Is there evidence of copied work, such as duplicated or partially duplicated tables? Inspecting individual participant data

Does the dataset contain repeated sequences of baseline values?



Results

- 71 participants 5 continents, but mostly Europe (55%), Australia/ NZ (21%), N America (14%).
- 25 pages of comments: 16 new checks proposed, many suggestions to modify existing checks (e.g. merging, splitting or rewording).

Domain	Number of checks
1. Inspecting results in the paper	28
2. Inspecting the research team and their work	19
3. Inspecting conduct, governance and transparency	22
4. Inspecting text and publication details	7
5. Inspecting individual participant data	41
	117



INveStigating ProblEmatic Clinical Trials in Systematic Reviews

Aim: To develop a tool for identifying problematic randomised controlled trials in the context of health systematic reviews.

- 1. Convene a panel of people with expertise and experience of investigating problematic studies.
- 2. Create an extensive list of methods for detecting problematic studies.
- 3. Apply the list to a sample of systematic reviews (feasibility, impact on review conclusions)
- 4. Enter the items into a Delphi process
- 5. Prospective testing in production and update of systematic reviews.





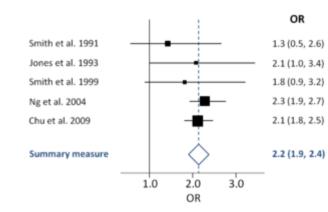


Approximately 50 researchers...

Domain	Number of checks
1. Inspecting results in the paper	28
2. Inspecting the research team and their work	19
3. Inspecting conduct, governance and transparency	22
4. Inspecting text and publication details	7
5. Inspecting individual participant data	41
	-117 76

Applying the long list of checks...





To RCTs in 50 Cochrane Reviews

- How often is each check failed?
- How feasible are the checks?
- What is the impact of removing flagged trials?

Interested?



- Need input and collaboration at all stages methodologists, trialists, systematic reviewers, editors, publishers, patients, research integrity professionals, or researchers with experience.
- Credible tool needs to be feasible, backed by broad consensus.
- INSPECT-SR Workshop at Colloquium: come and try an early draft version, feedback join us!
- Need people to participate in Delphi (methods experts and potential users of the tool)
- Need people who would be willing to test the tool while undertaking a systematic review.
- If you have any expertise, experience or interest, please contact me:



Available tools or frameworks



Experts identified warning signs of fraudulent research: a qualitative study to inform a screening tool

Lisa Parker • Stephanie Boughton • Rosa Lawrence • Lisa Bero 🙎 🖂

RESEARCH



Checklist to assess Trustworthiness in RAndomised Controlled Trials (TRACT checklist): concept proposal and pilot

Ben W. Mol^{1,2}, Shimona Lai¹, Ayesha Rahim¹, Esmée M. Bordewijk³, Rui Wang¹, Rik van Eekelen^{3,4}, Lyle C. Gurrin^{5*}, Jim G. Thornton⁶, Madelon van Wely^{3,4,7} and Wentao Li¹



Identifying and handling potentially untrustworthy trials in Pregnancy and Childbirth Cochrane Reviews

Alfirevic Z, Kellie FJ, Stewart F, Jones L, Hampson L, on behalf of Pregnancy and Childbirth Editorial Board Research Synthesis Methods

RESEARCH ARTICLE 👌 Open Access 💿 🛈 😒

Identifying and managing problematic trials: A research integrity assessment tool for randomized controlled trials in evidence synthesis

Stephanie Weibel 🔀, Maria Popp, Stefanie Reis, Nicole Skoetz, Paul Garner, Emma Sydenham

Individual participant data integrity assessment tool



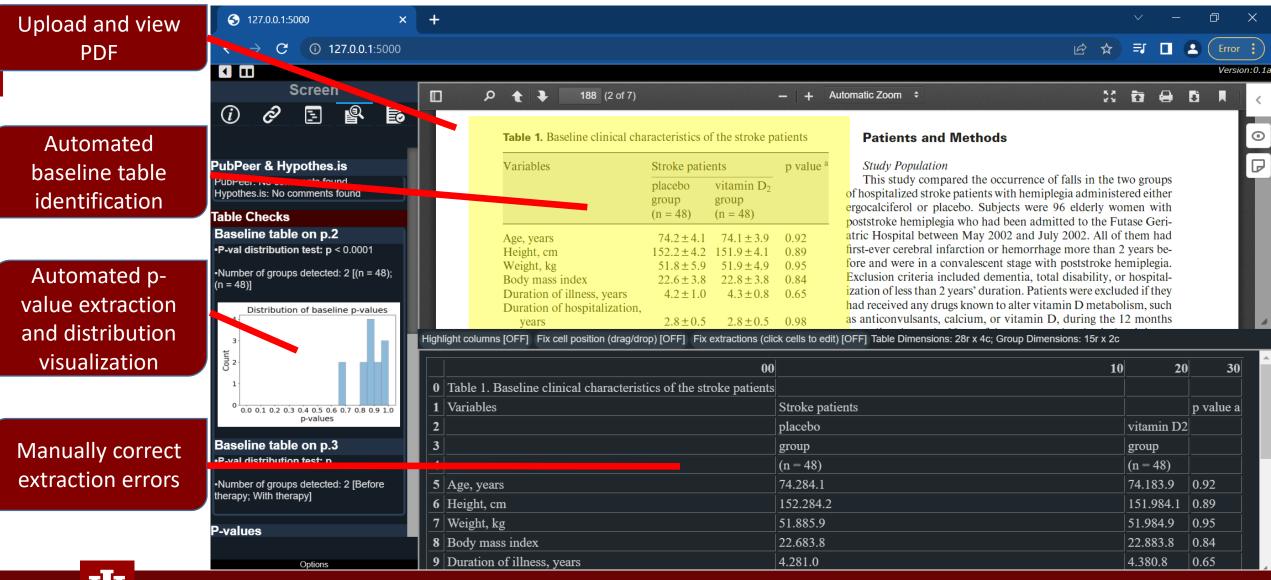
- Tool consisting of a checklist and semi-automated scripts to assess clinical trials for individual participant data meta-analyses (IPD-MA)
- Based on existing literature, mapping exercise and expert consensus
- Pilot tested and refined using two large IPD-MA in child health
- Developed by NextGen Evidence Synthesis Team at NHMRC Clinical Trials Centre, University of Sydney

For more information, or if you are interested in piloting the tool, please contact Kylie Hunter at <u>kylie.hunter@sydney.edu.au</u>





Semi-automated software for analytical forensics: RCT baseline tables as a proof of concept



Work in progress by Colby Vorland





2. Some principles for investigating potentially problematic RCTs.







• For this part of the talk, drawing on my experience investigating potentially problematic trials for journals and publishers over past four years. Trained by Stephen Evans.

• **Confidentiality** – following examples are illustrative – inspired by real cases, but I have changed details.

- Investigation involves a thorough examination of the manuscript, data, and other sources (registration, correspondence with authors, potentially other papers from the authors).
- Not trying to prove misconduct. Could these data have arisen from a genuine RCT?







The University of Manchester

• Conclusions based on a **holistic assessment** – not a single statistical test.

• Usually a day's work (at least).

• Illustrating some basic principles here – not comprehensive, not a tutorial!

human reproduction ORIGINAL ARTICLE Infertility

Endometrial scratching in women with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1,a}, F. Mol², J.S.E. Laven³, E.R. Groenewoud⁴, M.A.F. Traas⁵, C.A.H. Janssen⁵, G. Tekkenburg², J.P. de Bruin⁴, R.H.F. van Operraal¹, J.W. Maas⁶, E. Holl¹, K. Fleischer⁴, M.H.A. van Hooff¹, C.H. de Koning¹¹, A.E.P. Cantineau¹⁴, C.B. Lambaki⁴, M. Yecherg¹², A.H. van Houden¹⁴, A.P. Flanger¹⁴, E.A. Brinkhuis¹⁴, J.E. den Hartog¹³, M.Y. Glas¹⁴, N.F. Klijn¹⁷, S. van der Meer¹⁴, M.L. Bandell¹⁶, J.C. Boxmeer¹⁶, J. van Bosteldorp¹³, J. Smeenk¹³, M. van Wely¹³, M.J.C. Eijkemans^{13,4}, H.L. Torrance¹, and F.J.M. Broekmans¹

Sperment of propaging & Republics the States, Gaussia (Marci Care), Marcia Care, Marci Care, Marcia Care, Mar

"Commpondence address: Department of Garaccology & Reproductive Medicine: University Medical Centre Universe Pro Bee 85500, 3566 GA Universit, the Nethenhands: E-mail in as welf-one-gendraling/anameter.in Schemitter on June 6, 2020: enablement on July 27, 2020; extensiol decision and Anguest 28, 2020.

In Author(1) 2020. Mahate (p) Ontot Universe), Prior and Holf of European Source) of Neural Neproduction and Endopology. In a Oran Acoust Institute inducted under the series of the Control Acoustica University (Locate Regul/Constitutionerseinung/Deressit), in Art (B), who are non-control in exet, distribution, and repredention in any medium, provided the original work is preperly cited. For contenential music, ylease contact the premission acon. NIH U.S. National Library of Medicine

ClinicalTrials.gov

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

NSPECT SR

numan reproduction ORIGINAL ARTICLE Infertility

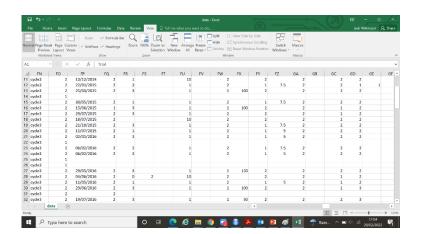
> Endometrial scratching in women with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1-a}, F. Moi², J.S.E. Laven³, E.R. Groenewoud⁴, M.A.F. Traas², C.A.H. Janssen⁶, G. Tekkenburg², J.P. de Bruin⁴, R.H.F. van Oppernzaij¹, J.W. Maas⁶, E. Melli¹, K. Fleichce¹⁻², M.H.A. van Hooff¹¹, C.H. de Koning⁴¹, A.E.P. Cantineau⁴, C.B. Lambala¹, M. Veberg²¹, A.H. van Housol⁴¹, K. A.Fleichce¹⁻², E.B. Lambala¹, M. Veberg²¹, A.H. van Housol⁴¹, A.F. P. Stener³², E.A. Brinkhuis⁴¹, J.E. den Harcog³¹, M.Y. Glas⁴², N.F. Klijn⁴⁷, S. van der Meer⁴⁷, M.L. Bandel⁴¹, J.C. Boxmeer³, J. van Disseldorp³¹, J. Smeenk¹², M. van Web³², M.J.C. Eijkemans^{13,4}, H.L. Torrance⁴, and F.J.M. Broekmans¹

Dependent of generating & Augustant Theora, Lasardy Telac Caro Maria, Usardy Michael, Nellow, 1976. Build 2016. Michael Reichter, Michael P. Michael S. Michael M. Sterner, of Computing per Heitel Depander of Demans Internet. Reichter J. Michael S. Bernetter, Michael M. Sterner, and Computing per Heitel Depander of Demans. Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Heitel Depander of Demans. Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuronal 2019. Heitel Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuronal 2019. Heitel Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuronal 2019. Heitel Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuronal 2019. Heitel Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuronal 2019. Heitel Demands Genet Angel, D. C. Antenn, et Reichter M. Disserver of Computing per Neuron 2019. Demands Genet Angel, D. C. G. C. Barrow, et Reichter, M. Disserver of Computing Personal Per Neuronal 2019. Heitel Demands J. Disserver, P. Barrow, D. K. Barrow, et Reichter, D. Sterner, B. Sterner, Sterner, Sterner, Sterner, B. Sterner, B. Sterner, B. Sterner, B. Sterner, B. Sterner, S

"Compositions address: Department of Genescology & Reproductive Medicine, University Medical Centre Utrecht, Utrecht, Utrecht, Email (Centre Utrecht, Email))))); (Centre Utrecht, Email (Centre Utrecht, Email)))); (Centre Utrecht, Email (Centre Utrecht,

Activity USD: Adulted by Oxder University Press on held of European Science of Homan Reproduction and Endopology. In 20m Anama and the Endolated which the weat of the Cassic Construction Net Construction Linear Reproduction and Reproduction and Reproduction and Reproduction and Reproduction and Reproduction in any readent free regimes work is properly toted. For conversion and Reproduction and Reproduction and Reproduction in any readent free regimes work is properly toted. For conversion and Reproduction and Reproduction in any readent free regimes work is properly toted. For conversion and readom sciences and Reproduction and



NIH U.S. National Library of Medicine

ClinicalTrials.gov

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

Stage 2: Detailed investigation

- Request additional documentation, individual participant data (IPD)
- Analysis of IPD

NSPECT SR

human reproduction ORIGINAL ARTICLE Infertility

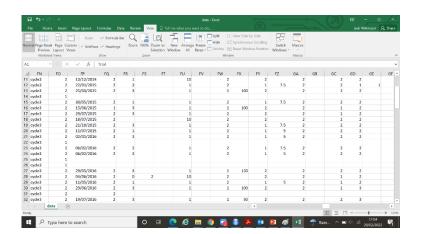
> Endometrial scratching in women with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1-a}, F. Moi², J.S.E. Laven², E.R. Groenewoud⁴, M.A.F. Traas², C.A.H. Janssen⁶, G. Tekkenburg², J.P. de Bruin⁴, R.H.F. van Oppernzaij, J.W.M. Mase⁶, E. Moll⁴, K. Fleicher²⁻, M.H.A. van Hooff⁴, C.H. de Koning⁴¹, A.E.P. Cantineau⁴, C.B. Lambala⁴, M. Vedberg²¹, A.H. van Housde⁴¹, A.F. Planger¹⁻, E.A. Brinkhuis⁴¹, J.E. den Hartog²¹, M.Y. Glaz²⁴, N.F. Klijn⁴⁷, S. van der Meer⁴⁴, M.L. Bandell⁴, J.C. Boxmeer⁴⁵, J. van Disseldorp²¹, J. Smeenk²¹, M. van Wely²¹, M.J.C. Eljkemans^{1,34}, H.L. Torrance⁴, and F.M. Broekman⁴

Dependent dissociation benches University Neter Caren Much, University Peter Sou 1999, 1816 de Marcines, Int. Method, Sou 2000, 1999, 199

"Compositence address: Department of Synaecology & Reproductive Medicine, University Medical Centre Univelyt, Univ PO Ban 85590, 1508 GA Univelut, the Nationalistic E-mail: na variehosperituipingBorncamedical Schmitted on June 6, 2020; resubmitted on July 27, 2020; eduarial decision on August 26, 2020

Autor(1) 2020. Addited by Called University Press on banked of European Society of Human Reproduction and Endopsidge. In Copin Allowing and the fact Addition for a similar of the Calles Constant Reproduction and Endopsidge. Legendralowing and the Society of the Called Constant Reproduction and Endopsidge. The constant of the sequence Legendralowing Constant, and Ingelescitics in our andianty messale the negative one's the property similar Texanson and provide the sequence one of the constant of the sequence status.



NIH) U.S. National Library of Medicine ClinicalTrials.gov

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

Stage 2: Detailed investigation

- Request additional documentation, individual participant data (IPD)
- Analysis of IPD

NSPECT SR

How?

How?

human reproduction ORIGINAL ARTICLE Infertility

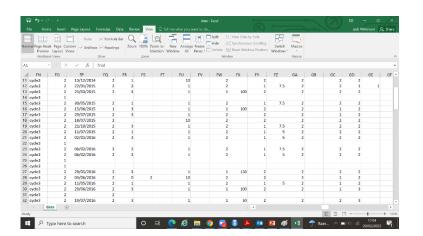
> Endometrial scratching in women with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1-a}, F. Mol², J.S.E. Laven², E.R. Groenewoud⁴, M.A.F. Traas⁵, C.A.H. Janssen⁵, G. Tekkenburg², J.P. de Bruin⁴, R.H.F. van Oppernaal¹, J.W. Mase⁶, E. Moll¹, K. Fleicher², M.H.A. van Hooff¹, C.H. de Koning⁴¹, A.E.P. Cantineau⁴, C.B. Lambaki⁴, M. Verberg²¹, A.H. van Housde¹¹, A.F. Planger¹³, E.A. Brinkhuis⁴¹, J.E. den Harog²¹, M.Y. Glas²⁴, N.F. Klijn⁷¹, S. van der Meer²⁴, M.L. Bander¹⁰, J.C. Boxmeer²⁵, J. van Disseldorp²¹, J. Smeenk¹², M. van Wely²¹, M.J.C. Eijkemans^{13,4}, H.L. Torrance¹, and F.J.M. Brockmans¹

Dependent dispersionel (Lesson Lesson Verlace) Network (Lesson Marchan Verlag), Network (Lesson Verlag), Network Verlag), Network Verlag), Network Verlag), Network (Lesson

"Compondence address: Department of Gynaecology & Reproductive Medicine, University Medical Centre Univelyt, Univ PO Ban 85500, 3588 GA Univelut, the Netherlands: E-mail: n.a. varbiospehhalpingBurncamedical Submitted on June 6, 2020; resubmitted on July 27, 2020; editorial decision on August 26, 2020

Anotorio 2021. Molecular (c) Outer University Press included of Elements Science of Homan Reproduction and Elements in Colom Maximum Line Elements and which reasons in Elem Cassing Constraints and Constraints. The Colom Maximum Line Element and Inspections on the Columnic Constraints and Constraint Linear Reproduction and Reproduction and Reportation rans incorrenal in suite, distribution and reproduction in any insidem, provided the original work is properly cred. For summerical are use plasmic context generationing and constraints.



NIH U.S. National Library of Medicine

ClinicalTrials.gov

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

Stage 2: Detailed investigation

- Request additional documentation, individual participant data (IPD)
- Analysis of IPD

NSPECT SR

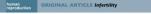
How?



How?



Human Reproduction, Vol.36, No.1, pp. 87–98, 2021 Advance Access Publication on December 8, 2020 doc10.1093/hummep/dess268



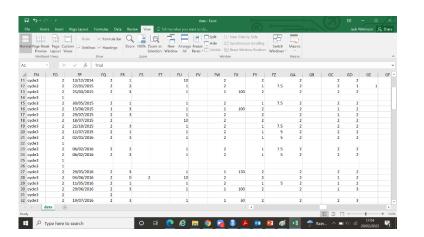
Endometrial scratching in women with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1,4}, F. Mol², J.S.E. Laven³, E.R. Groenewoud⁴, M.A.F. Traas², C.A.H. Janssen⁴, G. Teklenburg⁴, J.P. de Bruin¹, R.H.A. van Opport¹, J. W.M. Maral⁴, E.F. Effection the the C.B. Lamball⁴, M. Verberg¹⁴, A.M. van Heusden¹¹, A.P. Manger¹⁴, M.E. van Rumste³, L.F. van der Voet¹¹, Q.D. Pieterse²¹, J. Vinser²³, E.A. Brinkhuis⁴¹, J.E. den Hartog²¹, M.W. Glas³, N.F. Klijn⁷¹, S. van der Meer²⁴, M.L. Bandlel¹¹, C.E. Boxneer³², J. van Diselderp²¹, J.Smeenk²¹, M. van Web³², M.J.C. Eijkernans¹¹⁴, H.L. Torrance¹, and F.M. Brockmans¹

Dependent disponsing & Amperiases Messies Learning Metal Caret Messa, Usaria Messa, Weith Schmidt, Weith Schmidt, Weith Metal Messa, Weith Messa, We

Comportance address: Department of Genecology & Reproductive Medicine, University Medical Centre Utrecht, Utrec O Ben 1950; 1956 GA Uniceli, the Netheninini, E-mail n.e. variangerhaping/anametrical university on June 4: 2020 machimited and July 27: 2020.

Kellon() 2003. Hobbied by Celoted University Press on Select of European Society of Human Reproduction and Enterpology. In Opin Jonain motils distributed under the lamma of the Created Converses Reproduct Nets Converses Lamma Reproduction and Reproductions and Reproductions



NIH U.S. National Library of Medicine

ClinicalTrials.gov

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

Stage 2: Detailed investigation

- Request additional documentation, individual participant data (IPD)
- Analysis of IPD



How?



How?



- Want to ensure fabricated data cannot influence patient care (e.g. through meta-analysis).
- Want to avoid unintentionally removing genuine data from the literature.



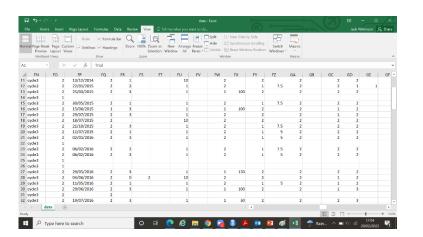
with one failed IVF/ICSI cycle outcomes of a randomised controlled trial (SCRaTCH)

N.E. van Hoogenhuijze^{1,4}, F. Moi², J.S.E. Laven³, E.R. Groenewoud⁴, M.A.F. Traas², C.A.H. Janssen³, G. Teklenburg⁴, J.P. de Bruin³, R.H.A. von Dieff⁴, C.J. M. Kong⁴, E.E. P.C. Christop¹¹ C.B. Lambalk⁴, M. Yerberg⁴, A.H. van Heusden¹⁸, A.P. Manger⁴, M.H.E. van Ruurs⁴, J.F. and en Hartog²¹, M.W. Glas²⁴, N.F. Kijn⁻¹, S. van der Meer¹⁴, M.L. Bandel¹¹, C.G. Sonce¹⁵, J. van Diselderp¹³, J. Smeenk¹³, M. van Web³, M.J. G. Eijkernans^{1,14}, H.L. Torrance¹, and F.J.M. Broekmans¹

Dependent of questions in the section of balance of the section of

Comegoridence address: Department of Genecology & Reproductive Medicine, University Medical Centre Utrecht, Utrecht, Utrecht, Email na varhaugerhausgehausg

Kebroft 2000 Hubbled by Cebrol University Pres on Select of European Society of Human Reproduction and Entrypology. In Opin Jossi motils distributed under the lamma of the Creater Converses Entrybutes Nets Converses Lamma Reproduction and Reproductions and Reproductions



NIH) U.S. National Library of Medicine

Stage 1: Concerns with study

 Usually on the basis of published or submitted information (manuscript, trial registration)

Stage 2: Detailed investigation

- Request additional documentation, individual participant data (IPD)
- Analysis of IPD



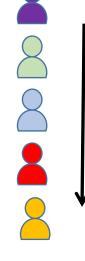
How?





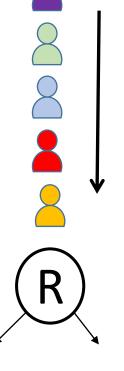
- Want to ensure fabricated data cannot influence patient care (e.g. through meta-analysis).
- Want to avoid unintentionally removing genuine data from the literature.

1. Potential participants present over time, have their baseline measurements taken before randomisation.



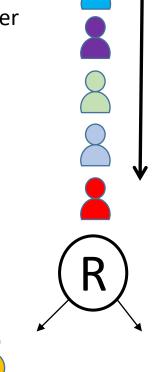
1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.



1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.



1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence. R

R

1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.

1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence. R

1. Potential participants present over time, have their baseline measurements taken before randomisation. Baseline measurements assessed 2. Eligible participants are sequentially allocated to study R arms according to a random sequence.

- No systematic differences between groups in baseline characteristics.
- Any patterns in baseline characteristics over time should appear in both groups...

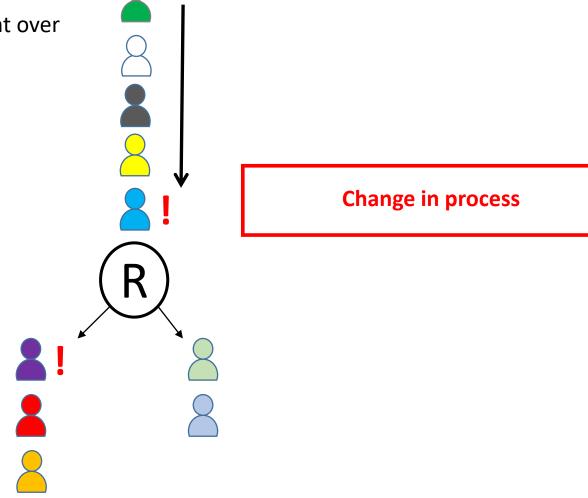
1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.

Change in process R

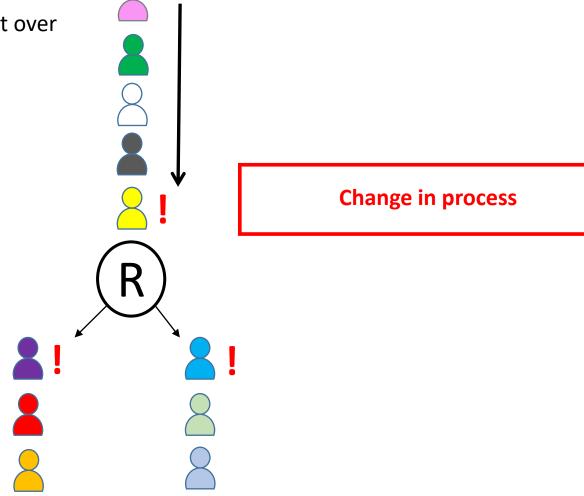
1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.



1. Potential participants present over time, have their baseline measurements taken before randomisation.

2. Eligible participants are sequentially allocated to study arms according to a random sequence.



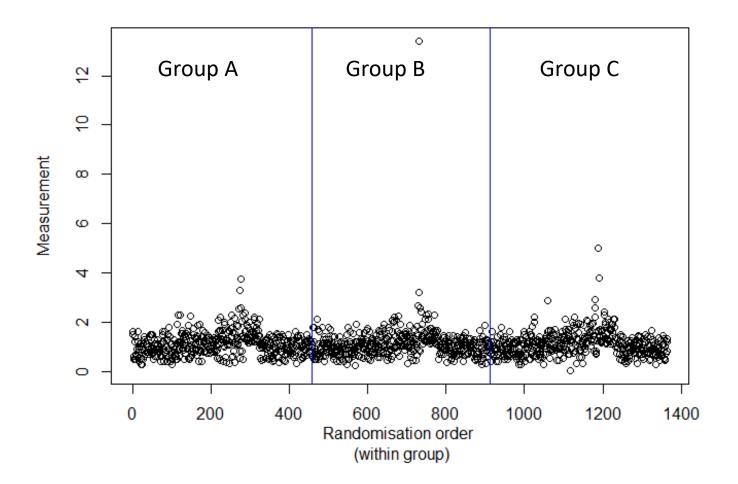
1. Potential participants present over time, have their baseline measurements taken before randomisation. **Change in process** 2. Eligible participants are sequentially allocated to study R arms according to a random sequence.

1. Potential participants present over time, have their baseline measurements taken before randomisation. **Change in process** 2. Eligible participants are sequentially allocated to study R arms according to a random sequence. 3. Patterns in baseline characteristics should be apparent in both arms.



Example – genuine data

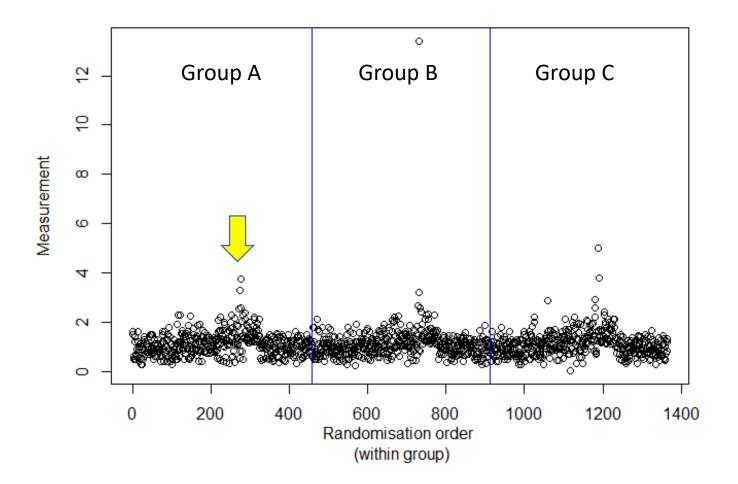
3-arm RCT. Blue lines divide into treatment groups – plotted in randomisation order.





Example – genuine data

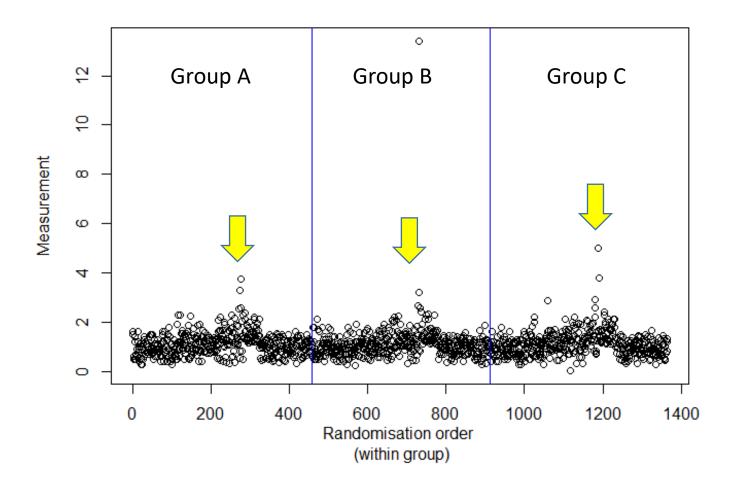
3-arm RCT. Blue lines divide into treatment groups – plotted in randomisation order.





Example – genuine data

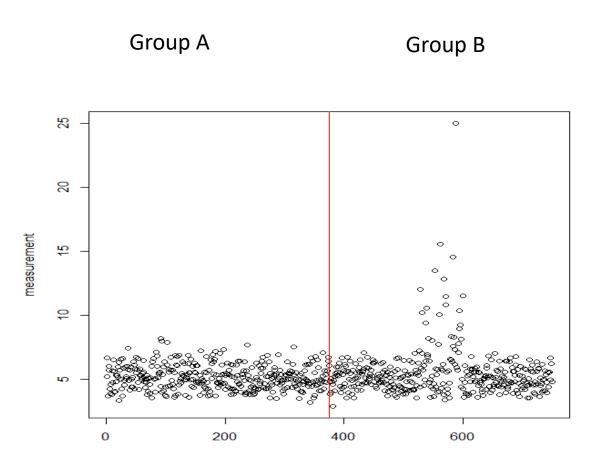
3-arm RCT. Blue lines divide into treatment groups – plotted in randomisation order.





Example – dubious data

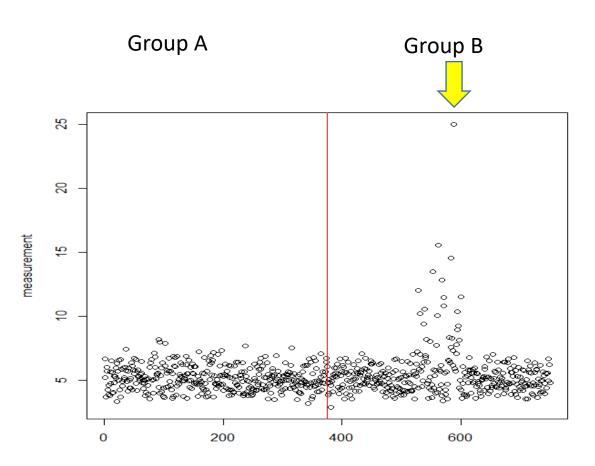
2-arm RCT





Example – dubious data

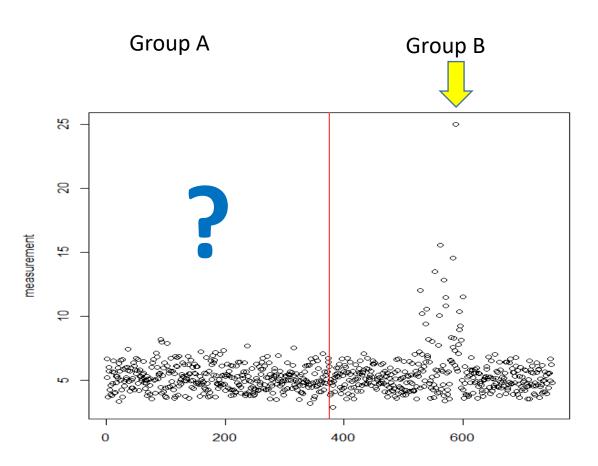
2-arm RCT





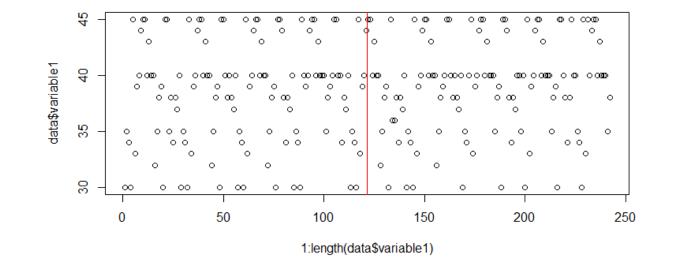
Example – dubious data

2-arm RCT



NSPECT SR

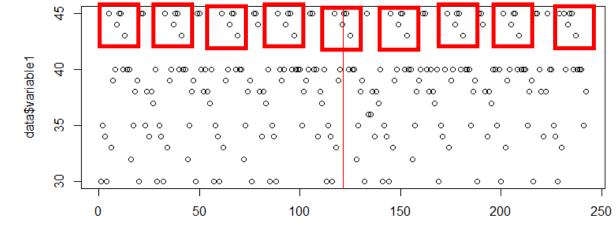
Other problems can be revealed by these plots



Take a moment – can you spot any problems?



Other problems can be revealed by these plots



1:length(data\$variable1)



Looking at outcome variables

• Outcomes are a bit different

• They are influenced by treatment, so we do expect to see differences

• But plotting against randomisation order can still reveal improbable patterns...



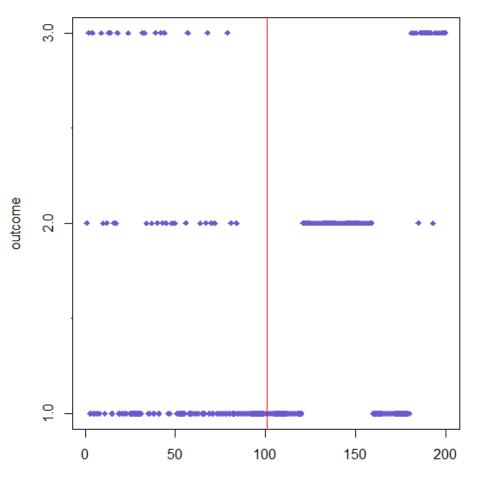
Looking at outcome variables

• Outcomes are a bit different

• They are influenced by treatment, so we do expect to see differences

• But plotting against randomisation order can still reveal improbable patterns...

Categorical outcome (1,2,3)



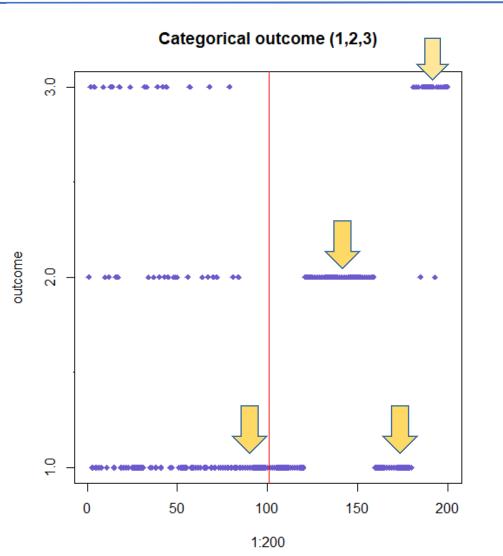


Looking at outcome variables

• Outcomes are a bit different

• They are influenced by treatment, so we do expect to see differences

• But plotting against randomisation order can still reveal improbable patterns...



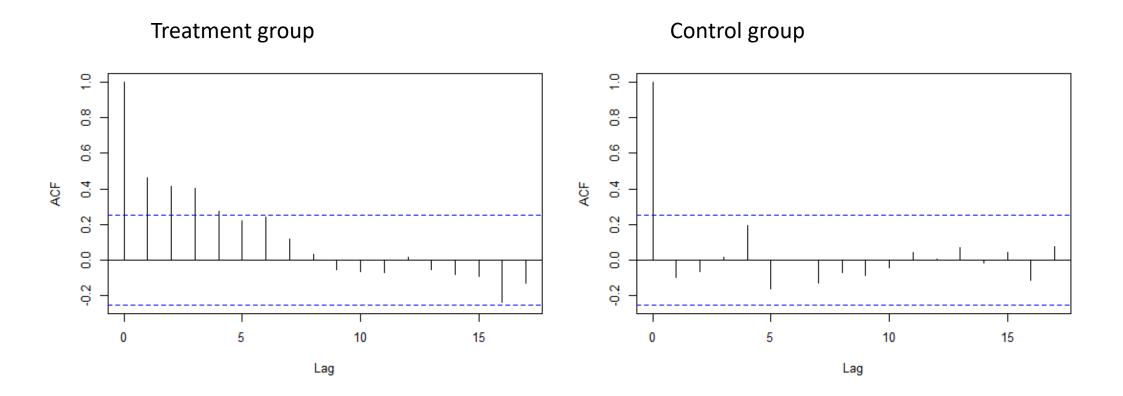
NSPECT SR

Correlation across rows in the dataset

- We don't expect to see substantial correlation between the baseline values of successive participants.
- E.g. each participant's duration of infertility shouldn't be related to the duration of infertility of the person recruited after them, or to the next person's, or the next person's etc.
- We do expect correlation across rows if someone has typed (fabricated) values into the column – people are poor random number generators.
- Certainly don't expect this to differ between randomised groups.

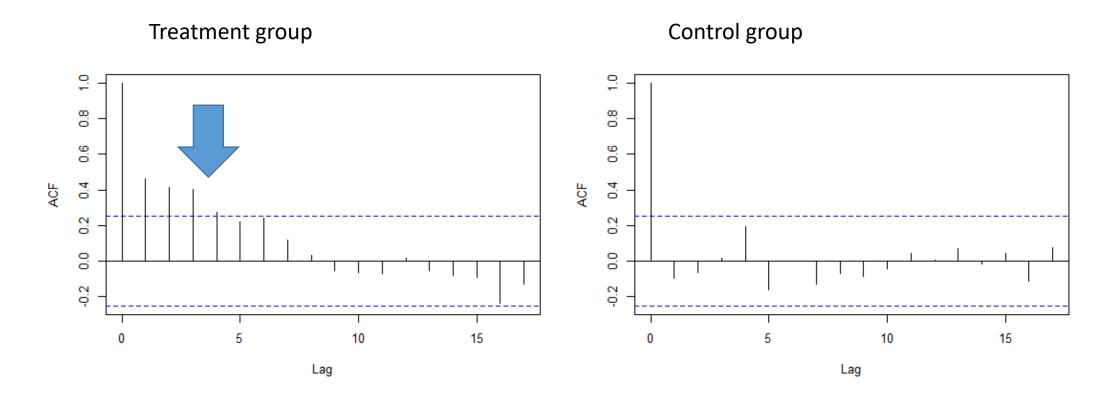
Study ID	InfertilityDuration
1	12
2	8
3	8
4	3
5	1
6	3
7	9
8	3
9	7
10	
11	5
12	1
13	4
14	6
15	2





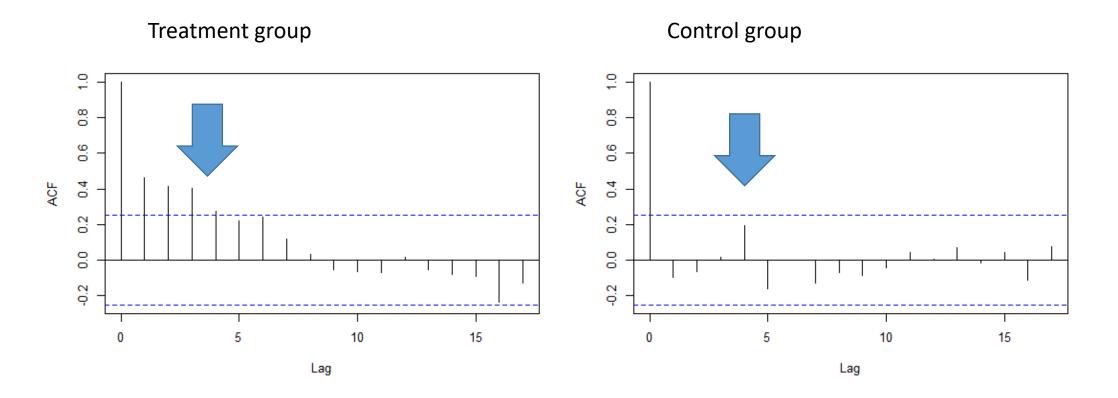
Plot of correlation between duration of infertility values 1 row apart (Lag = 1), 2 rows apart (Lag = 2), 3 rows apart etc.





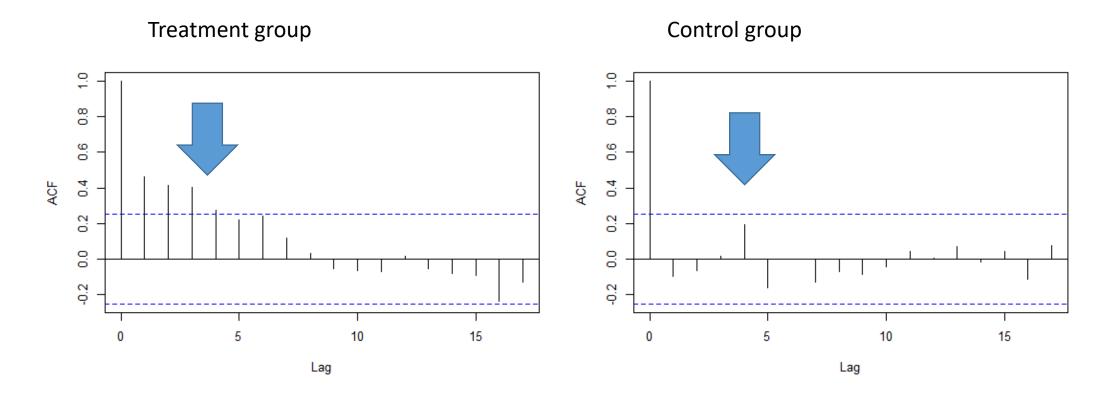
- Plot of correlation between duration of infertility values 1 row apart (Lag = 1), 2 rows apart (Lag = 2), 3 rows apart etc.
- In treatment group, there is a correlation between successive rows, which decays as we get further apart.





- Plot of correlation between duration of infertility values 1 row apart (Lag = 1), 2 rows apart (Lag = 2), 3 rows apart etc.
- In treatment group, there is a correlation between successive rows, which decays as we get further apart.
- Control group looks like genuine data no serial correlation between rows.



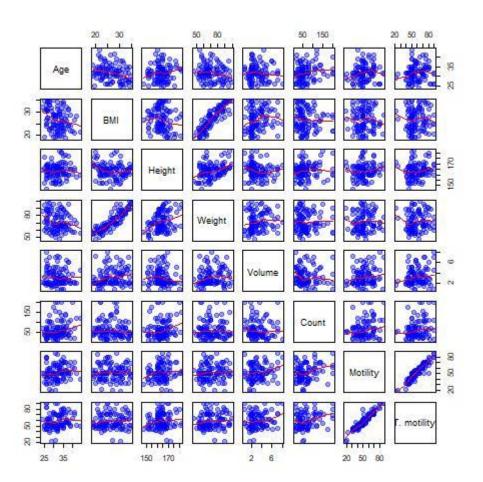


- Plot of correlation between duration of infertility values 1 row apart (Lag = 1), 2 rows apart (Lag = 2), 3 rows apart etc.
- In treatment group, there is a correlation between successive rows, which decays as we get further apart.
- Control group looks like genuine data no serial correlation between rows.
- The treatment group correlation is suspicious, and the difference between groups is more so.



Relationships between variables

- Are expected relationships between variables present?
- Hard to fake (unless you know what you are doing).
- Requires contextual knowledge (e.g. should we expect relationship between gestational age and birthweight in a particular trial).
- Don't expect multivariate distribution to differ between randomised groups.





Closing comments

- Have shown just a few basic checks here. Different approaches may be more or less appropriate for particular cases.
- We understand a lot about characteristics of data arising from RCTs. Can use this to assess whether data are (in)compatible with a genuine RCT.
- "Could there be an explanation for this?"
- Sometimes there is clear evidence of fabrication (e.g. certain cases with repeating sequences in the data). Other times, unclear whether misconduct or very poor conduct.
- Either way, may have reservations about using the data to decide how patients are treated.



Thanks to expert panel members

Elizabeth Loder	Toby Lasserson	Kyle Sheldrick	Andrew Grey	Susan Garfinkel
John Carlisle	Tianjing Li	Emily Lam	David Torgerson	Andreas Lundh
Karla Soares-Weiser	Neil O' Connell	Rebecca Jones	Esmée Bordewijk	Lyle Gurrin
Rita Redberg	Lisa Parker	Darren Dahly	Nick Brown	Lene Seidler
Jo Dumville	Virginia Barbour	Alison Avenell	Wentao Li	Kylie Hunter
Mike Clarke	Ben Mol	James Heathers	Richard Stevens	
Emma Sydenham	Barbara Redman	Gideon Meyerowitz- Katz	Rafael Perera- Salazar	
Jane Dennis	Jill Hayden	Madelon van Wely	Sarah Lensen	

- Need people to participate in Delphi.
- Need people who would be willing to test a tool while undertaking a systematic review.
- If you have any expertise, experience or interest, please contact me:
- jack.wilkinson@manchester.ac.uk or

