Afternoon timetable – further techniques and practical demonstrations 1pm – 4pm

1.00pm	**Alison Avenell**	**Introduction**
1.05pm	**Alison Avenell**	**Try out 'REAPPRAISED' checklist for evaluation of publication integrity of a published trial; basic statistical techniques for investigation.**
2.15pm	**Tea Break**	
2.30pm	**Jana Christopher** **Elisabeth Bik**	**Image manipulation: generative AI, examining raw data, tadpole paper mills. Hashing of raw images at acquisition, coupled with bespoke checking software to verify genuine images.** **Screening techniques: naked eye, Photoshop, automated tools (ImageTwin, Forensically, Fig Check, Proofig demonstrations), advantages and drawbacks.**
3.40 – 4.00pm	**All**	**Discussion**

HSRU
Promoting Excellence in Health Services Research

Checking the integrity of published biomedical research

Using the REAPPRAISED checklist on one trial report

Alison Avenell

Health Services Research Unit, University of Aberdeen, Scotland

No conflicts of interest declared

HSRU is core funded by the Chief Scientist Office of the Scottish Government Health and Social Care Directorates. The author accepts full responsibility for this talk.

CHIEF SCIENTIST OFFICE

1495 UNIVERSITY OF ABERDEEN

www.abdn.ac.uk/hsru

Follow us on @hsru_aberdeen

Acknowledgements

Andrew Grey, Mark Bolland, Greg Gamble – University of Auckland

David Cooper, Mari Imamura – University of Aberdeen

Anaesthesia 2021, 76, 472-479 doi:10.1111/anae.15263

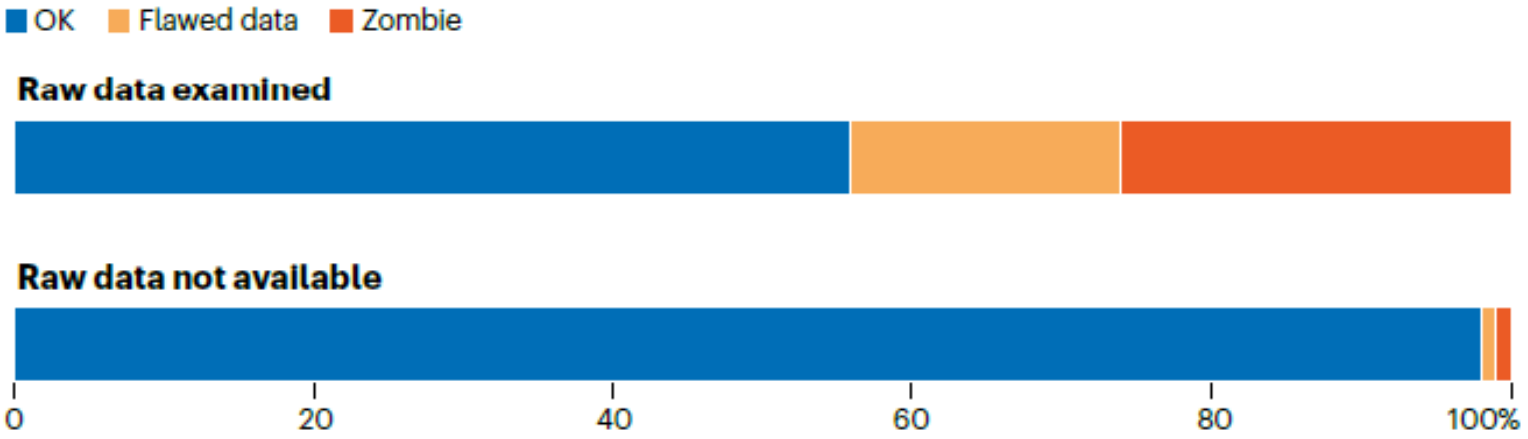
Original Article

False individual patient data and zombie randomised controlled trials submitted to *Anaesthesia*

J. B. Carlisle[1,2]

THE PREVALENCE OF 'ZOMBIE' TRIALS

More than one-quarter of a subset of manuscripts describing randomized clinical trials submitted to the journal *Anaesthesia* between 2017 and 2020 seemed to be faked or fatally flawed when their raw data could be examined, editor John Carlisle reported. He called these 'zombies'. But when their raw data could not be obtained, Carlisle could label only 1% as zombies.



https://www.nature.com/articles/d41586-023-02299-w

Things to consider first

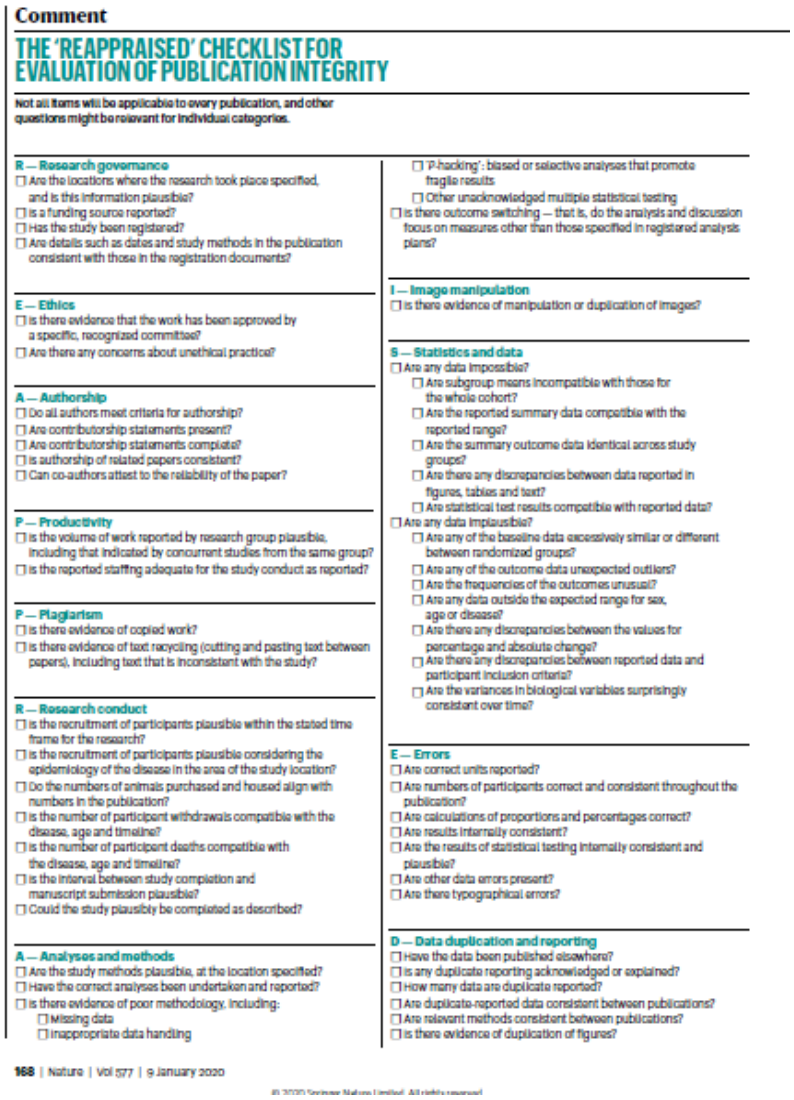
- Always consider what you read may have compromised integrity
- Check the author(s) for other publications, study registries, conference abstracts
- Look for correspondence on papers by citation searching, e.g. Web of Science, Scopus
- Check out Retraction Watch and its database of retractions

https://retractionwatch.com/retraction-watch-database-user-guide/

- Check out PubPeer and use its Plugin extension
- Use current Reference packages linked to RW's database:

  Edifix, EndNote, LibKey, Papers, and Zotero

REAPPRAISED checklist



Comment

THE 'REAPPRAISED' CHECKLIST FOR EVALUATION OF PUBLICATION INTEGRITY

Not all items will be applicable to every publication, and other questions might be relevant for individual categories.

R — Research governance
- ☐ Are the locations where the research took place specified, and is this information plausible?
- ☐ Is a funding source reported?
- ☐ Has the study been registered?
- ☐ Are details such as dates and study methods in the publication consistent with those in the registration documents?

E — Ethics
- ☐ Is there evidence that the work has been approved by a specific, recognized committee?
- ☐ Are there any concerns about unethical practice?

A — Authorship
- ☐ Do all authors meet criteria for authorship?
- ☐ Are contributorship statements present?
- ☐ Are contributorship statements complete?
- ☐ Is authorship of related papers consistent?
- ☐ Can co-authors attest to the reliability of the paper?

P — Productivity
- ☐ Is the volume of work reported by research group plausible, including that indicated by concurrent studies from the same group?
- ☐ Is the reported staffing adequate for the study conduct as reported?

P — Plagiarism
- ☐ Is there evidence of copied work?
- ☐ Is there evidence of text recycling (cutting and pasting text between papers), including text that is inconsistent with the study?

R — Research conduct
- ☐ Is the recruitment of participants plausible within the stated time frame for the research?
- ☐ Is the recruitment of participants plausible considering the epidemiology of the disease in the area of the study location?
- ☐ Do the numbers of animals purchased and housed align with numbers in the publication?
- ☐ Is the number of participant withdrawals compatible with the disease, age and timeline?
- ☐ Is the number of participant deaths compatible with the disease, age and timeline?
- ☐ Is the interval between study completion and manuscript submission plausible?
- ☐ Could the study plausibly be completed as described?

A — Analyses and methods
- ☐ Are the study methods plausible, at the location specified?
- ☐ Have the correct analyses been undertaken and reported?
- ☐ Is there evidence of poor methodology, including:
  - ☐ Missing data
  - ☐ Inappropriate data handling
  - ☐ 'P-hacking': biased or selective analyses that promote fragile results
  - ☐ Other unacknowledged multiple statistical testing
- ☐ Is there outcome switching — that is, do the analysis and discussion focus on measures other than those specified in registered analysis plans?

I — Image manipulation
- ☐ Is there evidence of manipulation or duplication of images?

S — Statistics and data
- ☐ Are any data impossible?
  - ☐ Are subgroup means incompatible with those for the whole cohort?
  - ☐ Are the reported summary data compatible with the reported range?
  - ☐ Are the summary outcome data identical across study groups?
  - ☐ Are there any discrepancies between data reported in figures, tables and text?
  - ☐ Are statistical test results compatible with reported data?
- ☐ Are any data implausible?
  - ☐ Are any of the baseline data excessively similar or different between randomized groups?
  - ☐ Are any of the outcome data unexpected outliers?
  - ☐ Are the frequencies of the outcomes unusual?
  - ☐ Are any data outside the expected range for sex, age or disease?
  - ☐ Are there any discrepancies between the values for percentage and absolute change?
  - ☐ Are there any discrepancies between reported data and participant inclusion criteria?
  - ☐ Are the variances in biological variables surprisingly consistent over time?

E — Errors
- ☐ Are correct units reported?
- ☐ Are numbers of participants correct and consistent throughout the publication?
- ☐ Are calculations of proportions and percentages correct?
- ☐ Are results internally consistent?
- ☐ Are the results of statistical testing internally consistent and plausible?
- ☐ Are other data errors present?
- ☐ Are there typographical errors?

D — Data duplication and reporting
- ☐ Have the data been published elsewhere?
- ☐ Is any duplicate reporting acknowledged or explained?
- ☐ How many data are duplicate reported?
- ☐ Are duplicate-reported data consistent between publications?
- ☐ Are relevant methods consistent between publications?
- ☐ Is there evidence of duplication of figures?

168 | Nature | Vol 577 | 9 January 2020

- **R**esearch governance
- **E**thics
- **A**uthorship
- **P**roductivity
- **P**lagiarism
- **R**esearch conduct
- **A**nalyses and methods
- **I**mage manipulation
- **S**tatistics and data
- **E**rrors
- **D**ata duplication and reporting

Try out REAPPRAISED on a trial report

PRELIMINARY COMMUNICATION

Effect of Folate and Mecobalamin on Hip Fractures in Patients With Stroke

A Randomized Controlled Trial

Yoshihiro Sato, MD
Yoshiaki Honda, MD
Jun Iwamoto, MD
Tomohiro Kanoko, PhD
Kei Satoh, MD

THE RISK OF A HIP FRACTURE IN patients after stroke is 2 to 4 times higher than that in age-matched healthy control patients.[1] These fractures usually occur relatively late after stroke onset and affect the paretic side of the body.[2,3] Hip fractures are associated with more deaths, disabilities, and medical costs than all other osteoporosis-related frac-

Context Stroke increases the risk of subsequent hip fracture by 2 to 4 times. Hyperhomocysteinemia is a risk factor for both ischemic stroke and osteoporotic fractures in elderly men and women. Treatment with folate and mecobalamin (vitamin B_{12}) may improve hyperhomocysteinemia.

Objective To investigate whether treatment with folate and vitamin B_{12} reduces the incidence of hip fractures in patients with hemiplegia following stroke.

Design, Setting, and Patients A double-blind, randomized controlled study of 628 consecutive patients aged 65 years or older with residual hemiplegia at least 1 year following first ischemic stroke, who were recruited from a single Japanese hospital from April 1, 2000, to May 31, 2001. Patients were assigned to daily oral treatment with 5 mg of folate and 1500 µg of mecobalamin, or double placebo; 559 completed the 2-year follow-up.

Main Outcome Measure Incidence of hip fractures in the 2 patient groups during the 2-year follow-up.

Results At baseline, patients in both groups had high levels of plasma homocysteine and low levels of serum cobalamin and serum folate. After 2 years, plasma homocysteine levels decreased by 38% in the treatment group and increased by 31% in

JAMA. 2005;293:1082-1088

R - Research governance

R — Research governance

- ☐ Are the locations where the research took place specified, and is this information plausible?
- ☐ Is a funding source reported?
- ☐ Has the study been registered?
- ☐ Are details such as dates and study methods in the publication consistent with those in the registration documents?

R – Research governance

Plausible??

628 participants with stroke recruited from Futase Social Insurance Hospital, Japan between April 1 2000 and May 31 2001

Poole et al. Letter to Neurology about another trial 2005; 65:1513

Re: Risedronate therapy for prevention of hip fracture after stroke in elderly women

'In our acute stroke service with a catchment area of 500,000, <10% of admissions would fulfill these criteria (15 in four months) and yet the authors recruited 374 patients in this period.'

Similar implausibility concerns from Halbekath et al. about another trial Re: Arch Intern Med 2007;167:513-4.

R – Research governance

Funding

Funding/Support: This study was not supported by any outside funds. We thank SOC Co Ltd, Tokyo, Japan, for providing the SAS software used in the data analysis.

How was the rest of the trial funded?

Not registered (ClinicalTrials.gov started 2000, ICMJE required trial registration 2005)

E - Ethics

E — Ethics

- ☐ Is there evidence that the work has been approved by a specific, recognized committee?
- ☐ Are there any concerns about unethical practice?

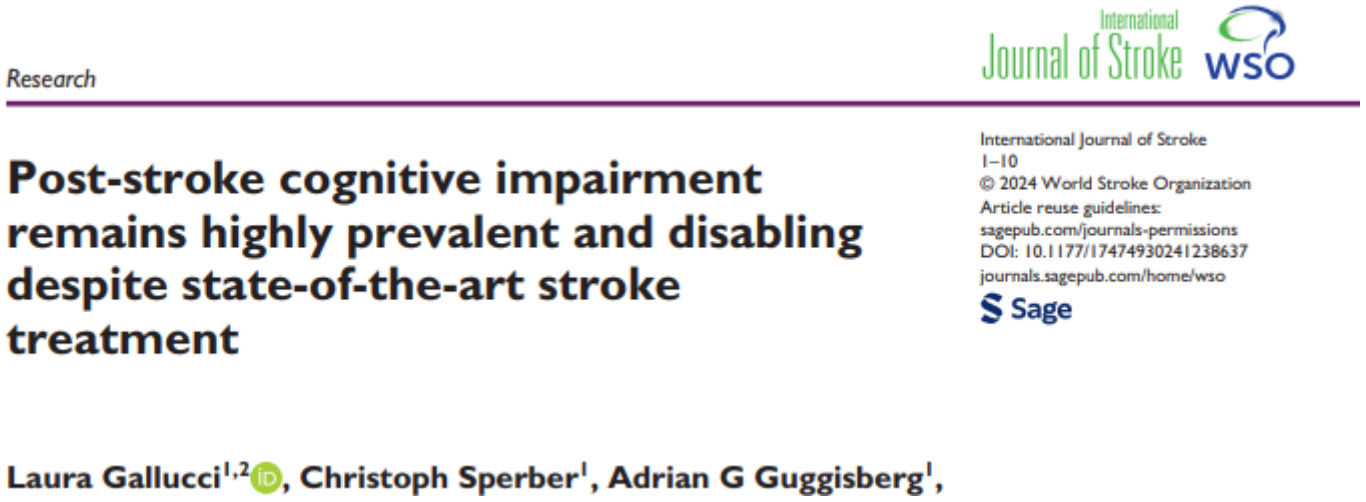
E- Ethics

'The study was approved by the local ethics committee, and written informed consent was obtained from all study participants in the presence of a witness.'

Not a specific, recognised ethics committee

Unethical practice?

What was the consent process for cognitive impairment/disability which was likely quite prevalent?



Research

International Journal of Stroke WSO

Post-stroke cognitive impairment remains highly prevalent and disabling despite state-of-the-art stroke treatment

International Journal of Stroke
1–10

DOI: 10.1177/17474930241238637
journals.sagepub.com/home/wso
Sage

Laura Gallucci[1,2], Christoph Sperber[1], Adrian G Guggisberg[1],

E - Ethics

Unethical practice? Over 2y:

'Patients were not allowed to take any other drugs that could affect bone and methionine metabolism.'

Drugs affecting bone and methionine metabolism:

'including corticosteroids, anticonvulsants, estrogens, calcitonin, bisphosphonates, calcium, folate, or vitamins B6, B12, D, and K.'

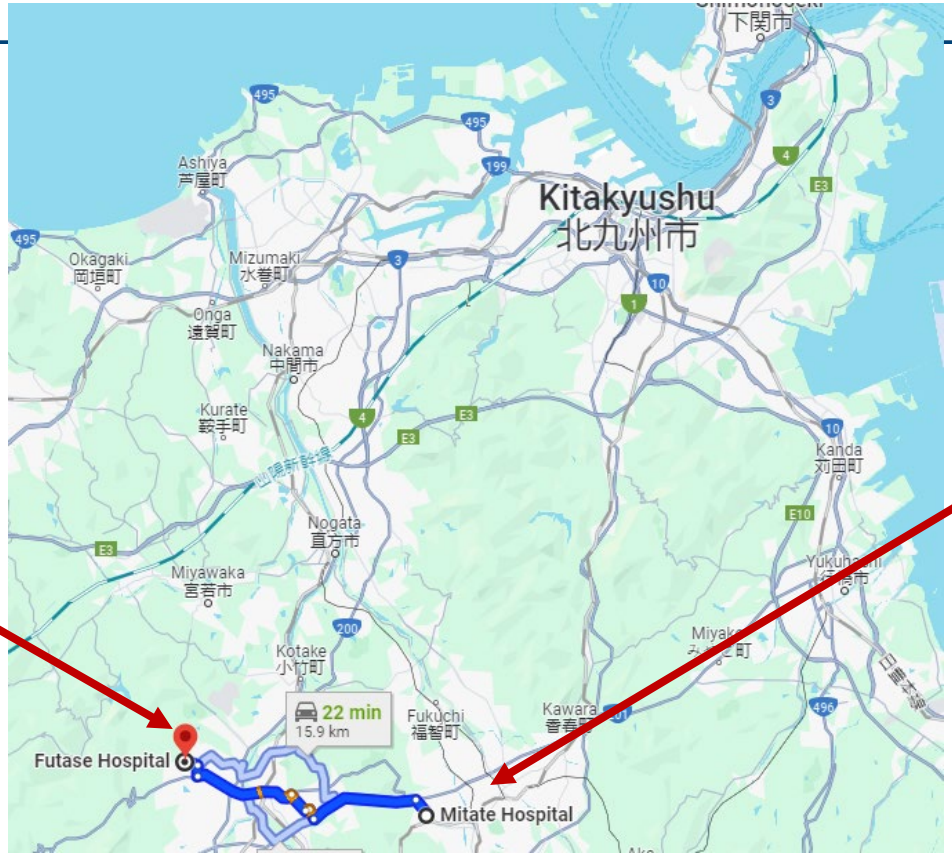
A - Authorship

A — Authorship

- ☐ Do all authors meet criteria for authorship?
- ☐ Are contributorship statements present?
- ☐ Are contributorship statements complete?
- ☐ Is authorship of related papers consistent?
- ☐ Can co-authors attest to the reliability of the paper?

Who are this group?

In original paper:
Recruitment at Futase Social Insurance Hospital
55 beds
3 doctors



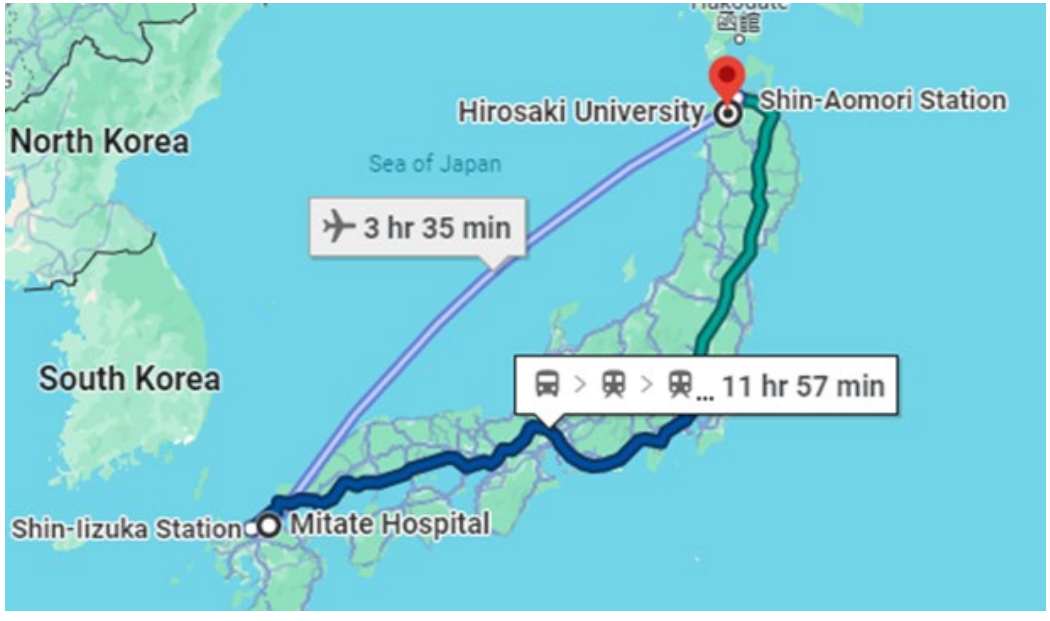
Y Sato
Neurologist
Mitate Hospital
410 beds
8 doctors
102 nurses

A- Authorship

Author Contributions: Dr Sato had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Satoh.
Acquisition of data: Sato, Honda, Kanoko, Satoh.
Analysis and interpretation of data: Sato, Iwamoto, Satoh.
Drafting of the manuscript: Satoh.
Critical revision of the manuscript for important intellectual content: Sato, Honda, Iwamoto, Kanoko.
Statistical analysis: Honda, Iwamoto.
Administrative, technical, or material support: Kanoko.
Study supervision: Sato.

Acquisition of data?

- **Sato and Honda at Mitate**
- **Kanoko and Satoh at Hirosaki (Satoh president of Hirosaki University) - 1564km away**



P - Productivity

P — Productivity

- ☐ Is the volume of work reported by research group plausible, including that indicated by concurrent studies from the same group?
- ☐ Is the reported staffing adequate for the study conduct as reported?

P - Productivity 2477 participants

8 first author trials published by Sato in 2005 Numbers randomised

Sato Y, Honda Y, Iwamoto J, Kanoko T, Satoh K. Amelioration by mecobalamin of subclinical carpal tunnel syndrome involving unaffected limbs in stroke patients. *J Neurol Sci* 2005;231:13-8. **135**

Sato Y, Honda Y, Iwamoto J, Kanoko T, Satoh K. Effect of folate and mecobalamin on hip fractures in patients with stroke: a randomized controlled trial. *JAMA* 2005;293:1082-8**. 628**

Sato Y, Iwamoto J, Kanoko T, Satoh K. Low-dose vitamin D prevents muscular atrophy and reduces falls and hip fractures in women after stroke: a randomized controlled trial. *Cerebrovasc Dis* 2005;20:187-92. **96**

Sato Y, Iwamoto J, Kanoko T, Satoh K. Risedronate sodium therapy for prevention of hip fracture in men 65 years or older after stroke. *Arch Intern Med* 2005;165:1743-8. **280**

Sato Y, Iwamoto J, Kanoko T, Satoh K. Amelioration of osteoporosis and hypovitaminosis D by sunlight exposure in hospitalized, elderly women with Alzheimer's disease: a randomized controlled trial. *J Bone Miner Res* 2005;20:1327-33. **264**

Sato Y, Iwamoto J, Kanoko T, Satoh K. Risedronate therapy for prevention of hip fracture after stroke in elderly women. *Neurology* 2005;64:811-6. **374**

Sato Y, Kanoko T, Satoh K, Iwamoto J. The prevention of hip fracture with risedronate and ergocalciferol plus calcium supplementation in elderly women with Alzheimer disease: a randomized controlled trial. *Arch Intern Med* 2005;165:1737-42. **500**

Sato Y, Kanoko T, Satoh K, Iwamoto J. Menatetrenone and vitamin D2 with calcium supplements prevent nonvertebral fracture in elderly women with Alzheimer's disease. *Bone* 2005;36:61-8. **200**

P - Productivity

Sato and Honda:

Only authors close to Futase Hospital

Between April 2000 and May 31 2001:

628 participants recruited: 45/month despite many exclusion criteria

No other staff mentioned, but for 2 years

'all patients were observed every 4 weeks in the outpatient clinic'

P - Plagiarism

P — Plagiarism

- ☐ Is there evidence of copied work?
- ☐ Is there evidence of text recycling (cutting and pasting text between papers), including text that is inconsistent with the study?

P - Plagiarism

- Use of templates for articles, making them all very similar
- Little depth to introduction or discussion
- But iThenticate shows no cause for concern

Methods very similar to these two papers, but no serious inconsistencies:

Sato Y, Iwamoto J, Kanoko T, Satoh K. Risedronate sodium therapy for prevention of hip fracture in men 65 years or older after stroke. Arch Intern Med 2005;165:1743-8

Sato Y, Kanoko T, Satoh K, Iwamoto J. The prevention of hip fracture with risedronate and ergocalciferol plus calcium supplementation in elderly women with Alzheimer disease: a randomized controlled trial. Arch Intern Med 2005;165:1737-42

R - Research conduct

R — Research conduct

- ☐ Is the recruitment of participants plausible within the stated time frame for the research?
- ☐ Is the recruitment of participants plausible considering the epidemiology of the disease in the area of the study location?
- ☐ Do the numbers of animals purchased and housed align with numbers in the publication?
- ☐ Is the number of participant withdrawals compatible with the disease, age and timeline?
- ☐ Is the number of participant deaths compatible with the disease, age and timeline?
- ☐ Is the interval between study completion and manuscript submission plausible?
- ☐ Could the study plausibly be completed as described?

R - Research conduct

R — Research conduct

- ☐ Is the recruitment of participants plausible within the stated time frame for the research? X
- ☐ Is the recruitment of participants plausible considering the epidemiology of the disease in the area of the study location? X
- ☐ Do the numbers of animals purchased and housed align with numbers in the publication? Not relevant
- ☐ Is the number of participant withdrawals compatible with the disease, age and timeline? X
- ☐ Is the number of participant deaths compatible with the disease, age and timeline? X
- ☐ Is the interval between study completion and manuscript submission plausible? √
- ☐ Could the study plausibly be completed as described? X

R - Research conduct

mg/dL (>5.70 mmol/L). The patients' clinical status was assessed at baseline, and all patients were observed every 4 weeks in the outpatient clinic, at which time all fractures were recorded. Falls

27 visits/patient over 2y

Only 4% died??

Only 7% other losses to follow-up??

A - Analyses and methods

A — Analyses and methods

- ☐ Are the study methods plausible, at the location specified?
- ☐ Have the correct analyses been undertaken and reported?
- ☐ Is there evidence of poor methodology, including:
  - ☐ Missing data
  - ☐ Inappropriate data handling
- ☐ '*P*-hacking': biased or selective analyses that promote fragile results
- ☐ Other unacknowledged multiple statistical testing
- ☐ Is there outcome switching — that is, do the analysis and discussion focus on measures other than those specified in registered analysis plans?

A – Analyses and methods

Plausible? Who did what? How were they funded?

Recruitment, visit scheduling, data entry?

Who provided placebos?

Metacarpal bone mineral density by computed x-ray densitometer?

Venous blood homocysteine by high-performance liquid chromatography?

Serum cobalamin and folate?

A - Analyses and methods

Fracture ascertainment?

Intervention group 6 hip/8 total fractures (75% hip)

Control group 27 hip/32 total fractures (84% hip)

Mortality 26/628 = 4% over 2y

Oral vitamin D3 and calcium for secondary prevention of low-trauma fractures in elderly people (Randomised Evaluation of Calcium Or vitamin D, RECORD): a randomised placebo-controlled trial

*The RECORD Trial Group**

Summary

Background Elderly people who have a fracture are at high risk of another. Vitamin D and calcium supplements are often recommended for fracture prevention. We aimed to assess whether vitamin D3 and calcium, either alone or in combination, were effective in prevention of secondary fractures.

Methods In a factorial-design trial, 5292 people aged 70 years or older (4481 [85%] of whom were women) who were mobile before developing a low-trauma fracture were randomly assigned 800 IU daily oral vitamin D3, 1000 mg calcium, oral vitamin D3 (800 IU per day) combined with calcium (1000 mg per day), or placebo. Participants who were recruited in 21 UK hospitals were followed up for between 24 months and 62 months. Analysis was by intention-to-treat and the primary outcome was new low-energy fractures.

Published online
April 28, 2005
DOI: 10.1016/S0140-6736(05)63013-9

See Comment

*Members of the group listed at end of paper

Correspondence to:
Prof Adrian M Grant, Health Services Research Unit, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK
a.grant@abdn.ac.uk

RECORD trial placebo group: 23% hip fractures Mortality 8% over 2y

I - Image manipulation

I — Image manipulation

☐ Is there evidence of manipulation or duplication of images?

Not here, but graphs may be copied

Numerical data from graphs may not match text or may not be given

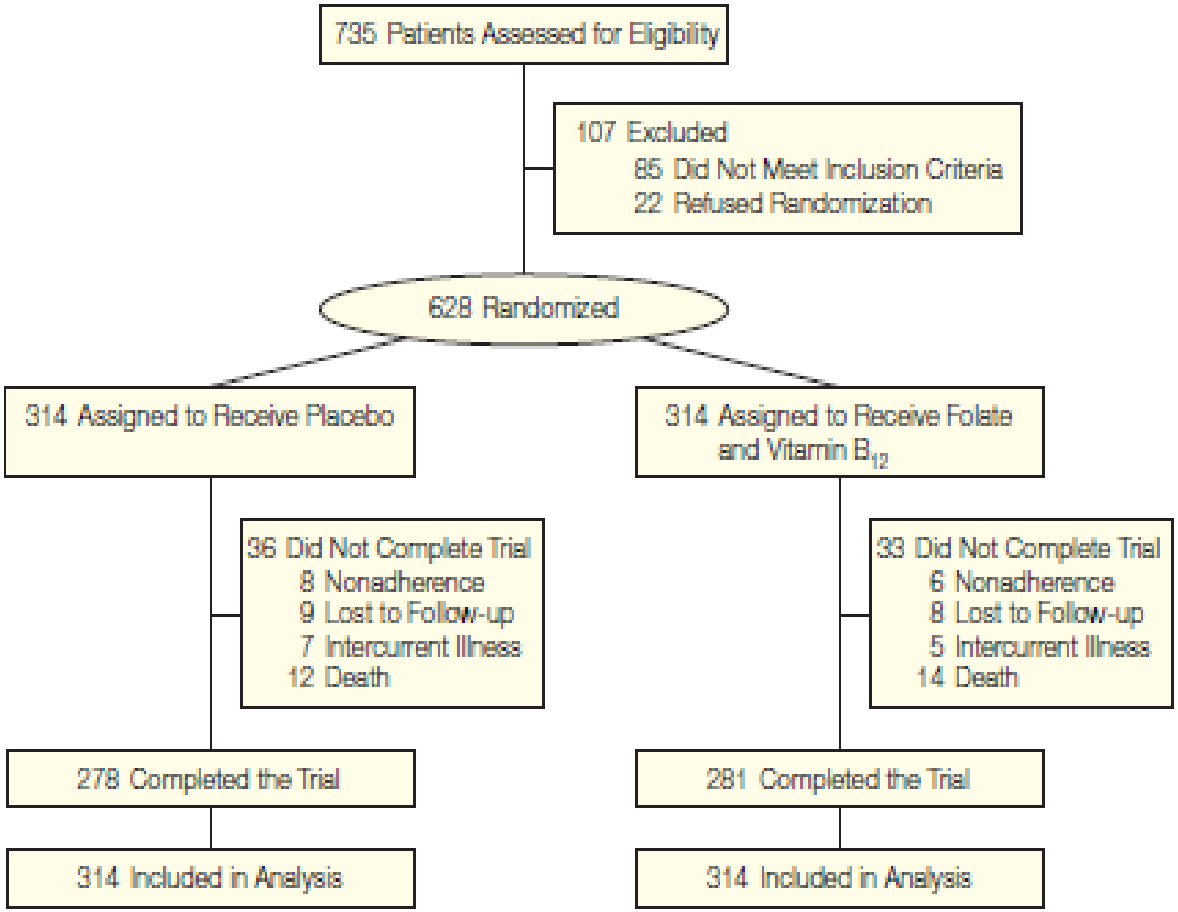
S - Statistics and data

S — Statistics and data

- ☐ Are any data impossible?
  - ☐ Are subgroup means incompatible with those for the whole cohort?
  - ☐ Are the reported summary data compatible with the reported range?
  - ☐ Are the summary outcome data identical across study groups?
  - ☐ Are there any discrepancies between data reported in figures, tables and text?
  - ☐ Are statistical test results compatible with reported data?
- ☐ Are any data implausible?
  - ☐ Are any of the baseline data excessively similar or different between randomized groups?
  - ☐ Are any of the outcome data unexpected outliers?
  - ☐ Are the frequencies of the outcomes unusual?
  - ☐ Are any data outside the expected range for sex, age or disease?
  - ☐ Are there any discrepancies between the values for percentage and absolute change?
  - ☐ Are there any discrepancies between reported data and participant inclusion criteria?
  - ☐ Are the variances in biological variables surprisingly consistent over time?

S – Statistics and data

Figure 1. Flow of Participants Through the Study



Do numbers add up?

Are numbers consistent throughout the paper?

22(3%) refused randomisation, which is unusually few

Randomised controlled trials

Randomisation of trial participants at the start of a trial:

- Reduces bias in assignment to treatments - selection bias
- We can test whether randomisation occurred properly by examining baseline data characteristics of participants at trial entry after randomisation
- Baseline characteristics of randomised groups should be balanced

S – Statistics and data

'Patients were randomised to the 2 groups by using permuted block size of 4.'

Implies 628/4 = 157 blocks filled = 314/group

But in 2006:

CORRECTION

Incorrect Description of Collaborating Hospitals: In the Preliminary Communication entitled "Effect of Folate and Mecobalamin on Hip Fractures in Patients With Stroke: A Randomized Controlled Trial" published in the March 2, 2005, issue of *JAMA* (2005;293:1082-1088), there was an incorrect description of the hospitals that collaborated in the study. The authors incorrectly stated that the study was performed in a single hospital (Futase Social Insurance Hospital). There were actually 3 additional collaborating hospitals, studying 205, 211, and 159 patients, respectively. The remaining 53 patients were from Futase Social Insurance Hospital, for a total of 628 patients.

JAMA 2006;296:396

'3 additional collaborating hospitals, studying 205, 211, and 159 patients, respectively. The remaining 53 patients were from Futase Social Insurance Hospital, for a total of 628'

Where were these three additional hospitals?

Central allocation of pills to sites after randomisation?

Baseline data - matching

Table 1. Baseline Characteristics of the Study Population*

Characteristic	Received Placebo (n = 314)	Received Folate and Vitamin B_{12} (n = 314)
Age, y	71.2 (4.2)	71.6 (5.1)
Sex, No. (%)		
Female	169 (53.8)	169 (53.8)
Male	145 (46.2)	145 (46.2)
Duration of illness, mo	16.9 (3.6)	16.9 (4.4)
Lacunar infarction/atherothrombotic infarction, No. of patients	217/97	212/102
Barthel Index†	67 (16)	68 (16)
Degree of hemiplegia‡		
Hand	4.5 (1.3)	4.5 (1.3)
Leg	4.5 (1.5)	4.5 (1.4)
Body mass index	22.3 (1.9)	22.3 (1.6)
Fallers, No. (%)§	70 (22)	70 (22)
Prevalence of vascular risk factors, No. (%)		
Hypertension	182 (58)	175 (56)
Diabetes mellitus	70 (22)	68 (22)
Hypercholesterolemia	59 (19)	59 (19)
Current smoker	64 (20)	64 (20)
Previous vascular event	62 (20)	58 (18)
BMD, mm Al		
Hemiplegic side‖	2.30 (0.24)	2.30 (0.25)
T score¶	−2.9 (1.0)	−2.9 (1.0)
Intact side	2.41 (0.24)	2.42 (0.24)
T score¶	−1.8 (1.0)	−1.8 (0.9)
Concentration levels		
Plasma homocysteine, μmol/L	19.9 (20.4)	19.9 (21.3)
Serum cobalamin, pg/mL	590 (332)	606 (382)
Serum folate, ng/mL	2.4 (1.5)	2.4 (1.5)

Baseline data - matching

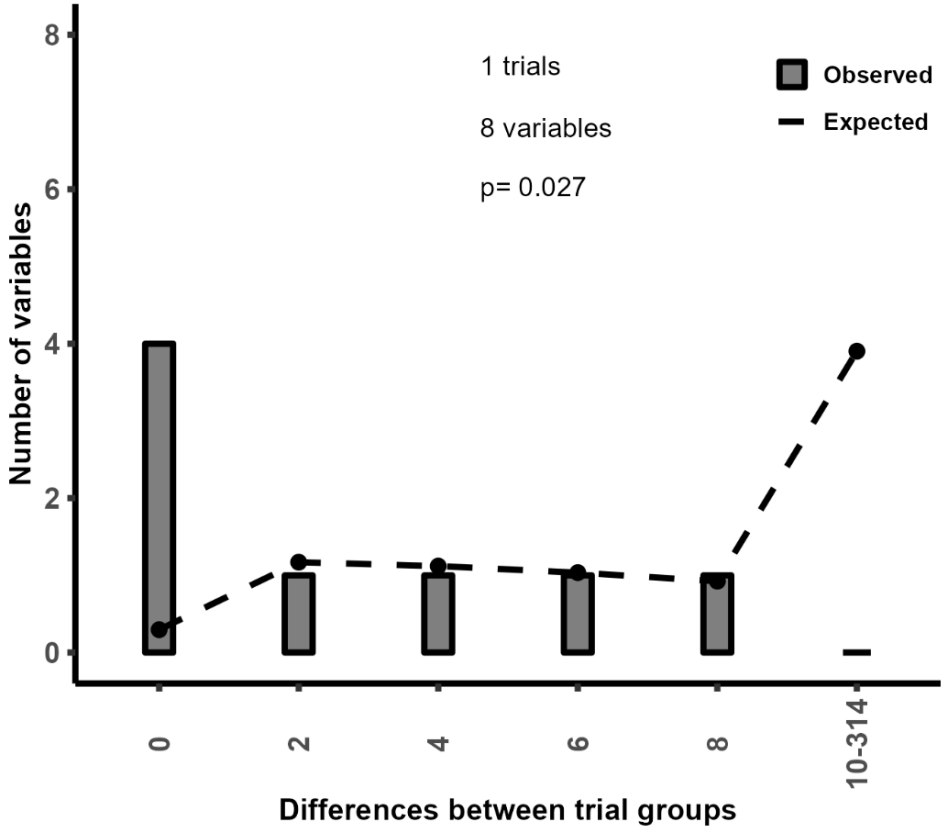
Table 1. Baseline Characteristics of the Study Population*

Characteristic	Received Placebo (n = 314)	Received Folate and Vitamin B_{12} (n = 314)
Age, y	71.2 (4.2)	71.6 (5.1)
Sex, No. (%)		
Female	169 (53.8)	169 (53.8)
Male	145 (46.2)	145 (46.2)
Duration of illness, mo	16.9 (3.6)	16.9 (4.4)
Lacunar infarction/atherothrombotic infarction, No. of patients	217/97	212/102
Barthel Index†	67 (16)	68 (16)
Degree of hemiplegia‡		
Hand	4.5 (1.3)	4.5 (1.3)
Leg	4.5 (1.5)	4.5 (1.4)
Body mass index	22.3 (1.9)	22.3 (1.6)
Fallers, No. (%)§	70 (22)	70 (22)
Prevalence of vascular risk factors, No. (%)		
Hypertension	182 (58)	175 (56)
Diabetes mellitus	70 (22)	68 (22)
Hypercholesterolemia	59 (19)	59 (19)
Current smoker	64 (20)	64 (20)
Previous vascular event	62 (20)	58 (18)
BMD, mm Al		
Hemiplegic side‖	2.30 (0.24)	2.30 (0.25)
T score¶	−2.9 (1.0)	−2.9 (1.0)
Intact side	2.41 (0.24)	2.42 (0.24)
T score¶	−1.8 (1.0)	−1.8 (0.9)
Concentration levels		
Plasma homocysteine, μmol/L	19.9 (20.4)	19.9 (21.3)
Serum cobalamin, pg/mL	590 (332)	606 (382)
Serum folate, ng/mL	2.4 (1.5)	2.4 (1.5)

S – Statistics and data

Identical numbers of women, fallers, patients with hypercholesterolemia, smokers in each group **at baseline**:

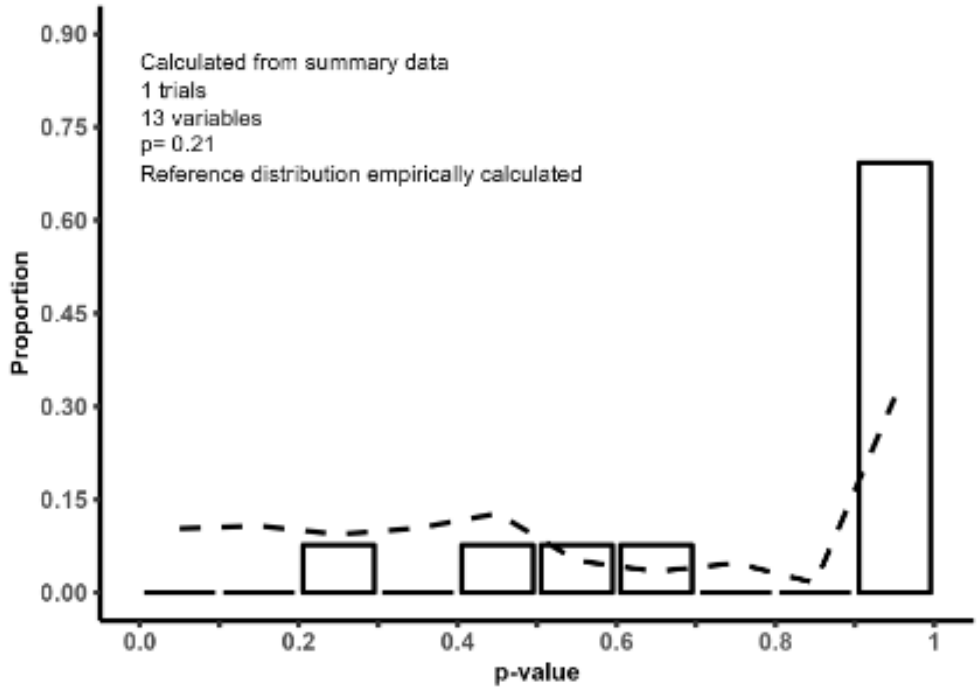
Identical numbers for one variable occur 6/100 times vs 94/100 for difference of ≥2 between groups. Examining all baseline categorical data for this trial:

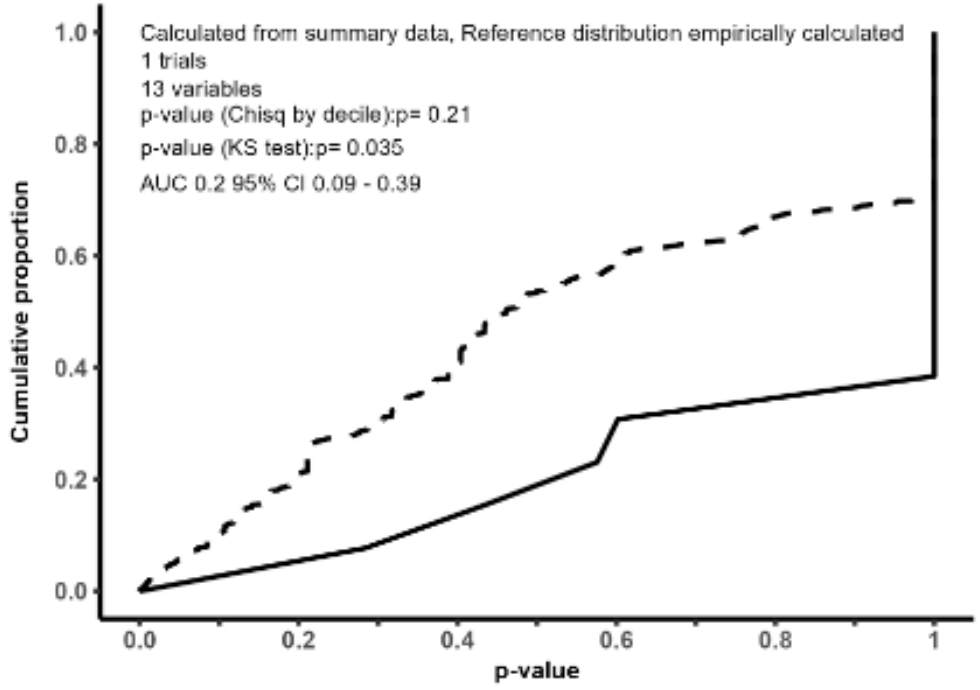


Matching

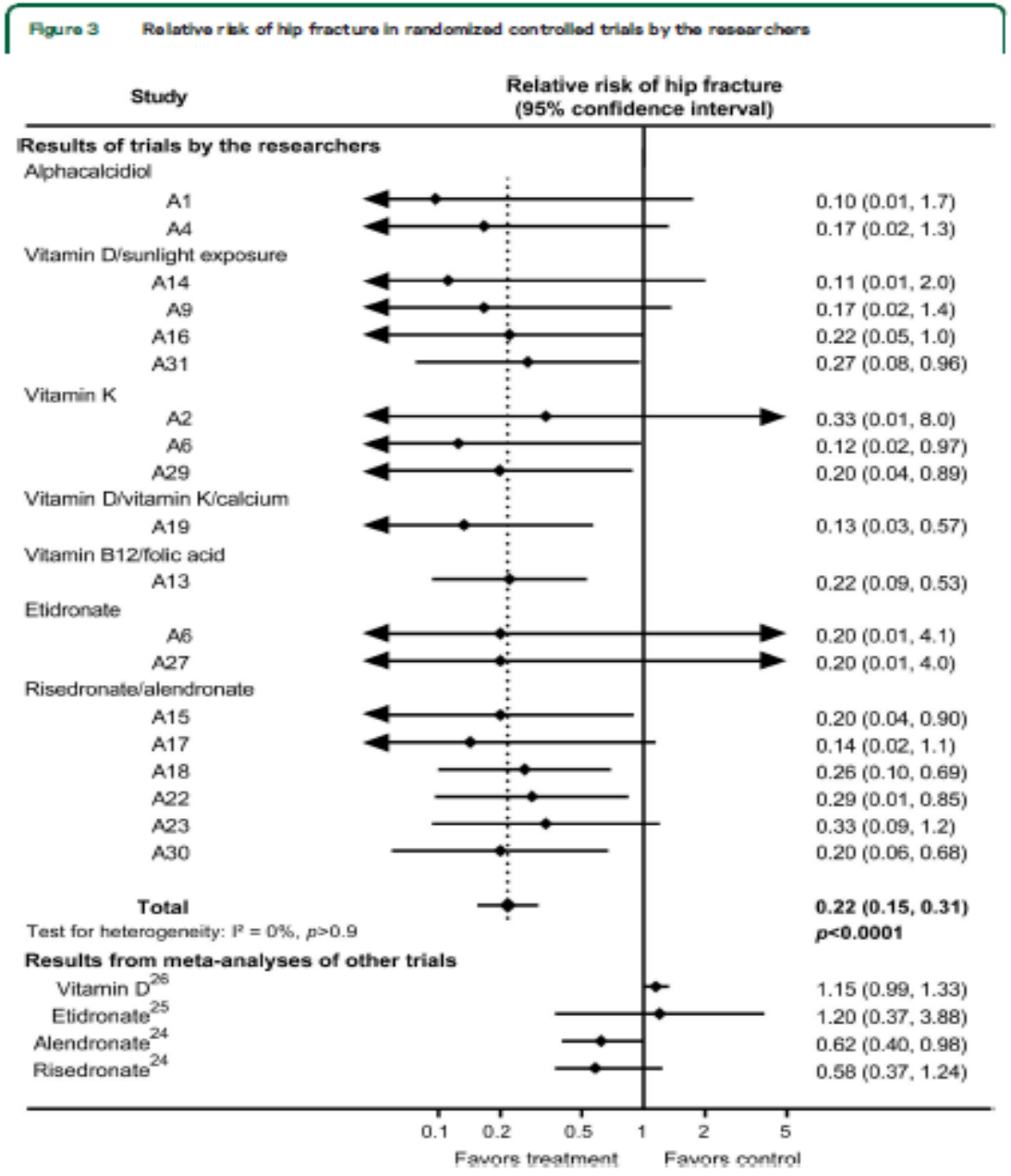
Variable	All variables	1 Sig. dig.	2 Sig. dig.	3 Sig. dig.	4 Sig. dig.	5 Sig. dig.
1 trials						
Number of variables, n (%)						
Mean	13	0 (0.0)	5 (38.5)	8 (61.5)		
SD	13	2 (15.4)	9 (69.2)	2 (15.4)		
Proportion of matching summary statistics in both treatment groups (%)						
Means match	69.2		80.0	62.5		
SDs match	38.5	50.0	44.4	0.0		
Both Mean and SD matches	23.1	50.0	22.2	0.0		
Reference data empirically generated						
1 trials, 10000 simulations						
Number of variables, n (%)						
Mean	130000	99 (0.1)	51745 (39.8)	78156 (60.1)		
SD	130000	20525 (15.8)	89652 (69.0)	19823 (15.2)		
Proportion of matching summary statistics in both treatment groups (%)						
Means match	30.2	100.0	47.0	19.0		
SDs match	35.1	76.2	32.9	2.4		
Both Mean and SD matches	13.2	33.3	11.5	0.0		

S – Statistics and continuous data for one trial report





S - Statistics and data



JAMA paper relative risk for hip fracture 0.22 (95% CI 0.09-0.53)

All Sato trials
Relative risk 0.10-0.33

Meta-analyses of other interventions
Relative risk 0.58 to 1.20

E - Errors

E — Errors

- ☐ Are correct units reported?
- ☐ Are numbers of participants correct and consistent throughout the publication?
- ☐ Are calculations of proportions and percentages correct?
- ☐ Are results internally consistent?
- ☐ Are the results of statistical testing internally consistent and plausible?
- ☐ Are other data errors present?
- ☐ Are there typographical errors?

E - Errors

Fracture rates wrongly calculated:

Hip fracture rates calculated as if there were no drop outs.

Ignored censoring mentioned in methods:

From Fig 2A for the placebo 25 were censored in year 1 and 38 in year 2 ie 63 dropout or fracture.

Assuming average follow-up of 6m and 18m respectively that means rather than 27/(314x2)x100 = 43/1000 patient years

Rates might be 27/(25x0.5+38x1.5+251x2)x1000 =

For placebo 47 hip fractures/1000 patient years

E - Errors

Table 2. Percentage Change of Bone Mineral Density, Plasma Homocysteine, and Serum Vitamins

	Percentage Change, Mean (SEM)					
	After 1 Year			After 2 Years		
	Received Placebo (n = 298)	Received Folate and Vitamin B_{12} (n = 299)	*P* Value*	Received Placebo (n = 278)	Received Folate and Vitamin B_{12} (n = 281)	*P* Value*
Bone mineral density†						
Hemiplegic side	–1.8 (0.2)	–1.7 (0.2)	.89	–3.0 (0.2)	–2.9 (0.2)	.69
Intact side	–1.0 (0.1)	–0.9 (0.1)	.85	–1.9 (0.2)	–1.9 (0.2)	.92
Concentration levels						
Plasma homocysteine	18.2 (1.1)	–36.1 (1.7)	<.001	31.2 (1.4)	–38.1 (1.7)	<.001
Serum cobalamin	–9.1 (3.2)	209.5 (14.6)	<.001	–20.5 (3.0)	214.4 (17.0)	<.001
Serum folate	–12.1 (1.4)	47.2 (3.2)	<.001	–30.1 (1.8)	51.2 (3.8)	<.001

*Comparison between treatment and placebo groups. Based on analysis of covariance model applied to rank-transformed data.
†Bone thickness as an aluminum equivalent measured by computed x-ray densitometry.

- **Why mention concentration levels, which imply a real value, when percentage changes are given?**
- **Note that percentage changes at 2y for plasma homocysteine, serum cobalamin and serum folate are higher than at 1y – we would expect them to be less as pill taking compliance falls with time**
- **Standard errors of the mean (SEM) very similar across time and treatments for bone mineral density**

E - Errors

A total of 628 consecutive poststroke outpatients were recruited from the Futase Social Insurance Hospita, Iizuka, Japan, from April 1, 2000, to May 31, 2001; follow-up occurred until May 30, 2003.

Abstract indicates 2-year follow-up

Why not last follow-up on May 31, 2003?

D - Data duplication and reporting

D — Data duplication and reporting

- ☐ Have the data been published elsewhere?
- ☐ Is any duplicate reporting acknowledged or explained?
- ☐ How many data are duplicate reported?
- ☐ Are duplicate-reported data consistent between publications?
- ☐ Are relevant methods consistent between publications?
- ☐ Is there evidence of duplication of figures?

D – Data duplication and reporting

JAMA paper:

There was no significant difference between the 2 groups in the number of falls per patient during the 2 years (mean, 2.2 [SD, 1.8] in the placebo group and mean, 2.3 [SD, 1.9] in the treatment group).

Original Contribution

Kurume Medical Journal, 57, 117-124, 2010

The Prevention of Hip Fracture with Menatetrenone and Risedronate Plus Calcium Supplementation in Elderly Patients with Alzheimer Disease: A Randomized Controlled Trial

YOSHIHIRO SATO, YOSHIAKI HONDA, KAZUO UMENO*, NORIMASA HAYASHIDA*, JUN IWAMOTO**, TSUYOSHI TAKEDA** AND HIDEO MATSUMOTO**

There was no significant difference between the two groups in the number of falls per subject during the 12 months (2.2±1.8 in the control group and 2.3±1.9 in the treated group).

- **Look for patterns**
- **Here the fall data are the same across two other trial reports**

We've looked at one trial report

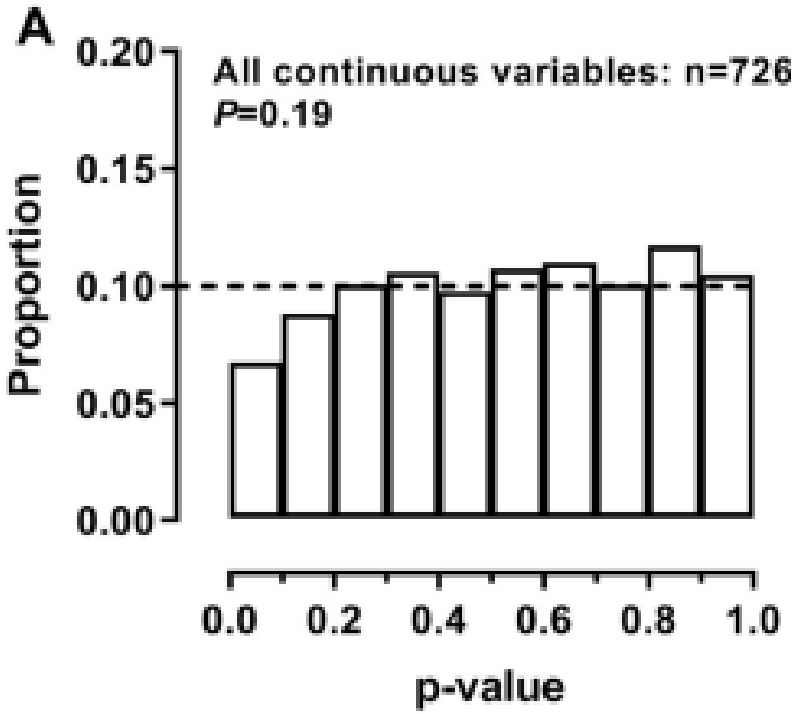
What do we find when we look across multiple trial reports from the same authors?

S - Statistics and data

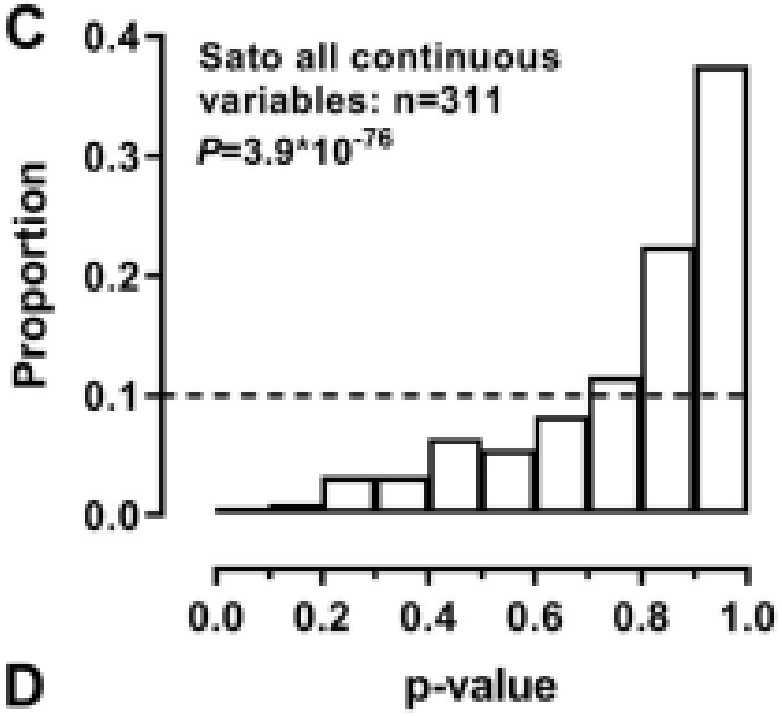
Bolland MJ et al. J Clin Epidemiol 2019;112:67-76; and 110;50-62

Correlation and non-normality of baseline variables or randomization methods do not appear to impact on baseline P-value distribution in genuine RCTs

But the distribution of P-values calculated from rounded summary statistics is not uniform.



13 RCTs Auckland, New Zealand

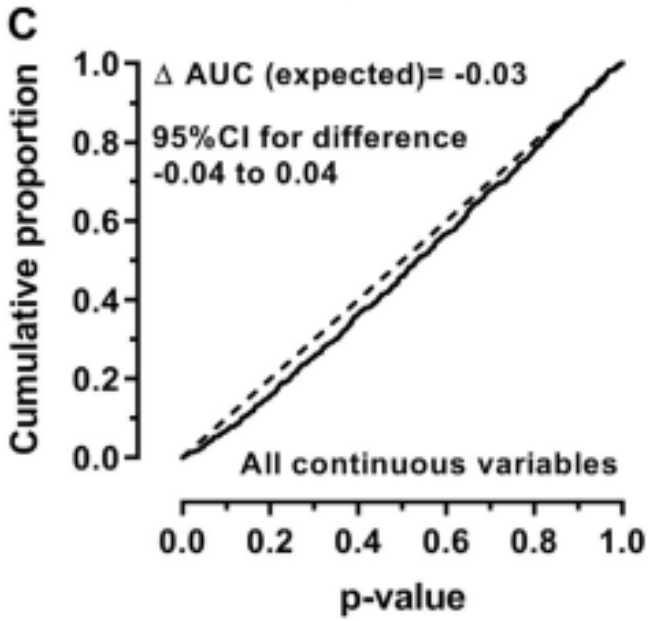


33 Sato/Iwamoto RCTs

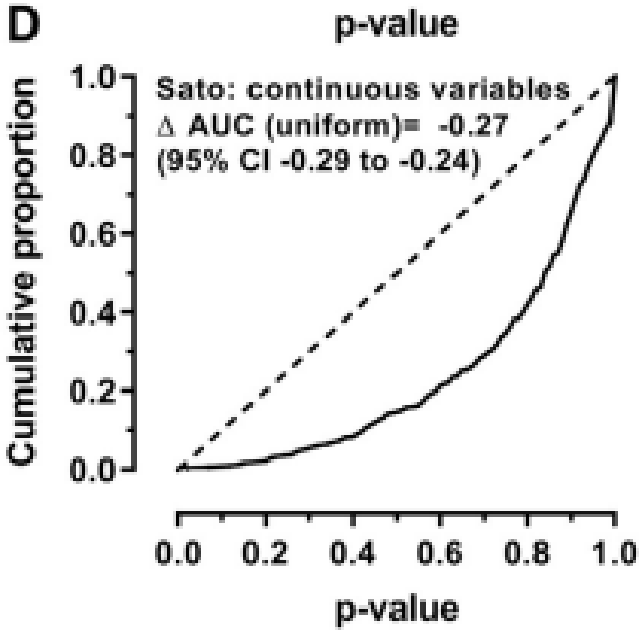
S – Statistics and data

Bolland MJ et al. J Clin Epidemiol 2019;112:67-76.

Cumulative distribution function



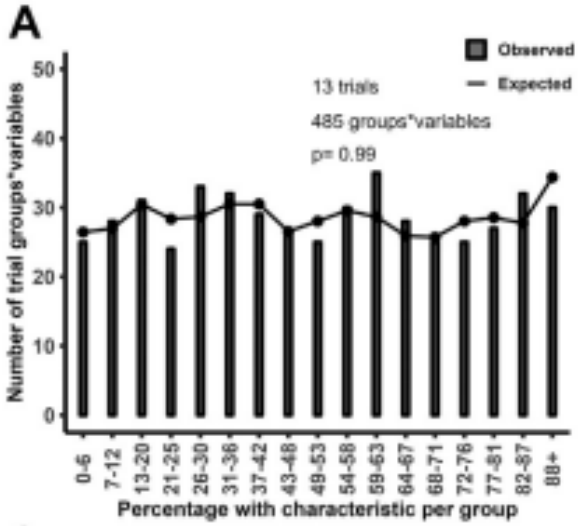
13 RCTs Auckland, New Zealand



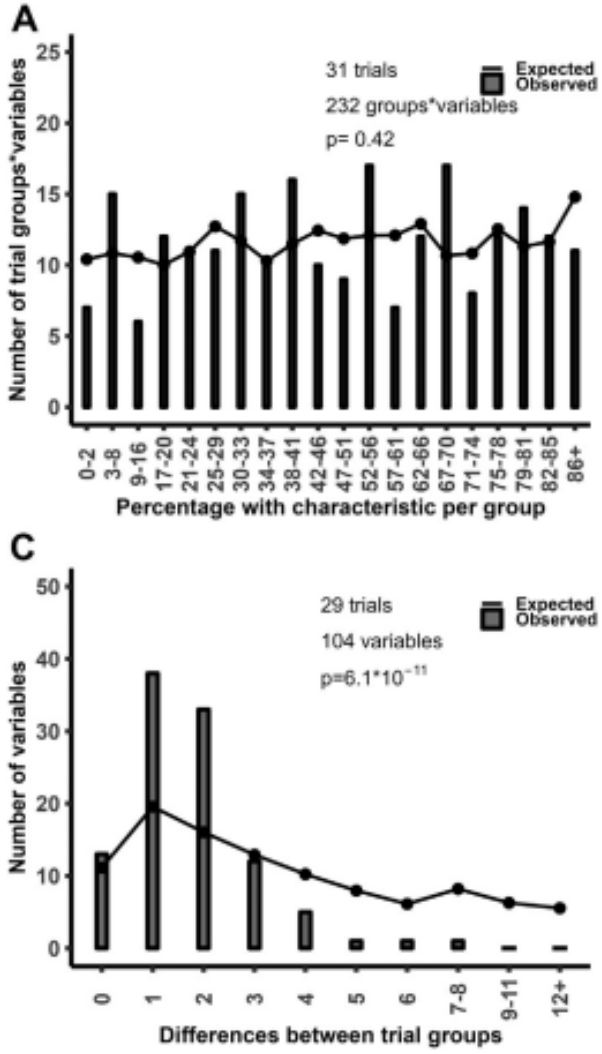
33 Sato/Iwamoto RCTs

S - Statistics and data

Bolland MJ et al. J Clin Epidemiol 2023;154:117-24. Baseline categorical data



13 RCTs Auckland, New Zealand



33 Sato/Iwamoto RCTs

S – Statistics and data

Freely available package “reappraised” for the R statistical programme

https://cran.r-project.org/web/packages/reappraised/index.html

Worked example in Bolland MB et al. J Clin Epidemiol 2024 https://doi.org/10.1016/j.jclinepi.2024.111365

Carlisle JB. Anaesthesia 2012;67:521e37.

Carlisle JB, et al. Anaesthesia 2015; 70:848e58.

Carlisle JB, Loadsman JA. Anaesthesia 2017;72: 17e27.

Bolland MJ, et al. J Clin Epidemiol 2019;112:67e76.

Bolland MJ, et al. J Clin Epidemiol 2019;110:50e62.

Bolland MJ, et al. J Clin Epidemiol 2021;131: 22e9

Bolland MJ, et al. J Clin Epidemiol 2021;136:180e8.

Bolland MJ, et al. J Clin Epidemiol 2023;154:117e24

S – Statistics and data – other tests

http://www.prepubmed.org/data_thugging/

Tools for Data Thugging

So you want to be a data thug? These tools will help. Although designed for **integer** data, as long as the data has a consistent granularity sometimes a simple transformation is all that is needed to get the tools to work with your data.

GRIM Test
Use this test if you want to check means to two decimals and your sample size is less than 100.

General GRIM Test
Have means reported to a different decimal precision? Have a composite measure? Then this test is for you.

GRIMMER Test
Want to check standard deviations, standard errors, or variances? Then go here.

SPRITE
Want to reconstruct the data? Then go here, or Nick Brown's app.

One-Way ANOVA (any data, not just integer data)
Want to check F statistics of one-way ANOVAs while allowing for rounding uncertainty? Then go here.

Two-Way ANOVA (any data, not just integer data)
Want to check F statistics of two-way ANOVAs while allowing for rounding uncertainty? Then go here.

Benford's Law
https://www.statology.org/benfords-law/

Questions

- **Assume that you are an editor receiving this information**
  - What would you do next?
- **How soon would you want to alert readers as to these issues?**
- **Which of all these pieces of evidence do you think most indicates compromised research integrity?**

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

HSRU
Promoting Excellence in Health Services Research

Comments please

Alison Avenell a.avenell@abdn.ac.uk

Twitter: @reappraisedblog

Blog: https:/reappraised.wordpress.com

REAPPRAISED CHECKLIST:
https://www.nature.com/articles/d41586-019-03959-6

Publication:

https://www.tandfonline.com/doi/full/10.1080/08989621.2022.2082290

- Bolland MJ, Avenell A, Gamble GD, Grey A. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. Neurology 2016; 87: 1-12.
- Bolland M, Avenell A, Gamble G, Buranyi S, Grey A. A randomised investigation of journal responses to academic and journalist enquiry about possible scientific misconduct BMC Research Notes 2018;521:https://doi.org/10.1186/s13104-018-3613-1
- Grey A, Bolland MJ, Gamble G, Avenell A. Quality of reports of investigations of research integrity by academic institutions. Res Integr Peer Rev 2019;4:3 https://doi.org/10.1186/s41073-019-0062-x
- Bolland MJ, Gamble GD, Avenell A, Grey A. Rounding, but not randomization method, non-normality, or correlation, affected baseline P-value distributions in randomized trials. J Clin Bolland MJ, Gamble G, Grey A, Avenell A, Lumley T. Baseline P value distributions in randomized trials were uniform for continuous but not categorical variables. J Clin Epidemiol 2019;112:67-76
- Avenell A, Stewart F, Grey A, Gamble G, Bolland M. An investigation into the impact and implications of published papers from retracted research: systematic search of affected literature. BMJ Open 2019; 9: e031909.
- Mistry V, Grey A, Bolland MJ. Publication rates after the first retraction for biomedical researchers with multiple retracted publications. Account Res 2019;26:277-287.
- Grey A, Avenell A, Gamble, G, Bolland M. Assessing and raising concerns about duplicate publication, authorship transgressions and data errors in a body of preclinical research. Sci Eng Ethics 2019;26:2069-9
- Bolland MJ, Gamble G, Grey A, Avenell A. Empirically generated reference proportions for baseline p-values from rounded summary statistics. Anaesthesia 2020; 75:1685-1687.
- Bolland MJ, Grey A, Avenell A, Klein AA. Correcting the scientific record - a broken system? Account Res 2021;28:265-79

- Bolland MJ, Gamble GD, Avenell A, Cooper D, Grey A. Participant withdrawals were unusually distributed in randomized trials with integrity concerns: a statistical investigation. J Clin Epidemiol 2021;131:22-29.
- Grey A, Portch R, Gaby A, Grey H, Bolland M. Clinical trial registry documents and publication integrity. Account Res 2021;28:149-161.
- Bolland MJ, Grey A, Avenell AA. Citation of retracted publications: a challenging problem. Account Res 2022;29:18-25.
- Grey A, Avenell A, Bolland M. Timeliness and content of retraction notices for publications by a single research group. Account Res 2022;29:347-78
- Bolland MJ, Gamble GD, Avenell A, Grey A. Identical summary statistics were uncommon in randomised trials and cohort studies. J Clin Epidemiol 2021;136:180-8
- Avenell A, Bolland M, Gamble GD, Grey A. A randomized trial alerting authors, with or without co-authors or editors, that research they cited in systematic reviews and guidelines has been retracted. DOI: 10.1080/08989621.2022.2082290
- Grey A, Avenell A, Bolland MJ. Procrastination and inconsistency: expressions of concern for publications with compromised integrity. Acc Res 2022; DOI 10.1080/08989621.2022.2112572
- Bolland MJ, Gamble GD, Avenell A, Cooper DJ, Grey A. Distributions of baseline categorical variables differ from the expected distributions in randomized trials with integrity concerns. J Clin Epidemiol 2023;154:117-24
- Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CJ. Check for publication integrity before misconduct. Nature 2020;577:167-9

Name		Presentation details
Leslie	McIntosh	Monday posters PP-208: **Identifying fabricated networks within authorship-for-sale enterprises**
Grainne	McNamara	Monday posters PP-166: **The prevalence of image duplications in pre-publication journal submissions**
Danielle	Oste	Monday posters PP-068: **Non-verifiable cell lines in cancer research papers describing human gene research**
Anna	Abalkina	Monday 15:30 - 16:30 OP06.1: **Challenges posed by hijacked journals in Scopus**
Tianyi	Hu	Monday 15:30 - 16:30 OP07.1: **Artificial Intelligence-based Feature Recognition of "Paper Mill" Patterns**
Svetlana	Kleiner	Monday 15:30 - 16:30 OP07.3: **Findings From a Paper Mill**
Daniel	Acuna	Monday 17:00 - 18:30 OP14.2: **Correlating Indicators of Research Integrity: An Analysis of Image Integrity, Statistical Inconsistencies, and Missing Citations**
Rene	Aquarius	Monday 17:00 - 18:30 OP14.1: **Uncovering Duplicated Images in Scientific Literature - Systematic Review as a Tool for Detection**
Guillaume	Cabanac	Monday 17:00 - 18:30 OP14.4: **Year after year: Tortured conference series thriving in computer science**
Gilles	Hubert	Monday 17:00 - 18:30 OP14.5: **Bad smells in reviewers' reports? Text-mining the MDPI Open Peer Review Corpus**
Jack	Wilkinson	Tuesday 10:30 - 12:00 OP17.3: **Development of a tool (INSPECT-SR) to identify problematic randomised controlled trials in systematic reviews of health interventions**
Anna	Abalkina	Tuesday posters PP-189: **Publication strategies of paper mills: a case-study from the Tanu.pro paper mill/brokerage company**
Elizabeth	Hay	Tuesday posters PP-185: **Integrating Publication Ethics and Research Integrity into your editorial submission and peer review workflows: tips and techniques to achieve effective communication**
Leslie	McIntosh	Tuesday posters PP-127: **A Taxonomy of Retractions: Enhancing Trust Assessment in Scholarly Research**
Elizabeth	Moylan	Tuesday 15:30 - 16:30 OP20.4: **How can human experience influence the development of integrity tools and workflows? A case study from image screening**
Lisa	Parker	Tuesday 13:30 - 15:00 Plenary D: **Paper mill submissions- identifying red flags for moving targets**
Thomas	Stoeger	Tuesday 13:30 - 15:00 Plenary D: **How research paper mills avoid consequences of journal de-indexing**
Deborah	Kahn	Tuesday 13:30 - 15:00 Plenary D: **United2Act on paper mills: an update**
Leslie	McIntosh	Tuesday 15:30 - 16:30 OP21.1: **A taxonomy of scholarly disinformation to enhance research integrity**
Joris	van Rossum	Wednesday 13:00 - 14:30 Plenary F: **Perspective of publishers and policy makers**