The role of iron status markers in predicting response to intravenous iron in haemodialysis patients on maintenance erythropoietin

Nicola Tessitore¹, Giovanni Pietro Solero², Giuseppe Lippi², Antonella Bassi², Giovanni Battista Faccini², Valeria Bedogna¹, Linda Gammouro¹, Giorgio Brocco², Giuseppe Restivo¹, Patrizia Bernich¹, Antonio Lupo¹ and Giuseppe Maschio¹, with the technical assistance of Enis Padovani

¹Divisione di Nefrologia and ²Laboratorio di Analisi Chimico-Cliniche, Azienda Ospedaliera di Verona, Verona, Italy

Abstract

Background. Iron deficiency (ID) is the main cause of hyporesponsiveness to erythropoietin in haemodialysis patients and its detection is of value since it is easily corrected by intravenous iron. Markers of iron supply to the erythron, including erythrocyte zinc protoporphyrin (Er-ZPP), percentage of hypochromic erythrocytes (Hypo), reticulocyte haemoglobin content (CHR) and soluble transferrin receptor (sTfR), may be more accurate predictors of ID than ferritin (Fe) and transferrin saturation (TSat), but relative diagnostic power and best threshold values are not yet established.

Methods. In 125 haemodialysis patients on maintenance erythropoietin, the diagnostic power of the above parameters was evaluated by ROC curve, multivariate regression, and stepwise discriminant analyses. Diagnosis of ID was based on haemoglobin response to intravenous iron (992 mg as sodium ferric gluconate complex over an 8-week period).

Results. Fifty-one patients were considered iron deficient (haemoglobin increase by 1.9 ± 0.5 g/dl) and 74 as iron replete (haemoglobin increase by 0.4 ± 0.3 g/dl). ROC curve analysis showed that all tests had discriminative ability with the following hierarchy: Hypo (area under curve W = 0.930, efficiency 89.6% at cut-off > 6%), CHR (W = 0.798, efficiency 78.4% at cut-off ≤ 29 pg), sTfR (W = 0.783, efficiency 72.4% at cut-off > 1.5 mg/l), Er-ZPP (W = 0.773, efficiency 73.0% at cut-off > 52 µmol/mol haem), TSat (W = 0.758, efficiency 70.4% at cut-off < 19%) and ferritin (W = 0.633, efficiency 64.0% at cut-off < 50 ng/ml). Stepwise discriminant analysis identified Hypo as the only variable with independent diagnostic value, able to classify 87.2% of patients correctly. Additional tests did not substantially improve diagnostic efficiency of Hypo > 6% alone.

Conclusions. In haemodialysis patients on maintenance erythropoietin, Hypo > 6% is the best currently available marker to identify those who will improve their response after intravenous iron. Cost-effectiveness suggests that this parameter should be a first-line tool to monitor iron requirements in clinical practice.

Keywords: erythropoietin; haemodialysis; iron deficiency; percentage hypochromic erythrocytes; reticulocyte haemoglobin content; soluble transferrin receptor

Introduction

Due to accelerated erythropoiesis coupled with ongoing dialysis-related blood losses, haemodialysis patients on maintenance erythropoietin (Epo) frequently develop iron-deficient erythropoiesis, resulting from depletion of iron stores (absolute iron deficiency) or failure to deliver adequate iron to bone marrow in spite of apparently appropriate iron stores (functional iron deficiency) [1].

The early and accurate detection of both absolute and functional iron deficiency (ID) is important because it is the most common cause of suboptimal response to Epo, and can easily be corrected by intravenous (i.v.) iron administration, which improves response to the hormone.

The response to i.v. iron, producing either an increased erythropoietic response at a constant dose of Epo or a decrease in Epo needs, is the best criterion for a 'gold standard' of iron-deficient erythropoiesis. It is additionally the most widely accepted reference standard for clinical practice, where the most...
important goal is to identify subjects who need iron to optimize Epo treatment [2].

Since traditional indices of iron status, including ferritin and transferrin saturation (TSat), are inaccurate in predicting iron-deficient erythropoiesis during Epo therapy at any proposed cut-off level [3], the usefulness of additional parameters, especially those allowing direct estimates of bone marrow iron availability, have been investigated.

Interest has been generated in the use of erythrocyte and reticulocyte indices, such as the percentage of hypochromic red blood cells (Hypo) and the reticulocyte haemoglobin content (Chr), made available by the flow cytometric haematological analysers Technicon H and ADVIA. Both assays provide direct insight into bone marrow iron supply and utilization by evaluating the degree of haemoglobinization. In addition, they can be performed simultaneously during routine blood counts with minimal or no incremental cost and no additional blood sampling.

An early study [4] in Epo-treated patients with Hypo >10% showing increased haemoglobin levels after i.v. iron with parallel reductions in hypochromia prompted the recently published European Best Practice Guidelines for the Management of Anaemia in Patients with Chronic Renal Failure to state that Hypo >10% is the most accurate marker for ID [5]. However, other studies, suggested that lower threshold values may be more appropriate [6–9].

Chr has also proved to be a reliable marker of iron-deficient erythropoiesis during Epo therapy, at cut-off values of <26 pg or <28 pg, and may be an early and more sensitive index than erythrocyte hypochromia due to the shorter life span of reticulocytes [9–11].

The role of erythrocyte zinc protoporphyrin (Er-ZPP), an abnormal end-product of haem synthesis that accumulates during limited bone marrow iron availability, has also been evaluated as a marker of iron status. Its determination is simple, rapid, and inexpensive, although its long elimination half-life limits its usefulness in detecting rapid changes in iron status. Er-ZPP has been shown to be a better marker of iron-deficient erythropoiesis than ferritin and TSat during Epo treatment, with a high sensitivity and specificity at a threshold value >90 µmol/mol haem [12]. However, recent studies have shown that Er-ZPP could not predict responses to i.v. iron supplementation over a wide number of cut-off levels, ranging from 50 to 110 µmol/mol haem [13].

In recent years the measurement of soluble transferrin receptor (sTfR) in serum has been proposed as a marker of adequacy of iron supply to the erythron [14,15]. However, it may be of less value in patients on Epo therapy because increased erythropoiesis itself raises sTfR levels [16].

Unfortunately, most studies examining serum sTfR in haemodialysis [14,17] failed to discriminate between the effects of Epo and iron deficiency, making it difficult to evaluate the usefulness of sTfR as a marker of ID during Epo treatment.

In summary, it has not been clearly established whether newer indices of iron status provide better results than the traditional tests [18,19]. In addition, it is unclear which test and which threshold value is most accurate, since studies addressing diagnostic accuracy in detecting ID during Epo therapy are limited in number [7–13,20,21] and report conflicting results.

The available data are difficult to interpret for two reasons. First, there is a bias in sample size, with populations too small for a meaningful statistical analysis [8–10,12], and second, these studies used different criteria for the diagnosis of ID, ranging from inaccurate tests such as low ferritin [19,20] and TSat [20,21] to the more appropriate haemoglobin [7,8], haematocrit [22], and reticulocyte response to i.v. iron challenge [10,11].

Although simultaneous evaluation of the different markers of bone marrow iron availability should provide better insight into the relative diagnostic values for ID, few studies have addressed this issue; almost all compared erythrocyte and reticulocyte indices, the majority favouring Chr [9,10,20], while some favoured Hypo [8] and others showed no preference [18,19].

Even fewer studies compared Er-ZPP and sTfR with erythrocyte and reticulocyte indices. Er-ZPP levels were shown to have no [13] or weak [22] correlations with Hypo and a weak correlation with sTfR [17]. The latter correlated well with Hypo [15].

The purpose of our study was (i) to compare the diagnostic power of established tests of bone marrow iron availability (Hypo, Chr, Er-ZPP, and sTfR) and of traditional indices (ferritin and TSat) to identify iron-deficient erythropoiesis subjects in a large unselected population of haemodialysis patients on maintenance Epo, and (ii) to identify best threshold values using the haemoglobin response to i.v. iron treatment as the ‘gold standard’.

To our knowledge, this is the first study to simultaneously compare the diagnostic power of all the available laboratory markers of iron status in haemodialysis patients on Epo.

Subjects and methods

Patients

A cohort of 125 chronic thrice-weekly haemodialysis patients on maintenance Epo therapy from two dialysis units in Verona (Policlínico and Ospedale Civile Maggiore Hospitals) and satellite dialysis units at San Bonifacio and Vareggio Hospitals were included in the study.

Seventy-one were males and 54 females, aged 31–84 years; 104 were on bicarbonate haemodialysis with low-flux cellulose or synthetic membranes, 19 were on acetate-free biofiltration with AN69 membranes, and two were on haemofiltration using PAN membranes. All had been treated for at least 5 months with recombinant human erythropoietin alpha, 70 intravenously and 55 by the subcutaneous route.
Patients with recent bleeding episodes, need of haemo-transfusions, clinically evident inflammatory and infectious diseases, malignancies, haemoglobinopathies, overt hyperparathyroidism (serum PTH above 500 ng/ml), aluminium overload (serum aluminium above 70 μg/ml), folate and vitamin B<sub>12</sub> deficiencies (serum folate below 7.0 nmol/l and B<sub>12</sub> below 156 pmol/l respectively), intercurrent events such as bleeding or infection during the iron challenge period, were excluded from the study.

Prior to the study a majority of patients received i.v. iron as sodium ferric gluconate complex in sucrose (Ferlitsi) on an as-needed basis according to TSat <20% and/or ferritin <100 ng/ml. No i.v. iron was provided in the 3 weeks prior to the study. The total amount of i.v. iron supplied in the 6 months prior to the study ranged from 0 to 1395 mg (median amount 475 mg).

Methods

Blood samples for laboratory testing were obtained prior to the first-of-the-week haemodialysis sessions. Standard haematology parameters, including reticulocyte count and erythrocyte and reticulocyte indices were assayed on the Advia 120 haematology analyser (Bayer Diagnostics, Germany). Samples were processed within 3 h in order to minimize storage-dependent changes in cell volume, which affect haemoglobin concentration and hence analytic variability of the Hypo assay. This instrument, using flow cytometry, allows determination of volume and haemoglobin concentration of individual erythrocytes by measuring the amount of laser light scattered at two different angles from each spherical cell. Erythrocytes with haemoglobin concentration <28 g/dl are considered to be hypochromic and values are expressed as percentages of total red blood cells (Hypo).

Reticulocytes, identified by staining RNA with the dye oxazine, are measured for volume and haemoglobin concentration. The reticulocyte haemoglobin content (CHR) is calculated as the product of volume and haemoglobin concentration, providing the mean of the whole population.

Serum ferritin, iron, and transferrin were measured by enzyme-linked immunosorbent assay, by photometry with Ferene-S as a chromogen after reduction with ascorbic acid (Bayer Diagnostic, UK) and by immunonephelometry respectively. Transferrin saturation (TSat in %) was calculated as serum iron (μg/dl)×70.9/serum transferrin (mg/dl).

Er-ZPP was measured by fluorometry (Shimadzu, RF-551) after erythrocyte lysis, deproteinization, extraction, and separation by high-performance liquid chromatography.

In a sub-population of 62 patients, serum stTFR levels were measured by a commercially available automated particle-enhanced immunonephelometric (PETIA) assay (Dade Behring, Marburg, Germany), using highly purified stTFR isolated from human serum as a calibrator. This group of patients was not different from the total population in terms of sex distribution (35 males and 27 females), age (28–78 years), haemoglobin concentrations, Epo dose, or prevalence of ID (33.9 vs. 40.8%, P = n.s.). Intra- and inter-assay variation coefficients for all iron status laboratory indices were always less than 4%.

All markers of bone marrow iron supply showed good biological and analytical reproducibility with the following coefficients of variation between values measured 2 days apart: CHr, 0.78±0.51% (n=30, range of values 24.9–34.1 pg); Er-ZPP, 3.63±1.54% (n=19, range of values 34–96 μmol/mol haem); stTFR, 4.55±3.55% (n=21, range of values 0.86–2.59 mg/l); and Hypo, 6.20±4.21% (n=62, range of values 2.5–14.7%).

Figure 1 shows the amplitude of limits of agreement of the two repeated measurements for Hypo and CHr according to the Bland and Altman analysis.

Serum aluminium, intact PTH, folate, and vitamin B<sub>12</sub> concentrations were measured by atomic absorption spectrometry, immunoradiometry, and radioimmunoassay respectively.

C-reactive Protein (CRP) levels were assessed in 87 patients and were measured by a commercially available PETIA assay (Dade Behring, Marburg, Germany).

Study design

The haemoglobin response to approximately 1 g i.v. iron at constant doses of Epo was used as the reference standard for ID [2] and, as proposed by Allegre et al. [23], subjects showing an increase in haemoglobin above or equal to 15% of baseline after i.v. iron were considered iron deficient, while those with lesser or no change in haemoglobin were considered iron replete.

Iron was administered as a slow (over 2 min) bolus injection at the end of each dialysis session as sodium ferric gluconate complex in sucrose, which was the only available i.v. iron preparation at our institution. Since this preparation is unstable, showing rapid release of iron, we used a single dose of 31 or 62 mg iron in relation to predialysis serum

![Image](image.png)

**Fig. 1.** Amplitude of limits of agreement of two repeated measurements at 2-day intervals of erythrocyte and reticulocyte indices according to Bland and Altman analysis (Lancet 1986; 1: 307–310). Dotted lines indicate 95% confidence interval.
transferrin levels (lower or higher than 170 mg/dl respectively), to avoid the risk of oversaturation of transferrin and ‘free iron’ reactions [24].

This protocol was based on results from a preliminary study examining the relationship between predialysis serum transferrin levels, dose of i.v. iron, and TSat, evaluated 10 min after i.v. iron administration. This study showed that 62 mg iron was consistently associated with transferrin oversaturation in subjects with predialysis transferrin <170 mg/dl (TSat 157.8 ± 37.8% at serum transferrin 147 ± 13 mg/dl, n = 6), but not in those with higher transferrin levels (TSat 78.5 ± 14.1% at serum transferrin 209 ± 31 mg/dl, n = 6). In contrast, the 31 mg dose did not lead to transferrin oversaturation at any transferrin level: TSat 62.3 ± 17.5% at serum transferrin 209 ± 32 mg/dl (n = 5) and 75.6 ± 20.3% at serum transferrin 141 ± 20 mg/dl (n = 6).

Haemoglobin and iron status indices, except Er-ZPP and sTfR, were evaluated every 2 weeks during the i.v. iron challenge period and treatment was continued until the end point for the diagnosis of ID was reached during at least two consecutive haemoglobin measurements, or if serum ferritin was above 600 ng/ml, a value associated with risk of iron overload, or if stable haemoglobin levels (i.e. less than 0.5 g/dl difference in two consecutive measurements) were achieved after a minimum amount of 1023 mg i.v. iron.

This protocol, resulting from concerns to avoid adverse effects associated with i.v. sodium ferric gluconate complex administration [24], could lead to differences in delivered iron and the length of the iron challenge period. However, this represents only a minor bias, since all subjects received at least 682 mg iron, an amount that is adequate for diagnostic purposes. In addition, frequent monitoring of haemoglobin levels allowed identification of events other than iron load that may have affected haemoglobin, leading to exclusion of patients from the study.

The amount of supplied iron ranged from 682 to 1770 mg (median of 992 mg) over a treatment period ranging from 6 to 19 weeks (median of 8 weeks).

After iron withdrawal, we re-evaluated Er-ZPP and sTfR levels and assayed haemoglobin every 2 weeks for an additional 6 weeks, to identify late responders.

The Epo dose was held constant throughout the iron challenge period in all except five patients, in whom it was halved due to brisk increase in haemoglobin. No adverse reactions were observed during the i.v. iron treatment period, except for a transient metallic taste reported by three patients.

Statistical analyses

To identify optimal tests and threshold values for predicting ID, receiver-operating characteristics (ROC) curve analysis was performed [25] by computing sensitivity and specificity of the different tests at various cut-off levels. Sensitivity was defined as the percentage of patients with ID having a positive test (true positive, TP), and specificity as the percentage of iron-replete patients having a negative test (true negative, TN). ROC curve analysis was performed by the Astute, DDU Software (The University of Leeds, UK). Diagnostic efficiency of different threshold levels, showing the probability that test and diagnosis agree, was calculated as the sum of TP and TN.

Since sensitivity, specificity, and efficiency are uncalibrated measures of test performance, and random values are strongly dependent on prevalence and level of the test, we also calculated their calibrated values as measures of test quality, called ‘kappa coefficients’: \( \kappa(1,0) \) was the calibrated form of sensitivity, \( \kappa(0,0) \) was the calibrated form of specificity, and \( \kappa(0,5) \) was the calibrated form of efficiency. Kappa coefficients were calculated as follows: \( \kappa(1,0) = \frac{SE - Q}{SE + Q} \), \( \kappa(0,0) = \frac{(SP - Q)}{(SP + Q)} \) and \( \kappa(0,5) = \frac{(EFP - P - Q')}{(1 - P - Q')} \), where \( P \) = prevalence of the disease, i.e. the percentage of patients with ID, \( P' = 1 - P \), i.e. the percentage of iron-replete patients, \( Q \) = level of the test, i.e. the sum of TP and FP (false positive, i.e. percentage of iron-replete patients who test positive) and \( Q' = 1 - Q \), i.e. the sum of TN and FN (false negative, i.e. percentage of patients with iron deficiency who test negative).

Multivariate regression analysis was performed to assess the predictive value of response to i.v. iron, the dose of Epo and amount of iron administered in the 6 months prior to the study. A stepwise discriminant function analysis was used to assess the diagnostic accuracy of the different tests and their associations. Data evaluation was carried out using the SPSS software package (SPSS Inc., Chicago, USA).

All data, based on single measurements, are expressed as means ± standard deviation unless otherwise noted. Differences between groups were evaluated by non-parametric Mann–Whitney U test for the analysis of discrete variables, and a P value < 0.05 was considered statistically significant.

Results

After the i.v. iron challenge, 51 patients (40.8%) showed an increase in haemoglobin equal or higher than 15% from baseline, increasing by 1.86 ± 0.49 g/dl, and were considered iron deficient (responders). The remaining 74 patients (59.2%) had no change or less than 15% increase from baseline haemoglobin, increasing by 0.43 ± 0.34 g/dl, and were considered iron replete (non-responders).

Responders and non-responders received the same median i.v. iron (990 mg, range 682–1440 mg, and 1020 mg, range 744–1770 mg respectively) over the same median period of time (7 weeks, range 6–12, and 8 weeks, range 6–19 weeks respectively).

Table 1 summarizes baseline haemoglobin, Epo dose, and iron status markers in the two groups of patients. Responders had lower haemoglobin levels at higher Epo doses. They also showed lower ferritin, Tsat, and CHr, and higher Hypo, Er-ZPP and sTfR concentrations than non-responders. Both groups had a similar prevalence of patients with ferritin

<p>| Table 1. Baseline haemoglobin, Epo dose, and laboratory markers of iron status |
|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Responders</th>
<th>Non-responders</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin (g/dl)</td>
<td>9.28 ± 0.73</td>
<td>10.36 ± 0.87</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Epo (u/week)</td>
<td>8431 ± 3219</td>
<td>6378 ± 2662</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Ferritin (ng/ml)</td>
<td>165 ± 115</td>
<td>225 ± 127</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Transferrin saturation (%)</td>
<td>17.7 ± 6.6</td>
<td>25.4 ± 8.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CHr (pg)</td>
<td>29.0 ± 2.6</td>
<td>31.6 ± 2.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hypochromic RBC (%)</td>
<td>12.1 ± 8.4</td>
<td>2.7 ± 1.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Er-ZPP (umol/mol haem)</td>
<td>68.7 ± 25.2</td>
<td>47.6 ± 18.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>sTfR (mg/l)</td>
<td>1.86 ± 0.65</td>
<td>1.38 ± 0.38</td>
<td>&lt;0.005</td>
</tr>
</tbody>
</table>
<100 ng/ml, i.e. with depleted iron stores. The prevalence was 35.3% in responders and 20.2% in non-responders, \( P = \text{n.s.} \).

The changes induced by i.v. iron challenge on haemoglobin levels and laboratory indices of iron status are shown in Figure 2 and Table 2 respectively. Haemoglobin significantly increased in responders from the 4th week of treatment but did not change in non-responders.

Laboratory markers of iron status changed in the expected direction in responders, except for Er-ZPP, which did not change, while ferritin, Tsat, and CHr only increased in non-responders.

Responders and non-responders had a high prevalence of elevated CRP levels (\( > 6 \) mg/l): 37.1% (13 of 35) and 26.9% (14 of 52) respectively. Subjects with normal and high CRP values had similar haemoglobin increases after the i.v. iron load in both responders (1.89 ± 0.51 vs 1.79 ± 0.5 g/dl, \( P = \text{n.s.} \)) and non-responders (0.44 ± 0.33 vs 0.42 ± 0.34 g/dl, \( P = \text{n.s.} \)). In addition, non-responders with normal and high CRP values had similar baseline haemoglobin, Epo dose, ferritin, Tsat, Hypo, CHr, Er-ZPP, and sTfR (10.3 ± 0.9 vs 10.4 ± 0.9 g/dl, 6236 ± 2842 vs 6142 ± 2656 u/week, 198 ± 112 vs 287 ± 153 ng/ml, 24.8 ± 8.6 vs 27.5 ± 9.4%, 2.9 ± 2.1 vs 3.0 ± 1.0%, 31.4 ± 1.9 vs 32.6 ± 2.2 pg, 47.5 ± 18.5 vs 45.2 ± 13.1 \( \mu \text{mol/mol haem} \) and 1.48 ± 0.40 vs 1.34 ± 0.21 mg/l respectively).

All tests showed discriminative ability, with an area under the curve (W) significantly different from 0.5. The test with the largest area under the curve was Hypo (\( W = 0.9296, SE = 0.0233, P < 0.0001 \)), followed by CHr (\( W = 0.7976, SE = 0.0425, P < 0.0001 \)), sTfR (\( W = 0.7834, SE = 0.0591, P < 0.0001 \)), Er-ZPP (\( W = 0.7734, SE = 0.0748, P < 0.0001 \)), Tsat (\( W = 0.7576, SE = 0.0429, P < 0.0001 \)) and ferritin (\( W = 0.6329, SE = 0.0605, P < 0.05 \)).

Statistically significant differences between the area under the curve were observed for Hypo vs ferritin (\( P < 0.001 \)) and Tsat (\( P < 0.05 \), and for CHr vs Ferritin (\( P < 0.05 \)).

The best threshold values for detecting ID were >6% for Hypo, ≤29 pg for CHr, >1.5 mg/l for sTfR, >52 \( \mu \text{mol/mol haem} \) for Er-ZPP, <19% for Tsat, and <50 ng/ml for ferritin.

Using multivariate regression analysis, the only significant predictors of ID were Hypo (\( P < 0.001 \)) and the amount of i.v. iron administered in the 6 months prior to the study (\( P < 0.05 \)). Stepwise discriminant analysis showed Hypo to be the only test with an independent diagnostic value for ID (Wilks’ lambda 0.611, \( F = 78.385, P < 0.0001 \)), being able to classify 87.2% of cases correctly.

In Table 3, individual tests and their combinations that showed the best diagnostic performances are listed in hierarchical order, according to the calibrated value for efficiency, \( \kappa(0.5,0) \). Data on the efficiency of the currently used cut-off levels of the various tests are also included.

A Hypo >6% was the best predictor of ID, either by itself or in combination with other tests. Its combination with CHr ≤29 pg produced a minor improvement in sensitivity and efficiency, while combination with...
Table 3. Diagnostic characteristics of the different tests and their associations in predicting iron deficiency

<table>
<thead>
<tr>
<th>Test</th>
<th>Efficiency (%)</th>
<th>$\kappa_{(0.5)}$ (%)</th>
<th>Sensitivity (%)</th>
<th>$\kappa_{(1.0)}$ (%)</th>
<th>Specificity (%)</th>
<th>$\kappa_{(0.0)}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypo &gt; 6 or CHR $\leq$ 29</td>
<td>90.4</td>
<td>80.0</td>
<td>86.3</td>
<td>77.5</td>
<td>93.2</td>
<td>82.6</td>
</tr>
<tr>
<td>Hypo &gt; 6</td>
<td>89.6</td>
<td>78.1</td>
<td>82.4</td>
<td>72.1</td>
<td>94.6</td>
<td>85.3</td>
</tr>
<tr>
<td>Hypo &gt; 6 or ferritin &lt; 50</td>
<td>86.4</td>
<td>71.7</td>
<td>82.4</td>
<td>70.7</td>
<td>89.2</td>
<td>73.0</td>
</tr>
<tr>
<td>Hypo &gt; 8</td>
<td>84.8</td>
<td>66.8</td>
<td>64.7</td>
<td>51.5</td>
<td>98.6</td>
<td>95.2</td>
</tr>
<tr>
<td>Hypo &gt; 6 or TSat $\leq$ 19</td>
<td>83.2</td>
<td>66.8</td>
<td>96.1</td>
<td>91.7</td>
<td>74.3</td>
<td>51.3</td>
</tr>
<tr>
<td>Hypo &gt; 4</td>
<td>81.6</td>
<td>63.0</td>
<td>86.3</td>
<td>73.7</td>
<td>78.4</td>
<td>55.0</td>
</tr>
<tr>
<td>Hypo &gt; 6 or Er-ZPP $\geq$ 52</td>
<td>80.0</td>
<td>60.5</td>
<td>94.4</td>
<td>88.3</td>
<td>71.9</td>
<td>46.0</td>
</tr>
<tr>
<td>Hypo &gt; 6 or sTTR $\leq$ 1.5</td>
<td>77.4</td>
<td>53.9</td>
<td>85.7</td>
<td>73.2</td>
<td>73.2</td>
<td>42.6</td>
</tr>
<tr>
<td>CHR $\leq$ 29</td>
<td>78.4</td>
<td>52.9</td>
<td>56.9</td>
<td>40.8</td>
<td>93.2</td>
<td>75.0</td>
</tr>
<tr>
<td>Hypo $\geq$ 10</td>
<td>77.6</td>
<td>49.3</td>
<td>45.1</td>
<td>32.7</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>sTTR $\leq$ 1.5 or TSat $\leq$ 19</td>
<td>74.1</td>
<td>49.3</td>
<td>90.5</td>
<td>79.7</td>
<td>65.8</td>
<td>35.7</td>
</tr>
<tr>
<td>sTTR $\leq$ 1.5</td>
<td>74.2</td>
<td>47.3</td>
<td>80.9</td>
<td>64.3</td>
<td>70.7</td>
<td>37.4</td>
</tr>
<tr>
<td>CHR $\geq$ 30</td>
<td>74.4</td>
<td>47.1</td>
<td>68.6</td>
<td>46.9</td>
<td>78.4</td>
<td>47.0</td>
</tr>
<tr>
<td>Er-ZPP $\geq$ 52</td>
<td>73.0</td>
<td>45.8</td>
<td>80.6</td>
<td>62.0</td>
<td>68.7</td>
<td>36.3</td>
</tr>
<tr>
<td>CHR $\geq$ 28</td>
<td>72.8</td>
<td>38.2</td>
<td>37.2</td>
<td>24.6</td>
<td>97.3</td>
<td>83.9</td>
</tr>
<tr>
<td>TSat $\leq$ 19</td>
<td>70.4</td>
<td>37.8</td>
<td>58.8</td>
<td>34.8</td>
<td>78.4</td>
<td>41.3</td>
</tr>
<tr>
<td>Ferritin $&lt; 100$ or TSat $&lt; 20$</td>
<td>64.0</td>
<td>28.3</td>
<td>68.6</td>
<td>35.6</td>
<td>60.8</td>
<td>23.4</td>
</tr>
<tr>
<td>Ferritin $\leq$ 50</td>
<td>64.0</td>
<td>16.3</td>
<td>19.6</td>
<td>9.5</td>
<td>94.6</td>
<td>51.8</td>
</tr>
<tr>
<td>Er-ZPP $&gt; 90$</td>
<td>67.0</td>
<td>15.1</td>
<td>13.9</td>
<td>7.4</td>
<td>96.9</td>
<td>55.7</td>
</tr>
<tr>
<td>CHR $\leq$ 26</td>
<td>63.2</td>
<td>15.1</td>
<td>9.8</td>
<td>6.0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ferritin $&lt; 100$</td>
<td>60.8</td>
<td>15.1</td>
<td>35.3</td>
<td>11.9</td>
<td>78.4</td>
<td>20.6</td>
</tr>
</tbody>
</table>

TSat $< 19\%$ and Er-ZPP $> 52 \mu$mol/mol haem produced a substantial increase in sensitivity, but at the expense of reduced specificity and efficiency.

**Discussion**

Iron deficiency has emerged as the major factor that limits response to Epo in haemodialysis patients, and aggressive treatment by i.v. iron has been advocated to optimize the use of the hormone [26]. This was confirmed by our finding of a high prevalence of ID (approximately 41%) in an unselected population of haemodialysis patients on maintenance Epo who were non-aggressively and unevenly treated with i.v. iron on an as-needed basis according to low ferritin and TSat values. These patients showed hyporesponsiveness to the hormone and improvement in their haemoglobin levels by 1 to 3 g/dl, after an average supplementation of 1 g i.v. iron.

Subclinical inflammatory states, as assessed by elevated CRP levels, did not appear to affect haemoglobin responses to i.v. iron in responders and non-responders and therefore did not interfere with our diagnostic approach. In addition, the finding that responses to Epo and that levels of markers of iron utilization at the erythron were independent of CRP levels in non-responders, suggests that even in patients with increased CRP concentrations, the lack of haemoglobin response to i.v. iron is not the result of an inflammatory process, but instead may identify a state of iron repletion.

An accurate identification of patients who will or will not respond to i.v. iron has relevant clinical and economical implications, since it allows improved response to Epo and avoidance of unnecessary risks associated with iron overload and overzealous therapy with parenteral iron [24]. Therefore, tests revealing threshold values with both acceptable sensitivity and acceptable specificity, i.e. efficient tests, are required, and efficiency should be the primary selection criterion.

While there is consensus that conventional indices of iron status are poorly correlated to ID, the diagnostic role of the newer indices of iron availability to bone marrow and their relative efficiency and best cut-off values remain to be established. This is because studies evaluating diagnostic power are limited and have biased designs, including the choice of reference standard, sample size, and patient selection criteria.

Our study, which allowed a simultaneous evaluation of four established markers of iron supply to bone marrow (Hypo, CHR, Er-ZPP and sTTR) was adequate in size for proper statistical analysis and utilized a widely accepted ‘gold standard’ for ID, should provide useful new information.

ROC curve analysis showed that all markers of iron status were able to discriminate iron-deficient from iron-replete patients. Curve analysis also indicated a hierarchy of tests, with Hypo having the wider area under the curve followed by CHR, sTTR, Er-ZPP, TSat and ferritin.

The notion that Hypo is the most accurate marker was confirmed by multivariate regression analysis as well as by stepwise discriminant regression analysis, which identified this as the only variable with independent diagnostic value for ID with the ability to correctly classify 87% of the patients, indicating also that test combinations would not improve diagnostic performance.

Our data are in agreement with the recent European Best Practice Guidelines for the Management of Anemia in Patients with Chronic Renal Failure [5],
which favour Hypo as the best marker of iron-deficient erythropoiesis. Indeed, Hypo was the only test that predicted the presence of ID with more than 80% certainty at cut-off values ranging from 4 to 8%.

The best combination of sensitivity and specificity was at a threshold > 6% (κ(0.5, 0) 78.1%) instead of the currently accepted cut-off > 10%, which showed absolute specificity but unacceptably low sensitivity. Our results are in agreement with studies suggesting a better performance of Hypo at threshold levels <10%, in the range of 2.5–5% [6,7,9,18]. However, threshold values <6% allowed only modest increases in sensitivity at the expense of a marked decrease in specificity.

The other markers of bone marrow iron availability showed similar diagnostic capacities with the most efficient threshold values significantly lower than Hypo (κ(0,5,0) ranging from 52.9 to 45.8%), including CHr, which has been proposed as the earliest and most sensitive marker for ID [9–11,20].

In this study, CHr was a reliable marker of ID, but was inferior to Hypo even at its best cut-off value ≤29 pg (κ(0,5,0) 52.9%, with low sensitivity and high specificity). This is at variance with the reports by Fishbane et al. [10] and Cullen et al. [9], who favoured CHr over Hypo. It should be noted that the results of these studies should be considered preliminary since the population sizes appeared insufficient for accurate statistical analysis. These differences can also be explained by the different criteria used to diagnose ID. Fishbane et al. [10] used an increase in corrected reticulocyte index of greater than 1 unit during 2 weeks after single dose i.v. iron infusion, and Cullen et al. [9] used TSat ≤15%. In contrast, we used the increase in haemoglobin after prolonged i.v. iron supply, which may identify different degrees and/or different stages of ID.

Our finding that the combination Hypo > 6% or CHr ≤29 pg showed the best efficiency value for ID (κ(0,5,0) 80.0%, with an improvement in sensitivity compared to Hypo alone) suggests that the two tests, which can be performed simultaneously, can be considered complementary. Hypo, as an indicator of chronic iron requirements, may better apply to patients on steady-state erythropoietic stimulation, as in the present study, while CHr may be more useful in reflecting acute changes in iron availability.

Both sTfR and Er-ZPP had similar discriminative abilities for ID, and were slightly better than TSat. Our data confirm that Er-ZPP has discriminative power (i.e. can be used as a marker) for ID during Epo treatment. Its best threshold value was > 52 μmol/mol haem, but it had no value that could be used to separate responders from non-responders, even over values ranging from 50 to 90 μmol/mol haem (efficiency ranging from 67.0 to 73.0%). This represents an important limitation to its utilization. In addition, its relative insensitivity to detect rapid changes in iron status was confirmed by our finding of no change in Er-ZPP in iron-depleted patients after acute i.v. iron challenge.

The usefulness of Er-ZPP is improved when it is used in association with Hypo. The combination Hypo > 6% or Er-ZPP > 52 μmol/mol haem significantly increased the sensitivity of Hypo > 6% (94.4%) but was associated with a substantial drop in specificity (71.9%) and efficiency (80.0%). Similar but somewhat better diagnostic performances can be obtained by the association Hypo > 6% or TSat <19%, which provides the highest sensitivity (96.1%). Both combinations of tests might be of value in clinical practice, where an optimally sensitive test is to be preferred. This is because the benefits of i.v. iron are well established whereas the risks linked to iron overload are still matters of debate.

Although others have suggested that sTfR is not a reliable marker of ID during Epo treatment [16], our study showed this assay to have an acceptable discriminative ability, with the most efficient cut-off at sTfR > 1.5 mg/l (κ(0,5,0) 47.3%). In our population of patients on maintenance erythropoietin, who are likely to be in steady-state erythropoietic stimulation, serum levels of sTfR may be useful in monitoring iron status. This conclusion is supported further by our finding of modest but significant decreases in serum sTfR levels in iron-depleted patients after i.v. iron challenge. This result clearly favours sTfR over Er-ZPP. The Er ZPP test shows a similar diagnostic performance but is more inert than sTfR in detecting iron-status changes.

In addition, the association sTfR > 1.5 mg/l or TSat <19% had a high sensitivity (90.5%) in detecting ID with acceptable efficiency, indicating that this combination may be a useful tool in identifying patients who will respond to i.v. iron when erythropoietin and reticulocyte indices are not available. This combination is preferred over TSat, which even at the most efficient threshold value (TSat <19%) showed inadequate efficiency, sensitivity, and specificity.

Our data confirmed that ferritin predicts ID poorly, even at the best cut-off value of <50 ng/ml (κ(0,5,0) 15.1%) and that the threshold values suggested by the NKF-DOQI guidelines for iron repletion therapy (ferritin <100 ng/ml and TSat <20%) [26] produced an unacceptably low accuracy for detecting responses to i.v. iron.

Since cost, availability, and reproducibility are also important for implementing tests in clinical practice, we evaluated the cost profile and both analytical and biological variability of the different markers. Hypo was the most convenient test. It was performed at no increased costs except for those of routine blood counts, whereas the other tests necessitated additional expenses (in Euros, 0.83 for CHr, 5.17 for sTfR, 7.69 for Er-ZPP, 10.48 for TSat, and 13.00 for ferritin).

All markers of iron supply to bone marrow showed good reproducibility. In our study, using fresh blood samples, CHr showed an extremely low coefficient of variation (mean value <1%), whereas Hypo, which has high variability when samples are stored for prolonged periods of time, showed an acceptable biological and analytical variability (coefficient
of variation 6.2%). The other tests had coefficients of variation less than 5%, indicating excellent reproducibility.

In conclusion, although all markers of iron status in haemodialysis patients on maintenance Epo show discriminative ability for iron-deficient erythropoiesis, the best laboratory marker for those who will improve haemoglobin response after i.v. iron treatment, and who are hence considered iron deficient, is Hypo at a threshold level >6%. Hypo demonstrated excellent diagnostic efficiency, inexpensiveness, and good reproducibility, indicating that it could be a first-line test for monitoring iron requirements in clinical practice.

The diagnostic accuracy of Hypo can be slightly improved by simultaneous evaluation of CHr, which itself could be a suitable second choice because of its acceptable diagnostic accuracy, low cost, and high reproducibility.

When erythrocyte and reticulocyte assays are not available, sTfIR and its combination with TSat appear to be useful alternatives.

Acknowledgements. Part of this work was presented at the 31st Annual Meeting of the American Society of Nephrology, Philadelphia, Pennsylvania, USA, 25–28 October 1998, and at the 57th ERA-EDTA Congress (First combined Congress with EKRA), Nice, France, 17–20 September 2000. The authors wish to thank the patients who took part in the study, the nursing staff of the dialysis units, Dr Paolo Beltrami for creative advice, and Dr Luisa Zanolla for statistical advice.

References


Received for publication: 01.10.00 Accepted in revised form: 31.1.01