physicians should encourage the use of this technique in their institutions and among their patients.

Mark H. Ebell, MD, MS
Wayne State University
Detroit, Michigan

LABORATORY DIAGNOSIS OF IRON DEFICIENCY ANEMIA

TITLE: Laboratory diagnosis of iron-deficiency anemia: an overview
AUTHORS: Guyatt GH, Oxman AD, Mahmoud A, et al.
JOURNAL: *Journal of General Internal Medicine*
DATE: March-April 1992; Volume 7:145–53.

Clinical question. **What is the most appropriate laboratory test for the evaluation of a patient with suspected iron deficiency anemia?**

Background. Iron deficiency is very common, but physicans' approaches to laboratory evaluation of this condition vary greatly. Physicians use many different strategies, including mean cell volume (MCV), transferrin saturation, serum ferritin, red cell volume distribution, red cell protoporphyrin, and red cell ferritin. The purpose of this overview was to compare these strategies and identify the most useful test for the evaluation of possible iron deficiency anemia.

Studies reviewed. The literature search was well described and thorough, employing a broad search for articles on MEDLINE between 1966 and 1990, including foreign language papers, and citations from those articles as well as reviews. A manual search of *Index Medicus* before 1966 would have been helpful, as would polling experts to see what is missing. Formal reviews such as this should also describe the authors' standards for including studies for review. Inclusion criteria for this article were: (1) age greater than 18 years old and hemoglobin below 13 for men and below 11 for women; (2) quantitation of at least one of the diagnostic methods; and (3) use of bone-marrow aspiration as a reference standard for comparison with the diagnostic test. These criteria are appropriate.

Study design and validity. Receiver-operator characteristic (ROC) curves, which plot the test's sensitivity against 1 minus its specificity, are the preferred method for comparing diagnostic tests. The area under an ROC curve (AUROCC) represents the probability, from 0 to 1, that a test will correctly identify the diseased person. A higher AUROCC, ie, close to 1.0, represents a better test, while an AUROCC of 0.5 describes a test that does not yield any useful information, ie, its ability to classify patients is no better than tossing a coin. In this overview article, the AUROCCs of different diagnostic tests were plotted and compared, and the pooled data were used to generate likelihood ratios for different test results. A likelihood ratio is the odds that a given level of a diagnostic test will be found in a patient with (as opposed to without) the disease in question.

A particular strength of this study was the authors' assiduity in evaluating study relevance and methodologic quality, ie, population, interventions, outcomes measurement. These were reviewed independently by two observers, and agreement was calculated by the weighted kappa statistic. Agreement for relevance was outstanding (κ=0.82), and for quality, very good. (0.40 to 0.63).

Outcomes measured. The authors used histologic examination of bone marrow as their principal outcome, which is a clinically appropriate standard. In some studies, bone marrow aspiration was not done on all subjects; these data were included if results on individual subjects could be obtained.

Results. The serum ferritin performed much better than any other diagnostic test or combination of tests for iron deficiency; however, serum ferritin performed differently in different populations: patients with "inflammatory disease," including infection, malignancy, connective tissue disease, and liver disease, and a group without inflammatory conditions. The cutpoint for iron deficiency, ie, the point at which the test defines iron deficiency, is higher in patients with "inflammatory" conditions, but the relationship is relatively predictable. For example, a serum ferritin of 30 would have a likelihood ratio of 2 in a mixed population and a ratio of 4 in patients with underlying inflammatory disease.

Clinical recommendations. **This overview provides strong evidence for using serum ferritin as an initial laboratory test for the evaluation of iron deficiency anemia. Ferritin retains its usefulness in patients with underlying inflammatory or liver disease, but the cutpoint should be altered. Assuming that a likelihood ratio of more than 4 indicates disease, while less than 0.25 rules out disease, the data in this article suggest that in patients with inflammatory disease, a ferritin test result of less than 30 indicates iron deficiency; greater than 130 rules out iron deficiency. For patients without inflammatory disease, a ferritin test result of less than 20 indicates iron deficiency; greater than 100 rules out this condition. Other testing, including bone marrow aspiration, should be consid-**

ered for patients in whom the involvement of iron deficiency remains uncertain.

Warren Newton, MD
University of North Carolina
Chapel Hill, North Carolina

C PNEUMONIAE IN ADULT ASTHMA

TITLE: Treatment of *Chlamydia pneumoniae* infection in adult asthma: a before–after trial
AUTHORS: Hahn DL
JOURNAL: *The Journal of Family Practice*
DATE: October 1995; Volume 41:345–351

Clinical question. **Is antibiotic therapy for *Chlamydia pneumoniae* an effective treatment for adult-onset asthma?**

Background. Recently, it has become clear that several chronic illnesses may be caused in part by chronic bacterial infection. For example, infection with *Helicobacter pylori* has been implicated in gastric and duodenal ulcer disease. Some physicians have anecdotally reported the onset of asthma in patients who have had a *Chlamydia pneumoniae* infection, as well as an improvement in asthma symptoms in adults and children treated with antibiotics for this infection.

Population studied. The author studied adult outpatients with moderate asthma (two or more symptomatic episodes per week) or moderately severe asthma (previous oral steroid use or hospitalization) who presented to his suburban family practice. The population was described as "mostly white and middle-class," with a mean age of 48 years (range 17 to 78 years) and a mean age of onset of asthma symptoms of 42 years. Almost one half of the study population were smokers. This group appears typical of patients cared for by family physicians.

Study design and validity. The study design used is called "pretest/posttest without a control group." That is, the FEV_1 level and subjective symptoms were measured antibiotics that are used to treat *Chlamydia pneumoniae* were given, and the FEV_1 level and subjective symptoms were reassessed after the antibiotic intervention. Patients were followed for an average of 6 months (range 1.5 to 36 months) after treatment, and the last recorded test results were used for the analysis.

This design has a number of possible biases. For example, patients are likely to present to the physician when their symptoms are worse than average. By simply waiting a few months, the symptoms are more likely to improve than worsen, a phenomenon known as "regression to the mean." (The same phenomenon is responsible for the so-called sophomore jinx in sports, in which an outstanding rookie season is followed by a less successful second year.) Patients who experienced improvement may have been more likely to take other medications, make lifestyle changes, mitigate environmental exposures, experience improvement due to seasonal change, or differ in some other way, than nonresponders. The variable follow-up period is also troubling, as is the lack of information on patients who were not asked to participate. The design could have been strengthened by enrolling all patients, adding a comparison group of patients who were randomized to receive a placebo, having a standard follow-up period, and using a previously validated symptom score.

Outcomes measured. The primary measures were the percent change in the FEV_1 level and the score on the Patient Global Improvement (PGI) scale. The latter is a Likert-type scale ranging from −4 (complete worsening) to 0 (no change) to +4 (complete improvement). It has not been previously validated, but did correlate with the changes in FEV_1 levels.

Results. The FEV_1 improved an average of 12.5% with treatment (95% confidence interval [CI], 4.6 to 20.3). Twenty-five patients were classified as "responders" with a posttreatment PGI score of +3 or +4, whereas 21 had little or no response to treatment and a posttreatment PGI of less than 3. Responders tended to have a shorter duration of asthma symptoms before treatment (3.1 years vs. 8.3 years, $P=.01$) and were less likely to have used inhaled steroids (24% vs 62%, $P=.015$).

Recommendations for clinical practice. **The results of this innovative practice-based trial are intriguing, and certainly warrant further study. It is tempting to respond by giving all one's newly diagnosed adult, nonatopic asthma patients a round of erythromycin, considering the possible benefit and short duration of therapy in this chronic illness. However, moving too quickly to advocate a new treatment has several disadvantages: if further study does not support the efficacy of this treatment, physicians will have to "unlearn" this behavior; using ineffective treatments drives up costs, including those of side effects; and widespread adoption may have the effect of stifling further research, as in the rapid adoption of bypass**