



Published in final edited form as:

Obstet Gynecol. 2009 October ; 114(4): 877–884. doi:10.1097/AOG.0b013e3181b5c9e8.

Methodology and Analytic Techniques Used in Clinical Research: Associations With Journal Impact Factor

Lindsay M. Kuroki, BA¹, Jenifer E. Allsworth, PhD², and Jeffrey F. Peipert, MD, Ph.D.²

¹Warren Alpert Medical School of Brown University, Providence, RI

² Division of Clinical Research, Department of Obstetrics and Gynecology, Washington University in St. Louis, School of Medicine, St. Louis, MO.

Abstract

Objective—To describe research methodology and statistical reporting of published articles in high impact factor, general medical journals compared to moderate impact factor obstetric and gynecology journals.

Methods—A cross-sectional analysis was performed on 371 articles published from January to June 2006, in 6 journals (high impact factor group—*Journal of the American Medical Association*, *Lancet*, the *New England Journal of Medicine*; moderate impact factor group—*American Journal of Obstetrics and Gynecology*, *British Journal of Obstetrics and Gynecology*, and *Obstetrics & Gynecology*). Articles were classified by level of evidence. Data abstracted from each article included: number of authors, clearly stated hypothesis, sample size/power calculations, statistical measures, and use of regression analysis. Univariable analyses were performed to evaluate differences between the high and moderate impact factor groups.

Results—The majority of published reports were observational studies (50%), followed by randomized controlled trials ([RCTs] 24%), case reports (14%), systematic reviews (6%), case series (1%), and other study types (4%). Within the high factor impact group, 35% were RCTs compared to 12% in the moderate impact group (RR 2.9, 95% CI 1.9-4.4). Recommended statistical reporting (eg point estimates with measures of precision) was more common in the high impact group ($P < 0.005$).

Conclusions—The proportion of RCTs published among the high impact factor group was 3 times that of the moderate group. Efforts to provide the highest level of evidence and statistical reporting have potential to improve the quality of reports in the medical literature available for clinical decision making.

INTRODUCTION

Evidence-based medicine continues to guide clinical decision making based on the best available evidence in the literature. The U.S. Preventive Services Task Force(1) and the American College of Obstetricians and Gynecologists(2) have adopted similar rating

Address for Correspondence: Jeffrey F. Peipert, M.D., Ph.D. Washington University in St. Louis School of Medicine Division of Clinical Research Department of Obstetrics and Gynecology Campus Box 8219 4533 Clayton Avenue St. Louis, MO 63110-1094 Phone: 314-747-4016 Fax: 314-747-4019 peipertj@wudosis.wustl.edu.

Financial Disclosure: The authors did not report any potential conflicts of interest.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

systems to characterize the quality of available evidence (Box 1). Under this hierarchical approach, randomized controlled trials (RCTs) remain the gold standard of research, as they provide the strongest evidence of causality and the best methodology to test the effectiveness of an intervention.(3, 4) Recently there has been a progressive effort by researchers and journal editors to assess and improve the quality of published studies.

Examples of such initiatives have been pursued within the field of Obstetrics and Gynecology.(5-9) A 2002 study by Welch et al.(6) revealed a significant improvement in the quality of statistical analysis in articles published over a 5-year interval by the *American Journal of Obstetrics and Gynecology*. Similar findings were also reported in *Obstetrics & Gynecology* by Dauphinee et al.(7) who also noted a trend towards increased observational studies and decreased anecdotal reports. The number of RCTs, however, did not change over their 10-year study period and has consistently remained behind the more abundantly published observational studies.(5, 7, 10)

In response to the evidence-based paradigm, quality assessment reviews of published articles have included evaluations of journal quality factors,(11) methodology(12) and bibliographic citations,(9) logistic regression reporting,(13) and the quality of discussion sections.(14) Findings from such reports have the potential to improve the quality of published literature. Not only are “periodic report cards”(10) cited as future goals, but there is agreement that clinicians, researchers, policy makers, and journal editors all have a fundamental responsibility to meet and contribute toward higher standards of research evidence.

Specifically focusing on quality assessment studies within obstetrics and gynecology (OBGYN) journals, there still remain gaps within the current literature that warrant addressing. Few study objectives compare OBGYN research methodology to those reported in other medical journals. Our study objective was to classify and compare research methodology and statistical reporting of published articles in 6 major medical journals, selected based on journal impact factor and medical specialty (3 from moderate impact factor OBGYN journals and 3 from high impact factor general medical journals). More specifically, we hypothesized that journals with a higher impact factor publish more RCTs and have higher quality of statistical reporting compared to moderate impact factor OBGYN journals.

MATERIALS AND METHODS

We performed a cross-sectional analysis of 371 published articles, indexed in the MEDLINE database, to evaluate the research methodology and statistical methods used for clinical research. All articles were published in 6 medical journals from January-June 2006. Selection of journals was based on their medical specialty and 2005 journal impact factor. The high impact group included *New England Journal of Medicine* (impact factor = 44.0), *Lancet* (IF = 23.4), and *Journal of the American Medical Association* (impact factor = 23.3). The moderate impact group consisted of *Obstetrics & Gynecology* (impact factor = 4.2), *American Journal of Obstetrics and Gynecology* (impact factor = 3.1), and the *British Journal of Obstetrics and Gynecology* (impact factor = 2.1).

All articles selected contained clinically-oriented research based on human subjects only. We began by generating a library of all 2006 publications for each journal using PubMed and a reference manager, EndNote 9.0. After a list for all 6 journals was generated, we performed an initial screening of articles to narrow our sample population to only include publications that involved original research, systematic reviews, case reports, and case series. We excluded commentaries, correspondences, clinical expert series, discussions,

editorials, and letters to the editor. However, to achieve this goal in each of the selected journals, we had to apply different “keyword” exclusion criteria according to the individual journal style (Figure 1). Traditional reviews that failed to provide original data or perform a systematic review of the literature were excluded as well. The number of articles which met eligibility criteria varied by journal. From this set, we randomly selected articles using a computer-generated random number sequence.

Articles were assigned a random number, ranked, and abstractions were carried out until we collected data from at least 50 articles per journal, all which met inclusion criteria. A standardized form was used to record the abstracted data. The methodology of each study was categorized according to its level of evidence based on guidelines (Box 1) set forth by *Obstetrics & Gynecology*.

We were also interested in gathering data regarding number of authors, whether the authors provided a clearly stated research hypothesis, whether attributable risk and/or number needed to treat were reported, and whether or not statistical tests/analyses were performed. Statistical measures from both the abstract and results sections of each article were also captured. Specifically, we noted: 1) use of effect measures (e.g. mean differences, odds ratio, relative risk, or other effect measures), confidence intervals (CI), and/or P-values; and 2) whether non-significant differences were reported using actual P-values versus stating “NS” (non-significant). Higher quality (or “recommended”) statistical reporting were reports that included point estimates with measures of precision, as defined by current guidelines. (15, 16) We noted whether study findings were positive or negative. A negative study was defined as any research whose primary analysis, based on the stated hypothesis, was not statistically significant. In those manuscripts where the hypothesis was missing or not clearly stated, the primary endpoint was derived from the study aim/objective or article title. Due to their anecdotal nature and lack of statistical reporting, case reports and case series were omitted from our analysis of quality statistical measures and other study characteristics listed in Table 2 (i.e. number of authors, study hypothesis, positive/negative study, sample size calculations, and regression analysis).

Institution of several steps was made to ensure the quality and accuracy of the data collected from each journal. All initial abstractions were completed by one investigator (LK) and a random sample of 10% of papers was re-abstracted by a second investigator (JA). No systematic errors were identified during this review.

Sample size calculations were predetermined to detect a 3-fold difference in the percentage of RCTs published in journals within the high impact factor group versus the moderate group (moderate group rate = 13%). We determined that a minimum of 50 articles per journal would be necessary to achieve a power of 80% ($\beta = 0.20$, $\alpha = 0.05$).

Univariable analyses evaluated the association between impact factor status and methodological study design, quality statistical measures, and other study journal factors (number of authors, presence of a clearly stated hypothesis, negative study, whether sample size/power calculations were provided, and presence of any regression analysis) using chi-square tests. All statistical analyses were completed using SAS version 9.1 (SAS Institute, Cary, NC). This study was approved by the Washington University in St. Louis Human Research Protection Office and was classified as exempt.

RESULTS

We reviewed a total of 371 articles published from 6 major medical journals, including 56 (15%) from *American Journal of Obstetrics & Gynecology* (AJOG); 59 (16%) *British Journal of Obstetrics & Gynecology* (BJOG); 66 (18%) *Obstetrics & Gynecology*; 58 (16%)

Journal of the American Medical Association (JAMA); 69 (19%) *Lancet*, and 63 (17%) *New England Journal of Medicine* (NEJM). Table 1 summarizes the methodological quality of the studies. The overall majority of published reports consisted of analytic observational studies (50%) followed by RCTs (24%), case reports (14%), systematic reviews (6%), case series (1%), and other (4%). Further breakdown of observational studies revealed a higher proportion of cohort (25%) and cross sectional studies (20%) than case-control studies (5%). The proportion of RCTs published among the high impact factor group (JAMA, *Lancet*, NEJM) was almost 3 times that of the moderate group (AJOG, BJOG, *Obstetrics & Gynecology*) (RR 2.9, 95% CI 1.9-4.4). Specifically within the high-impact group, 35% were RCTs and 35% observational studies vs. 12% and 67%, respectively, in the moderate impact group.

The following results are from analyses that excluded case reports and case series (See Table 2). Number of authors per article ranged from 1 to 47 with the majority of published articles containing between 5-9 authors. Eighty-five percent of published articles in the high impact group listed ≥ 5 authors versus 64% in the moderate impact group ($p < 0.0001$). A clearly stated hypothesis was reported in 21% of all studies reviewed, 21% of articles in the moderate impact factor group and 22% in the high impact group ($P = 0.6$). NEJM had the highest reporting rate at 33%, while all other journals ranged between 8-29%. Overall, negative studies were not common; *Obstetrics & Gynecology* had the lowest percentage of negative studies (8%). Specifically, 18% and 14% of published articles in the high and moderate impact factor groups, respectively, produced primary analyses that were not statistically significant.

In terms of sample size calculations, 47% of published articles in the high impact group provided power and/or sample size calculations versus 23% in the moderate group ($P < 0.0001$). Further breakdown of the latter, revealed a higher frequency of calculations in *Obstetrics & Gynecology* than other OBGYN journals, although not statistically significant (32% vs. 19%).

We were also interested in making other comparisons between journal impact factor and statistical measures. Overall, recommended statistical reporting (e.g. effect size/P-value, effect size/CI, or effect size/P-value/CI) in the abstract was more common in the high impact factor group (61% vs. 36%, $P < 0.0005$, Table 2, Figure 2A). Sixty-three percent of published articles in NEJM contained recommended statistical measures, followed by *Lancet* (62%), JAMA (59%), *Obstetrics & Gynecology* (55%), BJOG (31%), and AJOG (25%). In contrast, recommended statistical measures within the results section of the manuscript was present in 74% of all published articles in our review. Within the high impact group, 84% had recommended statistical reporting versus 65% of articles within the moderate impact group ($P = 0.002$, Table 2, Figure 2B). NEJM had the highest reporting rate at 90%, *Lancet* (81%), JAMA and *Obstetrics & Gynecology* (80%), AJOG (60%), and BJOG (58%). Interestingly, *Obstetrics & Gynecology* was more comparable to the higher impact group (80% vs. 84%) than the other OBGYN journals (80% vs. 58%).

The high impact factor group reported specific P-values for non-significant values more often than the moderate group (97% vs. 88%, $P = 0.04$). *Lancet* had the highest reporting rate at 100%, followed by NEJM, JAMA and *Obstetrics & Gynecology* all at 95%, AJOG 92% and BJOG 73%. Again, the reporting rate in *Obstetrics & Gynecology* was more reflective of the high impact group (95% vs. 97%) than its OBGYN counterparts (95% vs. 84%)

Lastly, we also captured the proportion of studies which performed regression analysis. Overall, articles published in the high impact factor group conducted more regression analyses than those published in the moderate impact group (60% vs. 48%, $P = 0.03$).

DISCUSSION

In our analysis of published articles from January to June 2006, RCTs represented nearly one fourth of all studies. The proportion of RCTs published among the high impact factor journal group was nearly 3 times that of the moderate group. Other studies evaluating research methodology have shown that observational studies are the most commonly published in the literature, (5, 7) and our results support this finding. We also noted that the rate of recommended statistical reporting (e.g. usage of effect size and its precision as well as specific P-values for NS values)(17) within the abstract and manuscript were highest in the high impact factor group. Other factors that were positively associated with high journal impact factor include number of authors (>5), sample size/power calculations, actual P-values given for NS differences, and presence of regression analysis

There are several potential explanations why the moderate impact factor group (e.g. OBGYN journals) contained fewer RCT publications. First, many issues in obstetrics and gynecology do not lend themselves easily to this study design (e.g. surgical interventions, contraceptive studies, etc.). Ethical concerns of RCTs also play an important role. Another final possible explanation may be that the highest quality RCTs in OBGYN are submitted to the general medical journals.

The quality of statistical reporting in *Obstetrics & Gynecology* was more reflective of the high impact factor group than the other moderate impact factor OBGYN journals with regards to two factors: 1) reporting specific P-values for NS values; and 2) reporting statistical measures with estimated effect sizes and estimates of precision within the results section of the manuscript. Reporting of statistical measures within the abstract was also more comparable to the high impact group than its OBGYN counterparts. We concur with previous findings(6, 7) that improved statistical reporting may be attributed to better listing and describing of statistical procedures in the instructions for authors, more strict editing criteria by journal editors, reviewers and statistical consultants, as well as increasing collaborations between researchers and epidemiologists/statisticians to design, analyze and present data. Recommended statistical reporting may be more reflective of journal editing of original manuscripts than authors discriminately providing higher quality statistical reports to higher impact factor journals.

Furthermore, among the moderate impact factor group, *Obstetrics & Gynecology* also had the greatest proportion of studies that reported sample size/power calculations. We also noted that *Obstetrics & Gynecology* had the lowest number of negative studies, regardless of impact factor status. A recent study in neonatology by Littner et al.,(18) showed that articles with negative results are more likely than those with positive findings to be published in lower impact factor journals. Publication bias may explain the low prevalence of negative studies in *Obstetrics & Gynecology*, but our research was not designed to assess either hypothesis. Interestingly, a recent study on predictors of publication showed that after adjusting for other study characteristics, having statistically significant results did not improve the chance of a study being published (odds ratio 0.83; 95% CI, 0.34-1.96).(19)

Our study recognizes the need to continue improving the quality of methodology and statistical reporting in the medical literature. The Consolidated Standards of Reporting Trials (CONSORT) statement(17) was developed as an evidence-based approach to improve the quality of reporting of RCTs and has been adopted by many journals including the six in our study. Since its debut in 1996, and subsequent revised version, the CONSORT checklist appears to have already led to some improvements.(20, 21) A similar set of guidelines for observational studies (STROBE statement) was created more recently in 2007, and also has potential to improve overall quality of observational research reports, especially given the

high prevalence of observational studies.(16, 21, 22) With regards to our study, 4 of the 6 quality variables we analyzed were listed in both CONSORT and STROBE guidelines. They included a clearly stated hypothesis, sample size calculations, a summary of results which included estimated effect size and its precision (e.g. 95% confidence interval), and reporting of regression analyses. We believe a clearly stated research hypothesis is an important quality criteria. It is quite surprising that a clearly stated hypothesis was not commonly cited in any of the journals included in our study, regardless of impact factor status.

Our study has several limitations. We attempted to quantify reporting of attributable risk and number needed to treat in our article abstractions, but due to the low number of studies which reported these measures, we were unable to assess their relationship with journal impact factor. We also recognize that the outcome of a study (positive or negative) and number of authors are not reliable markers of quality of research. We agree with previously reported claims that the high author count may instead reflect other factors such as increased collaborations, reports from networks of investigators, funding for research, complexity of research, and pressure on academic faculty to publish. (23, 24) Furthermore, our research has limited power to detect more modest differences between groups of journals and individual journals, given the relatively small number of articles abstracted per journal. Other important issues that remain unanswered include: (1) are the discrepancies in the quality of reporting between individual journals functions of the quality of the initial manuscript submitted or the strict editing criteria of individual journals; (2) do author's have a tendency/preference to submit higher-level studies (e.g. RCTs) to high-impact journals; and (3) does publication of RCTs lead to a high impact factor status, or does a high impact factor lead to submission of RCTs?

Critical evaluation of published medical literature is an important skill that all healthcare providers and researchers should acquire. Published articles that clearly describe the research hypothesis, study methodology, power/sample size calculations, and statistical analyses allow readers to better assess whether study results should influence clinical practice guidelines and policies. Therefore, journals should continue to strive for the highest level of methodologic design and highest quality of statistical reporting.

Box 1

Levels of Evidence for Original Research Articles

Rating	Description
United States Preventive Services Task Force ¹	
I	Evidence obtained from at least one properly randomized, controlled trial.
II-1	Evidence obtained from well-designed controlled trials without randomization.
II-2	Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.
II-3	Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) also could be regarded as this type of evidence.
III	Opinions of respected authorities based on clinical experience, descriptive studies and case reports, or reports of expert committees.
American College of Obstetricians and Gynecologists ²	
I	A randomized controlled trial.

Rating	Description
II	A cohort or case-control study that includes a comparison group.
III	An uncontrolled descriptive study including case series.

Acknowledgments

Supported in part by a Midcareer Investigator Award in Women's Health Research (K24 HD01298), by a Clinical and Translational Science Award (UL1RR024992), and by Grants K12RR023249 and KL2RR024994 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research.

Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

REFERENCES

- Preventive Services Task Force. Guide to clinical preventive services: report of the U.S. Preventive Services Task Force. 2nd ed.. Williams & Wilkins; Baltimore: 1996.
- Guide to Writing for Obstetrics & Gynecology. 4th ed. Available at: <http://edmgr.ovid.com/ong/accounts/guidetowriting.pdf>
- Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med.* Jun 22; 2000 342(25):1887–92. [PubMed: 10861325]
- Peipert JF, Gifford DS, Boardman LA. Research design and methods of quantitative synthesis of medical evidence. *Obstet Gynecol. Sep;* 1997 90(3):473–8. [PubMed: 9277665]
- Funai EF, Rosenbush EJ, Lee MJ, Del Priore G. Distribution of study designs in four major US journals of obstetrics and gynecology. *Gynecol Obstet Invest.* 2001; 51(1):8–11. [PubMed: 11150867]
- Welch GE, Gabbe SG. Statistics usage in the American Journal of Obstetrics and Gynecology: has anything changed? *Am J Obstet Gynecol.* Mar; 2002 186(3):584–6. discussion 339. [PubMed: 11904628]
- Dauphinee L, Peipert JF, Phipps M, Weitzen S. Research methodology and analytic techniques used in the Journal Obstetrics & Gynecology. *Obstet Gynecol.* Oct; 2005 106(4):808–12. [PubMed: 16203723]
- Grant A. Reporting controlled trials. *Br J Obstet Gynaecol.* Apr; 1989 96(4):397–400. [PubMed: 2751952]
- Roach VJ, Lau TK, Ngan Kee WD. The quality of citations in major international obstetrics and gynecology journals. *Am J Obstet Gynecol.* Oct; 1997 177(4):973–5. [PubMed: 9369854]
- Funai EF. Obstetrics & Gynecology in 1996: marking the progress toward evidence-based medicine by classifying studies based on methodology. *Obstet Gynecol.* Dec; 1997 90(6):1020–2. [PubMed: 9397124]
- Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA.* Jun 5; 2002 287(21):2805–8. [PubMed: 12038918]
- Grimes DA, Schulz KF. Methodology citations and the quality of randomized controlled trials in obstetrics and gynecology. *Am J Obstet Gynecol.* Apr; 1996 174(4):1312–5. [PubMed: 8623862]
- Mikolajczyk RT, DiSilvestro A, Zhang J. Evaluation of logistic regression reporting in current obstetrics and gynecology literature. *Obstet Gynecol.* Feb; 2008 111(2 Pt 1):413–9. [PubMed: 18238980]
- Clarke M, Alderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals. *JAMA.* Jun 5; 2002 287(21):2799–801. [PubMed: 12038916]

15. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. Apr 18; 2001 285(15):1987–91. [PubMed: 11308435]
16. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. Apr; 2008 61(4):344–9. [PubMed: 18313558]
17. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. Aug 28; 1996 276(8):637–9.
18. Littner Y, Mimouni FB, Dollberg S, Mandel D. Negative results and impact factor: a lesson from neonatology. *Arch Pediatr Adolesc Med*. Nov; 2005 159(11):1036–7. [PubMed: 16275793]
19. Lee KP, Boyd EA, Holroyd-Leduc JM, Bacchetti P, Bero LA. Predictors of publication: characteristics of submitted manuscripts associated with acceptance at major biomedical journals. *Med J Aust*. Jun 19; 2006 184(12):621–6. [PubMed: 16803442]
20. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA*. Apr 18; 2001 285(15):1992–5. [PubMed: 11308436]
21. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. Sep 4; 2006 185(5):263–7. [PubMed: 16948622]
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. Oct 20; 2007 370(9596):1453–7. [PubMed: 18064739]
23. Khan KS, Nwosu CR, Khan SF, Dwarakanath LS, Chien PF. A controlled analysis of authorship trends over two decades. *Am J Obstet Gynecol*. Aug; 1999 181(2):503–7. [PubMed: 10454707]
24. Levsky ME, Rosin A, Coon TP, Enslow WL, Miller MA. A descriptive analysis of authorship within medical journals, 1995-2005. *South Med J*. Apr; 2007 100(4):371–5. [PubMed: 17458396]



Fig. 1.

Selection process of published articles in six journals between January 2006 and June 2006. Exclusion criteria were article type (commentaries, correspondence, Clinical Expert Series, discussions, editorials, and letters) or journal-specific keywords using EndNote9 (EndNote, Carlsbad, CA). **American Journal of Obstetrics & Gynecology* (AJOG) keyword: “closing.” †The *British Journal of Obstetrics and Gynecology* (BJOG) did not have any EndNote9 keyword exclusions. ‡*Obstetrics & Gynecology* (Obstet Gynecol) keywords: “ACOG Committee Opinion,” “ACOG Practice Bulletin.” §*Journal of the American Medical Association* (JAMA) keywords: “a piece of my mind,” “the cover,” “interview,” “news,” “JAMA pt page,” “historical.” ||*Lancet* keywords: “interview,” “news,” “historical.” #*New England Journal of Medicine* (NEJM) keywords: “clinical problem,” “clinical practice,” “images in clinic,” “videos in clinical,” “historical.”

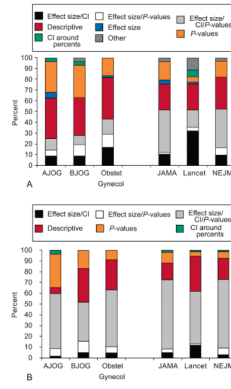


Figure 2. Breakdown of statistical reporting by Journal (A) in the abstract and (B) in the manuscript. Abbreviations: CI – confidence interval, P – P-values.

Table 1

Study Characteristics	Moderate-Impact-Factor Group					High-Impact-Factor Group				P
	AJOG (n=56)	BJOG (n=52)	Obstet Gynecol (n=51)	Total (n=159)	JAMA (n=51)	Lancet (n=52)	NEJM (n=51)	Total (n=154)		
Number of authors										
1-4	16 (29)	25 (48)	16 (31)	57 (36)	13 (25)	8 (15)	2 (4)	23 (15)	<.001	
5-9	34 (60)	27 (52)	33 (64)	94 (59)	17 (33)	25 (48)	19 (38)	61 (40)		
10 or more	6 (11)	0	2 (4)	8 (5)	21 (41)	19 (37)	30 (58)	70 (45)		
Hypothesis clearly stated?	16 (29)	4 (8)	13 (26)	33 (21)	11 (22)	6 (12)	17 (33)	34 (22)	.6	
Negative study	12 (21)	6 (12)	4 (8)	22 (14)	11 (22)	6 (12)	10 (20)	27 (18)	.4	
Sample size/power calculations	9 (16)	11 (21)	16 (32)	36 (23)	22 (43)	18 (35)	33 (65)	73 (47)	<.001	
Statistical measures										
Recommended reporting in abstract	14 (25)	16 (31)	28 (55)	58 (36)	30 (59)	32 (62)	32 (63)	94 (61)	<.001	
Recommended reporting in article	33 (60)	30 (58)	41 (80)	104 (65)	41 (80)	42 (81)	46 (90)	129 (84)	.002	
P-values given for NS differences*	36 (92)	21 (73)	41 (95)	99 (88)	38 (95)	33 (100)	42 (95)	113 (97)	.04	
Regression analysis performed?	29 (52)	19 (37)	28 (55)	76 (48)	27 (53)	29 (56)	36 (71)	92 (60)	.03	

Table 2

Methodology	Moderate-Impact-Factor Group			High-Impact-Factor Group			Total (n=190)
	AJOG (n=56)	BJOG (n=59)	Obstet Gynecol (n=66)	JAMA (n=58)	Lancet (n=69)	NEJM (n=63)	
Level I							
RCT without placebo	4 (7)	6 (10)	2 (3)	9 (16)	8 (12)	11 (17)	28 (15)
RCT with placebo	2 (4)	5 (8)	3 (5)	8 (14)	9 (13)	21 (33)	38 (20)
Controlled trial, other	0	0	0	2 (3)	0	1 (2)	3 (2)
Level II							
Cohort study	28 (50)	18 (31)	17 (26)	10 (17)	11 (16)	8 (13)	29 (15)
Prospective	17 (30)	14 (24)	10 (15)	8 (14)	11 (16)	8 (13)	27 (14)
Retrospective	11 (20)	4 (7)	7 (11)	2 (3)	0	0	2 (1)
Case-control study	5 (9)	5 (8)	4 (6)	0	1 (1)	5 (8)	6 (3)
Cross-sectional study	8 (14)	13 (22)	23 (36)	17 (29)	9 (13)	5 (8)	31 (16)
Survey	1 (2)	6 (10)	9 (14)	16 (9)	7 (12)	1 (1)	9 (5)
Database	3 (5)	3 (5)	10 (15)	16 (9)	6 (10)	5 (7)	11 (6)
Diagnostic	0	1 (2)	1 (2)	2 (1)	0	1 (2)	1 (1)
Other	4 (7)	3 (5)	4 (6)	11 (6)	4 (7)	3 (5)	10 (5)
Level III							
Case report	0	6 (10)	14 (21)	20 (11)	6 (10)	17 (25)	33 (17)
Case series	0	1 (2)	1 (2)	2 (1)	1 (2)	0	3 (2)
Systematic reviews	2 (4)	5 (8)	2 (3)	9 (5)	5 (9)	10 (14)	15 (8)
Meta-analysis	2 (4)	2 (3)	1 (2)	5 (3)	5 (9)	7 (10)	12 (6)
Other	7 (13)	0	0	7 (4)	0	4 (6)	4 (2)