2026 PCRS ANNUAL MEETING

REPRODUCTIVE FRONTIERS: BRIDGING BIOLOGY, PRACTICE, AND POSSIBILITY

MARCH 18-22 | RANCHO MIRAGE, CA

PACIFIC COAST REPRODUCTIVE SOCIETY

What gets Published and Why (critiquing a manuscript)

- Kurt Barnhart, MD, MSCE
- Editor in chief
- Fertility and Sterility
- University of Pennsylvania

2026 PCRS ANNUAL MEETING

REPRODUCTIVE FRONTIERS:
BRIDGING BIOLOGY, PRACTICE, AND POSSIBILITY

MARCH 18-22 | RANCHO MIRAGE, CA

Disclosure Slide

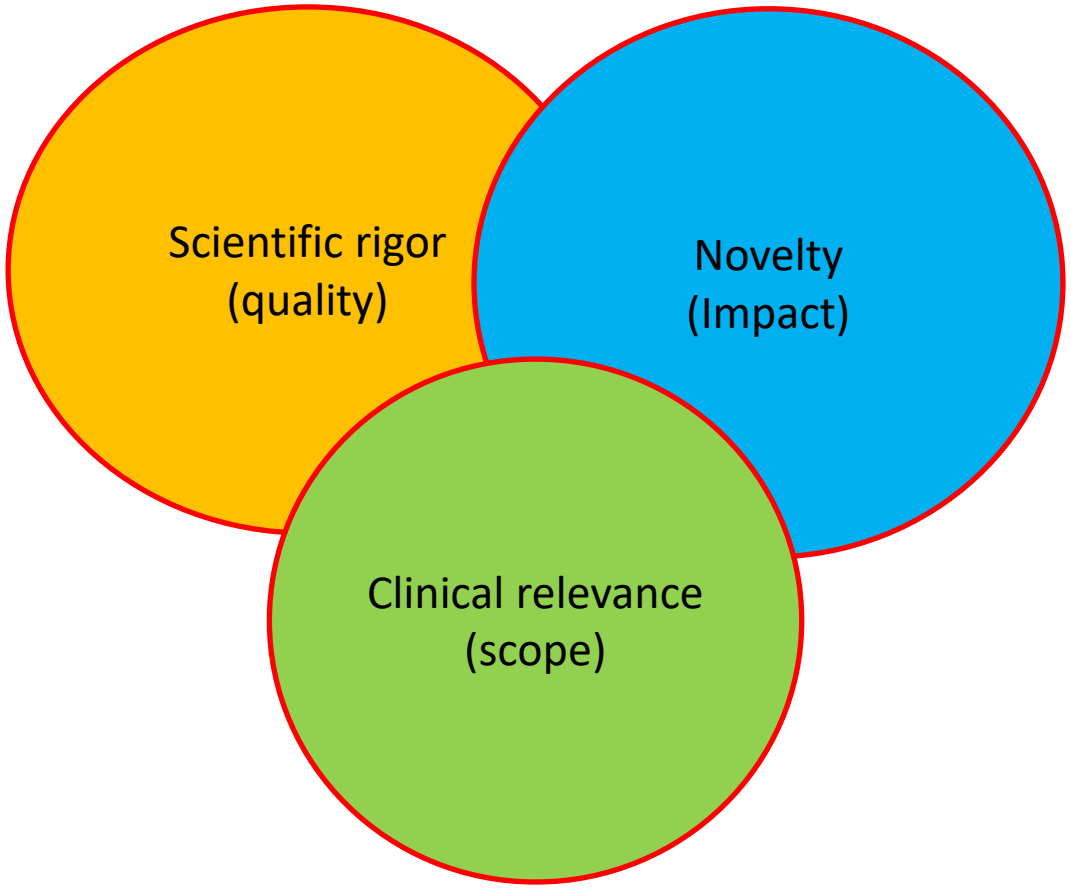
- Neither I nor members of my immediate family have any actual or potential financial interests to disclose relating to the content of this presentation.

Needs Assessment Statement and Expected Learning Outcomes

- The mainstay of medical publishing is scientific integrity. In this session we will review what type of articles are most likely to be published in Fertility and Sterility and how to assess them as a peer reviewer. One does not have to be a statistician to assess the credibility and value of a manuscript. Attention will be placed on how to assess the hypothesis and methods of a paper using real world examples.

- Objectives
- Demonstrate the balance of characteristics of a manuscript that has a high likelihood of acceptance in Fertility and Sterility.
- Assess the impact of components of a well written manuscript.
- Integrate scientific and clinal skills to assess the validity of a manuscript.

Desired Characteristics of a Manuscript



Why Publish?

Disseminate scientific information

(greater good)

Become an authority

(prestige)

Promotion

(publish or perish)

Revenue

(IP, marketing, bonus)

How to Ensure Quality

Peer review

Most rigorous review process

There is desire for improvement

Biased or careless review

Quality of the Journal

Reputation, rigor, selectivity

Editorial decisions need to be made

What makes good science?
(How do you get a good review?)

Hypothesis based

- Discovery

Understand the hierarchy of design

- The appropriate design for the question

Need to think about subtlety

- You are the expert and can educate

What is the significance of the finding?

- Novel, change practice, adds to debate

Can it be susceptible to confounding or bias?

- Alternative reasons
- Plausibility

Objective conclusions

Ability to reproduce

Can others check

Replication is the heart of science

Science gives with prove

Evidence based medical practice

How to Publish in a High Impact Journal

- Imagine your research telling a story (with a hypothesis)
  - Logic, not chronology (set up the story and provide evidence)
- Know your audience
  - Be clear in your context
- Be short, clean and clear
  - Do not cram the methodology or the lit review: Strong verbs and short sentences
- Make good figures/tables
  - They should stand alone
- Title and abstract matter
  - Need to attract readers

Rigor: Is your study better than previous?
Novelty: Is your study the first (or fill a gap)?
Scope: Does your study impact care?

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

PLoS Medicine | www.plosmedicine.org 0696 August 2005 | Volume 2 | Issue 8 | e124

1) Lack of confirmation: Replication is the foundation of science

Ill founded strategy of claiming conclusive findings based on a p value of less than 0.05

Type 1 error: 5%

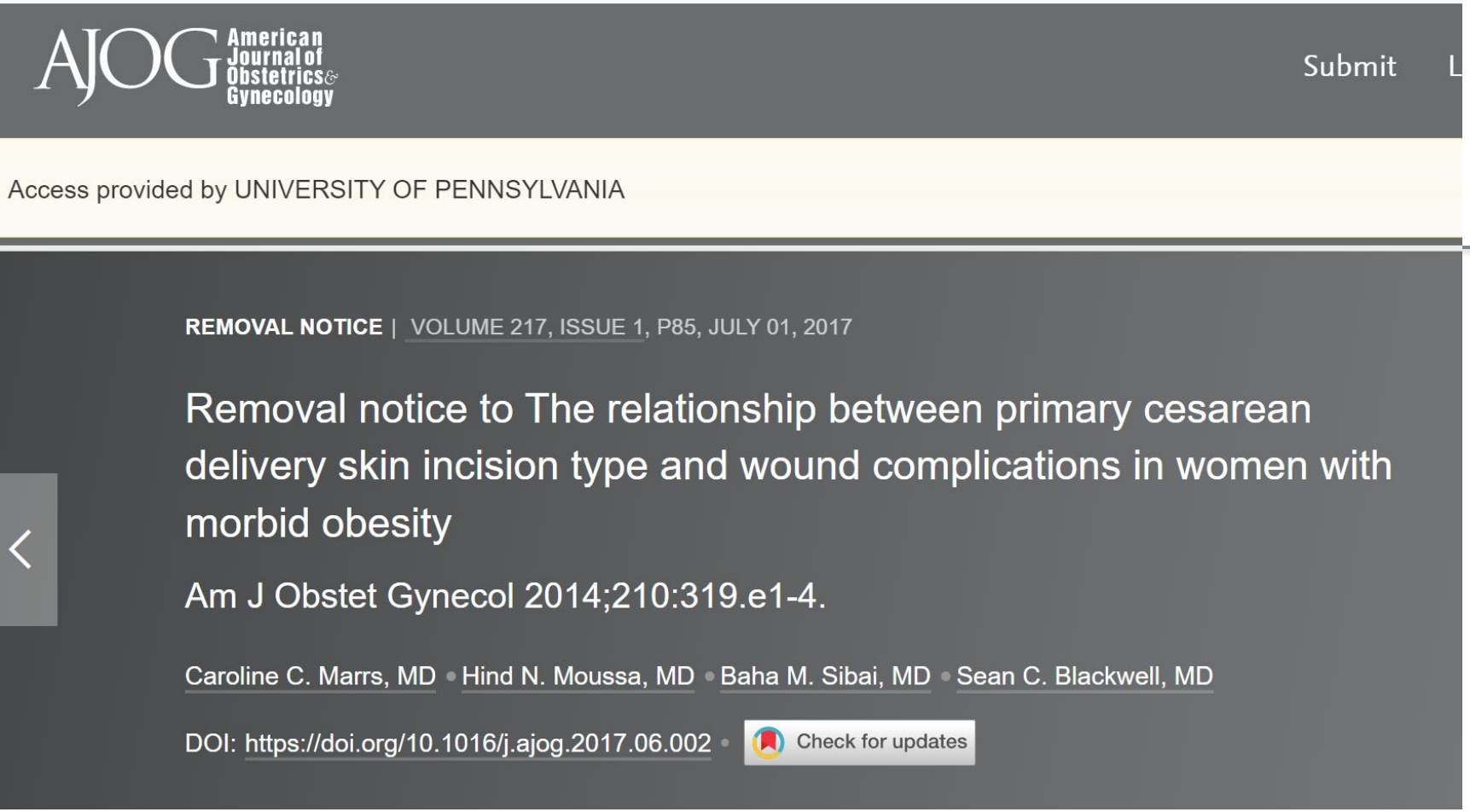
Type 2 error: 20%

2) Signs that a studies finding are **likely to be false**:

- Small sample size — **Stopped early? (regression to the mean)**
- Small effect size
- When there is a greater number of tested relationships (with lesser preselection) — **Fishing, data mining**
- Greater flexibility in study design — **Gerrymandering**
- Low precision in definitions and outcomes. — **Confounding**
- Great financial (or other) interest — **Bias**
- The “hotter” the field, the less likely research findings are true — **Exaggeration**

The Evolution of Bias

- **Transparency**
  - Improvement in methods, what was done, why, stats
- **Conflict of interest**
  - Cherry picking, secondary hypothesis, spin
- **Bias**
  - “survivor bias”, “bias of positive findings”
- **Fraud**
  - “FFP”: fabrication, falsification, and plagiarism (lying, cheating, stealing)
  - Past few years fraud in reproductive medicine (RCTs), >700 articles retracted (20 in F&S) 15 letters of concern, 6 editors notes



Are all errors Intentional?

Likely Error, not Fraud

The original publication reported that univariate analysis showed that a vertical skin incision in obese women undergoing Cesarean delivery was associated with a higher odds ratio for wound complications than a transverse skin incision. Multivariable analyses showed a reversal of the association (i.e. the odds of wound complications were lower in women with a vertical skin incision). However, **there was an error in the way the variable was entered in the logistic analysis**. Re-analysis with the correct coding of the variable indicates that a transverse skin incision is associated with decreased odds of wound complication compared to a vertical skin incision.

Fraud: Several Investigators have Come Under Scrutiny

"While performing data extraction, we noticed similarities between values in the baseline characteristics (i.e. age, Body Mass Index (BMI), parity) and outcome tables in different randomized controlled trials (RCTs) authored by Ahmed Badawy and Hatem Abu Hashim, from the Mansoura University in Egypt. Following these concerns, we undertook a systematic assessment of the RCTs published by these two authors."

- **Professor Ben Mol** and colleagues at Monash University, Victoria Australia

- 24 studies from Badawy (fellow of the Royal College of ObGyn) and 11 from Abu Hashim: 2006 – 2016 ***27 found to be problematic!***

Identical Table 1 in (2004)a & (2006)

3. Unfer (2004)a
4. Casini (2006)

3. (2004)a

TABLE 1

Characteristics of participants at baseline.

Characteristic	Group A (n = 179)	Group B (n = 197)
Age at start of treatment (years)	49 ± 4.3	50 ± 3.9
Height (cm)	165.5 ± 5.3	163.4 ± 5.6
Weight (kg)	67 ± 10.2	65.8 ± 10.7
Mean age at menopause (y)	50.2 ± 6.5	49.8 ± 6.3
Years of menopause	5.6 ± 4.3	5.8 ± 4.5
Parity (n)	2.8 ± 1.8	2.7 ± 1.7

Note: Unless otherwise specified, values are mean ± SD. The *P* value for all data was not significant.

Phytoestrogen long-term treatment. Fertil Steril 2004.

4. (2006)

TABLE 1

Characteristics of participants at baseline.

Characteristic	Group A (n = 39)	Group B (n = 39)
Age at start of treatment (y)	49 ± 4.3	50 ± 3.9
Height (cm)	165.5 ± 5.3	163.4 ± 5.6
Weight (kg)	67 ± 10.2	65.8 ± 10.7
Mean age at menopause (y)	50.2 ± 6.5	49.8 ± 6.3
Years of menopause	5.6 ± 4.3	5.8 ± 4.5
Parity (n)	2.8 ± 1.8	2.7 ± 1.7

Note: Values are mean ± SD. The *P* value for all data was not significant.

PHY and psychological assessment in menopause. Fertil Steril 2006.

- **ON TOP OF THAT, BUT NOT STATED BY THE JOURNAL** (2004)a[3] reports mean age at menopause is 50 & duration of menopause 5.6 and 5.8 years, suggesting age at study commencement would be between 55 - 56 years, however, it is 49 and 50; **therefore, either initial age commencement is incorrectly reported or these data points cannot be true**

J Assist Reprod Genet (2007) 24:400–405
DOI 10.1007/s10815-007-9159-0

ASSISTED REPRODUCTION

Efficacy of a combined protocol of urinary and recombinant follicle-stimulating hormone used for ovarian stimulation of patients undergoing ICSI cycle

Table 1 Demographic data and stimulation outcome

	uFSH/rFSH group A	rFSH group B	*p* value
Patients (*n*)	58	61	
Mean age (years) ±SD	34.1±2.5	35.1±3.1	NS
Mean BMI±SD	22.6±1.8	23.6±1.7	NS
Mean duration of sterility (years) ±SD	5.1±1.2	4.1±1.4	NS
Primary infertility % (*n*)	71.9 (41)	74.6 (44)	NS
Tubal factor % (*n*)	47.4 (27)	44.1 (26)	NS
Male factor % (*n*)	40.3 (23)	40.6 (24)	NS
Unexplained infertility % (*n*)	12.3 (7)	15.3 (9)	NS
Duration of stimulation (days)	11.4±2.1	13.1±2.2	NS
Estradiol level on hCG day (pg/ml)	2,056±560	1,987±699	NS
Endometrial thickness on hCG day (mm)	10.8±2.1	11.2±3.1	NS

NS: not significant

Recruitment June 2005 to March 2006.

Ovarian stimulation protocols based on follicle-stimulating hormone glycosylation pattern: impact on oocyte quality and clinical outcome

TABLE 1

Demographic data and stimulation outcomes.

	hFSH/rFSH Group A	rFSH Group B	hFSH Group C	*P* value
Patients (n)	63	65	60	
Mean age (y) ± SD	36.6 ± 3.23	34.9 ± 3.74	35.0 ± 4.09	.169
Mean BMI ± SD	23.6 ± 1.8	23.6 ± 1.7	23.1 ± 1.8	.289
Mean duration of sterility (y) ± SD	4.1 ± 1.2	4.1 ± 1.4	4.5 ± 1.3	.232
Cause of infertility				
Primary, % (n)	71.4 (45)	71.3 (46)	71.7 (43)	.321
Tubal factor, % (n)	42.8 (27)	43.0 (28)	43.4 (26)	.154
Male factor, % (n)	36.5 (23)	36.9 (24)	36.6 (22)	.469
Unexplained, % (n)	6.4 (4)	6.2 (4)	6.7 (4)	.443
Other	14.2 (9)	13.8 (9)	13.3 (8)	.411
Duration of stimulation (d)	11.6 ± 2.1	11.2 ± 2.2	11.4 ± 3.1	.334
Estradiol level on hCG day (pg/mL)	2056 ± 560	1987 ± 699	1985 ± 658	.877
Total FSH dose	2610 ± 472.5	2520 ± 495	2565 ± 697.5	.971
Endometrial thickness on hCG day (mm)	10.6 ± 2.1	10.4 ± 2.1	10.5 ± 2.4	.957

Note: No statistically significant differences observed between the groups. BMI = body mass index; hCG = human chorionic gonadotropin; hFSH = human-derived follicle-stimulating hormone; SD = standard deviation; rFSH = recombinant follicle-stimulating hormone.

Selman. FSH glycosylation pattern and oocyte quality. Fertil Steril 2010.

Recruitment January 2008 to February 2009

Identical number

1-2 digits difference

Effect of oxytocin infusion on reducing the blood loss during abdominal myomectomy: a double-blind randomised controlled trial. BJOG. 2017 Jan;124(2):292-298

Table 1. Patient's characteristics and preoperative variables in two groups

	Study group (n = 40)	Placebo group (n = 40)	P
Age (years)	38.08 ± 0.74	36.25 ± 0.74	0.62
Weight (kg)	67.35 ± 1.59	68.78 ± 1.56	0.99
Height (cm)	160.25 ± 0.75	160.65 ± 0.68	0.26
Gravidity (number)	0–4	0–3	0.85
Uterine size (week)	14.8 ± 0.25	15.05 ± 0.62	0.74
Largest fibroid	14.8 ± 0.25	15.05 ± 0.62	
Length	70.45 ± 4.15	72.80 ± 5.52	0.32
Width	66.33 ± 3.36	68.95 ± 5.61	0.45

Data are presented as mean ± SD or range, based on Student's t-test.

P < .0001
P = .0001
P = .0146
P = .0205
P = .0205
P = .0345
P = .0133

Remarkable similar standard deviations

Similar standard deviations

Only 3 and 6

Uterine size and largest fibroid similar

Data is incorrect AND does not follow expected distributions (forensic accounting)

What is Wrong With this Table

Polymorphism	Cases (n = 246), n (%)	Controls (n = 492), n (%)	Crude OR (95% CI)	*P* value	Adjusted OR[a] (95% CI)	*P* value
GNB3 C825T						
CC	44 (17.9)	162 (32.9)	1.0 (Referent)			
CT	144 (58.5)	240 (48.9)	2.65 (1.32–4.14)	.007	2.34 (1.24–3.86)	.008
TT	58 (23.6)	90 (18.3)	2.58 (1.12–4.32)	.008	2.14 (1.10–4.26)	.009
C allele	232 (47.2)	564 (57.3)	1.0 (Referent)			
T allele	260 (52.8)	420 (42.7)	2.72 (1.18–7.14)	.002	2.16 (1.16–6.98)	.004
Dominant model						
CC	44 (17.9)	162 (32.9)	1.0 (Referent)			
CT+TT	202 (72.1)	330 (67.1)	4.14 (1.05–15.6)	.027	3.74 (1.11–12.4)	.022

All of the data are EVEN numbers!

The probability that this happens by chance is 0.0977%, or 1 in 1,000

How do You Prove Fraud?

- **Accusations are serious and damaging**
- **Ask for the data**
- **Forensic statistics**
- **Common sense and verification**
  - **Time from study to submission (no time for birth)**
  - **No drop out, or incredible enrolment**
- **<u>Need objective data</u>**
  - **Billing for Hospitals**

    Number of subjects in a clinic

How do you respond to potential misconduct?

- **Reject the paper**
  - The most control is during the review process
- **If accepted the paper must be reviewed**
  - Editor, committee, publisher
  - Retracted
  - Letter of concern
  - Letter of note (erratum)

What about consequences to our patients?

Meta-Analysis > Int J Gynaecol Obstet. 2018 Jun;141(3):295-303. doi: 10.1002/ijgo.12473.
Epub 2018 Mar 15.

Systematic review and meta-analysis of single-dose and non-single-dose methotrexate protocols in the treatment of ectopic pregnancy

Jin-Sung Yuk [1], Jung Hun Lee [1], Won I Park [2], Hyeong Sik Ahn [3], Hyun Ju... [3]

Affiliations + expand

PMID: 29485731 DOI: 10.1002/ijgo.12473

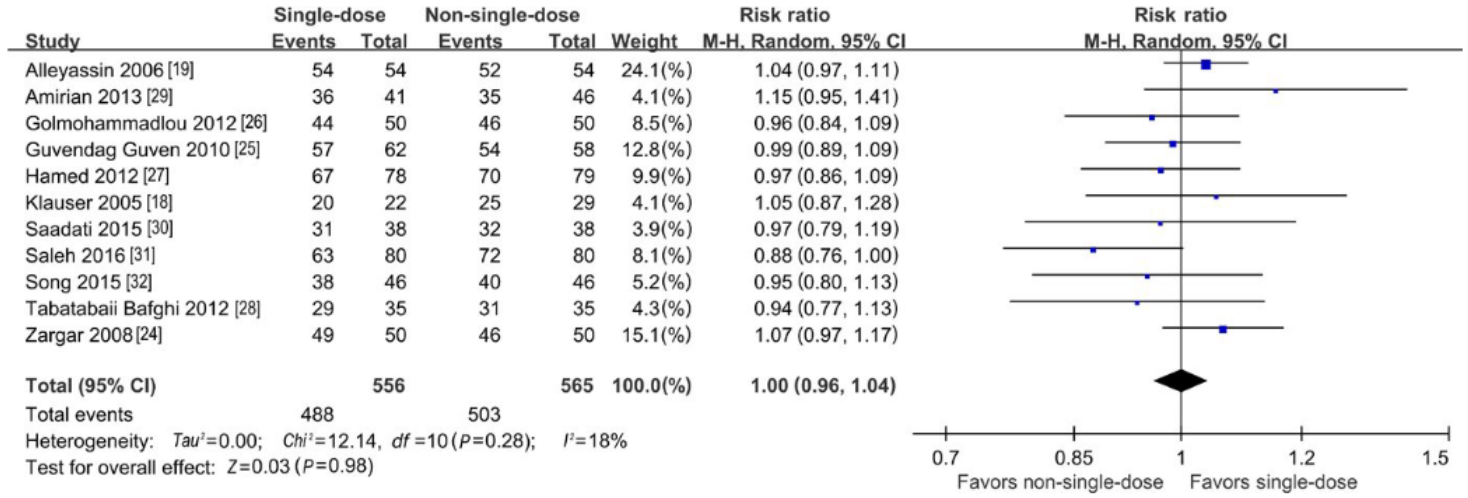
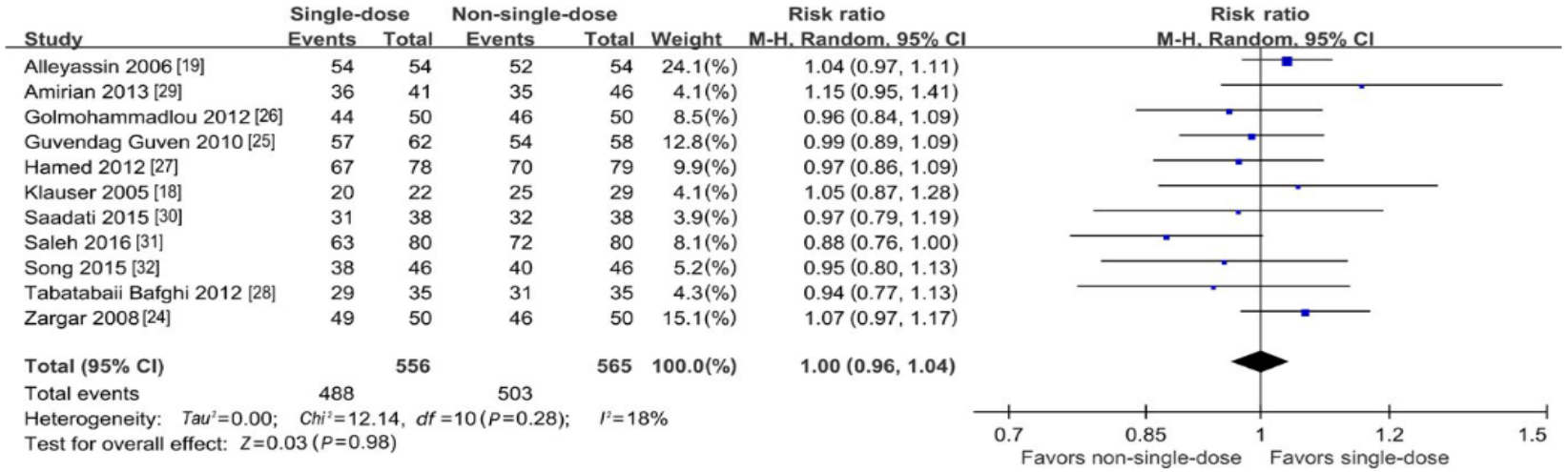


FIGURE 2 Success rate in the treatment of ectopic pregnancy. Abbreviations: CI, confidence interval; M-H, Mantel-Haenszel test; Random, random-effects model.

Main results: The single-dose and non-single-dose protocols had similar success rates (RR 1.00, 95% CI 0.96-1.04; 11 trials, 1121 patients, I^2 =18%). The non-single-dose protocols had a higher adverse effect rate than did the single-dose protocol (RR 0.73, 95% CI 0.59-0.91; nine trials, 934 patients, I^2 =0%).

Conclusions: The single-dose methotrexate protocol was the optimal protocol for the medical treatment of ectopic pregnancy.



Alleyassin : the numbers of subjects, success rate, and the RR of the paper are not correct in the meta-analysis

Amiria: this is a resident's thesis in Iran; it is not properly blinded and likely not randomized.

Golmohamadlou: The abstract of the paper is in English, the rest in Arabic. As far as one can tell there is no blinding.

Hamed: Incorrect numbers and RR are in the meta-analysis, and do not reflect the raw data of the original paper.

Klauser: This is an unpublished abstract at a meeting and is not a manuscript

Saadati: The number from the papers are not the same used in the meta-analysis

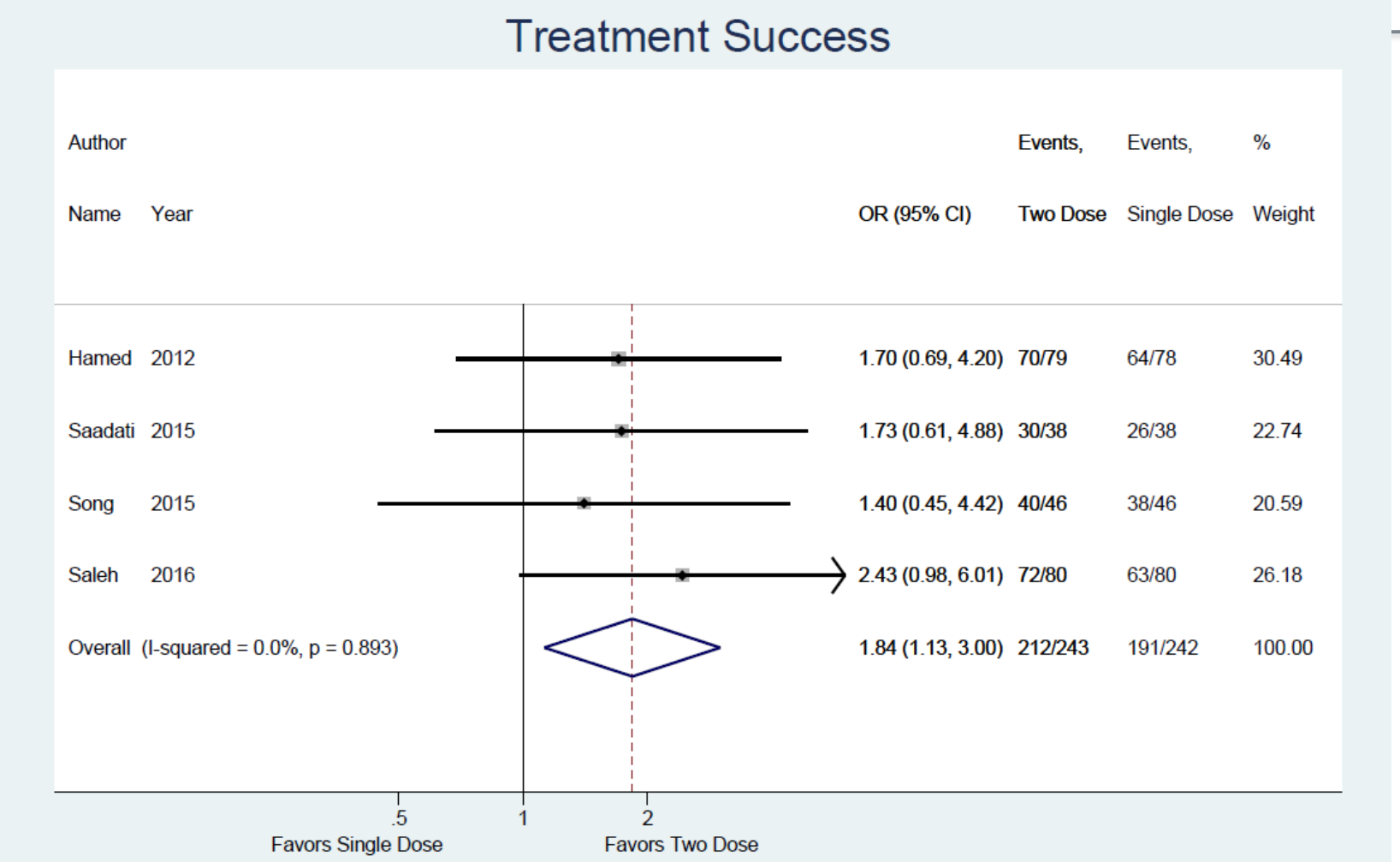
Song: this is a well-done study with the effect opposite of what the author of the meta-analysis claim is the summary finding. Of note this paper has one of the lowest weights and is not significantly smaller than other studies.

Zargar: not at randomized trial. Conducted in Iran with enrollment "one by one" with inclusion criteria vastly different form the use of MTX in other studies.

This paper at best is very, very sloppy. *However, given that all "errors" are in one directly, I think this is fraud.*

Of note, I reviewed this paper for two journals (rejected). This paper was simply readdressed and published unchanged.

Two dose vs Single dose: Meta-Analysis



Two-dose versus single-dose methotrexate for treatment of ectopic pregnancy: a meta-analysis

Snigdha Alur-Gupta, MD; Laura G. Cooney, MD; Suneeta Senapati, MD, MSCE; Mary D. Sammel, ScD; Kurt T. Barnhart, MD, MSCE

AJOG at a Glance

Why was this study conducted?

This study was conducted to compare the odds of treatment success, side effects, surgery for ruptured ectopic pregnancy, and length of follow-up of commonly used methotrexate protocols for the treatment of ectopic pregnancy.

Key findings

The 2-dose protocol was superior to the single-dose protocol in treatment success, including in women at higher risk for failure, such as those with high human chorionic gonadatropin and large adnexal mass.

What does this add to what is known?

This adds an updated meta-analysis of a 2-dose versus a single-dose protocol and additional analyses of a multi-dose versus a single-dose protocol using only quality randomized controlled trials.

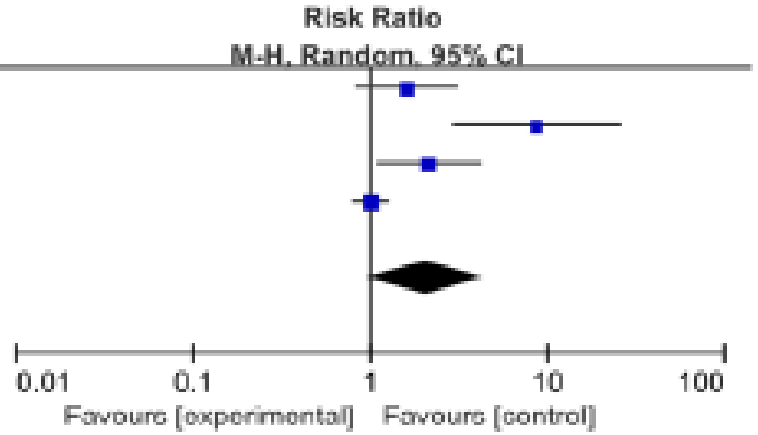
Treatment of women with subclinical hypothyroidism (Live Birth)

B

Study or Subgroup	LT4 supplementation Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Year
Negro 2005	16	43	10	43	25.3%	1.60 [0.82, 3.12]	2005
Abdel Rahman 2010	26	35	3	35	18.9%	8.67 [2.89, 26.02]	2010
Kim 2011	17	32	8	32	25.1%	2.13 [1.07, 4.21]	2011
Wang 2017	90	282	91	285	30.6%	1.00 [0.79, 1.27]	2017
Total (95% CI)		392		395	100.0%	2.05 [0.96, 4.36]	
Total events	149		112				

Heterogeneity: Tau² = 0.47; Chi² = 18.82, df = 3 (P = 0.0003); I² = 84%

Test for overall effect: Z = 1.86 (P = 0.06)



Rao et al. Reproductive Biology and Endocrinology (2018) 16:92

Table 2
Characteristics of 70 Women With Subclinical Hypothyroidism 1 Month into the Study

Characteristics	Group A (levothyroxine treatment) (n = 35)	Group B (placebo) (n = 35)	*P* value
Basal FSH, mean (SD), mIU/mL[a]	7.6 (2.2) (range, 3.2-12)	6.3 (2) (range, 2.3-10.3)	.285
Basal LH, mean (SD), mIU/mL[b]	7.4 (1.9) (range, 3.6-11.2)	8.3 (1.6) (range, 5.1-11.5)	.129
Basal prolactin, mean (SD), ng/mL[c]	21.4 (1.5) (range, 19.9-25.9)	24.2 (2) (range, 22.2-30.2)	.546
Basal thyrotropin, mean (SD), mIU/L[d]	1.1 (0.3) (range, 0.8-2)	4.9 (0.7) (range, 4.2-7)	.026
Basal free T_3, mean (SD), pg/mL[e]	3.1 (0.7) (range, 1.68-4.5)	3 (0.7) (range, 1.49-4.6)	.109
Basal free T_4, mean (SD), ng/dL[f]	0.95 (0.4) (range, 0.5-2.2)	1.01 (0.5) (range, 0.5-2.5)	.140
No. stimulation days, mean (SD)	12.8 (2) (range, 10.8-18.8)	13.4 (1.7) (range, 11.7-18.5)	.51[illegible]
Estradiol at aspiration, mean (SD), pg/mL	1750 (155) (range, 1440-2060)	1069 (56) (range, 957-1181)	.0[illegible]2
No. follicles punctured, mean (SD)	15.6 (1.3) (range, 13-18.2)	10 (2.2) (range, 5.6-14.4)	[illegible]029
No. oocytes retrieved, mean (SD)	6.2 (0.7) (range, 5.5-8.3)	6.1 (0.9) (range, 5.2-8.8)	.450
No. of metaphase II at time of injection	29	14	.019
Miscarriage, %	9	13	.031
Fertilization, %	51.9	18.8	.015
Pregnancy, %	35	10	.021
Delivery, %	26	3	.017

Abbreviations: FSH, follicle-stimulating hormone; LH, luteinizing hormone; T_3, triiodothyronine; T_4,

Note the dramatic difference in pregnancy rates

CORRECTION

It was brought to our attention that the data in the manuscript, "Improved In Vitro Fertilization Outcomes After Treatment of Subclinical Hypothyroidism in Infertile Women," by Rahman et al (*Endocr Pract.* 2010;16:792-797) had statistical errors, as well as errors in Table 2. Rather than percentage rates for miscarriage and pregnancy, the percentages should have been absolute numbers. In recalculating the values, levothyroxine therapy of subclinical hypothyroidism did not affect the miscarriage rate ($P = .44$), but did affect the pregnancy rate with a more significant *P* value of .0001. Our calculations are shown below.

Lewis Braverman, MD, FACP, FACE
Editor-in-Chief
Endocrine Practice

For the outcome of pregnancy:
2 × 2 contingency table

	Outcome 1	Outcome 2	Total
Group 1	35	0	35
Group 2	10	25	35
Total	45	25	70

Fisher exact test
The 2-tailed *P* value <.0001

For the outcome of miscarriage:
2 × 2 contingency table

	Outcome 1	Outcome 2	Total
Group 1	9	26	35
Group 2	13	22	35
Total	22	48	70

Fisher exact test
The 2-tailed *P* value = .4403

But numbers still do not add up (loss vs pregnancy)

Rahman *et al*, Endocrine Practice 2010;16:792-7

TABLE 2

Comparison of COS results and IVF/ICSI outcome.

Factor	LT4 treatment	Control	*P* value
No. of cycles initiated	32	32	
No. of cycles retrieved	32	32	
No. of ET cycles	32	32	
Cycles with ICSI, n (%)	14 (43.8)	14 (43.8)	NS[a]
Days of rhFSH	9.1 ± 1.2	9.0 ± 1.1	NS[b]
Total dose of rhFSH	1,880.6 ± 425.5	1,919.9 ± 397.7	NS[b]
Days of GnRH antagonist	4.3 ± 1.0	4.3 ± 1.0	NS[b]
On the day of hCG injection			
TSH (mIU/L)	2.9 ± 1.0	6.8 ± 1.9	<.001[a]
FT4 (ng/dL)	1.3 ± 0.1	1.2 ± 0.2	.017[a]
PRL (ng/mL)	15.8 ± 3.3	16.3 ± 3.5	NS[b]
No. of follicles ≥14 mm	8.9 ± 3.4	9.1 ± 3.3	NS[b]
EMT (mm)	10.1 ± 1.1	9.8 ± 1.2	NS[b]
On the day of β-hCG measurement			
TSH (mIU/L)	2.3 ± 0.4	6.9 ± 2.0	<.001[a]
FT4 (ng/dL)	1.4 ± 0.3	1.0 ± 0.2	<.001[a]
No. of oocytes retrieved	9.3 ± 3.9	9.2 ± 3.2	NS[b]
No. of mature oocytes	8.2 ± 3.4	7.5 ± 2.6	NS[b]
No. of fertilized oocytes	8.1 ± 3.4	7.2 ± 2.3	NS[b]
No. of grade I, II embryos	3.3 ± 1.6	2.2 ± 1.3	.007[b]
No. of embryos transferred	2.9 ± 0.5	2.9 ± 0.4	NS[b]
No. of embryos cryopreserved	2.5 ± 2.7	1.8 ± 2.3	NS[b]
Embryo implantation rate, % (n)	26.9 (25/93)	14.9 (14/94)	.044[a]
Clinical PR per cycle initiated, % (n)	53.1 (17/32)	37.5 (12/32)	NS[a]
Miscarriage rate, % (n)	0 (0/17)	33.3 (4/12)	.021[a]
Live birth rate per cycle initiated, % (n)	53.1 (17/32)	25.0 (8/32)	.039[a]

Note: Values are mean ± SD unless otherwise noted. PR = pregnancy rate; NS = not significant.

[a] Fisher's exact test or χ^2 test.

[b] Student's *t* test.

Kim. LT4 for subclinical hypothyroidism. Fertil Steril 2011.

This results of this paper were Driven by the 4 women with Miscarriage in the control group

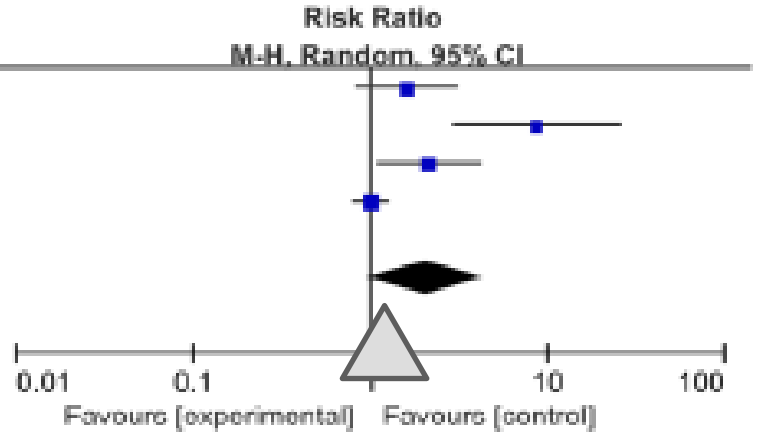
One different outcome and all statistical significance is lost

Treatment of women with subclinical hypothyroidism (Live Birth)

B

Study or Subgroup	LT4 supplementation Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Year
Negro 2005	16	43	10	43	25.3%	1.60 [0.82, 3.12]	2005
Abdel Rahman 2010	26	35	3	35	18.9%	8.67 [2.89, 26.02]	2010
Kim 2011	17	32	8	32	25.1%	2.13 [1.07, 4.21]	2011
Wang 2017	90	282	91	285	30.6%	1.00 [0.79, 1.27]	2017
Total (95% CI)		392		395	100.0%	2.05 [0.96, 4.36]	
Total events	149		112				

Heterogeneity: Tau² = 0.47; Chi² = 18.82, df = 3 (P = 0.0003); I² = 84%
Test for overall effect: Z = 1.86 (P = 0.06)



Rao et al. Reproductive Biology and Endocrinology (2018) 16:92

ASRM PAGES

Subclinical hypothyroidism in the infertile female population: a guideline

Practice Committee of the American Society for Reproductive Medicine

The American Society for Reproductive Medicine, Birmingham, Alabama

- It is recommended to counsel women that SCH is not associated with an increased risk of miscarriage (strength of evidence: B; strength of recommendation: moderate).
- It is recommended to counsel women that TSH levels between 2.5 and 4 mIU/L are not associated with an increased risk of miscarriage (strength of evidence: B; strength of recommendation: moderate).
- There is insufficient evidence to counsel women that SCH is associated with infertility (strength of evidence: C; strength of recommendation: weak).

Table 1

A taxonomy of research misconduct*

Research misconduct (in descending order of seriousness)
• Fabrication: Invention of data or cases
• Falsification: Wilful distortion of data
• Plagiarism: Copying of ideas, data, or words without attribution
• Failing to get consent from an ethics committee for research
• Not admitting that some data are missing
• Ignoring outliers without declaring it
• Not including data on side effects in a clinical trial
• Conducting research in humans without informed consent or without justifying why consent was not obtained to an ethics committee
• Publication of *post-hoc* analyses without declaration that they were *post hoc*
• Gift authorship
• Not attributing other authors
• Redundant publication
• Not disclosing a conflict of interest
• Not attempting to publish completed research
• Failure to do an adequate search of existing research before beginning new research

*Taken from Ref. 16: Evans S. How Common is it? Royal College of Physicians of Edinburgh. Joint Consensus Conference on Misconduct in Biomedical Research. Suppl. 7 2000;(30)1

asrm American Society for Reproductive Medicine

Google translate used to avoid plagiarism

Software to detect fraudulent figures

ASRM Research Integrity Committee

- Formed January 2023
- Received more than 57 allegations (17 2021, 14 2022, 16 2023, 10 2024)
  - Implausible enrollment/dropout/retention/etc. or discrepancy between abstract or trial registration and published paper
  - Methodological or data reporting error
  - Repeating numbers within or across manuscripts
  - Misrepresentation of data in conclusions
  - Duplicate publication or plagiarism
- 12 retractions
- 6 statement of concern
- 3 errata
- 3 cases declined to investigate
- Many still pending (only 2 before formation of the committee in 2023)

What Drives lack of scientific integrity?

- **Misunderstanding of science?**
  - **Randomization: "Starting whatever medicine my PI says to start"**
  - **"Missing data is a sign of a bad study"**
- **Financial gain?**
  - **Financial bonus for publication "two years salary", "If my study works, I get a new hospital system"**
  - **The prestige of Professorship**
  - **Tenure**
  - **Support of a product (marketing) "Boasts the immune system", "supports digestive health"**

- **Why do people commit fraud?**

  Knowingly!!!

Rise and Fall of Elizabeth Holmes (Theronos)

- Dropped out of Stanford to start a company: Theranos
- Recruited Henry Kissinger, James Mattis, George Schultz on Board
- Sold an idea (full set of labs on a finger prick of blood)
- Youngest self-made female billionaire
- Raised 700 million from private investors
- Convicted of criminal fraud
  - To investors, not patients

Fraud or Extreme Belief? (Fake It Until You Make It)

- 70+ tests from a finger stick (Edison)
- "Validated" by commercial companies (Pfizer)
- Claimed used on medivac helicopters in the field of battle
- Defense: "Believed it could be done and never took the money and walked away"
- Used modified commercial analyzers
  - "All tests performed with proprietary Theronos technology"
- Put the logo of Pfizer on Theranos internal documents and sent to investors. "Was that wrong?"
- "Military contacts are very important to us"
- Is "a true believer" a defense for fraud?

What Can we Learn from Elizabeth Holmes?

- Beware of lack of transparency
  - Look at the raw data (yourself)
  - Do not be bullied by your superior
- Need to be objective and realize when you (your hypothesis) may be wrong
  - Negative science has value
- Do not exaggerate the positive and discount the negative
  - Be carful when you throw out data (“outlier”), pick a secondary outcome post hoc (thousands of test results were “invalidated”)
- Conduct science for the right reasons

The Burden of Proof

Does a paper suggest a significant finding

- Are the methods convincing (and appropriate)
- Are there alternative explanations
- Does it conform to what is know

Is the paper trying to convince of effect, or no effect

- Absence of evidence is not evidence of absence

It is easy to lie with statistics
But it is far easier to lie without them

Extraordinary claims require extraordinary evidence

It is easier to argue against something than for something
(tobacco, global warming)

What am I looking for in Fertility and Sterility

- Quality science
  - Hypothesis driven, attention to study design, well conducted
- New, better, applicable
- Quality presentation
  - Clear, concise, attention to detail

- Barnhart Pet Peeves
  - Spontaneous conception, ICSI men
  - Acronyms (specifically when made up) and in tables, headings and titles
  - Using 20 words when 10 will do
  - You do not know your study design (or the editor of the Journal)

New Instructions to Authors

Each manuscript should be designated by study type:

- Clinical trial
- Cohort study or Case-control study
- Other observational study (specify)
    - Cost-effectiveness analysis
    - Decision analysis
    - Study of screening and diagnostic tests
    - Comparative effectiveness research
    - Genetic association study
    - Quality improvement study
    - Survey study
    - Cross sectional study
    - Mixed methods
    - Qualitative study
- Systematic review (with meta-analysis)
- Systematic review (without meta-analysis)
- Laboratory based study
- Research letter
- Letter to the editor

Video article

Instructions provide guide to reporting and review

Cohort and case control studies are observational research when subjects are characterized by a specific exposure and outcome without randomization. The study design, if prospective or retrospective, should be clearly stated along with the a priori hypothesis. A cohort study compares exposed to non-exposed participants to evaluate a specific outcome. A case-control study selects participants based on the outcome (for example pregnant versus non-pregnant participants) and then looks back at exposure, such as a treatment or a risk factor. Each study should include specified detail about the population studied including clear inclusion and exclusion criteria (including the study setting and dates). The exposure and outcome must be clearly defined. If there are multiple exposures or outcomes each must be clearly identified and primary, secondary or exploratory. The data source should be clearly described and data validation methods explicitly stated. Confounding variables available (and pertinent variables not available), should be described. The rationale for the sample size should be stated. The amount of missing data and how it was managed in the analysis should be stated. Results should be presented clearly and include both crude and adjusted associations with 95% confidence intervals. Statistical methods should be concisely and clearly reported. Causal language should not be used in describing results. The discussion section should place the results in context with the published literature and address study limitations.

- The EQUATOR Reporting Guidelines should be followed including the STROBE for an observational study

How to write a paper

- Elevator pitch clear, accurate, not overstated
- Introduce the problem, tell me why I should be interested and what is the hypothesis you will be testing.
- Tell what how you studied the issue, what was the population, why you selected them and who was excluded. Why you studied that number of subjects. What data did you collect, why, and how was it defined into exposure and outcome? How was it compared.
- Tell me what you found clearly and persuasively. A picture is worth a thousand words
- Discuss how the finding are novel, different, or the same as others. Consider alternative explanations, could they be due to chance bias, confounding. Why are the findings convincing why should we have caution.

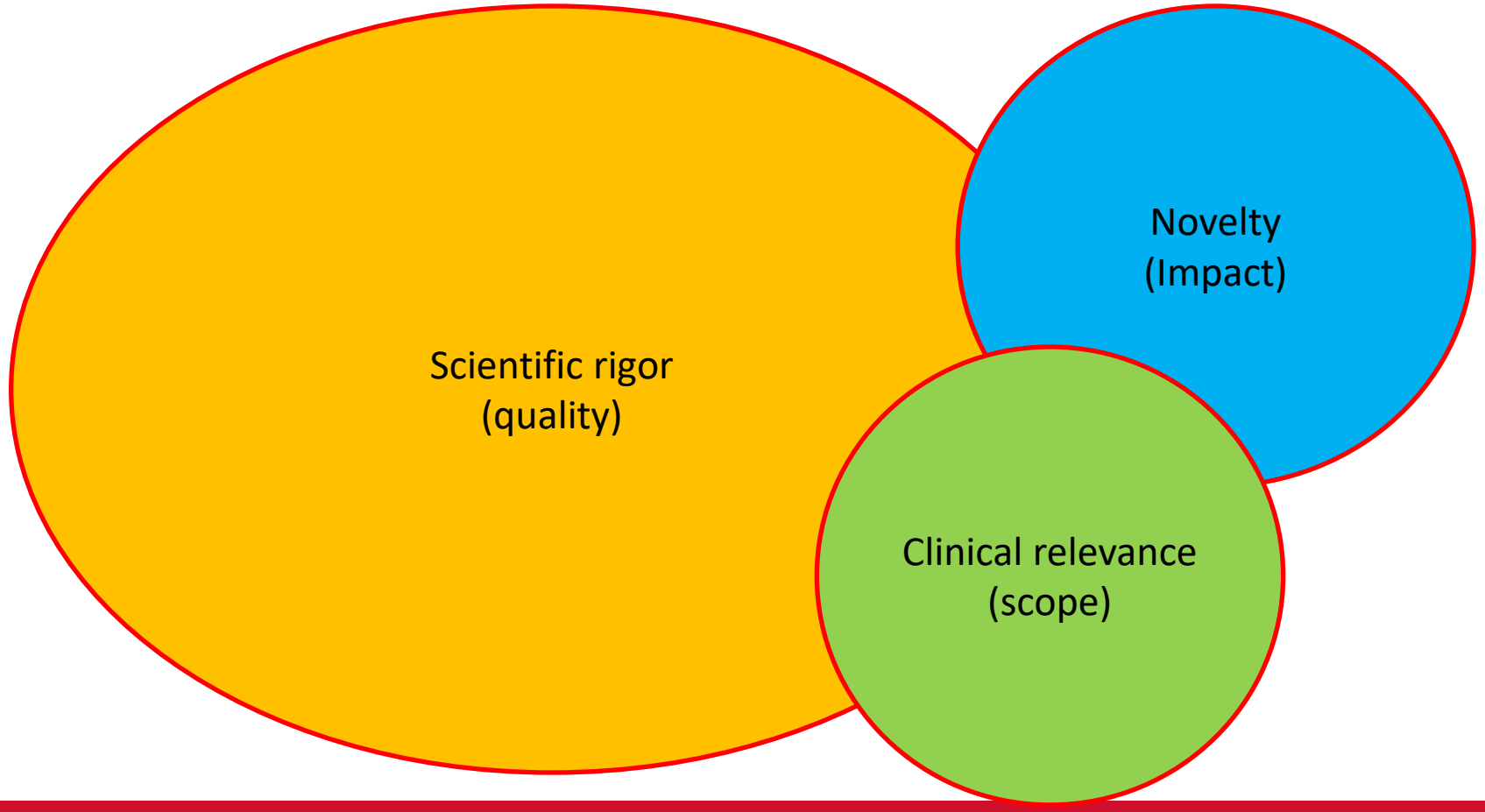
How to write a paper

- Elevator pitch (abstract, conclusion) clear, accurate, not overstated
- Introduce the problem, tell me why I should be interested and what is the hypothesis you will be testing.
- Tell what how you studies the issue (study design), what was the population, why you selected them and who was excluded. Why you studied that number of subjects (sample size). What data did you collected, why, and how was it defined into exposure and outcome? How was it compared (stats)
- Tell me what you found (results) clearly and persuasively. A picture (tables and figures) is worth a thousand words
- Discuss how the finding are novel, different, or the same as others. Consider alternative explanations, could they be due to chance bias, confounding? Why are the findings convincing (strengths) why should we have caution (limitations)

Overview of how to review methods

- Read methods (do you understand their study design)
  - RCT: randomization, flow chart, ITT, RR
  - Cohort: population, exposure, outcome RR, aRR (confounding)
  - Case control: outcome, exposure OR, aOR, (bias and confounding)
  - Diagnostic test (sensitivity, specificity, predictive value, assessment of precision (ie 95% CI)
- Tables
  - Do they stand alone and fit the study design, do number make sense (acronyms)
- Do not understand stats? (that is ok).
  - Crude results, adjustment, validation, precision, can you follow the “story”

Desired Characteristics of a Manuscript



There is hope

- **Medical literature is still the most objective presentation of science in the world**
  - Do not throw out the best source of truth because of bad actors, or distrust all literature (need to be more discerning, pick your trusted sources)
  - Be wary of “top line” summaries and social media
- **Journal are making tremendous strides in review of new science AND are correcting the literature with retractions and letters of concern**
  - Checklist to assess integrity of primary papers
  - Newer methods of meta-analysis to assess underlying studies
  - Be part of the peer review process!

The Practice of Medicine Should be Evidence Based!

You can and should trust the medical literature

(but not blindly)

You should applaud the work of our medical societies (like ASRM)

and consider joining and helping

When possible, you can (and should) debunk the latest internet fad and remind patients of the bias of social media

You need to establish trustworthiness as a medical care provider!!

Thank you very much for your submission. But there are already a large number of publications on this exact topic with variable outcomes already reported in the medical literature (see citations below) so adding another non-definintive retrospective study to this body of literature is not adding an novel finding. If you are able to perform a prospective RCT, that may warrant merit for publication in F&S. Thank you.

https://pubmed.ncbi.nlm.nih.gov/40398679/

https://pubmed.ncbi.nlm.nih.gov/39723883/

https://pubmed.ncbi.nlm.nih.gov/39614958/

https://pubmed.ncbi.nlm.nih.gov/39564835/

https://pubmed.ncbi.nlm.nih.gov/38367686/

https://pubmed.ncbi.nlm.nih.gov/37300647/

https://pubmed.ncbi.nlm.nih.gov/37109005/

https://pubmed.ncbi.nlm.nih.gov/35522187/

Table 1 Baseline characteristics of advanced-age patients in two groups

	GH group, N = 158	Control group , N = 158	p-value
Age(Mean ± SD)	39.94 ± 2.91	39.68 ± 2.96	0.356
BMI(kg/m2)	22.34 ± 2.12	22.84 ± 2.27	0.053
Mestruation duaration	27.50 ± 3.82	27.73 ± 2.51	0.409
AMH(ng/ml)	1.96 ± 1.70	1.94 ± 1.58	0.875
PRL（Prolactin）(ng/ml)	15.68 ± 7.97	15.64 ± 8.33	0.771
TSH（Thyroid stimulating hormone）(uIU/ml)	1.88 ± 0.93	1.89 ± 1.06	0.771
AFC（antral follicles count）	7.72 ± 5.24	7.86 ± 5.57	0.926
Fasting Blood glucose(mmol/ml)	5.35 ± 0.40	5.34 ± 0.46	0.374
Glycosylated hemoglobin（%）	5.01 ± 1.34	5.30 ± 0.56	0.738
Infertility Cause(%)			0.089
Primary infertility	7.0% (11 / 158)	13% (20 / 158)	
Secondary infertility	93% (147 / 158)	87% (138 / 158)	

Randomization

A randomization sequence was generated by an independent statistician for group allocation at a ratio of 1:1. This sequence contained the enrollment sequence number, randomization number, and allocation group. The sequence was uploaded into the Interactive Web Response System

Statistical Analysis

Continuous variables are expressed as mean ± standard deviation (× ± s), with between-group comparisons using independent-sample t tests. Categorical variables are expressed as percentages (%), with between-group comparisons using the χ^2 test, continuity-corrected χ^2 test, or Fisher's exact test. Statistical significance was set at $P < 0.05$.

Table 3 Laboratory results of patients in two groups

	GH Group, N = 158	Control Group, N = 158	p-value	Absolute difference (95% CI)
Number of oocytes retrieved	7.77 ± 5.17	7.85 ± 5.53	0.995	0.00 (-1.00, 1.00)
Number of MII oocytes	6.23 ± 4.14	6.55 ± 4.71	0.795	'-0.00 (-1.00, 1.00)
Number of 2PN fertilized oocytes	5.08 ± 3.71	4.99 ± 3.92	0.692	0.00 (-1.00, 1.00)
Number of high-quality cleavage embryos	3.48 ± 3.06	3.48 ± 3.03	0.902	'-0.00 (-1.00, 1.00)
Number of blastcysts	2.09 ± 2.11	1.96 ± 2.25	0.342	0.00 (-0.00, 1.00)
Rate of blastcysts	0.40 ± 0.30	0.36 ± 0.31	0.226	0.01 (-0.00, 0.12)
Number of high-quality blastcysts	1.09 ± 1.57	1.07 ± 1.55	0.653	0.00 (-0.00, 0.00)
Number of euploid blastcysts	0.98 ± 0.99	0.89 ± 0.99	0.534	0.00 (-0.00, 0.00)
Rate of euploid blastcysts	0.38 ± 0.37	0.32 ± 0.34	0.325	0.00 (-0.00, 0.17)

Major anomalies were then further classified by organ system: cardiac, renal/genitourinary, gastrointestinal, musculoskeletal, and other (Figure 1). Pregnancies and neonates with monochorionic twins, vanishing twins, triplets, fetal demise, conceived with other ART methods (ovulatory induction agents alone, intrauterine insemination, etc.), cases diagnosed with chromosomal abnormalities, or incomplete follow-up were excluded. Chi-squared and Fisher's exact tests were used to analyze the data. Multivariable logistic regression was then performed adjusting for age, race, BMI, and parity. The primary outcome was presence of a congenital anomaly in pregnancies conceived via IVF vs spontaneously.

Table 2: Congenital Anomalies

	Total (n=968) n=70 (%)	IVF (n=521) n=37 (%)	Spontaneous (n=447) n=33 (%)	Adjusted Odds Ratio (95% CI)[ǂ]
Any Anomaly	70[¥] (7.23)	37 (7.10)	33 (7.38)	0.97 (0.54-1.71)
Major Anomaly	55 (5.68)	28 (5.37)	27 (6.04)	0.93 (0.49-1.77)
Minor Anomaly	32 (3.31)	17 (3.26)	15 (3.36)	0.98 (0.43-2.2)

[¥]5 IVF and 3 spontaneous pregnancies had both a major and minor congenital anomaly

[ǂ]Adjusted for maternal age, race, BMI, and parity

Q&A