Diagnosis of Iron-Deficiency Anemia in the Elderly

GORDON H. GUYATT, M.D., CHRISTOPHER PATTERSON, M.D., MAHMOUD ALI, M.D., JOEL SINGER, PH.D., MARK LEVINE, M.D., IRENE TURPIE, M.D., RALPH MEYER, M.D., *Hamilton, Ontario, Canada*

PURPOSE: **To determine the value of serum ferritin, mean cell volume, transferrin saturation, and free erythrocyte protoporphyrin in the diagnosis of iron-deficiency anemia in the elderly.**

PATIENTS AND METHODS: **We prospectively studied consecutive eligible and consenting anemic patients over the age of 65 years, who underwent blood tests and bone marrow aspiration. The study consisted of 259 inpatients and outpatients at two community hospitals in whom a complete blood count processed by the hospital laboratory demonstrated previously undiagnosed anemia (men: hemoglobin level less than 12 g/dL; women: hemoglobin level less than 11.0 g/dL).**

RESULTS: **Thirty-six percent of our patients had no demonstrable marrow iron and were classified as being iron-deficient. The serum ferritin was the best test for distinguishing those with iron deficiency from those who were not iron-deficient. No other test added clinically important information. The likelihood ratios associated with the serum ferritin level were as follows: greater than 100 μg/L, 0.13; greater than 45 μg/L but less than or equal to 100 μg/L, 0.46; greater than 18 μg/L but less than or equal to 45 μg/L, 3.12; and less than or equal to 18 μg/L, 41.47. These results indicate that values up to 45 μg/L increase the likelihood of iron deficiency, whereas values over 45 μg/L decrease the likelihood of iron deficiency. Seventy-two percent of those who were not iron-deficient had serum ferritin values greater than 100 μg/L, and in populations with a prevalence of iron deficiency of less than 40%, values of greater than 100 μg/L reduce the probability of iron deficiency to under 10%. Fifty-five percent of the iron-deficient patients had serum ferritin values of less than 18 μg/L, and in populations with a prevalence of iron deficiency of greater than 20%, values of less than 18 μg/L increase the probability of iron deficiency to over 95%.**

CONCLUSION: **In a general geriatric medical population such as ours, with a prevalence of iron deficiency of 36%, appropriate use of serum ferritin determination would establish or refute a diagnosis of iron deficiency without a bone marrow aspiration in 70% of the patients.**

Anemia is an extremely common problem in the elderly, and next to anemia of chronic disease, iron deficiency is the most common cause. Iron-deficiency anemia is important to diagnose because appropriate iron therapy may improve symptoms, inappropriate iron therapy may cause clinically important side effects, and iron deficiency may be a marker for occult gastrointestinal pathology.

Although bone marrow aspiration provides a definitive diagnosis of iron-deficiency anemia, the value of less invasive tests of iron stores in general populations has been well established [1–9]. Serum ferritin and transferrin saturation are the tests most commonly used. Because bone marrow aspiration can be painful and is more expensive than laboratory tests, the procedure is often reserved for patients in whom the diagnosis remains in doubt after noninvasive test results are available.

Our interest in the investigation of iron deficiency in the elderly was stimulated by a clinical impression that application of cutoff points for laboratory tests for the diagnosis of iron deficiency derived from younger populations was misleading in a geriatric population. There are a number of reasons why results found in younger populations may not apply to the elderly. The iron-binding capacity decreases with aging [10,11], and is affected by factors such as malnutrition and chronic disease, which have a higher prevalence in the elderly [12]. Serum ferritin levels increase with aging [13], and may be elevated by acute and chronic inflammatory conditions [14,15]. One small study has suggested that measurements of transferrin saturation and serum ferritin in elderly anemic patients with and without iron deficiency differ significantly from those found in younger patients [16]. These problems have led to varying recommendations regarding the interpretation of results of noninvasive tests of iron stores in the elderly [16–18].

There are other reasons why further study of the diagnosis of anemia is warranted. First, investigations to date have generally used a single cut-point, and reported on the sensitivity and specificity of the tests. This approach discards valuable information. Use of multiple cut-points, with determination of likelihood ratios associated with each range of results, provides additional information for the clinician [19]. Second, statistically reliable determination of the best single test, and whether additional useful information could

From the Department of Medicine (GHG, CP, MA, ML, IT, RM), the Department of Clinical Epidemiology and Biostatistics (GHG, ML), and the Department of Family Practice (JS), McMaster University, Hamilton, Ontario, Canada. This work was supported in part by the Ontario Ministry of Health. Dr. Guyatt is a Career Scientist of the Ontario Ministry of Health. Requests for reprints should be addressed to Gordon H. Guyatt, M.D., Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, Room 2C12, 1200 Main Street West, Hamilton, Ontario, Canada, L8N 3Z5. Manuscript submitted July 13, 1989, and accepted in revised form November 14, 1989.

Current addresses: Department of Medicine, Chedoke Hospital, Hamilton, Ontario, Canada (CP); Department of Pathology, St. Joseph's Hospital, Hamilton, Ontario, Canada (MA); Department of Medicine, St. Joseph's Hospital, Hamilton, Ontario, Canada (IT); Department of Family Medicine, McMaster University Health Sciences Centre, Hamilton, Ontario, Canada (JS); and Department of Medicine, Henderson General Hospital, Hamilton, Ontario, Canada (ML, RM).

TABLE I

Reasons For Exclusion of Patients Found To Be Anemic on at Least One Hemoglobin Determination

Reason for Exclusion	Number of Patients Excluded
Patient judged too ill, demented, or terminal	212
Patient or family refused consent for bone marrow aspiration	200
Not anemic on second hemoglobin determination	200
Recent transfusion	152
Previous bone marrow aspiration had revealed diagnosis	108
Institutionalized	95
Miscellaneous	108
Total	1,075

TABLE II

Final Primary Diagnosis of Anemia

Diagnosis	Number of Patients
Iron-deficiency anemia	94
Anemia of chronic disease	113
Megaloblastic anemia	21
Multiple myeloma	4
Sideroblastic anemia	3
Dysmyeloplastic	3
Other*	21
Total	259

* Includes patients with leukemia, hemolytic anemia, hypoplastic and aplastic marrow, renal failure, and hypothyroidism, and those with inadequate information for definitive diagnosis.

be gained from performing a second or third test, has seldom been investigated. Third, other tests (including the free erythrocyte protoporphyrin) have been suggested as being potentially useful in confirming the diagnosis of iron-deficiency anemia, but have not been adequately studied.

Because of the frequency of anemia in the elderly, and because of the difficulties in performing bone marrow aspirations in all anemic patients, we believed it important to determine the accuracy of less invasive laboratory tests commonly used to assess iron stores. Our criterion or gold standard for the diagnosis of iron deficiency was the results of the bone marrow aspiration.

PATIENTS AND METHODS

Consecutive patients over the age of 65 years presenting to Chedoke Hospital in Hamilton, Ontario, between January 1984 and March 1988 with anemia (in men, hemoglobin 12.0 g/dL or less on two consecutive occasions; in women, 11.0 g/dL or less) were identified through the hospital laboratory. An additional much smaller group of patients admitted to St. Joseph's Hospital in Hamilton under one of the co-investigators and meeting study criteria were also included. We excluded institutionalized patients, those with recent blood transfusions or documented acute blood loss, or those whose participation in the study was judged unethical by their attending physician (for reasons such as impending death or severe dementia). Detailed criteria for definition of "too ill," "impending death," or "severe dementia" were not established. Rather, we relied on physician judgment in these areas. Similarly, we relied on physicians for the appropriate level of encouragement to patient participation when obtaining informed consent.

All patients had the following laboratory tests: hemoglobin, mean red cell volume (MCV), red cell distribution width (RDW), serum iron, iron-binding capacity, serum ferritin, and red cell protoporphyrin. The complete blood count was carried out using a Coulter S +*IV*™ (Coulter Electronics, Miami, Florida). Serum iron and iron-binding capacity were measured according to the methods of the International Committee for Standardization in Haematology [20]. Serum ferritin was determined using a radioimmunoassay described in detail previously [21]. Red cell protoporphyrin was measured using a previously described micromethod [22]. A bone marrow aspiration was undertaken and the findings were interpreted by a hematologist (M.A.) who was unaware of the results of the laboratory tests. The bone marrow slides were air-dried, fixed with methanol, and stained with Prussian blue [23]. Results of the first 65 marrow aspirations were also interpreted by a second hematologist (also unaware of the laboratory test findings), and discrepancies resolved by consensus. The results of the aspiration were classified as iron absent, reduced, present, or increased. After interpreting the marrow aspiration results, the hematologist reviewed all relevant clinical information and made a final decision regarding the cause(s) of the anemia. Anemia of chronic disease was diagnosed when the iron present in the reticuloendothelial cells (fragments) was increased and the number of sideroblasts (red cells containing iron granules) was decreased. The increase in reticuloendothelial iron was defined as iron granules covering 50% of all the fragments observed, and a decrease in sideroblasts was confirmed when iron granules were present in less than 20% of the red cells.

Statistical Methods

Receiver operating characteristic (ROC) curves for each test were generated. The area under the curves was compared using the method of Hanley and McNeil [24]. Since the ROC curves in this study were all generated from the same cohort of patients, we used the correction factor, which reflects the correlation between the tests [25]. Using the same cut-points, likelihood ratios for each category were calculated.

To determine the independent contribution of each test to the diagnosis, and whether a combination of tests could improve diagnostic accuracy, stepwise logistic regression procedures were used. The status of iron stores (present or absent) was used as the dependent variable, and the values of the diagnostic tests (dichotomized using the cut-point that maximized accuracy) as the independent variables.

Chance-corrected agreement between the two hematologists who interpreted the marrow aspiration results was calculated using a weighted kappa, with quadratic weights.

RESULTS

Aside from providing more precise estimates of the likelihood ratios, inclusion of 25 patients from St. Joseph's Hospital had no systematic effect on the results of any analysis. Thus, these patients will not be identi-

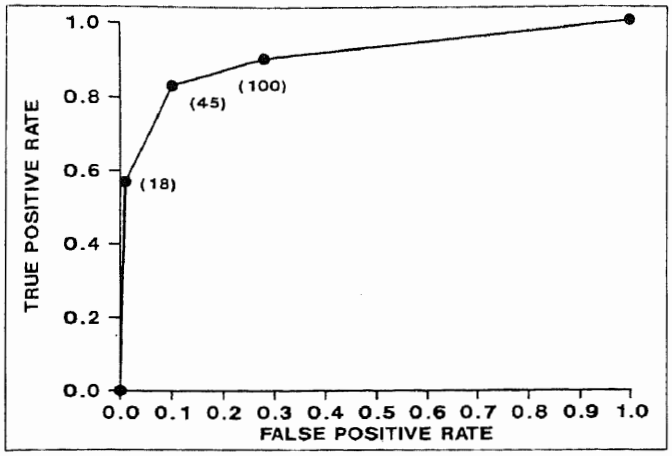

Figure 1. ROC curve for serum ferritin.

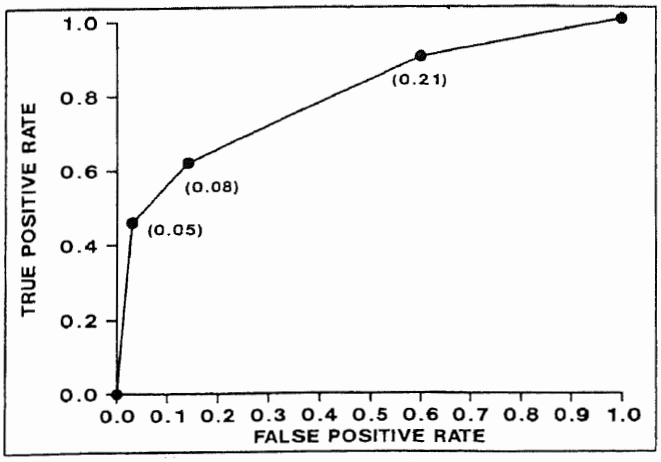

Figure 2. ROC curve for transferrin saturation.

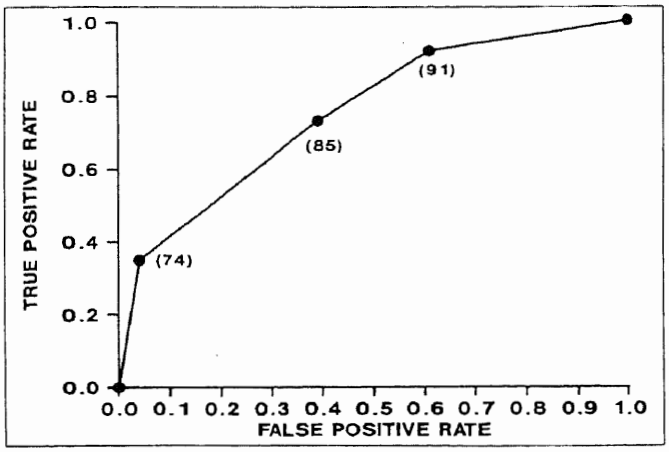

Figure 3. ROC curve for mean cell volume.

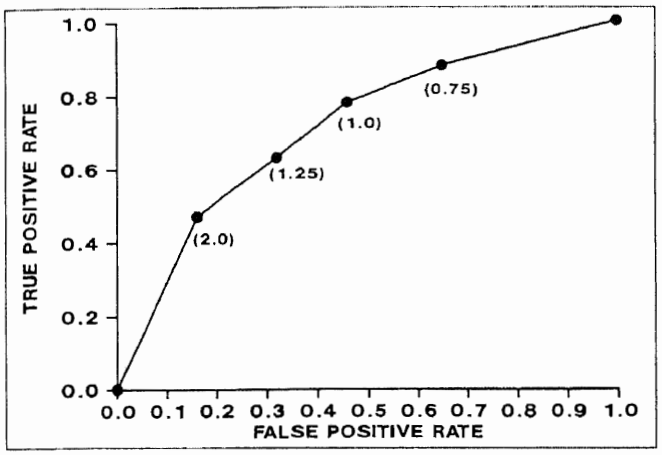

Figure 4. ROC curve for free erythrocyte protoporphyrin.

fied separately in the presentation of the results that follows.

A total of 1,334 patients over 65 years with anemia was identified. Of these, 259 proved eligible, participated in the study, and underwent bone marrow aspiration. Seventy-six participants were outpatients, and 183 were inpatients. Of the 259 bone marrow aspirates from the patients, 235 were interpretable (the quality being too poor in the others). The reasons for exclusion of anemic patients are presented in **Table I.** Most of the patients who recently received transfusions were postoperative patients, a large proportion of whom (at a hospital with a very busy orthopedic service) had undergone total hip replacement or had a recent hip fracture.

The mean age (± SD) of the participating patients was 79.7 ± 7.62 years; 119 (46%) were men. The mean hemoglobin level was 9.81 g/dL (± 1.39); 52.1% of the patients had a hemoglobin level less than 10.0 g/dL. A very wide variety of illnesses were not directly related to the anemia. Seventy-two patients had no medical diagnosis other than anemia (and its cause); 72 had one other diagnosis; 67 had two other diagnoses; 34 had three other diagnoses; and 14 had more than three other diagnoses. The most common medical diagnoses (aside from anemia), and the number of patients affected, were as follows: early dementia, 25; congestive heart failure, 25; chronic airflow limitation, 17; rheumatoid arthritis, 17; osteoarthritis, 14; pneumonia, 13.

The final diagnoses of anemia are presented in **Table II.** The weighted kappa-quantifying chance-corrected agreement for the 65 marrow aspirates that were interpreted by two hematologists was 0.84.

Figures 1 to 4 present the ROC curves for serum ferritin, transferrin saturation, MCV, and red cell protoporphyrin. Because, through administrative error and lost samples, all tests were not conducted in all subjects, the number of patients available for each analysis varied, and was sometimes less than 235. Examination of the ROC curves revealed that serum ferritin performed far better than any of the other tests. This was confirmed by the statistical analysis, which showed that the area under the ROC curves was 0.91, 0.79, 0.78, and 0.72 (respectively), for the four tests. Although the difference between the serum ferritin and the other three tests was statistically significant ($p \leq 0.001$ in each case), any differences seen in the other four curves can easily be explained by chance ($p \geq 0.1$).

Likelihood ratios for the four tests are presented in **Table III.** Consistent with the ROC curves, serum ferritin showed a far greater discriminative power than the other tests.

Likelihood ratios for RDW for distinguishing those with iron deficiency from those with anemia of chronic disease were examined. The likelihood ratios for RDW 0 to 15, 15 to 19, and greater than 19 were 0.39, 1.31, and 1.90, respectively. Ferritin proved a far more powerful predictor for differentiating iron deficiency from

TABLE III

Likelihood Ratios

Interval	Number Iron-Deficient	Number Not Iron-Deficient	Likelihood Ratio
Ferritin			
>100	8	108	0.13
>45 ≤ 100	7	27	0.46
>18 ≤ 45	23	13	3.12
≤18	47	2	41.47
Total	85	150	
Transferrin saturation			
>0.21	9	55	0.28
>0.8 ≤ 0.21	23	70	0.57
>0.05 ≤ 0.08	14	17	1.43
≤0.05	38	4	16.51
Total	84	146	
Mean cell volume			
>95	2	32	0.11
>91 – ≤ 95	5	26	0.34
>85 – ≤ 91	16	44	0.64
>74 – ≤ 85	32	42	1.35
≤74	30	6	8.82
Total	85	150	
Red cell protoporphyrin			
≥ 0 ≤ 0.75	10	53	0.34
>0.75 ≤ 0.1	8	28	0.51
>1 – ≤ 1.25	9	21	0.77
>1.25 ≤ 2	17	24	1.26
>2	40	24	2.98
Total	84	150	

TABLE IV

Likelihood Ratios from Logistic Regression Analysis

Interval	Number Iron-Deficient	Number Not Iron-Deficient	Likelihood Ratio
Ferritin			
Ferritin negative*†, transferrin saturation negative‡	13	126	0.18
Only ferritin positive	33	10	5.72
Ferritin positive, transferrin saturation positive	33	1	57.23
Total	79	137	

* Only four cases were serum ferritin-negative and transferrin saturation-positive.
† Cut-point for serum ferritin was 45 μg/L.
‡ Cut-point for transferrin saturation was 0.08.

TABLE V

Post-Test Probability of Iron Deficiency Given Varying Pre-Test Probabilities and Results of Serum Ferritin Determinations

	Pre-Test Probability			
	Low (5% – 20%)	Intermediate (40% – 60%)	High (80% – 95%)	Study Population (36%)
Serum ferritin result (μg/L)				
>100	0.6–3	8–16	34–71	7
45–100	2–10	24–41	39–90	21
18–45	14–44	68–82	93–98	64
<18	69–91	97–98	99–99.9	96

anemia of chronic disease, with likelihood ratios ranging from 0.05 to infinity. Finally, RDW added little to the predictive power of serum ferritin.

In the logistic regression model, ferritin was the best predictor of bone marrow iron stores. The only test that explained a statistically significant additional portion of the variance was the transferrin saturation. Using a cutoff of 45 μg/L for ferritin and 0.08 for transferrin saturation, likelihood ratios generated by using a combination of the tests are presented in **Table IV**. Little is gained by this model in comparison to serum ferritin: likelihood ratios greater than 1 are slightly higher, but the likelihood ratio less than 1 is not as low as the value obtained with a serum ferritin level of greater than 100 μg/L. Of patients with serum ferritin values of 18 to 100 μg/L, seven had transferrin saturation values of less than 0.05. All seven of these patients proved to be iron-deficient.

Other studies have reported elevated serum ferritin levels in patients with liver disease and inflammatory diseases, particularly rheumatoid arthritis [2,3,5,6,8,14,15]. Of the five patients with liver disease who were iron-deficient, three had serum ferritin values less than 18 μg/L. Of the six iron-deficient subjects with rheumatoid arthritis, five had a serum ferritin level less than 18 μg/L. Therefore, in our study, patients with liver disease or rheumatoid arthritis appeared to behave in a manner similar to that in the rest of the population. However, the numbers of patients with these conditions were insufficient to permit strong inference regarding the issue of differences among subgroups.

COMMENTS

Previous studies in younger subjects have consistently shown the usefulness of serum ferritin in the diagnosis of iron-deficiency anemia, and suggested that serum ferritin is more powerful than other blood tests [1–9]. Our results are consistent with these findings: in elderly patients with anemia, serum ferritin determination is by far the best test for diagnosis of iron deficiency. Other tests add only limited information in the diagnosis.

The MCV is ordinarily available with the complete blood count, and could thus influence the estimate of the probability of iron deficiency prior to ordering of other tests. However, in our population, even MCV values of less than 74 were not invariably associated with iron deficiency, and in many of the patients with iron deficiency the anemia was not microcytic. Only 6% of those with an MCV greater than 95 had iron deficiency; therefore, a very large MCV can be interpreted as virtually excluding iron deficiency.

The likelihood ratios for the possible ranges of results of serum ferritin determinations are presented in Table III. Previous studies in uncomplicated anemia have led to recommended cutoff points between normal and abnormal of 12 to 20 μg/L [1–9]. Using this approach, any value above 20 μg/L would be treated as a negative test result and as decreasing the likelihood of the patient having iron deficiency. In fact, in our population, ferritin values between 18 and 45 μg/L reflected an increase in the likelihood of iron deficiency (Table III), and the optimal cutoff in terms of maximizing accuracy was 45 μg/L (Figure 1). This result likely reflects the fact that serum ferritin levels in-

crease with age [13]. It may also reflect the high prevalence of chronic disease in the elderly, although only a small proportion of our population had inflammatory conditions thought to be associated with increased levels of serum ferritin.

Although these results might lead to the conclusion that a higher cutoff for serum ferritin should be used in the elderly, more information is to be gained by using multiple cut-points. The clinical usefulness of the likelihood ratios associated within different results of serum ferritin is illustrated in **Table V**. Table V examines four different scenarios: patients with low (5% to 20%), intermediate (40% to 60%), and high (80% to 95%) pre-test probability or prevalence of iron deficiency as an explanation for their anemia, as well as the population of the current study (in whom the prevalence of iron deficiency was 36%). The power of the serum ferritin level is made evident by examining the patients with intermediate probability, in whom the post-test probability of iron deficiency decreases to 8% to 16% if the serum ferritin level is greater than 100 μg/L, while a result of less than 18 μg/L increases the likelihood of iron deficiency to greater than 97%. Let us assume a physician is willing to diagnose a patient with a probability of 10% or less as not having iron-deficiency anemia, and a patient with a probability of 90% or more as having iron deficiency, without performing a bone marrow examination. Under these circumstances, a serum ferritin value of greater than 45 μg/L will obviate the necessity of a bone marrow aspiration in all patients with low prior probability; and a result of less than 18 μg/L in those with an intermediate prior probability, or less than 45 μg/L in those with a high prior probability, secures the diagnosis of iron deficiency.

The results depicted in the last column of Table IV suggest that, for clinicians dealing with populations similar to the one included in the present study, patients with values greater than 100 μg/L can be treated as not having iron deficiency, patients with values of less than 18 μg/L can be treated as having iron deficiency, and a bone marrow aspiration is necessary for diagnosis in those with intermediate values. Using this approach would lead to a diagnosis of iron deficiency in 21% of the patients, and exclusion of iron deficiency in 49%. Thus, bone marrow aspiration would be required in only 30%.

The present study has a number of strengths in comparison to previous investigations of the usefulness of laboratory tests in the diagnosis of iron deficiency. The sample represents a group of consecutive elderly patients presenting with anemia. We demonstrated the reproducibility of the interpretation of results of bone marrow aspiration, the procedure was undertaken in all patients, and the findings were interpreted by a hematologist unaware of the results of the laboratory investigations. We can therefore be confident of our conclusion that serum ferritin is the one peripheral blood test useful in the diagnosis of iron-deficiency anemia in the elderly; that the results should be interpreted differently from serum ferritin results in younger patients; and that when the information from the test is optimally utilized (by means of multi-level likelihood ratios), the test is extremely powerful in the diagnosis of iron-deficiency anemia.

ACKNOWLEDGMENT

We thank the following individuals for their help in data collection and preparation of this manuscript: Dr. Anne Benger for help with interpretation of bone marrow aspirates; and Sue Halcrow, Sandi Harper, Jenny Whyte, and Debbie Maddock for help with data collection, data processing, and manuscript preparation.

REFERENCES

1. Beck JR, Gibbons AB, Cornwell G, *et al:* Multivariate approach to predictive diagnosis of bone-marrow iron stores. Am J Clin Pathol 1979; 70: S665–S670.
2. Sheehan RG, Newton MJ, Frenkel EP: Evaluation of a packaged kit assay of serum ferritin and application to clinical diagnosis of selected anemias. Am J Clin Pathol 1978; 70: 79–84.
3. Krause JR, Stolc V: Serum ferritin and bone marrow biopsy iron stores. Am J Clin Pathol 1980; 74: S461–S464.
4. Ali MAM, Luxton AW, Walker WHC: Serum ferritin concentration and bone marrow iron stores: a prospective study. Can Med Assoc J 1978; 118: 945–946.
5. Mazza J, Barr RM, McDonald JWD, Valberg LS: Usefulness of the serum ferritin concentration in the detection of iron deficiency in a general hospital. Can Med Assoc J 1978; 119: 884–886.
6. Sorbie J, Valberg LS, Corbett WEN, Ludwig J: Serum ferritin, cobalt excretion and body iron status. Can Med Assoc J 1975; 112: 1173–1178.
7. Addison GM, Beamish MR, Hales CN, *et al:* An immunoradiometric assay for ferritin in the serum of normal subjects and patients with iron deficiency and iron overload. J Clin Pathol 1972; 25: 326–329.
8. Lipschitz DA, Cook JD, Finch CA: A clinical evaluation of serum ferritin as an index of iron stores. N Engl J Med 1974; 290: 1213–1216.
9. Walsh JR, Fredrickson M: Serum ferritin, free erythrocyte protoporphyrin, and urinary iron excretion in patients with iron disorders. Am J Med Sci 1977; 273: 293–300.
10. Pirie R: The influence of age upon serum iron in normal subjects. J Clin Pathol 1952; 5: 10–15.
11. Yip R, Johnson C, Dallman PR: Age-related changes in laboratory values used in the diagnosis of anemia and iron deficiency. Am J Clin Nutr 1984; 39: 427–436.
12. Powell DEB, Thomas JH: The iron binding capacity of serum in elderly hospital patients. Gerontol Clin (Basel) 1969; 11: 36–47.
13. Loria A, Hershko C, Konij N: Serum ferritin in an elderly population. J Gerontol 1979; 34: 521–525.
14. Nelson R, Chawla M, Connolly P, Laporte J: Ferritin as an index of bone marrow iron stores. South Med J 1978; 71: 1482–1484.
15. Bentley DP, Williams P: Serum ferritin concentration as an index of storage iron in rheumatoid arthritis. J Clin Pathol 1974; 27: 786–788.
16. Patterson C, Turpie ID, Benger AM: Assessment of iron stores in anemic geriatric patients. J Am Geriatr Soc 1985; 33: 746–767.
17. Awad MO, Berford AV, Grindulis KA, *et al:* Factors affecting the serum iron-binding capacity in the elderly. Gerontology 1982; 28: 125–131.
18. Lynch SR, Finch CA, Monsen ER, *et al:* Iron status of elderly Americans. Am J Clin Nutr 1982; 36: 1032–1045.
19. Department of Clinical Epidemiology and Biostatistics, McMaster University: Interpretation of diagnostic data: how to do it with simple maths. Can Med Assoc J 1983; 129: 22–29.
20. International Committee for Standardization in Haematology: The measurement of total iron and unsaturated iron binding capacity in serum. Br J Haematol 1978; 38: 281–287.
21. Luxton AW, Walker WHC, Gauldie J, Ali MAM, Pelletier C: A radioimmunoassay for serum ferritin. Clin Chem 1977; 23: 683–689.
22. Piomelli S, Young P, Gay G: A micro method for free erythrocyte protoporphyrin: the FEP test. J Lab Clin Med 1973; 81: 932–940.
23. Dacie SV, Lewis SM: Practical hematology, 6th ed. New York: Churchill Livingstone, 1984; 107–109.
24. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29–36.
25. Hanley JA, McNeil BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148: 839–843.