

Reviews Pooling Heterogeneous, Low-Evidence, High-Bias Data Result in Incorrect Conclusions: But Heterogeneity is an Opportunity to Explore



Abstract: Systematic Review submissions to our journal commonly pool heterogeneous studies of low levels of evidence and a high risk of bias. Pooling, or quantitative synthesis, of such study data regularly results in incorrect conclusions. We reject these submissions without peer-review (desk rejection), and typically invite authors to submit a new, subjective synthesis without pooling and to report ranges of the results of included studies rather than pooled values. Generally, quantitative synthesis, or meta-analysis, should restrict included studies to randomized controlled trials. However, systematic review with exploration of heterogeneity can result in valuable information toward determining strengths and deficiencies of current literature, and thus guide future research.

Systematic Review submissions to our journal commonly pool heterogeneous studies of low levels of evidence and a high risk of bias. Pooling, or quantitative synthesis of such study data regularly results in incorrect conclusions.¹⁻⁴ We reject these submissions without peer-review (desk rejection),⁵ and typically invite authors to submit a new, subjective synthesis without pooling and to report ranges of the results of included studies rather than pooled mean values. Generally, quantitative synthesis, or meta-analysis, should restrict included studies to randomized controlled trials.

However, systematic review with exploration of heterogeneity can result in valuable information.⁶⁻⁸ Shifting the analysis away from summary estimates to explore sources of heterogeneity can result in identification of the clinical and methodologic factors with which the heterogeneity is associated. Then, instead of the trite and low-value conclusion “future research is required,” a conclusion can pinpoint deficiencies in the literature to inform and direct future investigation.

A Hypothetical Example (Based on a Real Submission)

Consider a review of 10 case series (Level IV evidence) reporting reoperation rates after knee

arthroscopy. Data reported in the included studies is illustrated in Table 1, and a forest plot illustrating reoperation rates including 95% confidence intervals by study is illustrated in Figure 1.

Hypothetical *misinterpretation* based on improper pooling is as follows: the rate of reoperation after knee arthroscopy is 26% (effect size = 0.26). Readers will note that the 26% weighted mean value or summary estimate of reoperation rates is labeled as improper in Table 1 and thus not included in Figure 1.

But why do we deem this to be a misinterpretation based on “improper” pooling? The studies are wildly heterogeneous. Figure 1 shows that the I-squared statistic (how much of the difference in reoperation rates among the 10 studies is attributable to heterogeneity,

Table 1. Number of Reoperations, Reoperation Rate, and Sample Size Reported in 10 Studies on Knee Arthroscopy

	Number of Reoperations	Reoperation Rate	Sample Size
Study 1	1	2%	45
Study 2	41	46%	90
Study 3	5	8%	60
Study 4	32	44%	72
Study 5	95	53%	180
Study 6	2	1%	165
Study 7	4	4%	112
Study 8	12	19%	62
Study 9	53	43%	122
Study 10	5	11%	47
Total	250	26%	955

NOTE. Total reoperation rate of 26% represents weighted mean value of reoperation rates of the 10 studies as a result of improper (see text) pooling.

© 2018 by the Arthroscopy Association of North America
0749-8063/181233/\$36.00
<https://doi.org/10.1016/j.arthro.2018.10.005>

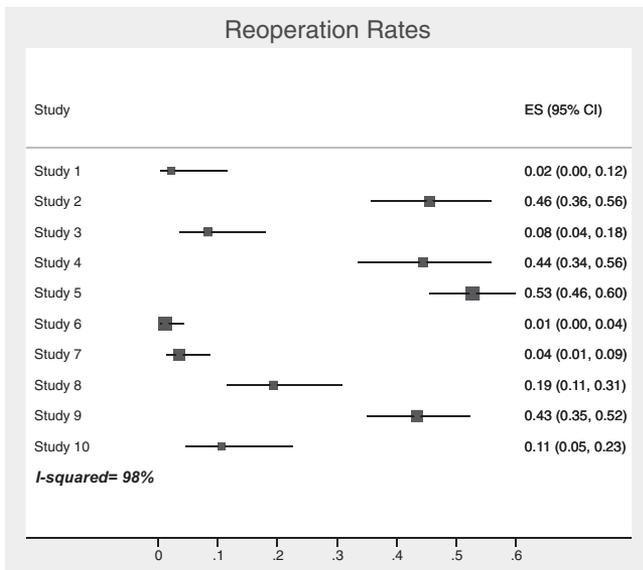


Fig 1. Forest plot displaying effect size (reoperation rates represented by squares, proportional in size to sample size of the study) and confidence intervals (horizontal lines) of reoperation by study. (ES, Effect Size; CI, Confidence Interval; I-squared, percentage of variation due to heterogeneity rather than chance.)

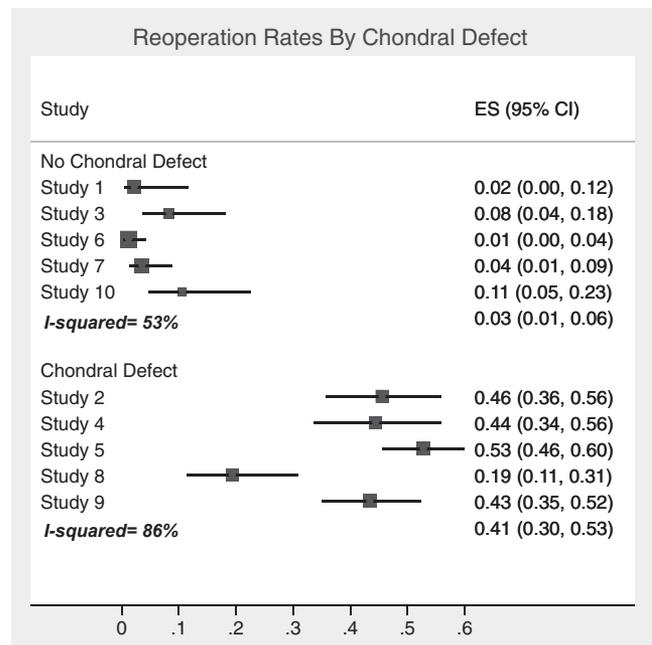


Fig 2. Forest plot displaying effect size (reoperation rates represented by squares, proportional in size to sample size of the study) and confidence intervals (horizontal lines) of reoperation by study and by chondral defects. (ES, Effect Size; CI, Confidence Interval; I-squared, percentage of variation due to heterogeneity rather than chance.)

rather than chance) is 98%. The large degree of statistical heterogeneity in a relatively small group of studies indicates underlying differences among the studies. Heterogeneity represents an opportunity to explore.

Table 2 illustrates additional data regarding the included studies. Four included patients with chondral defects. Stratifying these (Fig 2) reveals reoperation rates among studies excluding chondral defects ranges from 1% to 11% (I-squared, 53%) versus rates ranging from 19% to 53% (I-squared, 86%) among studies including chondral defects. Accounting for chondral defects begins to improve the consistency of the outcomes. Methodologic factors may also be examined. Stratifying by type of study (prospective versus retrospective, Fig 3) reveals reoperation rates ranging from 43% to 53% (I-squared, only 7%) in prospective studies versus 1% to 19% (I-squared, 75%) in

Table 2. Inclusion (Yes) or Exclusion (No) of Patients with Chondral Defects, and Type of Study (Prospective Versus Retrospective) Reported in 10 Studies on Knee Arthroscopy

	Chondral Defect	Type of Study
Study 1	No	Retrospective
Study 2	Yes	Prospective
Study 3	No	Retrospective
Study 4	Yes	Prospective
Study 5	Yes	Prospective
Study 6	No	Retrospective
Study 7	No	Retrospective
Study 8	Yes	Retrospective
Study 9	Yes	Prospective
Study 10	No	Retrospective

retrospective analyses. (Perhaps patients requiring reoperation after knee arthroscopy less commonly return to the same surgeon resulting in loss to research follow-up).

Considering both chondral defects and type of study provides additional clarity (Fig 4). All 4 prospective studies include patients with chondral defects, but only 1 of 6 retrospective studies includes patients with chondral defects (another explanation for the higher reoperation rates in the hypothetical prospective studies). Visual inspection of the forest plot in Figure 4 indicates that Study 8 (the sole retrospective study including chondral defects) is primarily responsible for the greater heterogeneity among retrospective studies, because the Study 8 reoperation rate of 19% deviates substantially from the rates reported in the other retrospective studies.

High-value Conclusions

As above, pooling low-evidence, high-risk-of-bias data from heterogeneous studies results in *misinterpretation*. We assert that the conclusion, “Knee arthroscopy has a 26% reoperation rate” is misconstrued, and the conclusion, “Future research is required” is pedestrian.

Exploration of heterogeneity leads to more meaningful, high-value conclusions: The (hypothetical) literature on knee arthroscopy is limited to Level IV evidence. The reported rates of reoperation are highly

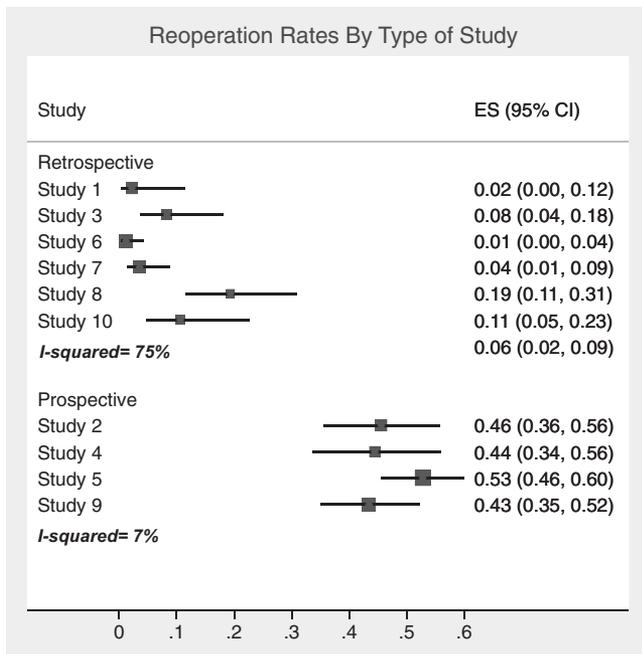


Fig 3. Forest plot displaying effect size (reoperation rates represented by squares proportional in size to sample size of the study) and confidence intervals (horizontal lines) of reoperation by study and by type of study (ES, Effect Size; CI, Confidence Interval; I-squared, percentage of variation due to heterogeneity rather than chance.)

inconsistent and range from 1% to 53%. Accounting for inclusion of patients with chondral defects at the time of surgery, and prospective study design improves

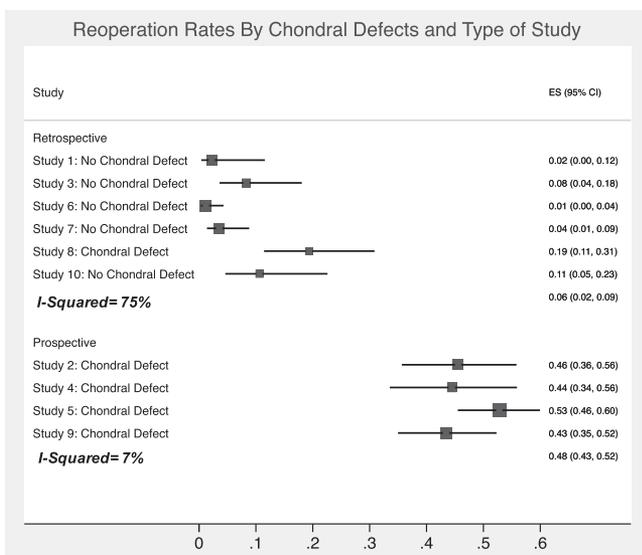


Fig 4. Forest plot displaying effect size (reoperation rates represented by squares, proportional in size to sample size of the study) and confidence intervals (horizontal lines) of reoperation by study and by chondral defects and by type of study. (ES, Effect Size; CI, Confidence Interval; I-squared, percentage of variation due to heterogeneity rather than chance.)

the consistency of reported reoperation rates. Future retrospective case series are not recommended; this type of study contributes to inconsistency in the literature. Future study should be prospective and should include explicit eligibility criteria with regard to the presence or absence of chondral defects.

Heterogeneity is an Opportunity to Explore

We hope that someday, arthroscopic and related surgical literature will be replete with high-quality, low risk of bias, homogeneous studies, and pooling study data on most any topic will well-inform patient care. Until then, quantitative synthesis of heterogeneous studies with a low level of evidence and a high risk of bias must be avoided in favor of subjective synthesis combined with an exploratory embrace of heterogeneity to determine strengths and deficiencies of current literature and unambiguously guide future research.

Mark P. Cote, P.T., D.P.T., M.S.C.T.R.
Associate Editor
James H. Lubowitz, M.D.
Editor-in-Chief
Michael J. Rossi, M.D., M.S.
Assistant Editor-in-Chief
Jefferson C. Brand, M.D.
Assistant Editor-in-Chief

References

- Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485-514.
- Cote MP, Apostolakos JM, Voss A, DiVenere J, Arciero RA, Mazzocca AD. A systematic review of meta-analyses published in *Arthroscopy: The Journal of Arthroscopic and Related Surgery*. *Arthroscopy* 2016;32:528-537.
- Harris JD, Brand JC, Cote MP, Faucett SC, Dhawan A. Research pearls: The significance of statistics and perils of pooling. Part 1: Clinical versus statistical significance. *Arthroscopy* 2017;33:1102-1112.
- Cote MP. Editorial commentary: Leveraging the literature: What can a half-century's worth of studies tell us about recurrent instability following primary shoulder dislocation? *Arthroscopy* 2017;33:1882-1883.
- Friedman PK. Desk reject. Available at, <https://savageminds.org/2012/11/08/desk-reject/>. Accessed October 9, 2018.
- Lubowitz JH, Brand JC, Rossi MJ. Where do arthroscopy editors stand on systematic reviews of the literature? *Arthroscopy* 2017;33:1753-1754.
- Dhawan A, Brand JC, Provencher MT, Rossi MJ, Lubowitz JH. Research pearls: The significance of statistics and perils of pooling. *Arthroscopy* 2017;33:1099-1101.
- Lubowitz JH, Brand JC, Provencher MT, Rossi MJ. Systematic reviews keep arthroscopy up to date. *Arthroscopy* 2016;32:237.