

Reassessing the value of varicocelectomy as a treatment for male subfertility with a new meta-analysis

Joel L. Marmar, M.D.,^a Ashok Agarwal, Ph.D.,^b Sushil Prabakaran, M.D.,^b Rishi Agarwal,^b Robert A. Short, Ph.D.,^c Susan Benoff, Ph.D.,^d and Anthony J. Thomas, Jr., M.D.^b

^a Division of Urology, Robert Wood Johnson Medical School at Camden, Camden, New Jersey; ^b Reproductive Research Center, Glickman Urological Institute and Department of Obstetrics-Gynecology, Cleveland Clinic, Cleveland, Ohio; ^c Health Research and Education Center, Washington State University, Spokane, Washington; and ^d North Shore-LIJ Research Institute, Manhasset, New York

Objective: To determine the efficacy of varicocelectomy as a treatment for male factor infertility by improving the chance of spontaneous pregnancy.

Design: Meta-analysis.

Setting: Cleveland Clinic's Glickman Urological Institute.

Patient(s): Infertile men with abnormal results on semen analyses and a palpable varicocele.

Intervention(s): Surgical varicocelectomy.

Main Outcome Measure(s): Spontaneous pregnancy outcome.

Result(s): The odds of spontaneous pregnancy after surgical varicocelectomy, compared with no or medical treatment for palpable varicocele, were 2.87 (95% confidence interval [CI], 1.33–6.20) with use of a random-effects model or 2.63 (95% CI, 1.60–4.33) with use of a fixed-effects model. The number needed to treat was 5.7 (95% CI, 4.4–9.5).

Conclusion(s): Surgical varicocelectomy in infertile men with palpable lesions and at least one abnormal semen parameter improves the odds of spontaneous pregnancy in their female partners. Five studies were included (two randomized, three observational). All were scored for bias. Our study suggests that varicocelectomy in selected patients does indeed have beneficial effects on fertility status. (*Fertil Steril*® 2007;88:639–48. ©2007 by American Society for Reproductive Medicine.)

Key Words: Varicocelectomy, pregnancy, meta-analysis, random effect

Is a varicocelectomy an effective treatment for male factor subfertility? It is a seemingly simple question that has been the focus of intense debate for nearly 50 years. Many studies report improvement after surgery (1–9), but other studies show no benefit (10–17). Clearly, there are conflicting opinions whether a varicocelectomy improves fertility.

These differences of opinion are most obvious in the clinical guidelines for male factor infertility that have been published by various professional groups. For example, the

Best Policies Practice Groups of both the American Urological Association and the American Society for Reproductive Medicine (18, 19) have stated jointly that correction of varicoceles is indicated for infertile men with palpable lesions and one or more abnormal semen parameters. However, they specifically noted that treatment of the varicocele is not indicated in patients with normal results of semen analyses or subclinical, nonpalpable varicoceles. In contrast, the National Collaborating Centre for Women's and Children's Health, 2005 (11) stated, "Men should not be offered surgery for varicoceles as a form of fertility treatment because it does not improve pregnancy rates." The European Urological Association Guidelines on Male Infertility simply concluded that treatment of varicoceles to achieve pregnancies remains controversial (20). Thus, the fundamental question remains whether the existing literature on varicoceles is reliable enough to resolve these differences of opinion and serves as the basis of a new meta-analysis.

Although randomized controlled trials remain the "gold standard" or level I evidence, the current group of randomized

Received January 11, 2006; revised December 1, 2006; accepted December 5, 2006.

Supported by funds from the Reproductive Research Center and the Glickman Urological Institute, Cleveland Clinic, and by grant number ES10496 to S. B. from the National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, Maryland.

Reprint requests: Ashok Agarwal, Ph.D., Professor, Lerner College of Medicine, Case Western Reserve University, and Director, Reproductive Research Center, Glickman Urological Institute and Department of Obstetrics-Gynecology, Cleveland Clinic, 9500 Euclid Ave., Desk A19.1, Cleveland, OH 44195 (FAX: 216-445-6049; E-mail: agarwaa@ccf.org).

controlled trials on the subject of varicocelectomy and pregnancy outcome has been criticized for several reasons. The critics state that some randomized controlled trials included men with subclinical varicoceles or normal semen analyses and others had significant dropouts after randomization (11, 21–23). A recent critique concluded that analysis of the randomized controlled trial data in the current literature does not allow us to draw any favorable or adverse conclusions to the treatment of varicoceles in infertile couples (24).

Despite these criticisms, our group still believed that it might be possible to develop a new and reliable meta-analysis from the existing literature; however, such a meta-analysis would require several considerations that were fundamentally different. First, we would select only articles that included infertile men with palpable varicoceles who had at least one low semen parameter on three samples. Second, we would limit this study only to surgical repairs because there has been a difference of opinion regarding the outcomes with surgery compared with embolization (25, 26). Third, we would include only articles that reported data on the relationship of “surgical” varicocelectomies and “spontaneous or natural” pregnancy rates.

Furthermore, we would blind the articles for the reviewers and score them with our new scoring system to evaluate four types of study bias. Still further we would follow the lead of some investigators who have suggested that it is reasonable to include level II evidence from observational or case-controlled studies in a meta-analysis so long as the observational studies have considered carefully the potential for bias (21, 27). In various clinical situations, the existing evidence from randomized controlled trials addressing the effectiveness of specific interventions may be quite limited; therefore, it may not be inappropriate for systematic reviews to include carefully considered nonrandomized studies to provide a detailed picture of our current knowledge and limitations (28). Therefore, in the present meta-analysis on surgical varicocelectomy and spontaneous pregnancy, we intended to include both randomized controlled trials and observational studies (Fig. 1).

MATERIALS AND METHODS

Types of Patients

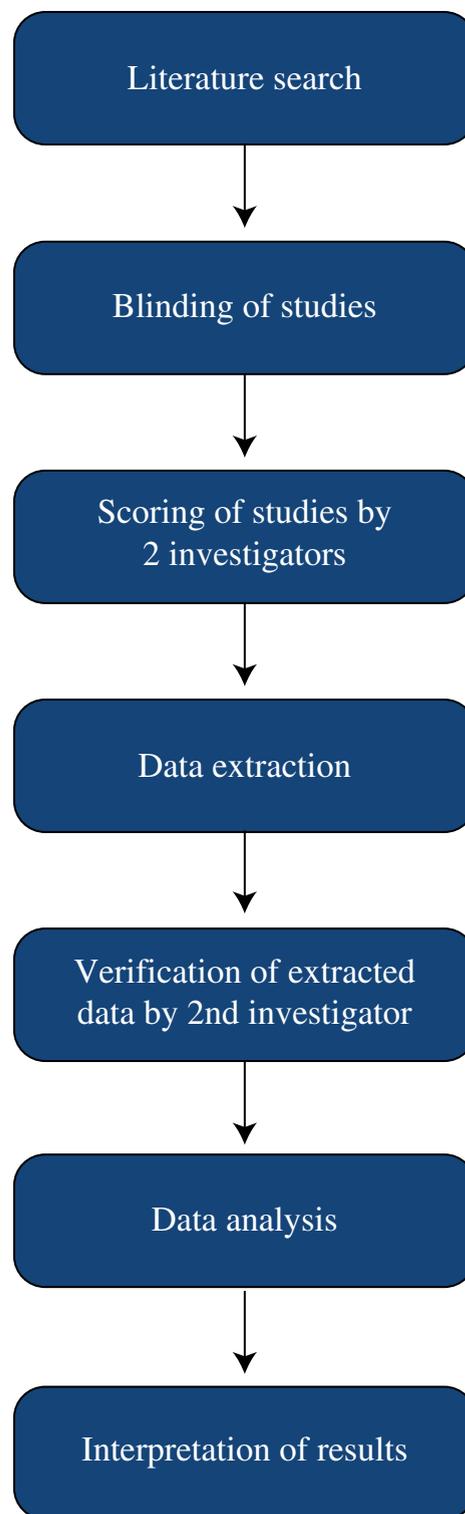
Studies of infertile men with a diagnosis of unilateral or bilateral palpable varicoceles and at least one abnormal semen parameter were included. Control groups were composed of infertile men with varicoceles who declined to undergo surgical repair and who were randomly assigned either to no treatment or to medical treatment.

Types of Intervention

Surgical varicocelectomy (high ligation, inguinal, or microsurgery) was reviewed.

FIGURE 1

Flow diagram on selection of studies for the meta-analysis.



Marmar. Varicocelectomy—a new meta-analysis. *Fertil Steril* 2007.

Types of Outcome Measure

The outcome measure was the effect of surgical varicocelectomy on natural or spontaneous pregnancy outcome during follow-up of up to 24 months.

Search Strategy for Identification of Studies

Studies were identified by performing an extensive search with BIOSIS, EMBASE, and Medline (from 1985 to present) with the help of a professional librarian, as well as by hand-searching review articles and cross-references. The overall strategy for study identification and data extraction is outlined in Figure 2. The following key words were used to search the databases: varicocelectomy, microsurgery, high ligation, infertility, semen parameters, and pregnancy rate or outcome. No exclusions were made on the basis of

language. Articles were evaluated for relevance by examining titles and abstracts. Studies were excluded if there were patients with subclinical varicoceles only or subclinical varicoceles combined with clinical varicocele and if the effect of treatment was examined only in an adolescent population.

Evaluation of Relevant Studies by Blinding and Scoring of Studies

All articles and reviewers were blinded during the evaluation period. The methods, results, tables, and figures from each study were extracted, and each article was assigned an identification number by an individual other than the two scorers. Actual quantitative or qualitative report of results was blacked out in each article to enable unbiased scoring of study quality. Data points or graphs were blacked out of figures, whereas axes and captions were still included for evaluation. Summary statistics, *P* values, or descriptions were blacked out of tables and texts, whereas labels such as comparison groups and parameters measured were left viewable. Two evaluators blinded to the concluding results, authors, journal, and year of the articles evaluated each study on its methodologic merits.

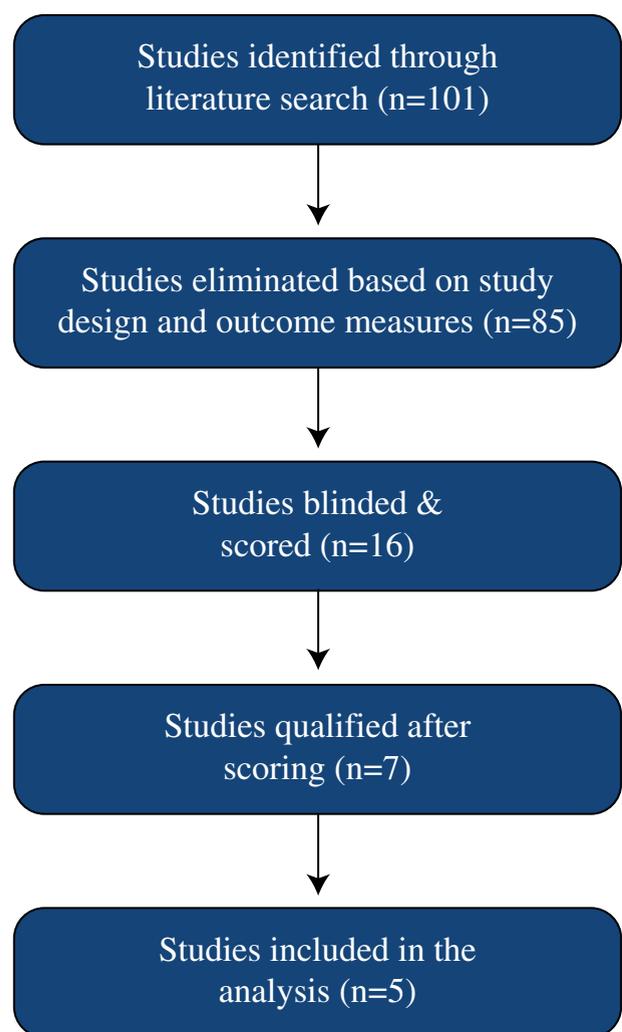
Articles with both preoperative and postoperative repeated measures of semen parameters were evaluated for methodologic quality by our initial scoring system (Appendix). The questions and scores were developed to evaluate four categories of bias: selection or follow-up bias, confounding bias, information or detection bias, and other sources of bias such as misclassification. Each study was scored by using the same set of questions for each type of bias (29, 30). Specific answers for different questions were given more weight than others as evidenced in the point system used to total the scores for each category of bias.

A higher score indicated that the study met most of the criteria required to avoid introducing bias in the study. If the point total for more than one category of bias was below an acceptable range, the study automatically was excluded from the final analysis. If the points for only one category totaled below the acceptable range, the study was reexamined to determine whether, indeed, the overall study was likely to be biased and, if not, whether it could be included in the meta-analysis.

The point ranges for exclusion or inclusion were determined by the epidemiologic importance of each study, its scientific quality, and the possibility of the article reaching a biased conclusion. For example, in the case of selection or follow-up biases, if a large proportion of subjects were lost to follow-up, then it becomes difficult to determine whether those who selectively dropped out may have been the ones with least improvement or whether the losses were simply too few in numbers to have biased the results. Thus, to deal with this potential quandary, a positive answer (no loss to follow-up over 10%) was given 2 points and a negative answer (loss to follow-up in more than 10% of participants) was given 1 point (Appendix). The same rationale was

FIGURE 2

Flow diagram on selection of studies evaluating the effect of varicocelectomy on the pregnancy outcome.



Marmar. Varicocelectomy—a new meta-analysis. *Fertil Steril* 2007.

carried out for all other sources of bias: confounding bias, information or detection biases, and other sources of bias such as misclassification. If information for a particular question was not stated, the study was given only 1 point for that question. Furthermore, the category of confounding was designed to include studies that made a comparison between the same subjects but not over more than a 2-year period. If the follow-up time was more than 2 years after the surgery, or with no follow-up within this time period, or if the study did not account for time-varying confounders, then it was likely that the study would be excluded.

This method of scoring studies was used rather than a simple checklist because the latter may produce bias (29). As an alternative, our scoring plan was intended to identify and quantify potential sources of bias. Two reviewers scored each study independently, and the final decision on whether a study was to be included was determined by discussion between the two reviewers.

Data Extraction

Data were extracted by one of the authors on a preformatted data extraction sheet. Population information (i.e., primary versus secondary infertility) and study characteristics such as the specific intervention (high ligation, microsurgery, and laparoscopy) were listed. These data were available for subsequent subgroup analyses.

The data were then entered in the RevMan software (version 4.2.8) developed by the Cochrane Collaborative for the purpose of meta-analysis (<http://www.cochrane.org>).

Effect of Varicocelectomy on Pregnancy Outcome

To examine the effect of varicocelectomy on “spontaneous or natural” pregnancy, we studied cohorts within a 2-year follow-up after varicocelectomy was performed in one male cohort and no, delayed, or medical treatment in the other cohort. Studies were excluded if they had men with subclinical varicoceles. Also, patients who had undergone assisted reproductive techniques (ART) such as IVF or IUI were not included in the analysis. Studies that used embolization or sclerosis techniques for varicocele corrections were excluded. Pregnancy data were recorded for the 24-month interval after surgery, and the overall odds were calculated by random-effects and fixed-effects models. The number needed to treat was calculated and evaluated by 95% confidence intervals (CI) (31). The number needed to treat was recalculated after removal of the most favorable study (1). All data were verified by a second investigator.

RESULTS

Of the 101 articles retrieved from the search containing pregnancy data, further elimination because of the study designs and relevance of outcomes measured was conducted, yielding 16 studies to be blinded. These 16 studies were then scored and assessed for quality. Two of these studies (15, 32) were excluded because they had both patients with surgical ligation and patients who had sclerosis or radiologic emboli-

zation. Patients whose partners achieved pregnancy with ART (IVF and/or IUI) were excluded from our analysis, but one of the included studies followed patients for spontaneous pregnancy after ART (4). The remainder of these studies were excluded on the basis of their scores for bias (Appendix). Therefore, our meta-analysis was limited to five surgical studies that included data on spontaneous pregnancy rates.

The mean age of the male cohorts was 31.2 years (range 20–46 years). Laterality was reported in four of the five studies. Left varicoceles were noted in 67.4% to 81.5% of the patients, bilateral varicoceles in 14.0% to 30.4%, and right varicoceles in 2.1% to 5.5%. The varicoceles were all palpable, but the specific sizes were recorded in only two of the five studies. Grade III (large) varicoceles were recorded in 9.5% and 34.3% of the patients, grade II (moderate) in 54.4% and 21.8%, and grade I (small) in 36.2% and 43.7%. The controls in four of the studies had no treatment, whereas in one study the controls used clomiphene citrate (Clomid).

The odds of spontaneous pregnancy after varicocelectomy compared with no or medical treatment for clinical varicocele were 2.87 (95% CI, 1.33–6.20, $P=.007$) with use of the inverse variance random-effects model (Table 1). A fixed-effects model also yielded a significant odds ratio of 2.63 (95% CI, 1.60–4.33], $P=.00001$). Results of the test for the presence of heterogeneity between study measures was not significant ($P=.17$).

Pregnancy outcome was also evaluated on the basis of the number needed to treat. Within the five studies there were 396 patients who underwent operation and who had 131 pregnancies (33.0%) versus 174 controls with 27 pregnancies (15.5%). The number needed to treat was 5.7 (95% CI, 4.1–9.5). When the data were recalculated after removal of the figures from the most favorable study (1), the number needed to treat was 6.6 (95% CI, 4.4–13.3), which represents comparable results following surgery.

DISCUSSION

A group of 20 scientists from nine countries, known as the Potsdam Consultation, convened to develop guidelines for the conduct and interpretation of meta-analyses. They added to the experience of earlier investigators and defined a meta-analysis as a systematic review that uses statistical methods to combine and summarize the results of several studies (29), and they listed 13 specific methodologic principles for the performance of these studies. In a separate study, Thacker et al. noted the benefits of a meta-analysis and stated that “decisions about clinical practice should be based on the combined weight of the evidence from available reports” (33). However, they warned that there are systematic rules for conducting a meta-analysis that include an explicit description of methodology so that results can be interpreted in light of any biases or limitations.

In the present meta-analysis, we examined the effect of surgical varicocelectomy on spontaneous pregnancy

TABLE 1

Confidence intervals and odds ratio for pregnancy after varicocelectomy among men with palpable lesions and at least one abnormal semen parameter.

Study	Varicocelectomy n/N	Control n/N	OR (random) 95% CI	OR (random) 95% CI
Grasso et al 2000	1 / 34	2 / 34		0.48 [0.04, 5.61]
Madgar et al 1995	16 / 25	2 / 20		13.50 [2.55, 71.40]
Marmar et al 1994	66 / 186	3 / 19		2.93 [0.82, 10.44]
Okuyama et al 1988	43 / 141	15 / 83		1.99 [1.02, 3.86]
Onozawa et al 2002	6 / 10	5 / 18		3.90 [0.76, 19.95]
Total (95% CI)	396	174		2.87 [1.33, 6.20]
Total events: 131 (Varicocelectomy), 27 (Control)				
Test for heterogeneity: Chi ² = 8.47, df = 4 (P = 0.17), I ² = 38.1%				
Test for overall effect Z = 2.68 (P < 0.00001)				

Note: OR = odds ratio; n = number of couples achieving pregnancy with male partners diagnosed with clinical varicoceles; N = total number of cases.

Marmar. Varicocelectomy—a new meta-analysis. Fertil Steril 2007.

outcome. We developed a different methodology than in previous meta-analyses: we adhered to the principles of the Potsdam guidelines, we “blinded” the reviewers during the evaluation process, and we developed a scoring system to quantify bias that was used to evaluate the literature on infertile men who had “surgical” correction of varicoceles. With this approach, the data from our meta-analysis led us to conclude that a surgical varicocelectomy improved the spontaneous pregnancy rates for infertile men with low semen parameters and palpable varicoceles. However, we believe that a critical discussion of our methods is necessary to understand the process.

A primary cause of bias in a meta-analysis can come from the reviewers. It is feasible that a reviewer may be influenced by the knowledge of institutions, investigators, and the concluding results associated with other aspects of the study (34). This has been termed inclusion bias (35). The mechanisms for inclusion bias have been reported (36), but they may be minimized by “blinding” the examiners and the documents during the evaluation process. Historically, the majority of meta-analyses have not been blinded. However, in the present meta-analysis we chose to make the effort and take particular precautions to proceed with blinded evaluations. In addition, we used other strategies to evaluate bias.

To evaluate the literature for bias, we developed a scoring system rather than a checklist (29). It is possible that a well-designed study with a traditionally acceptable checklist for inclusion (statement of randomization procedure, stated assessment of confounding, etc.) still may be biased because of unequal weighting. Some items on the checklist may value individual studies inappropriately over the others. Our scor-

ing plan was intended to adjust for this type of bias. In some instances, specialty groups have developed and used standardized protocols for scoring literature in their fields (37, 38). In the absence of standard methods in the present field, we developed our own scoring system to evaluate several types of bias. Although the scoring system was not validated statistically, during several meetings before the initiation of this meta-analysis we discussed and adopted a set of specific questions and point scores to limit quantitatively against potential bias in the literature under consideration. The new system was applied to the blinded manuscripts to determine their inclusion into the meta-analysis. Two reviewers scored each study, independently, and the final decision on inclusion or exclusion was determined during a discussion between the two reviewers. This approach has been considered and accepted in other fields for the development of other meta-analyses (38, 39).

Some investigators have been critical of the fact that most varicocele studies have been uncontrolled and not randomized controlled trials. In an attempt to address these matters, two recent meta-analyses have been published on varicoceles (10, 12) with only level I evidence from a group of randomized controlled trials. Essentially, these two meta-analyses included the same group of randomized controlled trials, and both reviews came to the conclusion that varicocele repairs do not improve subfertility. However, these meta-analyses, themselves, have been the subject of several critiques that have cited methodologic flaws that may have biased the results. Specifically, the National Collaborating Centre for Women’s and Children’s Health, 2005 report (11) was critical for several reasons: these meta-analyses

had clinical heterogeneity in the subjects selected, and there were differences in the mean age of the male partners, differences in the duration of infertility, and high dropout rates after randomization. In another critique, Templeton (22) commented that Evers and Collins (12) elected to exclude a large, hitherto-unpublished World Health Organization (WHO) study that appeared in abstract form. These WHO data suggested that varicocelectomy appeared to improve pregnancy, and, according to Templeton, “exclusion of the multi-centre data is important and could have made a difference” (22), which raises concerns for publication bias. In still other critiques, some investigators pointed out that four of the eight randomized controlled trials included in these meta-analyses had men with subclinical varicoceles, and two had men with normal semen parameters (23). For example, Ficarra et al. (24) reevaluated data from a prior meta-analysis (12) by removing studies that included men with subclinical varicoceles and normal semen parameters. When they recalculated the data for the “as-treated” groups, the pregnancy rates were 36.4% for the surgically treated group and 20% for controls ($P=0.009$). Therefore, these past meta-analyses have not resolved the issues surrounding varicocelectomy and subfertility. Nevertheless, our group believed that it was still possible to develop a new and more inclusive meta-analysis from the existing literature that may lead to valid conclusions.

In the present meta-analysis, we included both randomized controlled trials and observational studies. Although this approach may be controversial to some, we believe that it is sound for the current subject matter primarily because of the lack of reliable randomized controlled trial data. Several studies suggest that randomized controlled trials provide the highest level of evidence for causation, but they are known to be costly and difficult to complete, particularly without experiencing a significant number of dropouts after randomization (21, 27, 28). Furthermore, in some instances, the ethics of the randomized controlled trials may be viewed as borderline because the randomization and informed consent may not reach international standards, and in most cases they are done without peer review (40). Still further, some randomized controlled trials may be particularly unfair to infertile couples who are offered no treatment in one arm of the trial, when alternative treatment is available such as IVF (22). In these instances, treatment delays may expose these couples to the negative influence of advancing age on pregnancy outcome.

Because of these realities, we chose to include observational studies in addition to randomized controlled trials in the present meta-analysis. The Potsdam Consultation noted that observational studies should not be abandoned, but they may be included after critical appraisal, empirical study, and methodologic evaluation (29). Oftentimes, combining data from several smaller observational studies may be an efficient, effective, and perhaps the only means of reaching a conclusion (41). The challenge lies in developing a methodology for evaluating these observational studies and deciding whether to include a particular study. During the evaluation

period, the observational studies were scrutinized with the same scoring system for bias as randomized controlled trials. Further, we analyzed the data with a random-effects analysis to accommodate for heterogeneity. One other approach to reduce heterogeneity is to remove primary studies selectively from consideration, but this method opens the door for other bias. If removal of studies from the analysis is not based on biological and clinical differences of study design or specific interventions, removal may shift the weight of the evidence inappropriately on the measured outcome.

Although our approach was not totally free of problems, it contained other safeguards against methodologic bias that plagued earlier meta-analyses. Our scoring system excluded studies during the evaluation phase with large numbers of individual dropouts after randomization, which was different from other meta-analyses that included these types of studies (42). The individuals who dropped out may differ systematically from those who stayed in, and, if studies with excessive dropouts are included, then the meta-analysis may be influenced by confounding bias. Not all the studies included in prior meta-analyses (even randomized trials) examined the distribution of age even though it is known to be a confounder, and some studies demonstrated detection bias because they included individuals with subclinical varicoceles. The five studies in the present meta-analysis all had men with palpable varicoceles only, including Grasso et al. (14), who studied men with ultrasound and a scoring system that included small, palpable, grade I lesions. Although it takes careful planning and a greater work effort, it seems important to evaluate for all of these types of bias during the evaluation phase to include the most reliable data in the meta-analysis.

With regard to spontaneous pregnancy outcome, some studies found that there is no difference in the odds of pregnancy for men who underwent varicocelectomy in comparison with those who did not. However, this lack of difference may be due to the fact that the researcher was not seeking to record pregnancy as a main outcome variable. If large studies are included in a meta-analysis with only partial pregnancy data, they may be given more weight, despite that the study did not aim to measure the odds of pregnancy. This would cause the overall conclusion to be weighted toward no effect. If the patients' cases had been followed up thoroughly, there might have been a different observed effect of treatment on the outcome.

In our meta-analysis, we used an approach that was consistent with the guidelines for meta-analyses of the Potsdam Collaboration (29). We included only studies that had spontaneous pregnancy data as an intended outcome. We evaluated five studies that led us to conclude that a surgical varicocelectomy improved spontaneous pregnancy outcome on the basis of the odds ratio and the number needed to treat. Furthermore, even after removal of the figures of the most favorable study (1), the number needed to treat results from the remaining four studies seemed comparable to those of the original five.

In a side study, we found that the pregnancy results after varicocelectomy were usually associated with the improvement in sperm density. The four studies with improved pregnancy rates all reported statistically significant increases in postoperative sperm density, whereas the one study that reported no improved pregnancy rates had no improvement in the postoperative sperm density. Presently, these findings may have increased clinical relevance, because recent reports have linked sperm density to fertility in other situations. For example, a study of fertile and infertile populations reported that the mean sperm densities for these groups were 19.5 versus 8.5 $10^6/\text{mL}$, respectively ($P < .001$) (43). Another study reported that doubling the sperm concentration, for example, from 4 to $8 \times 10^6/\text{mL}$, increased the monthly conception rate by a factor of 2.8, and doubling the concentration from 8 to $16 \times 10^6/\text{mL}$ or from 16 to $32 \times 10^6/\text{mL}$ increased fecundability by factors of 1.34 and 1.32, respectively (44). Still other studies showed that the odds ratio for subfertility was 5.6 (3.3–8.3) for those men with sperm densities $< 13.5 \times 10^6/\text{mL}$, compared with an odds ratio of 1.3 (1.2–2.2) for men with sperm densities between 13.5 and $48 \times 10^6/\text{mL}$ (45). Thus, the measurement of sperm density continues to be important in the evaluation of male factor infertility. However, there is usually great variability in the parameters from consecutive semen studies, and a panel of at least three preoperative and three postoperative semen analyses is needed to avoid the statistical phenomenon of regression toward the mean (3). Although semen data were not included in the present manuscript, future varicocele studies probably should include this type of panel to evaluate both patients and controls.

Recently, other new molecular and genetic markers have been used to stratify patients with varicocele (46, 47). For example, some studies have documented increased DNA damage to the sperm in the semen of infertile men with varicoceles; others demonstrate the presence of oxidative stress (48, 49). In still other studies, there was no improvement in semen parameters after varicocelectomy among men with Y-chromosome microdeletions or abnormal karyotypes (50), and reduced pregnancy outcomes after varicocelectomy among men with increased testis tissue cadmium, and microdeletions in the sequence of the L-type, voltage-dependent calcium channel (51). Thus, the presence or absence of these markers may explain why some men with varicoceles are fertile and others do not improve after varicocelectomy. Furthermore, future studies in men with varicoceles may be more selective by stratification with these markers. This approach may identify those patients with a realistic opportunity to benefit from these procedures versus those in whom varicocelectomy is likely to fail on a molecular/genetic basis.

CONCLUSION

On the basis of the data from current literature, we conclude from this meta-analysis that a surgical varicocelectomy is an effective treatment for improving the spontaneous pregnancy rate for couples with an infertile male partner who has low semen parameters and a palpable varicocele. In the future,

randomized controlled trials should include stratification with a panel of semen analyses and molecular/genetic markers. Furthermore, the control groups should be offered some meaningful treatment such as IVF to avoid unfair conditions for participation in a research study.

Acknowledgments: Di He, B.S. (summer student, volunteer), provided statistical assistance. The authors would like to thank Andrew C. Novick, M.D., Chairman, Glickman Urological Institute, Cleveland Clinic, for his encouragement and support for research.

REFERENCES

1. Madgar I, Weissenberg R, Lunenfeld B, Karasik A, Goldwasser B. Controlled trial of high spermatic vein ligation for varicocele in infertile men. *Fertil Steril* 1995;63:120–4.
2. Schlesinger MH, Wilets IF, Nagler HM. Treatment outcome after varicocelectomy. A critical analysis. *Urol Clin North Am* 1994;21:517–29.
3. Marmar JL, Kim Y. Subinguinal microsurgical varicocelectomy: a technical critique and statistical analysis of semen and pregnancy data. *J Urol* 1994;152:1127–32.
4. Onozawa M, Endo F, Suetomi T, Takeshima H, Akaza H. Clinical study of varicocele: statistical analysis and the results of long-term follow-up. *Int J Urol* 2002;9:455–61.
5. Okuyama A, Fujisue H, Matsui T, Doi Y, Takeyama M, Nakamura N, et al. Surgical repair of varicocele: effective treatment for subfertile men in a controlled study. *Eur Urol* 1988;14:298–300.
6. Cayan S, Kadioglu TC, Tefekli A, Kadioglu A, Tellaloglu S. Comparison of results and complications of high ligation surgery and microsurgical high inguinal varicocelectomy in the treatment of varicocele. *Urology* 2000;55:750–4.
7. Goldstein M, Gilbert BR, Dicker AP, Dwosh J, Gnecco C. Microsurgical inguinal varicocelectomy with delivery of the testis: an artery and lymphatic sparing technique. *J Urol* 1992;148:1808–11.
8. Sayfan J, Soffer Y, Orda R. Varicocele treatment: prospective randomized trial of 3 methods. *J Urol* 1992;148:1447–9.
9. Segenreich E, Israilov SR, Shmueli J, Niv E, Servadio C. Correlation between semen parameters and retrograde flow into the pampiniform plexus before and after varicocelectomy. *Eur Urol* 1997;32:310–4.
10. Kamischke A, Nieschlag E. Varicocele treatment in the light of evidence-based andrology. *Hum Reprod Update* 2001;7:65–9.
11. Clinical guidelines: Fertility: assessment and treatment for people with fertility problems. National Collaborating Centre for Women's and Children's Health 2004;54–5.
12. Evers JL, Collins JA. Surgery or embolisation for varicocele in subfertile men. *Cochrane Database Syst Rev* 2004;CD000479.
13. Nilsson S, Edvinsson A, Nilsson B. Improvement of semen and pregnancy rate after ligation and division of the internal spermatic vein: fact or fiction? *Br J Urol* 1979;51:591–6.
14. Grasso M, Lania C, Castelli M, Galli L, Franzoso F, Rigatti P. Low-grade left varicocele in patients over 30 years old: the effect of spermatic vein ligation on fertility. *BJU Int* 2000;85:305–7.
15. Breznik R, Vlaisavljevic V, Borko E. Treatment of varicocele and male fertility. *Arch Androl* 1993;30:157–9.
16. Rageth JC, Unger C, DaRugna D, Steffen R, Stucki D, Barone C, et al. Long-term results of varicocelectomy. *Urol Int* 1992;48:327–31.
17. Krause W, Muller HH, Schafer H, Weidner W. Does treatment of varicocele improve male fertility? Results of the "Deutsche Varikozelenstudie," a multicentre study of 14 collaborating centres. *Andrologia* 2002;34:164–71.
18. Jarow JP, Sharlip ID, Belker AM, Lipshultz LI, Sigman M, Thomas AJ, et al. Best practice policies for male infertility. *J Urol* 2002;167:2138–44.
19. Sharlip ID, Jarow JP, Belker AM, Lipshultz LI, Sigman M, Thomas AJ, et al. Best practice policies for male infertility. *Fertil Steril* 2002;77:873–82.

20. Dohle GR, Colpi GM, Hargreave TB, Papp GK, Jungwirth A, Weidner W. EAU Guidelines on Male Infertility. *Eur Urol* 2005;48:703–11.
21. Williams JK. Understanding evidence-based medicine: a primer. *Am J Obstet Gynecol* 2001;185:275–8.
22. Templeton A. Varicocele and infertility. *Lancet* 2003;361:1838–9.
23. Marmar J, Benoff S. New scientific information related to varicoceles. *J Urol* 2003;170:2371–3.
24. Ficarra V, Cerruto MA, Ligouri G, Mazzano G, Minucci S, Tracia A, et al. Treatment of varicocele in subfertile men: The Cochrane review—a contrary opinion. *Eur Urol* 2006;49:258–63.
25. Yavetz H, Levy R, Papo J, Yogev L, Paz G, Jaffa AJ, et al. Efficacy of varicocele embolization versus ligation of the left internal spermatic vein for improved sperm quality. *Int J Androl* 1992;15:338–44.
26. Shlansky-Goldberg RD, VanArsdalen KN, Rutter CM, Soulen MC, Haskal ZJ, Baum RA, et al. Percutaneous varicocele embolization versus surgical ligation for the treatment of infertility: changes in seminal parameters and pregnancy outcomes. *J Vasc Interv Radiol* 1997;8:759–67.
27. Peipert JF, Phipps MG. Observational studies. *Clin Obstet Gynecol* 1998;41:235–44.
28. Norris SL, Atkins D. Challenges in using nonrandomized studies in systematic reviews of treatment interventions. *Ann Intern Med* 2005;142:1112–9.
29. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *J Clin Epidemiol* 1995;48:167–71.
30. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions 4.2.5*. In: *The Cochrane Library*, 2005. Chichester: Wiley.
31. Bender R. Calculating confidence intervals for the number needed to treat. *Controlled Clin Trials* 2001;22:102–10.
32. Nieschlag E, Hertle L, Fishedick A, Abshagen K, Behre HM. Update on treatment of varicocele: counselling as effective as occlusion of the vena spermatica. *Hum Reprod* 1998;13:2147–50.
33. Thacker SB, Stroup DF, Peterson HB. Meta-analysis for the practicing obstetrician gynecologist. *Clin Obstet Gynecol* 1998;41:275–81.
34. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987;316:450–5.
35. Egger M, Smith GD. Bias in location and selection of studies. *BMJ* 1998;316:61–6.
36. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 1997;350:185–6.
37. Beck CT. The effects of postpartum depression on maternal-infant interaction: a meta-analysis. *Nurs Res* 1995;44:298–304.
38. Dennis LK, Snetselaar LG, Nothwehr FK, Stewart RE. Developing a scoring method for evaluating dietary methodology in reviews of epidemiologic studies. *J Am Diet Assoc* 2003;103:483–7.
39. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;282:1054–60.
40. Comhaire F. Clinical andrology: from evidence-base to ethics. The “E” quintet in clinical andrology. *Hum Reprod* 2000;15:2067–71.
41. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008–12.
42. Evers JL, Collins JA. Assessment of efficacy of varicocele repair for male subfertility: a systematic review. *Lancet* 2003;361:1849–52.
43. Ombelet W, Bosmans E, Janssen M, Cox A, Vlasselaer J, Gyselaers W, et al. Semen parameters in a fertile versus subfertile population: a need for change in the interpretation of semen testing. *Hum Reprod* 1997;12:987–93.
44. Bonde JP, Ernst E, Jensen TK, Hjøllund NH, Kolstad H, Henriksen TB, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* 1998;352:1172–7.
45. Guzik DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, et al. Sperm morphology, motility, and concentration in fertile and infertile men. *N Engl J Med* 2001;345:1388–93.
46. Benoff S, Gilbert BR. Varicocele and male infertility: part I. Preface. *Hum Reprod Update* 2001;7:47–54.
47. Marmar J. Varicocele and male infertility: part II. *Hum Reprod Update* 2001;7:461–742.
48. Saleh RA, Agarwal A, Sharma RK, Said TM, Sikka SC, Thomas AJ Jr. Evaluation of nuclear DNA damage in spermatozoa from infertile men with varicocele. *Fertil Steril* 2003;80:1431–6.
49. Zini A, Blumenfeld A, Libman J, Willis J. Beneficial effect of microsurgical varicocelectomy on human sperm DNA integrity. *Hum Reprod* 2005;20:1018–21.
50. Cayan S, Lee D, Black LD, Reijo Pera RA, Turek PJ. Response to varicocelectomy in oligospermic men with and without defined genetic infertility. *Urology* 2001;57:530–5.
51. Marmar JL, Benoff S. The safety of ultrasonically guided testis aspiration biopsies and efficacy of use to predict varicocelectomy outcome. *Hum Reprod* 2005;20:2279–88.

APPENDIX

Initial scoring sheet used for evaluating studies that included men with palpable varicoceles, at least one low semen parameter, and pregnancy data after varicocelectomies

Study number _____ Reviewer Initials _____

Selection/Follow-up _____ Total _____

From what, if any, underlying cohort is the study population derived?

- 3 From a geographical cohort
- 3 From a community
- 2 From a clinic population
- 1 Unable to answer

How were subjects recruited?

- 3 All cases in the population were included
- 2 Cases were recruited consecutively over a period of time
- 3 Cases were randomly selected
- 1 Unable to answer

Was there loss of follow-up or lack of participation greater than 10% of those sampled initially?

- 1 Yes
- 2 No
- 1 Unable to answer

Did the investigators restrict against participants based on infection, previous treatment, and female factor infertility or conditions related to ART outcome and sperm parameters?

- 3 Yes
- 2 No
- 1 Unable to answer

Confounding

Total _____

Was the time between the two follow-up periods short enough to allow for no confounding by age within subjects (under 2 years)?

- 3 Yes
- 1 No
- 1 Unable to answer

Did they evaluate and account for potential confounders that may vary over time?

ie, amount of follow-up time, season, smoking, alcohol consumption, original sperm count, time-varying exposures, etc.

- 1 No
- 2 Yes, but they do not adjust
- 3 Yes, and they adjust for them when necessary
- 1 Unable to answer

Did the investigators pre-specify the same procedures for analysis for before and after the intervention?

- 2 Yes
- 2 Not applicable
- 1 No
- 1 Unable to answer

Information/Detection Bias

Total _____

Was the method of follow-up the same before and after treatment?

- 3 Yes
- 2 No
- 1 Unable to answer

If blinding was possible, were those evaluating outcomes blinded to the patient's intervention/disease status?

- 2 Blinding was not possible
- 1 Blinding was possible but not done for all/some investigators
- 3 Blinding was performed
- 1 Unable to answer

Was the measurement of outcome(s) objective?

Objective meaning medical records or diagnostic test, not objective/subjective meaning recall, etc.

- 3 Yes
- 2 No
- 1 Unable to answer

Was ascertainment of outcome performed at the same location both before and after treatment?

- 4 Yes
- 2 No
- 1 Unable to answer

Other

Total _____

Does the study combine outcomes across groups with very heterogeneous histories/durations of infertility and across different interventions?

- 2 Yes, together

- 3 No
- 1 Unable to answer

Was severity/grade of varicocele evaluated both before and after the intervention?

- 3 Yes
- 2 No
- 1 Unable to answer

Did investigators use an established set of guidelines for semen analysis?

- 4 Yes
- 1 No
- 1 Unable to answer

Exclusion Criteria

Category	Maximum Score	Minimum Score	Include Score	Exclude Score
Selection	11	4	11-7	6-4
Confounding	8	3	8-5	4-3
Information	13	4	13-12	11-4
Other	10	3	10-8	7-3

- Any study will be excluded if 2 or more categories score in the “exclude range”.
- Any study will be re-reviewed if only 1 category scores in the “exclude range”.