



Published in final edited form as:

Plast Reconstr Surg. 2011 March ; 127(3): 1029–1044. doi:10.1097/PRS.0b013e3182043630.

The Effect of Study Design Biases on the Diagnostic Accuracy of Magnetic Resonance Imaging to Detect Silicone Breast Implant Ruptures: A Meta-Analysis

Jae W. Song, MD¹, Hyungjin Myra Kim, ScD², Lillian T. Bellfi, BA³, and Kevin C. Chung, MD, MS⁴

¹ Research Fellow, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System, Ann Arbor, Michigan

² Associate Research Scientist, Center for Statistical Consultation and Research, The University of Michigan, Ann Arbor, Michigan

³ Research Assistant, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System, Ann Arbor, Michigan

⁴ Professor, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System, Ann Arbor, Michigan

Abstract

Background—All silicone breast implant recipients are recommended by the US Food and Drug Administration to undergo serial screening to detect implant rupture with magnetic resonance imaging (MRI). We performed a systematic review of the literature to assess the quality of diagnostic accuracy studies utilizing MRI or ultrasound to detect silicone breast implant rupture and conducted a meta-analysis to examine the effect of study design biases on the estimation of MRI diagnostic accuracy measures.

Method—Studies investigating the diagnostic accuracy of MRI and ultrasound in evaluating ruptured silicone breast implants were identified using MEDLINE, EMBASE, ISI Web of Science, and Cochrane library databases. Two reviewers independently screened potential studies for inclusion and extracted data. Study design biases were assessed using the QUADAS tool and the STARDS checklist. Meta-analyses estimated the influence of biases on diagnostic odds ratios.

Results—Among 1175 identified articles, 21 met the inclusion criteria. Most studies using MRI (n= 10 of 16) and ultrasound (n=10 of 13) examined symptomatic subjects. Meta-analyses revealed that MRI studies evaluating symptomatic subjects had 14-fold higher diagnostic accuracy estimates compared to studies using an asymptomatic sample (RDOR 13.8; 95% CI 1.83–104.6) and 2-fold higher diagnostic accuracy estimates compared to studies using a screening sample (RDOR 1.89; 95% CI 0.05–75.7).

Conclusion—Many of the published studies utilizing MRI or ultrasound to detect silicone breast implant rupture are flawed with methodological biases. These methodological shortcomings may result in overestimated MRI diagnostic accuracy measures and should be interpreted with caution when applying the data to a screening population.

Corresponding author and reprint requests sent to: Kevin C. Chung, MD, MS, Section of Plastic Surgery, The University of Michigan Health System, 1500 E. Medical Center Drive, 2130 Taubman Center, SPC 5340, Ann Arbor, MI 48109-5340, Phone: 734-936-5885, Fax: 734-763-5354, kechung@med.umich.edu.

Disclosures: None of the authors has a financial interest in any of the products, devices, or drugs mentioned in this manuscript.

Keywords

silicone breast implant; magnetic resonance imaging; screening; diagnostic accuracy study; meta-analysis

Substantial adverse public and media attention was directed toward the use of silicone gel breast implants in the early 1990s when concerns linking implants with connective tissue disorders,¹ cancer,² and neurologic sequelae³ resulted in a 15 year ban by the US Food and Drug Administration (F.D.A.) on April 16, 1992.⁴ During this ban, many rigorous clinical and epidemiological studies were conducted but failed to show compelling associations between implant rupture and autoimmune diseases.⁵ The lack of evidence persuaded the F.D.A. to re-approve the use of silicone breast implants in 2006 with the recommendation to screen all silicone breast implant recipients with magnetic resonance imaging (MRI) 3 years after implantation and biannually thereafter.⁶ These recommendations affect a large number of women undergoing breast augmentation. An estimated 1 million women underwent augmentation with silicone gel implants between 1963 and 1988 prior to the ban,⁷ and from 2008–2009, 265,074 women used silicone implants, which comprise nearly 50% of all women undergoing breast augmentation in the US.⁸

With a growing number of women being implanted with silicone gel implants,⁸ serial MRI screening throughout the implant's lifetime raises concerns regarding MRI as an optimal screening modality. Important criteria to be considered when choosing an optimal screening test include characteristics of the condition of interest (i.e., prevalence of the detectable condition) and test characteristics (i.e., sensitivity and specificity). The F.D.A.'s concern for screening is to detect silent ruptures, a rupture in a clinically asymptomatic patient.⁹ The range of rupture characteristics extend from large visible tears or focal ruptures through pin-sized holes to gel bleeds, which are microscopic silicone leaks through an otherwise intact implant envelope.¹⁰ Microscopic leaks may be caused by degenerating silicone elastomers and may evolve into larger leaks with migrating free silicone. How sensitive MRIs are in detecting gel bleeds, compared to intracapsular and extracapsular ruptures manifesting with clinical symptoms, remains unknown. In addition, the prevalence of gel bleeds, given the subclinical presentation, is difficult to assess. The most recent study using MRI to screen for ruptures reports a prevalence of 8% among asymptomatic women for implants of median age 11 years.¹¹ This study was restricted to implants 10–13 years of age and specific manufacturing styles of Inamed silicone breast implants (Inamed Corp., Santa Barbara, CA) and did not verify subjects with explantation. Nonetheless, the low prevalence of a silent rupture questions the utility of MRI in a screening population. Moreover, the manufacturing of silicone gel implants has been improving with more durable shells and more cohesive gel materials,¹² which will potentially further decrease the prevalence of rupture.

Screening test characteristics are fundamental in choosing the optimal test. Several authors have shown inaccurate diagnostic accuracy measures in studies flawed with study design biases, such as spectrum bias or partial verification bias.^{13, 14} Spectrum bias occurs when a study sample is comprised of a clinically restricted spectrum of patients. For example, symptomatic subjects are more likely to have a ruptured implant, resulting in higher sensitivity and specificity estimates. Furthermore, studies evaluating the accuracy of a screening test are particularly subject to partial verification bias. This bias occurs when not all subjects who are screened undergo the reference test, in particular, those with a negative screening test result. This bias can markedly reduce the apparent specificity and increase the sensitivity of the test.¹³

We examined the study quality of diagnostic accuracy studies using MRI to detect silicone breast implant rupture, given the current controversy about the MRI screening recommendation by the F.D.A.¹⁵ All silicone breast implant recipients are recommended by the F.D.A. to undergo serial MRI screening to detect implant rupture despite a lack of evidence showing serious consequences from a ruptured implant. Because some physicians believe that ultrasound is a more acceptable screening modality,¹⁶ we were also interested in the quality of diagnostic accuracy studies using US to detect implant rupture. We performed a systematic review of this literature and identified the most common biases and examined reporting quality using validated checklists. Next, we used this identified literature to perform a meta-analysis,^{17, 18} a statistical method commonly used to combine results from multiple studies. Figure 1 is a schematic illustrating the steps for a meta-analysis. The meta-analysis was conducted to also quantify the effect of biases on the reported MRI diagnostic accuracy measures.

METHODS

Search Strategy

Four databases (MEDLINE, EMBASE, ISI Web of Science, and Cochrane) were searched using patient (breast and silicone), intervention, outcome, and diagnostic accuracy entry terms up to April 2010 (Table 1). Two searches from each database were performed using first, a combination of patient, intervention and outcome terms and second, a combination of patient, intervention, outcome and diagnostic accuracy terms.¹⁹ There were no language restrictions but searches were limited to human studies. The standards of quality for reporting meta-analyses of observational studies were reviewed during the planning, conducting, and reporting of this meta-analysis.²⁰

Eligibility Criteria

Study inclusion criteria are shown in Table 2. First, titles and abstracts were screened independently by two reviewers. Selected articles from this screen underwent subsequent independent full-text reviews. The references of all articles selected for full text review were manually reviewed. Foreign language articles were translated and reviewed for inclusion.

Quality Assessment and Data Extraction

Following the Cochrane Collaboration recommendations,²¹ methodological quality was assessed independently by two reviewers using the Quality of Diagnostic Accuracy Studies (QUADAS) tool.^{22, 23} Completeness of reporting was assessed by the Standards for Reporting of Diagnostic Accuracy Studies (STARD) checklist.^{24, 25} The QUADAS tool is a 14-point assessment instrument developed for systematic reviews of diagnostic accuracy studies. Studies were checked as yes, no, or unclear. We defined a representative spectrum of patients to be both asymptomatic and symptomatic, to reflect a screening population, given the current context of the MRI as a screening tool. Authors were contacted when information was inadequate in the report. The STARD statement is a 25-item checklist aimed to improve the completeness of reporting diagnostic accuracy studies. Studies were checked as yes, no, incomplete or unclear. Inter-rater agreement was calculated by the kappa-statistic. Discrepancies were resolved by consensus.

Data extraction also included type of study (prospective or retrospective) and sample and implant characteristics. Numbers of true-positives, false-negatives, false-positives, and true-negatives were extracted, and the sensitivity (number of true positives divided by the number of true positives and false negatives) and specificity (number of true negatives divided by the number of true negatives and false positives) were calculated by the authors

to confirm the reported values. Data extraction was performed independently by 2 reviewers. Discrepancies were resolved by consensus.

Statistical Analysis

Pooled test sensitivity and specificity values were obtained using multi-level mixed-effects logistic regression models. A forest plot was generated to graphically represent heterogeneity across individual studies. A forest plot is a pictorial representation of each study's sensitivity and specificity bounded by the 95% confidence intervals. It is useful for visualizing how similar or dissimilar the reported measures are amongst studies and to identify studies with outlying values of the measures.¹⁸

We next assessed heterogeneity in the reported sensitivity and specificity across the studies using the Q and I² statistics²⁶ to determine whether the measures from the different studies are similar enough to be combined into a pooled summary measure. A small p-value (< 0.05) from the Q statistic suggests statistically significant heterogeneity among studies. The I² statistic quantifies the amount of heterogeneity among studies, and by convention, low, moderate, and high values of heterogeneity are indicated by I² values of 25%, 50%, and 75%, respectively.²⁶ If substantial statistical heterogeneity among studies emerges, sources of heterogeneity should be identified.²⁶ The variations in the reported sensitivities and specificities may be due to differing sample characteristics or differences in the way each study was conducted. To identify potential sources of heterogeneity, we performed subgroup analyses to evaluate if any of the various sample and study characteristics affected the sensitivities and specificities.

The effect of study characteristics was further examined using diagnostic odds ratio (DOR) as previously described by Lijmer *et al.*¹⁴ The DOR is useful because it is a single summary statistic for diagnostic accuracy incorporating both sensitivity and specificity (Figure 2A). It is the odds of a positive test in a diseased person relative to the odds of a positive test in a non-diseased person, in which a large DOR indicates a high sensitivity and specificity.²⁷ Briefly, the effect of study characteristics on diagnostic accuracy can be assessed using a regression model with the logarithm of the DOR as the dependent variable. From the regression model, we can obtain an estimate for relative DOR (RDOR), which is interpreted as a ratio of the DORs with versus without the study characteristic (Figure 2B). Thus, a RDOR of 1 indicates that the study characteristic does not influence the overall DOR, whereas a RDOR greater than 1 indicates that studies with the characteristic yield larger estimates of DOR than studies without the characteristic.²⁷

Publication bias was examined by construction of a funnel plot, and statistical significance of asymmetry was assessed by the Egger's test.²⁸ Stata 11.1 (StataCorp, College Station, TX) was used for statistical analyses.

RESULTS

Search Strategy and Study Selection

The initial search dated up to April 2010 using 4 databases (MEDLINE, EMBASE, ISI Web of Science, and Cochrane) identified 1175 articles (Figure 3). A total of 311 articles were duplicates. Of the remaining 864 articles, 768 were excluded upon review of the title or abstract based on inclusion and exclusion criteria, leaving 96 articles. Forty-two additional articles were included after a manual bibliography search from the 96 included articles, totaling 138 articles for full-text review. After full-text review, 117 articles were excluded leaving 21 articles that evaluated the diagnostic accuracy of the ultrasound and/or MRI for silicone breast implant rupture. Reasons for further exclusion are described in Figure 3.

Eight studies examined both ultrasound and MRI,²⁹⁻³⁶ 5 evaluated US only,³⁷⁻⁴¹ and 8 evaluated MRI only.⁴²⁻⁴⁹

MRI and Ultrasound Study Characteristics

All 21 studies were diagnostic cohort studies.²⁹⁻³⁶ 42-49 Table 3 summarizes characteristics of the included studies. Among the MRI studies, 2 studies were duplicated: Scaranelo *et al.*³⁵ examined breast and body coil MRIs separately, and Gorczyca *et al.*⁴⁴ compared Fast Spin-Echo and 3-Point Dixon MRIs separately. In total, 1,098 silicone breast implants in 615 women were examined with MRI, and 1,007 silicone breast implants in 577 women were examined with ultrasound.

Quality and Reporting Assessment

We assessed 21 studies using the QUADAS instrument and STARD checklist. Inter-rater agreement for the total QUADAS and STARD assessments was good for the 21 studies (76.8%, kappa statistic=0.58). More than 50% of the 16 MRI studies used a sample that was not representative of a screening sample (10 studies examined only symptomatic patients²⁹⁻³¹, 33, 34, 36, 44, 45, 47, 48, 2 studies examined only asymptomatic patients³⁵, 42), 9 did not explain reasons for individuals withdrawing from the study²⁹⁻³², 36, 42, 44-46, and 11 did not report uninterpretable results.²⁹, 30, 32-34, 43-46, 48, 49 The reference test diagnostic criteria were not specified in 7 studies (43.8%)²⁹, 34, 35, 42-44, 48, and 7 studies (43.8%)³¹, 34, 42-44, 48, 49 had partial verification bias (Figure 4A-B). More than 50% of the 13 ultrasound studies did not use a screening sample (10 studies examined only symptomatic patients²⁹⁻³¹, 33, 34, 36-39, 41, 1 study examined only asymptomatic patients³⁵), and 11 did not explain reasons for individuals withdrawing from the study²⁹⁻³³, 36-41 (Figure 5A-B).

Using the STARD checklist, we identified important specifications for diagnostic accuracy studies that were inconsistently addressed across studies. Only 4 of 16 MRI²⁹, 35, 36, 43 and 5 of 13 ultrasound²⁹, 35, 36, 40, 41 studies reported in their title or abstract that the reported sensitivity and specificity were applicable to the specified studied sample. In reporting the test results, only 31.3% of MRI (5 of 16)³¹, 35, 46, 47, 49 and 15.4% (2 of 13)³¹, 35 of ultrasound studies reported a time interval from the index test to explantation. This time interval ranged from 1 week³⁵ to 297 days⁴⁷ with a median of 3 months among the 5 MRI studies. This information is important especially for screening tests, given the possibility that rupture may have occurred during the interim period before explantation; this context is also known as disease progression bias. In addition, a few MRI (3 of 16)³³, 42, 47 and ultrasound (2 of 13)³³, 36 studies discussed the possibility of rupture at the time of explantation, an important detail given a surgical reference test.

Gel bleeds were inconsistently addressed across studies. Most studies (10 of 16 MRI; 7 of 13 ultrasound) did not address gel bleeds or did not include them in calculating sensitivity and specificity (Table 3). Five MRI³⁰, 32, 34, 35, 46 and 5 ultrasound studies³⁰, 32, 34, 35, 38 considered gel bleeds as not ruptured, and 1 MRI⁴⁷ and 1 ultrasound⁴¹ study considered gel bleeds as ruptured.

Because observer variability between radiologists can be particularly problematic with imaging tests, it is important to report estimates of test reproducibility. Although 87.5% of MRI (14 of 16) and 76.9% of ultrasound (10 of 13) studies reported the number of radiologists reading the films, only 7 MRI³¹, 33, 44, 45, 47-49 and 5³¹, 33, 37, 38, 40 ultrasound studies had 2 or more radiologists. Furthermore, only 4 MRI³³, 44, 48, 49 and 2 ultrasound³⁴, 37 studies discussed inter-observer agreement. Less than half of the studies

discussed indeterminate or inconclusive findings (5 MRI^{31, 32, 36, 44, 47} and 4 ultrasound studies^{31, 32, 36, 37, 40}).

Pooled Estimates and Heterogeneity

Forest plots for the 18 MRI studies are illustrated in Figure 6. Significant heterogeneity was present across studies for sensitivity and specificity (sensitivity, Q-statistic $p < 0.01$; $I^2 = 64.7$; specificity, Q-statistic $p < 0.01$; $I^2 = 84.9$). The pooled sensitivity and specificity for MRI were 87.0% (95% CI 81–91%) and 89.9% (95% CI 82–94%), respectively. Though not shown as a forest plot, the sensitivity for the 13 ultrasound studies ranged from 30.0%³² to 77.0%³², with significant heterogeneity across studies (Q-statistic 28.0, $p = 0.01$; $I^2 = 57.2$; 95% CI 31–84). The reported specificity was also highly variable, ranging from 55.0%⁴⁰ to 92.0%^{29, 39} with significant heterogeneity (Q-statistic 57.0, $p < 0.01$; $I^2 = 78.9$; 95% CI 68–90). The pooled sensitivity and specificity were 60.8% (95% CI 53–68%) and 76.3% (95% CI 68–83%), respectively.

Subgroup Analyses

Potential sources of heterogeneity and tests of heterogeneity are summarized in Table 4. We examined subgroups of study design biases, methodological characteristics, and test execution characteristics. Not included in the table are subgroups in which only 1 study was categorized to a group because statistical tests could not be done. For example, among the ultrasound studies, only 1 study used an asymptomatic sample,³⁵ 1 study was retrospectively conducted,³² and 1 study used a consecutive recruitment method.³⁷ The sensitivity and specificity in MRI studies that used a symptomatic sample were higher (sensitivity 88%; specificity 94%) compared to studies using an asymptomatic sample (sensitivity 76%; specificity 68%), although the differences were not statistically significant. Of note, ultrasound studies without partial verification bias had significantly higher specificity (81%) than studies with partial verification bias (67%, $p < 0.001$).

Influence of Biases on the Diagnostic Accuracy of MRI Studies

Regression analyses to quantify the influence of biases on diagnostic accuracy are illustrated in Figure 7. MRI studies using symptomatic samples had a DOR that was nearly 14-fold greater compared to the DOR of studies with asymptomatic samples (RDOR 13.8). MRI studies that used symptomatic samples had a DOR that was 1.89 times greater than studies that used a screening sample (i.e., symptomatic and asymptomatic samples). Studies that did not evaluate the condition of the explanted silicone implants in all subjects who underwent MRI evaluations (i.e. studies with partial verification bias) had a DOR 2.49 times greater than studies that surgically evaluated all implants in study subjects who underwent MRI evaluations.

Publication Bias

Two funnel plots were constructed to assess publication bias among MRI and ultrasound studies. Significant publication bias was detected in MRI studies, indicated by an asymmetric distribution of studies (Figure 8A, $p = 0.01$). No significant publication bias was detected among ultrasound studies (Figure 8B, $p = 0.87$).

DISCUSSION

Our review reveals several new insights about the current literature using MRIs or ultrasounds to detect silicone gel implant ruptures. Although the pooled summary measures across the studies indicate relatively high accuracy of MRI in detecting breast implant rupture with a pooled sensitivity of 87% and a specificity of 89.9%, the majority of the

current literature examined only symptomatic patients. This leads to a higher prevalence of silicone breast implant rupture and higher diagnostic accuracy estimates. We found the DOR, a measure of overall diagnostic test performance, of MRI to be 14-fold greater in symptomatic samples than in asymptomatic samples and 2-fold greater in symptomatic samples than in screening samples. This was shown in the subgroup analyses with higher sensitivity and specificity of the MRI in studies examining symptomatic samples than in studies using asymptomatic and screening samples (Table 4). These findings have widespread health policy implications given the F.D.A. recommendations to repeatedly screen silicone breast implant recipients with serial MRI exams.

Instituting a screening program requires careful consideration of several issues. First, the disease should have serious consequences.⁵⁰ Currently, the morbidity associated with silicone breast implant rupture remains unclear and is still under study. Second, the disease must have a pre-clinical yet detectable stage.⁵⁰ For silicone breast implant ruptures, this stage may be considered gel bleeds. Our results show a lack of consistency in addressing gel bleeds (Table 3). Third, a high prevalence of the pre-clinical stage among the screening population is optimal for a successful screening program. To date, there is a lack of evidence about the prevalence of subclinical gel bleeds. In addition, many studies report the mean age of implant at time of rupture to be greater than 10 years,^{29-31, 34-36, 47} suggesting that perhaps this group may consist of the high-risk sample that should garner directed attention. In light of this and the possibility of very low prevalence, adherence to the F.D.A. recommendation to screen with MRI at least 4 times within the first 10 years of silicone breast implantation will be costly and may potentially result in over-detection and over-treatment of a questionable non-life-threatening condition.

In addition, important screening test characteristics to consider include the sensitivity and specificity of the screening modality. Our results reveal many methodological flaws in the current literature, which may result in higher MRI sensitivity and specificity estimates (Figure 7). We showed that most of the included MRI studies reported diagnostic accuracy measures on symptomatic samples, which had a DOR that was nearly 14-fold and 2-fold greater than the DOR of detecting silicone breast implant ruptures in asymptomatic samples and screening samples, respectively. Thus, although MRI's diagnostic performance in detecting silicone breast implant ruptures in a symptomatic sample may be quite good, we find that the MRI's accuracy is magnitudes lower in detecting rupture in asymptomatic and screening samples.

These compelling results are noteworthy given the frequency of the F.D.A. recommendations for serial MRI exams as a screening test to detect silicone breast implant ruptures. As a screening program, these recommendations have been received with widespread controversy. There is a lack of high level evidence establishing serious health consequences from a ruptured silicone breast implant, and yet adherence to these recommendations result in substantial cost and use of resources. In particular, the benefits of screening within the first 10 years are unclear, and the effectiveness of such a screening program warrants further investigation. Moreover, screening programs should take into account patient preferences.⁵¹ Patient acceptability influences adherence to recommendations and have been important topics for other screening programs such as colorectal cancer^{52, 53} and prostate cancer screening.⁵⁴

The main strength of this review is our rigorous compliance with the recommended methods for carrying out and reporting a systematic review, specifically for diagnostic accuracy studies. Systematic reviews of diagnostic accuracy studies differ from intervention studies in 3 ways: the inclusion of diagnostic accuracy search terms,¹⁹ assessment of study quality and

completeness of reporting,²⁰ and meta-analysis methods. Additionally, we attempted to identify sources of heterogeneity across these studies.⁵⁵

There are several limitations to our study. First, the small number of studies may explain the lack of statistical significance of our results. The wide confidence intervals of the RDORs indicate low statistical power. Eight studies were excluded because they lacked sufficient data to construct 2×2 tables, which are essential in obtaining data to conduct meta-analyses. Efforts to contact authors were largely unsuccessful. This limitation emphasizes the importance of reporting diagnostic accuracy studies using the STARD checklist.²⁴ Another limitation is the possibility of publication bias among MRI studies, despite an extensive search through 4 databases without any language restrictions. A possible explanation for the asymmetric funnel plot may be the small number of available included studies.⁵⁶

CONCLUSION

In summary, many of the MRI and ultrasound diagnostic accuracy studies examining silicone breast implant ruptures are methodologically flawed, particularly because of the use of only symptomatic samples. The reported MRI sensitivity and specificity estimates may be high if applied to asymptomatic or screening samples. Given the current policy recommendations to screen asymptomatic women, further research is needed to investigate and identify long-term disease consequences of rupture, the effectiveness of MRI or other more optimal screening tests in an appropriate sample, the cost of screening strategies, and patient preferences for screening.

Acknowledgments

The authors thank David B. Lumenta, MD for his help with foreign language article translations. This project was supported in part by a Ruth L. Kirschstein National Research Service Awards for Individual Postdoctoral Fellows (1F32AR058105-01A1) (to Dr. Jae W. Song) and by a Midcareer Investigator Award in Patient-Oriented Research (K24 AR053120) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (to Dr. Kevin C. Chung).

References

1. Spiera H. Scleroderma after silicone augmentation mammoplasty. *JAMA*. 1988; 260:236–8. [PubMed: 3385898]
2. Silverstein MJ, Gierson ED, Gamagami P, et al. Breast cancer diagnosis and prognosis in women augmented with silicone gel-filled implants. *Cancer*. 1990; 66:97–101. [PubMed: 2162244]
3. Sanger JR, Matloub HS, Yousif NJ, et al. Silicone gel infiltration of a peripheral nerve and constrictive neuropathy following rupture of a breast prosthesis. *Plast Reconstr Surg*. 1992; 89:949–952. [PubMed: 1313982]
4. Hilts, PJ. FDA restricts use of implants pending studies. *New York Times Online*; [Accessed May 10, 2010]. Available at: <http://www.nytimes.com/1992/01/07/science/fda-seeks-halt-in-breast-implants-made-of-silicone.html?scp=1&sq=F.D.A.%20Seeks%20Halt%20in%20Breast%20Implants%20Made%20of%20Silicone&st=cse&pagewanted=2>
5. McLaughlin JK, Lipworth L, Murphy DK, et al. The safety of silicone gel-filled breast implants: A review of the epidemiologic evidence. *Ann Plast Surg*. 2007; 59:569–580. [PubMed: 17992155]
6. U.S. Food and Drug Administration. 2006 guidance for industry and FDA on saline, silicone gel, and alternative breast implants; 8.5 safety assessment – rupture. [Accessed May 3, 2010]. Available at: <http://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm133361.htm#85>
7. Terry MB, Skovron ML, Garbers S, et al. The estimated frequency of cosmetic breast augmentation among US women, 1963 through 1988. *Am J Public Health*. 1995; 85:1122–1124. [PubMed: 7625509]

8. American Society of Plastic Surgeons. Report of the 2009 statistics: National clearinghouse of plastic surgery statistics. 2010 [Accessed May 3, 2010]. Available at: <http://www.plasticsurgery.org/Media/Statistics.html>
9. Silicone gel-filled breast implants approved. *FDA Consum.* 2007; 41:8–9.
10. Brown SL, Silverman BG, Berg WA. Rupture of silicone-gel breast implants: Causes, sequelae, and diagnosis. *Lancet.* 1997; 350:1531–1537. [PubMed: 9388410]
11. Heden P, Nava MB, Van Tetering JPB, et al. Prevalence of rupture in Inamed silicone breast implants. *Plast Reconstr Surg.* 2006; 118:303–8. [PubMed: 16874191]
12. Adams, WP.; Potter, JK. Breast implants: Materials and manufacturing past, present, and future. In: Spear, SL., editor. *Surgery of the Breast: Principles and Art.* Vol. 1. Philadelphia: Lippincott Williams & Wilkins; 2006. p. 424-437.
13. Cronin AM, Vickers AJ. Statistical methods to correct for verification bias in diagnostic studies are inadequate when there are few false negatives: A simulation study. *BMC Med Res Methodol.* 2008; 8:75. [PubMed: 19014457]
14. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA.* 1999; 282:1061–1066. [PubMed: 10493205]
15. Singer, N. Implants are back, and so is debate. *New York Times Online*; [Accessed May 3, 2010]. Available at: http://www.nytimes.com/2007/05/24/fashion/24skin.html?_r=1
16. Chung KC, Greenfield ML, Walters M. Decision-analysis methodology in the work-up of women with suspected silicone breast implant rupture. *Plast Reconstr Surg.* 1998; 102:689–695. [PubMed: 9727432]
17. Chung KC, Burns PB, Kim HM. A practical guide to meta-analysis. *J Hand Surg Am.* 2006; 31:1671–1678. [PubMed: 17145390]
18. Borenstein, M.; Hedges, LV.; Higgins, JPT., et al. *Introduction to Meta-Analysis.* United Kingdom: John Wiley & Sons, Ltd; 2009. p. 1-406.
19. Deeks, JJ. Systematic reviews of evaluations of diagnostic and screening tests. In: Egger, M.; Smith, GD.; Altman, DG., editors. *Systematic Reviews in Health Care: Meta-Analysis in Context.* 2. London: BMJ Publishing Group; 2001. p. 248-282.
20. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA.* 2000; 283:2008–2012. [PubMed: 10789670]
21. Reitsma, H.; Rutjes, A.; Whiting, P., et al. Chapter 9: Assessing methodological quality. In: Deeks, J.; Bossuyt, PM.; Gatsonis, C., editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*, Vol. Version 1.0.0. Oxford, UK: Cochrane Collaboration; 2009. p. 1-27. Available from: <http://srdta.cochrane.org/>
22. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003; 3:25. [PubMed: 14606960]
23. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol.* 2006; 6:9. [PubMed: 16519814]
24. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ.* 2003; 326:41–44. [PubMed: 12511463]
25. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Ann Intern Med.* 2003; 138:W1–12. [PubMed: 12513067]
26. Bruce, N.; Pope, D.; Stanistreet, D. *Quantitative Methods for Health Research.* West Sussex, England: John Wiley & Sons, Ltd; 2008. p. 1-529.
27. Glas AS, Lijmer JG, Prins MH, et al. The diagnostic odds ratio: A single indicator of test performance. *J Clin Epidemiol.* 2003; 56:1129–1135. [PubMed: 14615004]
28. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997; 315:629–634. [PubMed: 9310563]

29. Ahn CY, DeBruhl ND, Gorczyca DP, et al. Comparative silicone breast implant evaluation using mammography, sonography, and magnetic resonance imaging: Experience with 59 implants. *Plast Reconstr Surg.* 1994; 94:620–627. [PubMed: 7938284]
30. Beekman WH, Joris Hage J, Taets Van Amerongen AHM, et al. Accuracy of ultrasonography and magnetic resonance imaging in detecting failure of breast implants filled with silicone gel. *Scand J Plast Reconstr Surg Hand Surg.* 1999; 33:415–18. [PubMed: 10614750]
31. Berg WA, Caskey CI, Hamper UM, et al. Single-lumen and double-lumen silicone breast implant integrity - prospective evaluation of MR and US criteria. *Radiology.* 1995; 197:45–52. [PubMed: 7568852]
32. Di Benedetto G, Cecchini S, Grassetti L, et al. Comparative study of breast implant rupture using mammography, sonography, and magnetic resonance imaging: Correlation with surgical findings. *Breast J.* 2008; 14:532–37. [PubMed: 19054000]
33. Everson LI, Parantainen H, Detlie T, et al. Diagnosis of breast implant rupture - imaging findings and relative efficacies of imaging techniques. *Am J Roentgenol.* 1994; 163:57–60. [PubMed: 8010248]
34. Ikeda DM, Borofsky HB, Herfkens RJ, et al. Silicone breast implant rupture: Pitfalls of magnetic resonance imaging and relative efficacies of magnetic resonance, mammography, and ultrasound. *Plast Reconstr Surg.* 1999; 104:2054–62. [PubMed: 11149768]
35. Scaranelo AM, Marques AF, Smialowski EB, et al. Evaluation of the rupture of silicone breast implants by mammography, ultrasonography and magnetic resonance imaging in asymptomatic patients: Correlation with surgical findings. *Sao Paulo Med J.* 2004; 122:41–47. [PubMed: 15257358]
36. Reynolds HE, Buckwalter KA, Jackson VP, et al. Comparison of mammography, sonography, and magnetic resonance imaging in the detection of silicone-gel breast implant rupture. *Ann Plast Surg.* 1994; 33:247–255. [PubMed: 7985960]
37. Caskey CI, Berg WA, Anderson ND, et al. Breast implant rupture: Diagnosis with US. *Radiology.* 1994; 190:819–23. [PubMed: 8115633]
38. Chung KC, Wilkins EG, Beil RJ Jr, et al. Diagnosis of silicone gel breast implant rupture by ultrasonography. *Plast Reconstr Surg.* 1996; 97:104–9. [PubMed: 8532766]
39. DeBruhl ND, Gorczyca DP, Ahn CY, et al. Silicone breast implants: US evaluation. *Radiology.* 1993; 189:95–98. [PubMed: 8372224]
40. Venta LA, Salomon CG, Flisak ME, et al. Sonographic signs of breast implant rupture. *AJR Am J Roentgenol.* 1996; 166:1413–1419. [PubMed: 8633455]
41. Medot M, Landis GH, McGregor CE, et al. Effects of capsular contracture on ultrasonic screening for silicone gel breast implant rupture. *Ann Plast Surg.* 1997; 39:337–341. [PubMed: 9339274]
42. Collis N, Litherland J, Enion D, et al. Magnetic resonance imaging and explantation investigation of long-term silicone gel implant integrity. *Plast Reconstr Surg.* 2007; 120:1401–06. [PubMed: 17898620]
43. Frankel S, Occhipinti K, Kaufman L, et al. MRI findings in subjects with breast implants. *Plast Reconstr Surg.* 1995; 96:852–9. [PubMed: 7652059]
44. Gorczyca DP, Schneider E, DeBruhl ND, et al. Silicone breast implant rupture: Comparison between three-point dixon and fast spin-echo MR imaging. *AJR Am J Roentgenol.* 1994; 162:305–310. [PubMed: 8310916]
45. Gorczyca DP, Sinha S, Ahn CY, et al. Silicone breast implants *in vivo* - mr imaging. *Radiology.* 1992; 185:407–410. [PubMed: 1410346]
46. Herborn CU, Marincek B, Erfmann D, et al. Breast augmentation and reconstructive surgery: MR imaging of implant rupture and malignancy. *Eur Radiol.* 2002; 12:2198–206. [PubMed: 12195470]
47. Holmich LR, Vejborg I, Conrad C, et al. The diagnosis of breast implant rupture: MRI findings compared with findings at explantation. *Eur J Radiol.* 2005; 53:213–225. [PubMed: 15664285]
48. Monticciolo DL, Nelson RC, Dixon WT, et al. MR detection of leakage from silicone breast implants: Value of a silicone-selective pulse sequence. *AJR Am J Roentgenol.* 1994; 163:51–56. [PubMed: 8010247]

49. Quinn SF, Neubauer NM, Sheley RC, et al. MR imaging of silicone breast implants: Evaluation of prospective and retrospective interpretations and interobserver agreement. *J Magn Reson Imaging*. 1996; 6:213–218. [PubMed: 8851430]
50. Cole P, Morrison AS. Basic issues in population screening for cancer. *J Natl Cancer Inst*. 1980; 64:1263–1272. [PubMed: 6767876]
51. Walsh JM, Terdiman JP. Colorectal cancer screening: Clinical applications. *JAMA*. 2003; 289:1297–1302. [PubMed: 12633192]
52. Dominitz JA, Provenzale D. Patient preferences and quality of life associated with colorectal cancer screening. *Am J Gastroenterol*. 1997; 92:2171–2178. [PubMed: 9399747]
53. Nelson RL, Schwartz A. A survey of individual preference for colorectal cancer screening technique. *BMC Cancer*. 2004; 4:76. [PubMed: 15533242]
54. Flood AB, Wennberg JE, Nease RF Jr, et al. The importance of patient preference in the decision to screen for prostate cancer. prostate patient outcomes research team. *J Gen Intern Med*. 1996; 11:342–349. [PubMed: 8803740]
55. Egger, M.; Smith, GD.; Schneider, M. Systematic reviews of observational studies. In: Egger, M.; Smith, GD.; Altman, DG., editors. *Systematic Reviews in Health Care: Meta-Analysis in Context*. Vol. 2. London: BMJ Publishing Group; 2001. p. 211-227.
56. Song F, Khan KS, Dinnes J, et al. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol*. 2002; 31:88–95. [PubMed: 11914301]

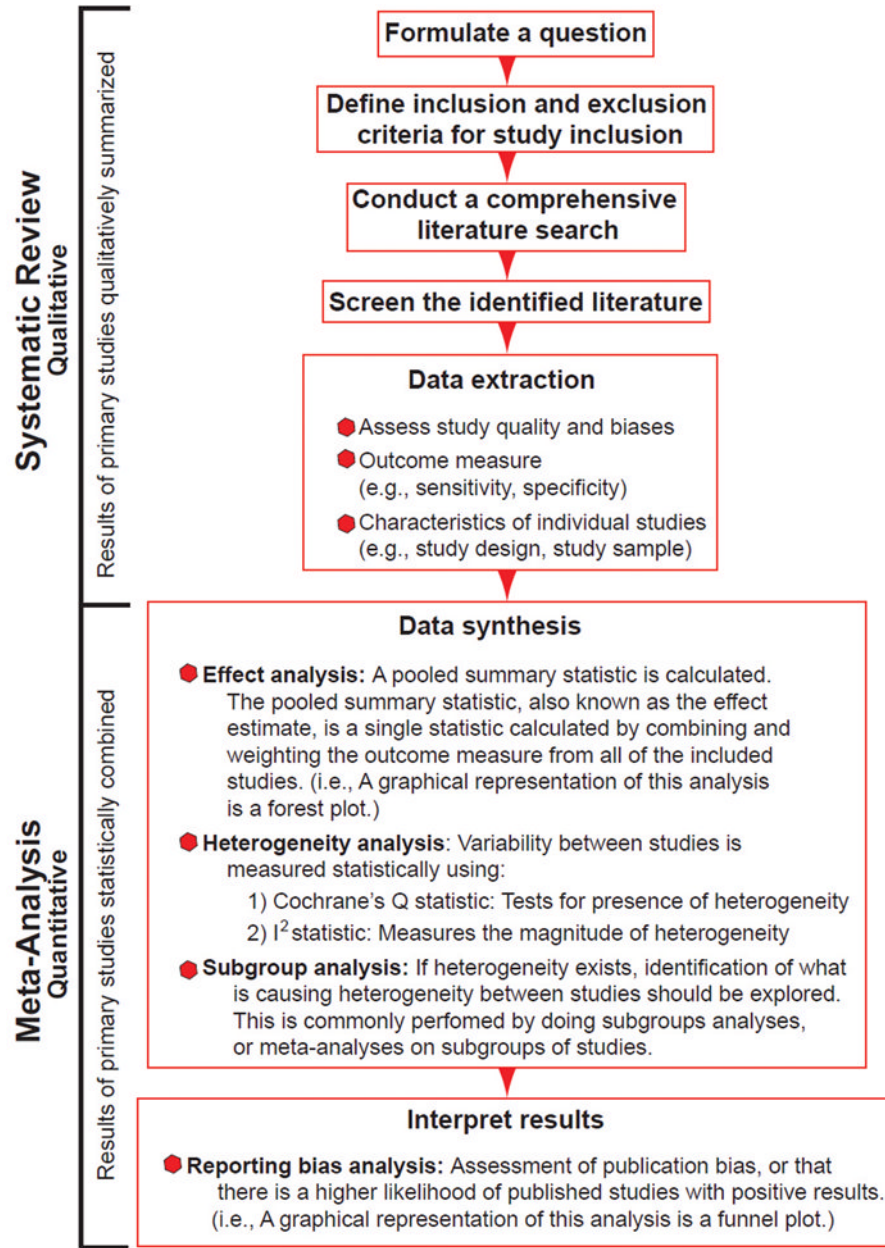


Figure 1.
 Steps to a Meta-Analysis

$$\text{Diagnostic Odds Ratio} = \frac{\frac{\text{Sensitivity}}{1-\text{Sensitivity}}}{\frac{1-\text{Specificity}}{\text{Specificity}}} = \frac{\text{Odds of a positive test | disease}}{\text{Odds of a positive test | no disease}}$$

$$\text{Relative Diagnostic Odds Ratio} = \frac{\text{Diagnostic Odds Ratio (with covariate}^\dagger)}{\text{Diagnostic Odds Ratio (without covariate)}}$$

$$= \frac{\text{Diagnostic Odds Ratio (symptomatic samples)}}{\text{Diagnostic Odds Ratio (asymptomatic samples)}}$$

Figure 2.**Diagnostic Odds Ratio**

(A) The diagnostic odds ratio (DOR) is a measure of the overall accuracy of a positive test and combines a test's sensitivity and specificity. It is interpreted as the odds of a positive test given disease, divided by the odds of a positive test given no disease. A large DOR of a test means the test has a high sensitivity and specificity for detecting a disease.

(B) The relative diagnostic odds ratio (RDOR) is a ratio of 2 diagnostic odds ratios.

[†]Covariates that will be examined are the different types of study or sample characteristics such as sample characteristics (i.e., symptomatic, asymptomatic, screening), partial verification bias, among others (See Figure 6).

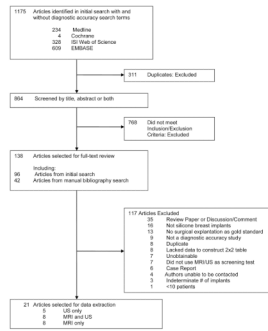


Figure 3.
 Selection of Studies for Meta-Analysis
 A trial flow diagram shows the number of identified, screened, and included studies.

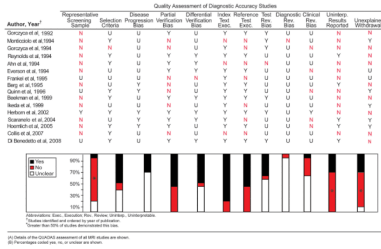


Figure 4. Quality Assessment of Magnetic Resonance Imaging Studies (A) Details of the QUADAS assessment of MRI studies are shown. (B) Percentages coded yes, no, or unclear are shown.

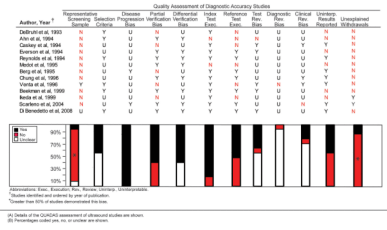
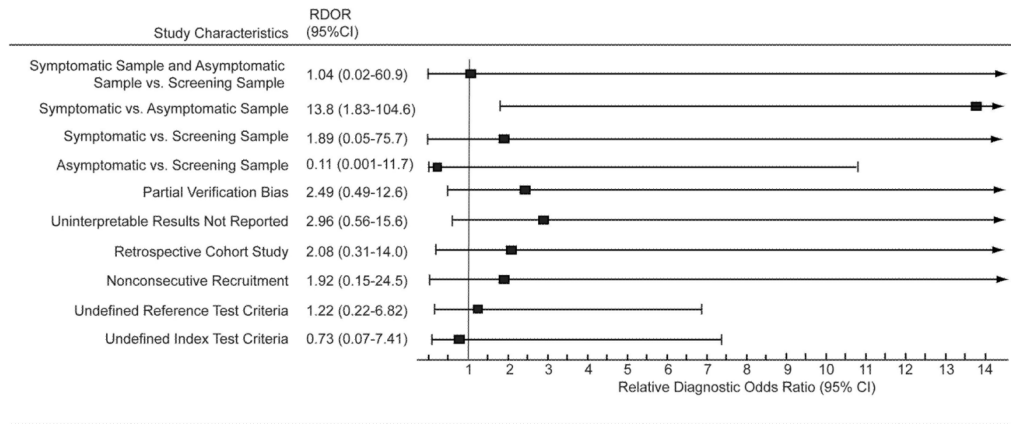


Figure 5. Quality Assessment of Ultrasound Studies (A) Details of the QUADAS assessment of ultrasound studies are shown. (B) Percentages coded yes, no, or unclear are shown.

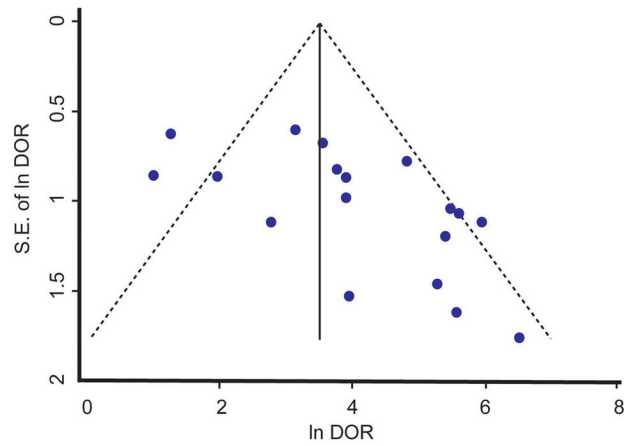


Abbreviations: CI, Confidence Interval; RDOR, Relative Diagnostic Odds Ratio; vs., versus

Figure 7. Relative Diagnostic Odds Ratios of Study Design Characteristics and Biases Examined with Univariate Regression Analysis

In both panels, the black squares indicate the relative diagnostic odds ratios for each individual study. 95% confidence intervals are indicated by the horizontal line.

Funnel Plot of 18 Magnetic Resonance Imaging Studies



Funnel Plot of 13 Ultrasound Studies

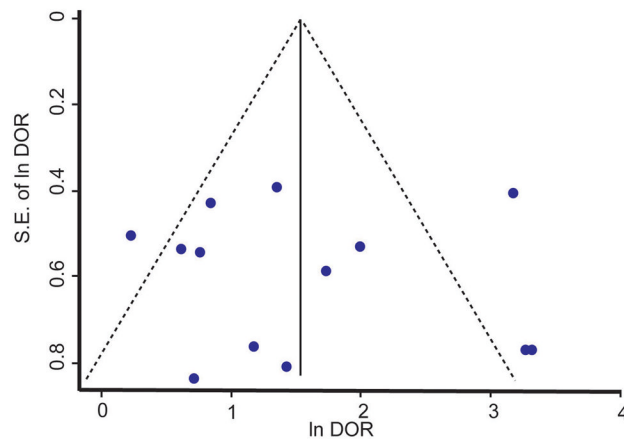


Figure 8.
 Funnel Plots to Assess Publication Bias
 (A) A funnel plot of the 18 MRI studies illustrates an asymmetric distribution of studies suggesting publication bias (Egger’s test, $p=0.01$)
 (B) A funnel plot constructed for the 13 ultrasound studies did not reveal a statistically significant distribution ($p=0.87$).

Table 1

Details of Search Strategy

| | MEDLINE ^a (1950-April 2010) | EMBASE ^b (1974-April 2010) | ISI (1948-April 2010) | COCHRANE (1966-April 2010) |
|------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Patient | | | | |
| Breast | Mammaplasty, breast implants, breast, or breast implantation | Breast, breast endoprosthesis, endoprosthesis, breast prosthesis, breast reconstruction, or breast augmentation | Breast implant, silicone breast implant, silicone gel breast implant, mammaplasty, breast augmentation, breast reconstruction, breast implantation | Mammaplasty, breast implants, breast, or breast implantation |
| Silicone | Silicone gels or silicones | Gel, silicone gel, silicone derivative, or silicone prosthesis | Silicone or silicone gel | Silicone gels or silicones |
| Intervention | Diagnostic tests, routine, ultrasonography, mammary, magnetic resonance imaging, or MRI, breast MR | Echo planar imaging, nuclear magnetic resonance imaging, echomammography, ultrasound, diagnostic imaging, diagnostic test, diagnosis, or radiodiagnosis | Diagnostic test, MRI, breast MR, magnetic resonance imaging, ultrasound, ultrasonography, reference standard, or breast MRI | Diagnostic tests, routine, ultrasonography, mammary, magnetic resonance imaging, MRI, or breast MR |
| Outcome | Prosthesis failure, postoperative complications or implant rupture | Prosthesis failure, postoperative complications, rupture, or diagnosis | Implant rupture, breast implant rupture, silicone gel breast implant failure, gel bleed, prosthesis failure, postoperative complication rupture, silent rupture, implant complication, or extracapsular | Prosthesis failure, postoperative complications, or implant rupture |
| Diagnostic Accuracy | Sensitivity and specificity, mass screening, predictive value of tests, ROC curve, false negative reactions, or false positive reactions | Sensitivity and specificity, risk, prediction and forecasting, false positive result, diagnostic accuracy, roc curve, or mass screening | Sensitivity, specificity, false negative, true positive, predictive value, accuracy, likelihood ratio, screening, or ROC curve | Sensitivity and specificity, mass screening, predictive value of tests, ROC curve, false negative reactions, or false positive reactions |
| Combination of Search Terms | Patient and Intervention and Outcome (and Diagnostic Accuracy) | Patient and Intervention and Outcome (and Diagnostic Accuracy) | Patient and Intervention and Outcome (and Diagnostic Accuracy) | Patient and Intervention and Outcome (and Diagnostic Accuracy) |

^aMedical Subject Headings (MeSH) terms and accompanying entry terms were exploded in search.

^bEntree preferred terminology; all terms exploded in search.

Table 2

Study Inclusion Criteria

| |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (1) Magnetic resonance imaging or ultrasound used to establish the diagnosis of implant rupture |
| (2) Sufficient data reported to compute sensitivity and specificity values |
| (3) Number of patients and implants with silicone breast implants reported (or if a mix of silicone and saline breast implants was used, the number of silicone breast implants was extractable) |
| (4) Explantation as the reference standard |
| (5) At least 10 patients examined |

Table 3

Characteristics of Diagnostic Accuracy Studies to Detect Silicone Breast Implant Rupture

| Source, Year | Location | Recruitment Dates | Recruitment Method | Study Design | Study Sample [§] | Indication for Surgery | Mean Subject age (range), y | Mean Implant age, (range), y | Type of implant [†] | No. of Patients | | No. of Implants | |
|---------------------------------------|--------------|-------------------|--------------------|--------------|---------------------------|--------------------------------------------------|-----------------------------|------------------------------|------------------------------|-----------------|----------|-----------------------|----------------------|
| | | | | | | | | | | In Study | Operated | In study | Surgically evaluated |
| MRI Studies | | | | | | | | | | | | | |
| Gorezyca et al., 1992.42 | USA | 1990–1992 | NC | Retro | Symp | NS | 45 (23–71) | NS | Both | 143 | 70 | 281 | 140 ^{††} |
| Monticciolo et al., 1994.45 | USA | NS | NC | Pros | Symp | imaging findings | 45 (30–68) | NS | Both | 35 | 28 | 46 | 38 |
| Gorezyca et al., 1994.41 [‡] | USA | 1992–1993 | NC | Pros | Symp | NS | 47 (28–72) | NS | NS | 82 | 41 | NS | 81 |
| Frankel et al., 1995.40 | USA | 1992–1999 | NC | Pros | NS | patient desire, clinical findings | 47 (24–77) | 8 (0.08–20) | Both | 68 | 20 | 119/107 ^{††} | 34/31 ^{‡‡} |
| Quinn et al., 1996.46 | USA | NS | NC | Pros | NS | NS | NS | NS | NS | 78 | 54 | 154 | 108 |
| Herbom et al., 2002.43 | Europe | 1996–1999 | C | Retro | Both | NS | 51 (24–71) | NS | Single | 110 | 25 | 81 | 41 |
| Hölmich et al., 2005.44 | Europe & USA | 1999–2002 | Unclear | Pros | Symp | imaging findings, clinical findings [§] | 46.1 (32–63) | 14.7 (6–25) | Both | 64 | 64 | 118 | 118 |
| Collis et al., 2007.39 | Europe | 1997–1998 | C | Pros | A symp ^d | imaging findings | 41 (23–62) | 8.8 (4.8–13.5) | Single | 149 | 21 | 298 | 42 |
| US Studies | | | | | | | | | | | | | |
| Debruhl et al., 1993.36 | USA | NS | NC | Pros | Symp | imaging findings, clinical findings | 46 (31–72) | NS | Both | 74 | 28 | 139 | 57 ^b |
| Caskey et al., 1994.34 | USA | 1991–1992 | C | Pros | Symp | clinical findings, patient desire | 46.5 (31–68) | 9.89 ^c (NS) | Both | 119 | 31 | 221 | 59 |
| Chung et al., 1996.35 | USA | 1991–1993 | Unclear | Pros | Symp | clinical findings, imaging findings | NS | NS | NS | 98 | 98 | 192 | 192 |
| Venta et al., 1996.37 | USA | NS | NC | Pros | Both | imaging findings | 45 (26–72) | 9.99 (1–22) | Both | 126 | 43 | 236 | 78 |
| Medot et al., 1997.38 ^d | USA | 1993 | NC | Unclear | Symp ^d | imaging findings, patient desire | 46 (30–65) | 3–25 ^e | NS | 65 | 65 | 122 | 122 |
| MRI and US Studies | | | | | | | | | | | | | |
| Ahn et al., 1994.26 | USA | 1992 | NC | Pros | Symp | NS | 46 (31–72) | 13 (6–19) | NS ^f | | | 24 | 24 |
| Reynolds et al., 1994.33 | USA | NS | NC | Pros | Symp | patient desire | 48.9 (37–63) | 12 (7–22) | Both | 13 | 13 | 24 | 24 |

| Source, Year | Location | Recruitment Dates | Recruitment Method | Study Design | Study Sample [§] | Indication for Surgery | Mean Subject age (range), y | Mean Implant age (range), y | Type of implant [†] | Gel bleed considered | No. of Patients | | No. of Implants | |
|------------------------------------------|----------|-------------------|--------------------|--------------|---------------------------|-----------------------------------|-----------------------------|-----------------------------|------------------------------|---------------------------|-----------------|----------|-------------------------------|----------------------|
| | | | | | | | | | | | In Study | Operated | In study | Surgically evaluated |
| Everson et al., 1994. ³⁰ | USA | NS | Unclear | Pros | Symp ^g | patient desire | 29–63 ^e | NS | NS | NS | 32 | 32 | 59(MRI), 61(US) | 63 |
| Berg et al., 1995. ²⁸ | USA | 1993–1994 | Unclear | Pros | Symp ^h | clinical findings, image findings | 47 (29–71) | 11.9 (0.7–26) | Single | Not Included [†] | 282 | 77 | 453 [†] | 122 [†] |
| Beekman et al., 1999, ²⁷ | Europe | 1996–1997 | Unclear | Pros | Symp | NS | 47 (25–71) | 16 (4–22) | Single | Not R | 18 | 18 | 35 | 35 |
| Ikeda et al., 1999. ³¹ | USA | NS | NC | Pros | Symp | NS | 45 (30–71) | 12 (3–25) | Both | Not R | 30 | 16 | 59 | 31 |
| Scaranelo et al., 2004. ^{32k} | Brazil | 1993–1996 | NC | Pros | Asymp | NS | NS | 11.9 (NS) | Single | Not R | 44 | 44 | 77(MRI) [†] , 83(US) | 83 |
| Di Benedetto et al., 2008. ²⁹ | USA | 2002–2006 | NC | Retro | NS | NS | 51 (24–71) | 7.8 (4–25) | NS | Not R | 63 | 63 | 82 | 82 |

Abbreviations: MRI, magnetic resonance imaging; US, ultrasound; y, year; NS, not stated; NC, Nonconsecutive; C, Consecutive; Retro, Retrospective; Pros, Prospective; Symp, Symptomatic; Asymp, Asymptomatic; Not R, Not Rupture; R, Rupture

[§] Descriptions of “concerned” patients were also included as “symptomatic” patients; “Both” includes symptomatic and asymptomatic patients.

[†] “Both” indicates single and double lumen implants were included in study. “Single” indicates single lumen implants.

^{††} Six patients had one, 61 patients had 2, and 3 patients had 4 silicone implants removed.

[‡] Authors compared Fast-Spin Echo & 3-Point Dixon MRI diagnostic accuracies, separately.

^{‡‡} Study reports 119 implants examined total, but only 107 implants were silicone implants; 34 implants were explanted but only 31 were silicone implants.

[†] Hoelrich *et al.* reports the primary indication for explantation in 52 (81%) women was the MRI findings; in “several cases” the indication was severe capsular contracture; for 7 women other reasons were indications, and in 3 women, no reason was given.

^a All patients were asymptomatic with no clinical findings (Nick Collis, MD, written communication, November 5, 2009).

^b Twenty-six patients had 2 implants removed, 1 had 1 implant removed, and 1 had 4 implants removed.

^c Range of implant age not provided in report. Mean age of ruptured implants was 13.9 years; mean age of not ruptured implants was 7.5 years.

^d All patients in this case-control study were described to be “concerned” or with chest pain; patients clinically diagnosed with capsular contracture were designated as “cases.” For meta-analysis, all numbers (from patients with capsular contracture and without) were combined.

^e Only ranges reported.

^f Only commercial manufacture of implants described.

^g Among 32 patients, only 1 was asymptomatic.

^h Only 6 out of 122 breast implants were described as being asymptomatic.

ⁱ Gel bleeds addressed in report, but designated as indeterminant and not included in analysis by Berg *et al.*

^j Numbers indicate total number of single lumen silicone implants extrapolated from study.

^k Authors compared MRI with breast coil and MRI with body coil separately.

^L Among the 77 implants evaluated for rupture by MRI, 20 implants were evaluated with MRI with breast coil and 57 implants with MRI with body coil.

Table 4
 Subgroup Analyses Based on Study Design Biases and Characteristics of Methodological Study Design and Text Execution

| Characteristic | No. Studies | Sn, % (95% CI) | p value | Sp, % (95% CI) | p value |
|----------------------------------------|-------------|----------------|---------|----------------|---------|
| Biases | | | | | |
| Spectrum Bias | | | | | |
| Magnetic Resonance Imaging | | | | | |
| Only symptomatic patients [†] | 11 | 88 (82–95) | 0.82 | 94 (90–98) | 0.09 |
| Only asymptomatic patients | 3 | 76 (56–96) | | 68 (41–95) | |
| Partial Verification Bias | | | | | |
| Magnetic Resonance Imaging | | | | | |
| Partial verification bias | 9 | 89 (83–96) | 0.13 | 91 (84–98) | 0.40 |
| No partial verification bias | 9 | 85 (77–93) | | 88 (79–97) | |
| Uninterpretable results reported | 12 | 88 (82–94) | 0.09 | 92 (87–98) | 0.97 |
| Uninterpretable results not reported | 6 | 86 (76–95) | | 84 (70–98) | |
| Ultrasound | | | | | |
| Partial verification bias | 5 | 61 (48–74) | 0.41 | 67 (55–79) | <0.001* |
| No partial verification bias | 8 | 61 (51–70) | | 81 (74–88) | |
| Uninterpretable results reported | 6 | 61 (49–73) | 0.39 | 83 (75–91) | 0.56 |
| Uninterpretable results not reported | 7 | 61 (51–71) | | 70 (60–81) | |
| Methodological Study Design | | | | | |
| Magnetic Resonance Imaging | | | | | |
| Retrospective | 4 | 88 (79–97) | 0.11 | 93 (85–100) | 0.87 |
| Prospective | 14 | 87 (80–93) | | 89 (82–96) | |
| Nonconsecutive [‡] | 12 | 83 (76–91) | 0.21 | 89 (82–96) | 0.23 |
| Consecutive | 2 | 90 (77–100) | | 70 (35–100) | |
| Test Execution | | | | | |

| Characteristic | No. Studies | Sn, % (95% CI) | p value | Sp, % (95% CI) | p value |
|-----------------------------------|-------------|----------------|---------|----------------|---------|
| Magnetic Resonance Imaging | | | | | |
| Undefined reference test criteria | 8 | 88 (80–96) | | 90 (82–99) | |
| Defined reference test criteria | 10 | 87 (80–93) | 0.05 | 90 (82–97) | 0.28 |
| Undefined index test criteria | 3 | 88 (79–100) | | 86 (68–100) | |
| Defined index test criteria | 15 | 86 (81–92) | 0.45 | 90 (85–96) | 0.43 |
| Ultrasound | | | | | |
| Undefined reference test criteria | 6 | 61 (50–73) | | 76 (82–99) | |
| Defined reference test criteria | 7 | 60 (50–71) | 0.42 | 77 (67–87) | 0.09 |
| Undefined index test criteria | 2 | 65 (79–100) | | 83 (68–98) | |
| Defined index test criteria | 11 | 60 (52–68) | 0.76 | 75 (66–83) | 0.75 |

Abbreviations: Sn, Sensitivity; Sp, Specificity

[†] Studies designated as unclear were not included in the individual subgroup analyses.

* Denotes $p < 0.001$