

The elusive diagnostic test for endometriosis

Tommaso Falcone, M.D.,^a and Edward Mascha, M.S.^b

Cleveland Clinic Foundation, Cleveland, Ohio

The prevalence of endometriosis in the general population is hard to define. However, it is noted to be quite high in patients with infertility or chronic pelvic pain. Laparoscopy is considered to be the “gold” standard for the diagnosis of endometriosis. Although many attempts have been made to use clinical, imaging, and serum markers to arrive at a diagnosis without surgery, so far none have proven clinically useful on their own. The value of a nonsurgical diagnostic test is even more critical if medical therapy is considered to be as effective as surgical therapy. In this way, a patient can avoid surgery altogether. There is a certain degree of misclassification with all diagnostic tests. The clinician must decide if the clinical and economic consequences of such misclassification are acceptable.

The evaluation of a diagnostic test requires some familiarity with the measurements used (1). The variability of test results could be due to intrinsic inaccuracies in the technical part of performing the test or the observer’s assessment of the result. Standard reports of the value of a diagnostic test include sensitivity (proportion of patients with the disease that have a positive test) and specificity (proportion of patients without the disease that have a negative test). Those that are incorrectly labeled as having the disease are the false positives and those that are incorrectly labeled as not having the disease are the false negatives. Sensitivity and specificity are measures of diagnostic accuracy that are not affected by the prevalence of the condition.

Another set of measurements that are important to consider in a test are the predictive values of a negative or positive test result (1). The positive predictive value (PPV) is the proportion of those with a positive test that have the disease. The negative predictive value (NPV) is the proportion of those with a negative test that do not have the disease. In interpreting the value of the predictive values of a test, it is important to know the pretest probability of the disease, that is, the frequency of the disease in the patient population. If the pretest probability is very high, such as in a subset of patients with chronic pelvic pain or infertility, the predictive value of a positive test will be high. Since these results are influenced by the frequency of disease, they may not necessarily be applicable to a different patient population.

Bayes’s theorem can be used to estimate the PPV and NPV of a test for varying populations if the sensitivity, specificity, and probability of the disease in the target population are known or can be well estimated. Using this method, however, implies that the sensitivity and specificity of the test do not change for various populations of interest. Although sensitivity and specificity are not affected by prevalence, they may well be affected by the spectrum of disease. It may be the case, for example, that the sensitivity is higher (easier to detect) in a population with more advanced disease (2). This would be the case in endometriosis where it is possible to miss the diagnosis of minimal disease but unlikely with more advanced disease.

Two other measurements may also be used to evaluate a test: likelihood ratios (LRs) and receiver operator characteristic (ROC) curves (1). The LR is a ratio of the true positives (sensitivity) to the false positives (1-specificity) at a particular value of a test. Alternatively, these quantities can be plotted for all observed values of a predictor or test and the curve generated (ROC). The area under the ROC curve (AUC) is a common measure of the

Received June 10, 2003;
revised and accepted June
10, 2003.

Reprint requests:
Tommaso Falcone, M.D.,
Department of Obstetrics
and Gynecology-A81,
Cleveland Clinic
Foundation, 9500 Euclid
Avenue, Cleveland, Ohio
44195 (FAX: 216-445-5526;
E-mail: falcont@ccf.org).

^a Department of Obstetrics
and Gynecology.

^b Department of
Biostatistics and
Epidemiology.

The opinions and
commentary expressed in
Editor’s Corner articles are
solely those of the author.
Publication does not imply
endorsement by the Editor
or American Society for
Reproductive Medicine.

0015-0282/03/\$30.00
doi:10.1016/S0015-0282(03)
01161-0

diagnostic accuracy of a test. Reporting LR without the ROC curve can be misleading since similar LR values are possible with quite different ROC areas (2).

Many diagnostic tests have been proposed. The most popular is the CA-125. A recent meta-analysis was performed to assess the diagnostic performance of serum CA-125 in detecting endometriosis (3). The ROC curve showed a poor diagnostic performance. At a specificity of 90%, a sensitivity of 28% was reported. If the sensitivity was increased to 50% the specificity dropped to 72%.

A significant role of the immune system in the pathogenesis of endometriosis has been recently documented. Based on these recent findings, immune markers have been proposed as potential diagnostic tests. Peritoneal fluid tumor necrosis factor- α (TNF- α) concentrations are elevated in patients with endometriosis. At a cut-off peritoneal fluid concentration of 20 pg/mL, a 96% sensitivity and a 95% specificity (positive LR of 19.2 and negative LR of 0.04) were reported (4). However, this may not be a practical test because it would require an intervention to obtain the peritoneal fluid.

Having a gold standard method to determine true disease status that is completely independent of the test itself is a critical part of determining the value of a test. Endometriosis is typically diagnosed visually at laparoscopy. Several articles, however, report that this is often an inaccurate determination. Histologic confirmation of a presumed endometriotic lesion observed at laparoscopy ranges from 60% to 85% (5, 6). This is principally due to the vast diversity in the visual appearance of endometriosis. Therefore, a diagnostic laparoscopy without histologic confirmation of the disease may lead to incorrect assignment of the diagnosis of endometriosis. Diagnosing very subtle or minimal endometriosis may be more difficult by laparotomy and this may lead to an underestimate of the prevalence of disease. Within the context of evaluating a diagnostic test, these issues are critical because it could lead to incorrect assignment of patients into cases and controls.

In the study by Gagne et al. (7) in this issue, a number of clinical and laboratory factors were used to predict the likelihood that a patient has endometriosis. As such, this is predictive model rather than a simple blood test. The predictive model includes the length of menses, gravidity, and histologic dating of the endometrium, but interestingly no pain parameter or infertility influenced the predictive value of the model. The model also included a variety of white blood cell serotypes that were determined by flow cytometry. The diagnosis of endometriosis was not confirmed by histology and relied only on visual observation. Although this may be appropriate in clinical practice, it may introduce a bias into a clinical study. Imaging was not used as part of the model. In certain cases, such as with ovarian cysts (8) or rectovaginal disease (9), pelvic imaging may improve the diagnostic probability of endometriosis.

The study reports that the test has a sensitivity of 61%. This means that almost 40% of patients with the disease will be incorrectly labeled as disease free (false negative). In clinical practice this may be insufficient for the management of patients. A negative test may inappropriately delay further diagnosis or specific treatment. Furthermore, the investigators state that the test is not designed to replace diagnostic laparoscopy. In this respect, if a clinician is faced with symptoms that require treatment, this test may not be practical in determining the next step. Interestingly, the diagnostic performance is identical for less extensive disease (stage 1 and 2) and more extensive disease (stage 3 and 4). The situation is quite different when CA-125 is used alone, where the ROC curve shows a better performance for advanced disease (3).

In the study by Gagne et al., (7) a prevalence of disease (endometriosis) was assumed to be 45% for the calculation of predictive values. The predictive value of this test would be different in clinical practices with a different frequency of disease; it would depend on the specific population of patients with infertility or chronic pelvic pain that a clinician deals with. This pretest frequency of disease may not apply to an infertile patient who is at the beginning of an investigation.

If the pretest probability of disease was lower, then the predictive value of a positive test would be lower. Even within the context of this study, where the pretest probability of endometriosis was high, the positive and negative predictive values of this test were 91% and 75%, respectively. Therefore, a positive test result is highly indicative of endometriosis, but a negative result cannot rule out endometriosis with a high level of clinical confidence. The ROC curves and the diagnostic odds ratio (similar to LR) in this study show good discriminating ability.

If the test is validated in different study populations, it could provide clinically useful information. If a clinician is contemplating a laparoscopy for the investigation of chronic pelvic pain, a positive test could allow the patient to proceed with medical management without surgery. In the context of chronic pelvic pain, medical treatment without laparoscopic confirmation has already been proposed (10). Therefore, this diagnostic test may aid in the empirical treatment of patients with chronic pelvic pain. A negative result means that the patient can still have endometriosis, although with a lower probability. In these cases, patients still have the option of proceeding to surgical evaluation.

In the context of the infertile patient, a laparoscopy may not be avoided, regardless of the test result. A laparoscopy in an infertile patient is not performed solely for the diagnosis of endometriosis. If the test shows a high probability of endometriosis, surgical treatment has been shown to be of clinical value (11). If the test is negative, a patient may still want to proceed with surgery to rule out pelvic adhesions. Adhesiolysis has been shown to increase the spontaneous

pregnancy rate (12). However, a nonsurgical diagnosis of endometriosis is useful even if the management is surgical rather than medical. An early diagnosis of endometriosis may avoid unnecessary testing or empirical treatment for other causes of infertility as well as inform the patient of the success of different options.

The development of a diagnostic test for endometriosis is of major clinical importance for the evaluation and management of patients with chronic pelvic pain and infertility. Each diagnostic test must undergo rigorous scrutiny to properly evaluate its role in our clinical practice.

References

1. Riegelman RK. Studying a study and testing a test. How to read the medical evidence. Philadelphia: Lippincott Williams & Wilkins, 2000: 127–66.
2. Zhou X, Obuchowski N, McClish D. Statistical methods in diagnostic medicine. New York: John Wiley & Sons, 2002:15–47.
3. Mol BW, Bayram N, Lijmer JG, Wiegerink MAH, Bongers MY, van der Ven F, et al. The performance of CA-125 measurement in the detection of endometriosis: a meta-analysis. *Fertil Steril* 1998; 70:1101–8.
4. Bedaiwy MA, Falcone T, Sharma RK, Goldberg JM, Attaran M, Nelson DR, et al. Prediction of endometriosis with serum and peritoneal fluid markers: a prospective controlled trial. *Hum Reprod* 2002;17:426–31.
5. Stratton P, Winkel C, Premkumar A, Chow C, Wilson J, Hearn-Stokes R, et al. Diagnostic accuracy of laparoscopy, magnetic resonance imaging, and histopathologic examination for the detection of endometriosis. *Fertil Steril* 2003;79:1078–85.
6. Mettler L, Schollmeyer T, Lehmann-Willenbrock E, Schüppler U, Schmutzler A, Shukla D, et al. Accuracy of laparoscopic diagnosis of endometriosis. *JSL* 2003;7:15–8.
7. Gagne D, Rivard M, Page M, Lepine M, Platon C, Shazand K, et al. Development of a nonsurgical diagnostic tool for endometriosis based on the detection of endometrial leukocyte subsets and serum CA-125 levels. *Fertil Steril* 2003;80(4):876–85.
8. Guerriero S, Mais V, Ajossa S, Paoletti AM, Angiolucci M, Melis GB, et al. Transvaginal ultrasonography combined with CA125 plasma levels in the diagnosis of endometrioma. *Fertil Steril* 1996;65:293–8.
9. Fedele L, Bianchi S, Portuese A, Boruito F, Dorta M. Transrectal ultrasonography in the assessment of rectovaginal endometriosis. *Obstet Gynecol* 1998;91:444–8.
10. Ling FW. Randomized controlled trial of depot leuprolide in patients with chronic pelvic pain and clinically suspected endometriosis. *Pelvic Pain Study Group. Obstet Gynecol* 1999;93:51–8.
11. Marcoux S, Maheux R, Berube S, and the Canadian Collaborative Group on Endometriosis. Laparoscopic surgery in infertile women with minimal or mild endometriosis. *N Engl J Med* 1997;337:217–22.
12. Tulandi T, Collins JA, Burrows E, Jarrell JF, McInnes RA, Wrixon W, et al. Treatment-dependent and treatment-independent pregnancy among women with periadnexal adhesions. *Am J Obstet Gynecol* 1990;162:354–7.