Health-related quality of life in patients with iron deficiency anemia: impact of treatment with intravenous iron

William E Strauss1
Michael Auerbach2
1AMAG Pharmaceuticals, Inc., Waltham, MA, USA; 2Auerbach Hematology and Oncology, Baltimore, MD, USA

Abstract: Most physicians appear to be aware of the health consequences of advanced anemia, especially in the acute setting, frequently responding with a not inconsequential therapeutic default of transfusion. In contrast, the profound impact that chronic anemia, of any degree, may have on a patient’s performance is underappreciated. The focus of this review is to 1) delineate the consistent and broad impact of anemia on patient quality of life as documented by multiple well-validated patient-reported outcome instruments and 2) demonstrate the essential normalization of the debilitation as assessed by these instruments following the administration of intravenous iron.

Keywords: anemia, health-related quality of life, intravenous iron, iron deficiency anemia

Introduction

Anemia, and specifically its most common cause – iron deficiency (ID), is a frequent cause of morbidity in both developed and developing nations alike. Yet, generally, the primary focus is on the “degree of anemia”. Too frequently there is an underappreciation of the impact ID has on patients. Attention to quality of life (QOL) by assessment of patient-reported outcome (PRO) instruments permits an appropriate refocusing of attention. In the following review, we will aim to first provide a brief overview of such instruments which have been utilized in the evaluation of health-related QOL (HRQOL) both at baseline and in response to treatment of IDA, followed by a summation of currently available data on the impact of treatment with intravenous (IV) iron on QOL.

While there are many causes of anemia, it is estimated that at least 50% of cases are due to ID.1 Additionally, iron deficiency anemia (IDA) may coexist with many other causes of anemia, such as chronic disease or inflammation, referred to as iron-restricted erythropoiesis or functional IDA.2 Irrespective of the underlying cause, signs and symptoms of anemia are consistent. Although anemia may impact any organ or tissue, such as skeletal muscle reducing the exercise capacity,3,4 fatigue or a similar term is the most common and consistent symptom.5

Before delving into a discussion of the impact of anemia on patient QOL and normalization of the debilitation upon treatment with IV iron, it is useful to review the instruments and tools that have been employed to assess HRQOL, as well as evidence establishing their validity. Fatigue is not specific to anemia, but is a common symptom associated with many chronic diseases.4 In addition to being multifactorial, fatigue is also challenging to define. The Center for Biobehavioral Studies of Fatigue Management defines fatigue as “the awareness of a decreased capacity of physical
and/or mental activity”. However, despite these challenges, almost all QOL instruments assessing HRQOL in patients with anemia focus on fatigue or a similar term because it is the symptom that has the greatest impact. These divergent aspects require clarity in the description of fatigue within the instrument. Many instruments use the term “fatigue” alone; others use “weariness”, “tiredness”, or “feeling worn out”. Most useful is the construct of a continuum with fatigue at one end and energy or vitality at the other. As noted later in this review, while such fatigue or tiredness has a direct impact, such as reduced activity level, exercise capacity, and difficulty concentrating, its effects are broader, with measurable aspects on overall physical and emotional well-being.

HRQOL instruments

The following is not intended as a comprehensive review of all HRQOL instruments, but rather as a representative sample of major well-validated tools assessing fatigue and other HRQOL symptoms of anemia. Such instruments can be divided into “condition-specific” and “generic” QOL assessments.

Condition specific

Functional Assessment of Cancer Therapy (FACT)-Anemia and FACT-Fatigue

Examples of PRO instruments are the FACT Measurement System and its subscales, the FACT-Anemia (FACT-An) scale and the FACT-Fatigue scale. Yellen et al developed the subscales while recognizing there was not a validated questionnaire comprehensively assessing the full spectrum of anemia symptoms. The FACT measurement system is a set of questions assessing varying aspects of QOL associated with cancer and other chronic illnesses. The basic FACT questionnaire contains >250 questions tailored to the specific condition being assessed, with individuals rarely asked >60 questions. FACT-An and FACT-Fatigue were developed to assess anemia in patients with cancer; however, they have subsequently been used for other conditions. FACT-Fatigue was developed and validated as a subcomponent of FACT-An, which consists of the FACT-Fatigue subscale (13 items) and 7 additional questions on nonfatigue items combined with 28 core FACT-General questions. The FACT-Fatigue subscale was anticipated by its designers to become a reliable and valid stand-alone instrument to assess fatigue. The condition-specific subscales (FACT-Fatigue and FACT-An) were developed using semi-structured interviews with patients who had anemia associated with cancer and oncologists. These subscales were subsequently validated in an additional sample of 50 patients with cancer who had hemoglobin levels between 7 and 15.9 g/dL. Both the FACT-An and the FACT-Fatigue subscales showed high internal consistency and test–retest reliability. The FACT-Fatigue subscale score differed according to hemoglobin levels (P≤0.05) and patient-rated performance status (P≤0.001).

Since their initial development and validation, the FACT-An and the FACT-Fatigue have been used extensively in the oncology field. A direct correlation between the improvement of hemoglobin levels and HRQOL, measured by the FACT-An, has been observed in both clinical trials and prospective community-based studies treating cancer-related anemia. Positive correlations between change in the FACT-An and change of hemoglobin levels have been reported (FACT-General: r=0.26, P=0.006; FACT-An: r=0.34, P<0.001; FACT-Fatigue: r=0.22, P=0.017). Over the last two decades, FACT has evolved into the Functional Assessment of Chronic Illness Therapy (FACIT) to reflect expansion of original questionnaires focused on cancer into other chronic illnesses (note: throughout this review, we will refer to either FACT or FACIT depending on the terminology used in the original publication). The association between hemoglobin concentration and FACT-Fatigue scores has been observed in studies involving other patient populations, such as those with chronic kidney disease (CKD) and primary hip arthroplasty. In a randomized trial assessing the efficacy of darbepoetin to reverse chemotherapy-induced anemia in patients with cancer, improvement in hemoglobin level was associated with score changes on the FACIT-Fatigue. A linear relationship was observed, with the smallest mean change in FACT-Fatigue score observed in the subgroup with no improvement in hemoglobin level and the largest mean change in FACT-Fatigue score in those with the most improvement in hemoglobin level. The responsiveness of the FACT-Fatigue scores to changes in hemoglobin level was also observed in patients with CKD.

To confirm the validity of the FACT-Fatigue scale before and after treatment of IDA due to multiple causes with IV iron, a series of exercises were conducted to confirm content and psychomotor validity. Qualitative interviews showed that fatigue is a central concern to patients with anemia, and the FACT-Fatigue scale sufficiently assessed this problem. Psychometric assessment demonstrated that the FACT-Fatigue scale was stable over time (intraclass correlation coefficient, 0.87) and internally consistent (α=0.93). The scale demonstrated convergence with other conceptually
relevant scales such as the 36-Item Short Form Health Survey (SF-36) Vitality (VT) scale ($r=0.74$).

**Generic PRO instruments**

**36-Item Short Form Health Survey**

Clearly, the most widely used generic PRO tool is the SF-36. This is a well-standardized generic measure of HRQOL used in healthy and various diseased populations. It consists of 36 questions measuring eight health domains that include physical and emotional aspects. These eight domains are combined into two component summary measures: the Physical Component Summary (PCS) and the Mental Component Summary scores. Given the widespread use of the SF-36 in various populations, multiple exercises to assess internal consistency and test–retest validity have been performed.

Previous studies of patients with anemia have documented that almost all of the SF-36 domains are impacted, with the greatest perturbation involving those related to physical activity, such as VT and physical function. However, as noted below, even those domains focusing on role limitations due to emotional health and overall general health perception are affected. The domain scores correlate with hemoglobin scores and response to treatment. Certain subscales of “generic” HRQOL instruments have been used as condition-specific assessments in patients with anemia. The foremost example is the VT scale of the SF-36, which includes four questions exploring the frequency and intensity of tiredness, essentially a continuum between “tiredness” and “full of energy”. The VT scale has been used along with a number of other HRQOL tools, such as the FACIT-Fatigue scale, in patients with anemia and has demonstrated excellent correlation with similar instruments measuring fatigue, such as a fatigue scale developed and published by Wolfe et al. Using a visual analog scale (VAS) to measure the degree of fatigue. Multiple studies have demonstrated the correlation of the SF-36 VT scale and the degree of anemia, with increasing scores tightly correlated with increasing hemoglobin levels in various populations.

The SF-36 is also useful in deriving population-based utility values by applying the Short Form Six Dimension (SF-6D) algorithm. Utility values represent preference-based weights of a health state with values ranging from 1 (full health) to 0 (dead), with some health states categorized as worse than dead. The SF-6D is based on eleven items from the SF-36, representing six multilevel dimensions that describe 18,000 health states. The SF-6D can be combined with survival data to produce a measure of quality-adjusted life years. Any patient who completes the SF-36 can be uniquely classified according to the SF-6D for economic analyses. The SF-6D can be regarded as a continuous outcome scored on a scale of 0.29–1.00, where 1.00 indicates “full health”, with a reported mean minimally important difference (MID) of 0.041. In the US, population norms of 0.75–0.80 (varying by age group and sex) have been reported. The global impact of IDA on HRQOL was reflected in the SF-6D utility value of 0.62, −0.15 points below the population mean, exceeding a clinically meaningful decrement of 0.041.

**Linear analog scale assessment**

Another generic PRO assessment using a different approach is the linear analog scale assessment (LASA). For this instrument, patients are asked to rate their perception of three measures (energy level, ability to perform activities of daily living [ADL], and overall QOL) over the past week on a 100 mm VAS, where 0 is worst and 100 is best. The LASA has been used to evaluate chemotherapy-associated anemia as well as the response of cancer-associated anemia to erythropoietin. Such studies have confirmed the correlation of change in hemoglobin levels and LASA scores.

One advantage of generic instruments is applicability across multiple conditions and diseases. Another equally important facet, as noted, is their ability to assess the impact these conditions have on broader aspects of HRQOL, such as overall physical functioning and emotional well-being. These latter aspects may not be optimally assessed by condition- or disease-specific instruments.

**Level of disability or decrement in function associated with anemia**

The usual focus of the impact of any therapeutic intervention for anemia, whether by erythropoiesis-stimulating agents (ESAs) or iron, is how much it made things better, be it by change in hemoglobin or markers of HRQOL. More important perhaps is an understanding of the severity and breadth of the impact on HRQOL anemia has.

The relationship between the SF-36 domains and hemoglobin levels (<11, 11–<12, 12–13, and >13 g/dL) was assessed in 1,200 patients with stage 3, 4, and 5 CKD from seven centers. There was a statistically significant increase in all four physical domains, the energy/VT domain, and the physical composite score of the SF-36 with increasing hemoglobin levels. The largest increases occurred between the <11 and 11–12 g/dL groups.

The impact of anemia in a population of independent elderly patients was assessed in a cross-sectional observational study conducted at three sites: an academic geriatric
practice, a hospital-based geriatric outpatient unit, and a community-based multispecialty internal medicine group.\textsuperscript{19} HRQOL and functional status were measured using the SF-36 and the FACT-An. Disability and depression were assessed using the Instrumental Activities of Daily Living and the Geriatric Depression Scale questionnaires, respectively. The mean SF-36 PCS score was 38.9 in patients with anemia (hemoglobin level <12 g/dL for women and <13 g/dL for men) and 44.1 in patients without anemia ($P<0.001$). Anemia was associated with greater fatigue ($P<0.001$), lower handgrip strength ($P=0.014$), increased number of disabilities ($P=0.005$), and more depressive symptoms ($P=0.002$). Multivariate regression analysis, adjusted for demographic and clinical characteristics, demonstrated strong associations for reduced hemoglobin and poorer HRQOL across multiple domains.

Minimally clinically important difference

How does a clinician interpret whether the changes in a value for a given HRQOL instrument are clinically important? This question is relevant whether the assessment determines a response to a therapeutic intervention or variation from what is considered “normal” for a representative cross section, where “normal” does not imply “healthy or without disease”, but rather what a large unselected population of individuals would report. One frequently used definition when applying the concept of MID to interventions is “the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate… a change in the patient’s management”.\textsuperscript{27} An alternative to the MID is a distribution-based approach based on calculating the effect size. However, this approach potentially underestimates a clinically important change and is more appropriately applied to meta-analyses.\textsuperscript{28} For this reason, we favor the MID approach. Norman et al\textsuperscript{29} conducted a systematic analysis of studies using multiple HRQOL instruments assessing a wide variety of chronic conditions and found a reliable and consistent finding that the MID was half the standard deviation (SD).

Cella et al\textsuperscript{30} conducted a series of analyses including three samples of patients: 50 patients with a variety of cancers undergoing treatment, 131 patients with cancer participating in a longitudinal evaluation of fatigue and QOL during chemotherapy, and 2,402 patients with a variety of cancers who were enrolled in an open-label nonrandomized clinical trial of a treatment for anemia in patients with cancer. Clinical indicators including hemoglobin level, performance status, and response to therapy were used to determine anchor-based differences. These analyses supported that the minimally clinically important difference for FACT-Fatigue was 3.\textsuperscript{30}

To place in context what deviations of HRQOL values may represent, a study by Bjorner et al\textsuperscript{31} anchors such results in clinical outcomes. A series of analyses on population data of the Medical Outcomes Study (N=3,445) focusing on the VT domain of the SF-36 were performed to evaluate the association between decrement in the scale from the population norm (50) and clinical outcomes. The RAND Medical Outcomes Measures of Quality of Life Core Survey (Medical Outcomes Study) was a 4-year study of patients with chronic conditions who completed the SF-36. The goal was to establish an MID for the SF-36 VT scale by evaluating the association of score differences with clinical conditions and functional outcomes. An initial analysis assessed VT scores (0–100 scale) on a number of chronic conditions that cause fatigue to determine the impact of each condition on VT. Scores were significantly reduced with anemia (5.4), similar to the scores for congestive heart failure (5.8) or COPD (5.7).

A second set of analyses examined the relationship between baseline VT scores and other outcomes at baseline, 1-year follow-up, and 7-year follow-up. Decreases in VT score were significantly associated with increased odds of negative outcomes, including being unable to work because of health at baseline (5 points below the norm: odds ratio [OR], 1.27; 10 points below the norm: OR, 1.62), job loss in the subsequent year (5 points below the norm: OR, 1.13; 10 points below the norm: OR, 1.28), and hospitalization at 1 year (5 points below the norm: OR, 1.08; 10 points below the norm: OR, 1.17). Both short- and long-term mortality were increased (5 points below the norm [0–18 months]: hazard ratio [HR] range, 1.10–1.71; 10 points below the norm [0–18 months]: HR range, 1.21–2.39; 5 points below the norm [≥19 months]: HR range, 1.05–1.31; 10 points below the norm [≥19 months]: HR range, 1.10–1.54). Thus, differences of 5–10 points in the VT score were associated with a statistically significantly increased risk of negative outcomes.

QOL in patients with anemia and cancer

Most practitioners would readily accept the diminished QOL in patients dealing concurrently with cancer and anemia. The following values have been reported with both generic and condition-specific instruments for these patients. A meta-analysis of 23 randomized controlled trials (RCTs) or single-arm studies was undertaken to evaluate the effectiveness of epoetin in over 11,000 anemic patients with cancer.\textsuperscript{32} The scores for each subscale of the SF-36 were converted to norm-based scores (based on the 1998 US general population), with
a mean of 50 and an SD of 10. A score of 100 represents the best health. If we focus on the baseline data of both treated patients and untreated controls, there were markedly diminished scores: VT (43.8), role limitations due to physical health (28.5), and general health (40.9). A clinically meaningful difference or MID of 5.0 has been reported for all SF-36 domains. Thus, at baseline, such patients had values that ranged from 6.2 to 21.5 points below the normalized mean of 50 for the SF-36 domains.32

A review of five trials of patients with anemia and cancer examined the scores for the condition-specific FACT-Fatigue. Mean (SD) FACT-Fatigue scale scores were 27.2 (11.8), 30.9 (10.8), and 27.2 (12.8) in the solid tumor, lymphoproliferative malignancy, and non-chemotherapy groups, respectively. These baseline FACT-Fatigue scale scores are 1.5–2 SD below those observed in the general population.33 The validation study of FACT-An and FACT-Fatigue scales demonstrated broad consistency of both these instruments with the overall ability of patients with cancer to perform everyday activities as measured by the Eastern Cooperative Oncology Group performance status rating.8 The performance status rating is a single-item rating of the degree to which patients are able to participate in customary activities without the need for rest, with scores ranging from 0 (“I have normal activity without symptoms”) to 4 (“I am unable to get out of bed”).

**QOL in patients with anemia alone**

Most of the studies discussed so far assess HRQOL in patients with anemia and cancer. The results of a double-blind RCT of IV ferumoxytol vs placebo illustrate what has been observed in patients primarily without cancer who have IDA due to any cause.34 This study used three different PRO instruments: the generic SF-36 and LASA and the condition-specific FACT-Fatigue. Some of the strengths of this study are its relatively large size, use of a blinded placebo control group, and that the FACT-Fatigue scale has been validated.19 The RCT with IV iron included 812 adults who had either experienced treatment failure with or were intolerant of oral iron therapy and were randomized to blinded treatment with either placebo (saline) or ferumoxytol.34 At baseline, the mean hemoglobin level was 8.8 g/dL. While the underlying cause of IDA could be any disorder other than stage ≥4 CKD, the most commonly observed causes were abnormal uterine bleeding and gastrointestinal disorders; 4.8% of patients had a diagnosis of cancer.

Figure 1 shows the baseline values of all SF-36 domains, which were lower than the population norm, exceeding the minimally clinically important difference of 5.0. Almost all domains, with the exception of bodily pain, were at least twice the MID (~1 SD). Notably, the scores for the VT domain were >10 points below the norm, which represents significant association with the clinical outcomes as noted by Bjorner et al.31

Although the MID for LASA has been reported,35 no population norms are available. In the IV iron placebo-controlled RCT, the baseline LASA energy (36), ability to perform ADL (43), and QOL (47.5) mean scores were similar to those reported by patients with anemia and cancer who were undergoing cytotoxic chemotherapy in two large clinical trials involving >3,500 patients (range, 38.6–39.4, 38.8–40.8, and 45.4–46.4, respectively).10,11

Finally, in the IV iron placebo-controlled RCT, the mean score of the FACIT-Fatigue scale at baseline was 24.4.34 To place this value in the appropriate context, Cella et al10 explored the responses to the FACIT-Fatigue questionnaire among three cohorts: 2,369 patients with nonmyeloid malignancy undergoing concurrent chemotherapy with a hemoglobin level ≤11 g/dL, 113 patients with similar cancers who had hemoglobin level >12 g/dL, and 1,010 persons representing the general US population collected from telephone interviews. The maximum score on the FACIT-Fatigue is 52. Predictably, the mean score in patients with cancer (40.0) was lower than that reported by the general population (43.6). While the 3.6-point difference is greater than what is considered the MID, most notable is the 23.9 score in those with anemia and cancer. The baseline score of 24.4 reported by Vadhan-Raj et al34 in a population of patients with IDA without cancer is comparable to the mean value previously reported in patients with anemia and cancer.

**Benefit of treatment**

The early efforts to establish the impact of anemia treatment on HRQOL largely involved use of ESAs. While the focus of this review is the impact of treatment with IV iron, it is helpful to briefly review the data following ESA therapy to reaffirm both the validity of the instruments and the relationship between hemoglobin and HRQOL. Although a number of instruments have been employed, the most commonly used were the FACT-An and/or the FACT/FACIT-Fatigue, LASA, and SF-36 or the SF-36 VT scale. The findings of the relationship between HRQOL markers and changes in hemoglobin held true whether these studies were small RCTs or large community-based programs. For example, 375 patients with solid or nonmyeloid hematologic malignancies and hemoglobin concentrations ≤10.5 g/dL were randomized to either epoetin alfa or placebo; hemoglobin concentrations increased.
significantly with epoetin alfa compared with placebo. In addition, there was improvement of all anemia-related QOL domains, including energy level, ability to perform ADL, and fatigue.36 Similarly, Dammacco et al37 performed a placebo-controlled study of 145 patients with multiple myeloma and anemia (hemoglobin level <11 g/dL). While the primary end point was transfusion requirements, also assessed was HRQOL using LASA. Patients receiving epoetin had a 1.8 g/dL increase in hemoglobin at week 12, the final assessment during the double-blind phase, whereas those who received placebo had no hematopoietic response. There was a significant improvement in LASA energy level ($P=0.01$) and ability to perform ADL ($P<0.001$), but the overall QOL ($P=0.07$) did not reach statistical significance.37 Such findings correlating hemoglobin level and fatigue as measured by the FACT-Fatigue scale were similarly demonstrated in an analysis based on a series of five placebo-controlled RCTs in various tumor types and whether or not patients were undergoing chemotherapy.33

Larger community-based, open-label evaluations of ESA treatment similarly have demonstrated the correlation between increasing hemoglobin levels following ESA therapy and improvement in HRQOL. For example, 2,370 patients with nonmyeloid malignancies receiving chemotherapy were evaluated in a prospective open-label, nonrandomized study conducted in >600 community-based practices.10 Epoetin alfa therapy resulted in significant increases in hemoglobin levels and improvements in functional status and fatigue as assessed by the LASA (energy level, ability to perform ADL, and overall QOL) and the FACT-An questionnaire. Improvements in QOL parameters correlated significantly ($P<0.001$) with increased hemoglobin levels, irrespective of the degree of tumor response to therapy.10 Of note, a subsequent analysis of these data demonstrated that the relationship between changes in hemoglobin and improvement in QOL was nonlinear. Increases in LASA scores resulted from increasing hemoglobin levels across a range of hemoglobin levels (8–14 g/dL). Beyond a hemoglo-
bin level of 12 g/dL, subsequent increases continued to lead to gains in QOL, but at a decreasing rate. A 1 g/dL increase in hemoglobin from 11 to 12 g/dL was associated with the greatest increase in QOL.9

These findings are consistent with the meta-analysis of 23 trials in patients with anemia and cancer noted previously.12 The baseline HRQOL of those who were treated with ESAs demonstrated markedly diminished FACT-Fatigue scores (mean, 26.7; range, 24.8–28.6), well below the population norm of 40.1. Similarly, subscale scores of the SF-36 were below the norm of 50 (role limitations due to physical health: mean, 28.5 [range, 24.3–32.6]; general health perception: mean, 40.9 [range, 34.2–47.6]; VT: mean, 43.8 [range, 38.9–48.7]). Following treatment with epoetin alfa, there was a consistent improvement in LASA, FACT-Fatigue, and FACT-An scores, as well as the role limitations due to physical health and VT subscales. The three LASA subscales increased 9–10.4 points on the 100 mmVAS, all larger than the previously reported MIDs (9.61, 8.74, and 9.81 for energy, ADL, and overall QOL, respectively). FACT-An and FACT-Fatigue increased to 5.3 and 4.6 points, respectively, each above the MID of 3 points. In contrast, untreated controls reported reduction in FACT scores over the same time period as the actively treated patients. Role limitations due to physical health increased by 11 points and VT increased by 5.8 points, with established MIDs of 5 points. Interestingly, role limitations due to emotional health also increased by 5.4 points, once again highlighting the breadth of the impact that anemia and its treatment can have on these HRQOL measurements.

While most data on HRQOL following treatment with ESAs have been on patients with anemia and cancer, similar results have been reported in other settings such as advanced CKD. Alexander et al14 studied 62 nondialysis-dependent patients with CKD treated with darbepoetin alfa. Improvement in physical activities, VT, and fatigue domains on the SF-36, FACT-An, and FACT-Fatigue showed clinically and statistically significant increases with treatment.14

Changes in HRQOL following treatment with IV iron
The focus on addressing anemia with IV iron came about 15–20 years ago, years after the widespread adoption of ESAs. This was, in part, because of the advent of IV iron dextran, which many believed to be associated with high rates of serious adverse events, such as hypersensitivity leading to anaphylaxis. Therefore, the extent of studies evaluating IV iron that included instruments to assess QOL is relatively small. However, some recent studies, both by their size and rigor, provide important insights.

PROs following oral iron
Two studies assessed HRQOL with oral iron supplementation.18,39 Researchers from the Women’s and Children’s Hospital in North Adelaide, Australia, assigned 430 pregnant women to receive either oral iron or placebo from week 20 to delivery.39 While the 20 mg daily dose of oral iron was low compared with the 100 mg dose previously investigated by this group, comparable increments in hemoglobin and ferritin levels were observed. The primary end point was change in iron parameters and hemoglobin level. SF-36 was measured at week 36 of gestation and at 6 weeks and 6 months post-delivery. At delivery, fewer women treated with iron had IDA (6/198) or ID (65/186) than those who received placebo (20/185 and 102/176, respectively; each comparison was statistically significant). However, there were no significant differences in any of the eight SF-36 domains at any time point.39

A second study similarly used the SF-36 to assess HRQOL in 92 Japanese women with newly diagnosed IDA (mean hemoglobin level 9.2 g/dL and ferritin level 3 ng/mL) treated with 100 mg of oral iron daily.38 The questionnaire was self-administered at baseline and at 1 and 3 months after starting therapy. At baseline, scores normalized for a Japanese population were below the population norm; however, they were only 1–3 points below the norm and, therefore, not reduced as much as reported in other studies. No data on hemoglobin response were provided, but there were statistically significant improvements in a number of HRQOL scores reported. Only the increase in score for physical function exceeded the MID (5), generally considered clinically significant.

PRO response following IV iron in association with ESA treatment
Most of the initial studies assessing HRQOL benefits associated with IV iron investigated the efficacy of IV iron added to ESA treatment in both ESA-responsive and -hyporesponsive patients. For over two decades, it has been recognized that patients receiving ESAs could experience overtax of their marrow’s ability to keep up with the increased iron demands of ESA treatment, which has been termed iron-restricted erythropoiesis or functional ID. While initially noted in end-stage renal disease,40 it can occur in any patient receiving ESAs.
A German group studied patients with rheumatoid arthritis and anemia of chronic disease, in whom they assessed QOL with the SF-36 VT scale and an instrument called the Multidimensional Assessment of Fatigue (MAF). The MAF is a questionnaire covering four dimensions of fatigue (severity, distress, timing, and interference with ADL) scored on a scale ranging from 0 to 50.31 Thirty patients were treated with open-label recombinant human erythropoietin. Although patients had anemia at baseline (hemoglobin level 10.7 g/dL), the article does not provide information to differentiate anemia of chronic disease from IDA, or what the iron indices were at baseline. While 23 of 28 patients who completed the trial developed functional ID (defined as having two of the following: ferritin level <50 ng/mL, total iron binding capacity <15%, or hypochromic red blood cells <10%), how these patients differed from baseline was not specified. The SF-36 VT and MAF both improved markedly during the study (VT increased from 28.2% to 47.1%, while fatigue, as measured by the MAF, decreased from 34.7% to 20.8%; \( P < 0.001 \) for both). These changes paralleled the increase in hemoglobin level. Unfortunately, it was impossible to discern the influence that IV iron had on these results.

More information can be found in a subsequent RCT that evaluated the response to IV iron, oral iron, or no iron in 157 patients with chemotherapy-associated anemia (hemoglobin level \( <10.5 \) g/dL, ferritin level \( \leq 200 \) or \( <300 \) ng/mL, and transferrin saturation \( \leq 19\% \)) treated with once-weekly recombinant human erythropoietin. The iron was administered as either a 100 mg IV bolus or a total dose infusion of iron dextran, each to the calculated dose for iron replacement.34 HRQOL was assessed at baseline and weekly with the three LASA scales (energy, ability to perform ADL, and overall QOL). Patients were not aware of their hemoglobin levels at any time during the study; however, the study was open label. In all four treatment groups, statistically significant increments in hemoglobin level were observed; however, mean increases in hemoglobin were significantly greater for the two IV iron groups than the no iron or oral iron group (\( P < 0.02 \)). For all three LASA domains, both IV iron regimens were associated with increases in score, all well above the previously reported MIDs (9.61, 8.74, and 9.81, respectively). The oral iron group had an increase in energy score from baseline, which did not exceed the MID, and very small changes in ability to perform ADL or overall QOL scores. The no iron group had further decreases in scores (Figure 2). Upon pooling the data from all treatment groups, there was a statistically significant correlation between an increase in hemoglobin level and changes in LASA domain scores (energy: \( r = 0.32, P < 0.0001 \); ability to perform ADL: \( r = 0.30, P < 0.0002 \); overall QOL: \( r = 0.31, P < 0.0001 \)).

Another study reported on the impact of adding IV iron, in this case iron sucrose, to ESAs in patients with either cancer or chemotherapy-associated anemia.42 Inclusion criteria were hemoglobin level \( <10 \) g/dL and no evidence of iron depletion (level not stated). In the initial stage of this two-stage trial, all patients received a fixed dose of ESAs.
After 8 weeks, patients were stratified as responsive to ESAs or not and then randomized to receive a relatively high dose of iron sucrose (7 mg/kg, up to a maximum of 500 mg IV over 4 hours). FACIT-Fatigue was measured at baseline, twice during the 8-week ESA-alone period, and at weeks 3, 6, 9, and 12 on both iron and ESA. The authors reported a greater change in the level of fatigue for those treated with iron sucrose vs no iron, which was statistically significant in the intention-to-treat population (3.3 vs 2.1, respectively; P=0.022), but not in the per-protocol population (3.0 vs 1.7, respectively; P=0.058). Unfortunately, no further information was provided.

Finally, Henry et al\textsuperscript{43,44} examined the impact of IV iron on the FACT-Fatigue score in an RCT of patients with anemia and cancer receiving chemotherapy.\textsuperscript{44,45} Inclusion criteria were a hemoglobin level <11 g/dL and being iron replete as defined by a ferritin level >100 ng/mL or transferrin saturation >15%. All patients received epoetin alfa 40,000 units subcutaneously once per week for the first 4 weeks, then dose adjusted per protocol. Patients were randomized to the following: no iron (ESA alone), oral iron sulfate 325 mg three times daily, or IV ferric gluconate 125 mg once weekly for 8 weeks. Change in HRQOL was measured 4 weeks after study start (week 5) and at the end of study (week 10) using the FACT-An. Hemoglobin responses were significantly greater by the end of study for the 129 evaluable patients who received IV ferric gluconate compared with oral iron or no iron. Only those who received ferric gluconate reported a significant improvement in the FACT-Fatigue scale (MID=3). Of note, the difference began at week 5, at which time the hemoglobin response was still similar among the groups. No significant improvement in fatigue was reported in either the no iron (ESA alone) or the oral iron (Table 1) group.\textsuperscript{41}

**Table 1** Change in hemoglobin and FACT-Fatigue from baseline

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean change from baseline at 4 weeks</th>
<th>Mean change from baseline at the end of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intravenous iron (n=41)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FACT-Fatigue score</td>
<td>4.9\textsuperscript{a}</td>
<td>6.4\textsuperscript{a}</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>1.0</td>
<td>2.4\textsuperscript{a}</td>
</tr>
<tr>
<td>Oral iron (n=44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FACT-Fatigue score</td>
<td>0.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>1.1</td>
<td>1.6</td>
</tr>
<tr>
<td>No iron (n=44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FACT-Fatigue score</td>
<td>2.4</td>
<td>2.3</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>1.0</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Notes: \textsuperscript{a}P<0.01 by paired t-test. \textsuperscript{b}P<0.01 vs oral iron and vs no iron (by analysis of covariance). Data originally reported by Henry et al.\textsuperscript{44}

Abbreviation: FACT, Functional Assessment of Cancer Therapy.

**HRQOL in patients with IDA treated with IV iron alone**

van Wyck et al\textsuperscript{43} reported on 477 women with heavy uterine bleeding treated with either ferric carboxymaltose or oral iron. The major focus of the study was efficacy, with the primary end point being the proportion of patients with a hemoglobin increase of 2 g/dL following treatment. HRQOL was assessed by the SF-36 and LASA. Most patients achieved a >2 g/dL increase in hemoglobin by the end of study at week 6 (82% vs 62% for ferric carboxymaltose and oral iron, respectively). The authors stated, “patients treated with ferric carboxymaltose compared to those prescribed ferrous sulfate reported greater gains in VT and physical function and experienced greater improvement in symptoms in fatigue,” but no numbers were provided. Visual examination of data in a graph revealed that, while both groups had significant improvement in a number of SF-36 domains (physical function, role limitations due to physical health, PCS score, VT, and mental health), only the PCS score (44 points) and VT (~38 points) were below the population norm at baseline (50 points). There does not appear to be any difference between the two cohorts in HRQOL scores at the end of study, with the exception of the PCS score; however, the change from baseline was well below what is considered clinically significant. It is challenging to reach conclusions based on visual evaluation of data alone; however, any improvement seems modest.

In opposition to the challenges of convincingly extracting HRQOL results from the above study, two more recent evaluations provide more rigorous evidence for the strong, beneficial impact of IV iron on HRQOL. Two RCTs using similar designs were conducted as part of a regulatory program for ferumoxytol. One RCT compared ferumoxytol 510 mg with placebo (normal saline) in a blinded fashion among 812 patients (of whom 808 received study treatment) intolerant of oral iron with IDA of any cause, except for advanced CKD.\textsuperscript{34} The other RCT evaluated a similar cohort of 605 patients intolerant of oral iron; however, in that trial, the comparator was 1.0 g of iron sucrose administered as five doses of 200 mg each.\textsuperscript{46} The design of the studies was identical, except that treatments were not blinded in the second study because of the disparate number of doses administered. Each study evaluated HRQOL by means of both condition-specific and general QOL instruments. Given the central role fatigue plays in patients with anemia, the primary assessment tool was the FACIT-Fatigue scale; however, both the LASA and SF-36, including the SF-6D, were also measured.
While we have noted the profound and consistent perturbations in HRQOL experienced with IDA, reviewing the status of those who participated in these two trials provides additional value. At baseline, the mean hemoglobin level was \(~8.9\) g/dL in the two studies. Across the 1,407 patients who entered and had baseline values obtained, there was a consistent and markedly abnormal HRQOL reported for each instrument (data on file). In the ferumoxytol groups, the FACIT-Fatigue scores were 24.1 and 27.8 in the two studies; the placebo group in one study had a score of 24.7 and the iron sucrose group in the other study had a score of 28.7 (data on file). All scores were far below the population norm of 40.1. Similarly, the baseline SF-36 VT scores in both treatment groups were comparable between the two studies (study 1: 38.0 points in the ferumoxytol group and 38.5 points in the placebo group; study 2: 41.1 points in the ferumoxytol group and 41.5 points in the iron sucrose group), which were 8.5–12 points below the population norm for the SF-36 of 50 points (data on file). Note that a 5-point change has been found to be associated with a 27% difference in the risk of being unable to work, a 13% difference in the risk of job loss, and an 8% difference in the risk of hospitalization at 1 year. A change of 10 points below the norm was found to be even more deleterious, with increases of 62% in the inability to work due to health, 28% in job loss at 1 year, 17% in hospitalization at 1 year, and 10%–54% in long-term mortality. While it would be obvious that the SF-36 VT subscale most closely parallels other assessments of fatigue or energy, the other baseline scores demonstrate how deeply and broadly IDA affects both mental and physical QOL (Figure 1).

The goal of this review is to not only emphasize how impaired patients with IDA feel, but also to review the evidence on whether IV iron provides clinical benefit. We believe the strongest evidence comes from the first of these two studies that evaluated ferumoxytol vs placebo. One of the major limitations of all evaluations of HRQOL is the placebo effect. Using a double-blind design, this large study overcomes this issue. In this trial, the FACIT-Fatigue scale was measured at baseline and every week thereafter during the 5-week study. The LASA and SF-36 questionnaires were administered at baseline and at weeks 3 and 5 only. Highlighting the critical importance of the placebo effect is that both treatment arms demonstrated improvement in FACIT-Fatigue scores at weeks 1 and 2. However, only the ferumoxytol-treated cohort continued to show an increase in scores from weeks 3 to 5, whereas the placebo group showed no further improvement (Figure 3). The mean increase in the FACIT-Fatigue score by week 5 was greater in ferumoxytol-treated patients (11.7) than in placebo-treated patients (6.8), which was both statistically significant \((P<0.001)\) and greater than

![Figure 3 Mean FACIT-Fatigue scores.](https://www.frontline-medical-communications.com/)

**Notes:** \(P\) values represent between-group differences in change scores and were derived from the least-squares mean and analysis of covariance model, adjusted for baseline hemoglobin and underlying condition. Error bars represent the 95% CI (1.96 × standard error). \(*P<0.05\); \(*P<0.01\); \(*P<0.001\).}

**Abbreviation:** FACIT, Functional Assessment of Chronic Illness Therapy.
the MID (4.9). Nonetheless, the response in placebo-treated patients demonstrates the impact of the placebo effect.

Similarly, a review of the SF-36 domains shows that both groups had some improvement from baseline, with an increase in SF-36 VT score at week 5 (5.4 for those who received placebo, just exceeding the MID of 5). In contrast, only in ferumoxytol-treated patients did the level of change reported exceed clinically meaningful improvement in almost all domains (VT, physical function, role limitations due to physical health, social function, role limitations due to emotional health, mental health, and the Mental Component Score), each greater than the MID. The increase in VT was approximately twice the MID (Figure 4). The changes for ferumoxytol-treated patients likewise increased in each of the LASA scales at both weeks 3 and 5, with magnitudes exceeding the MIDs for each. This was not observed in the placebo group (Figure 5).

The rigor of the placebo-controlled trial supports the beneficial impact of IV iron therapy on HRQOL. We believe that the case should be made that these beneficial effects could be anticipated with any IV iron, as the class achieves similar increases in hemoglobin level after administration of at least 1 g of iron. The companion study to this trial was essentially identical, with the exceptions that the comparator was iron sucrose and the design was open label.46 Ferumoxytol, given as two doses of 510 mg, was compared with five doses of 200 mg iron sucrose in 605 patients intolerant of oral iron who had IDA of any cause. Patients with grade ≥4 CKD were excluded. The baseline FACIT-Fatigue scores were comparable to those of the placebo-controlled trial (ferumoxytol: 27.8; iron sucrose: 28.7). By week 5, the FACIT-Fatigue score in the ferumoxytol group had increased 13.1 points to 41.7, commensurate with the reported general population mean
The changes from baseline continuously increased throughout the study period and were statistically significant \( (P<0.0001) \) at each follow-up visit. Each was greater than the MID \( (3) \). The greatest increase (mean change \( =5.4 \)) was observed from baseline to week 1. Similar increases were seen with iron sucrose, with an increase by week 5 from 10.6 points to 39.3 points. Consistent with the observations for ferumoxytol, the greatest weekly change occurred between baseline and week 1 (data on file).

Parallel improvements were noted with the other measures of HRQOL. Baseline scores for SF36 VT were comparable \( (41.1 \) in the ferumoxytol group and 41.5 in the iron sucrose group), well below the established population norm. Increases greater than the MID were observed with each IV iron formulation at weeks 3 and 5 (ferumoxytol: 10.6 and 13.0; iron sucrose: 9.9 and 12.5, respectively). Finally, all three LASA scales were confirmatory. The least-squares mean increases in LASA energy were 16.7 and 21.9 points at weeks 3 and 5 for ferumoxytol; 10.6 and 13.0; iron sucrose: 9.9 and 12.5, respectively). Finally, all three LASA scales were confirmatory. The least-squares mean increases in LASA energy were 16.7 and 21.9 points at weeks 3 and 5 for ferumoxytol, respectively, and were 13.0 and 21.3 points for iron sucrose, respectively. All were well above the reported MID. LASA ability to perform ADL and overall QOL scales were similar. The least-squares mean increases for ferumoxytol at week 3 were 16.9 and 15.9 for ability to perform ADL and overall QOL, respectively, and for iron sucrose were 12.6 and 12.4, respectively; at week 5, least-squares mean increases for ferumoxytol were 22.5 and 20.6, respectively, and for iron sucrose were 19.4 and 20.3, respectively.

A final point for consideration regarding the overall impact of ID is that iron plays an important role in multiple enzymatic processes. ID without anemia causes functional impairment, adversely affecting the use of energy sources by muscles, which negatively affects physical capacity and work performance.\(^{48,49}\) In the two ferumoxytol studies, we noted that the greatest increases in FACIT-Fatigue score occurred during the first week. It is reasonable to posit that by week 1, following only partial doses of either ferumoxytol or iron sucrose, there would be minimal, if any, change in hemoglobin concentration. In each of these studies, however, hemoglobin was not measured until week 3. Support for the concept that changes in HRQOL can occur prior to major increases in hemoglobin can be inferred from the study by Henry et al\(^{44}\) (Table 1). Following initial treatment with an ESA, patients with anemia and cancer were randomized to receive ferric gluconate, oral iron, or no iron. At week 4, the increases in hemoglobin were essentially identical \( (1.0, 1.1,\) and \( 1.0 \text{ g/dL}, \) respectively), yet the IV iron-treated patients reported an increase in FACT-Fatigue score that was greater.
than the MID – an observation not seen in the other cohorts. It was only at the end of the 10-week study that the hemoglobin response differed. This is emphasized by the title of the report “Does quality of life improvement precede anemia correction in patients with chemotherapy-induced anemia treated with intravenous iron?” Additional critical support for the importance of the effect of ID on QOL measures is provided by recent studies in patients with heart failure in whom the benefits of treatment with an IV iron, including improvement of exercise tolerance, were noted, irrespective of whether the patients had anemia or were solely iron deficient at study entry.1

Conclusion
We believe we have made a valid case for the marked and broad impact that anemia and IDA have on patients. It appears that, too often, clinicians do not appreciate how “mild” anemia may affect patients’ lives. We hope we have provided support for the improvement in QOL that occurs with treatment with IV iron, as opposed to that with oral iron, and the idea that practitioners give greater consideration to administering a treatment that can provide profound benefits to their patients’ well-being.

Acknowledgments
The authors wish to thank Drs David Henry and Robert Kaper for their review of the manuscript and insightful suggestions, as well as Dr Naomi Dahl for the editorial support. Disha Patel, PhD, of inScience Communications, Springer Healthcare (New York, NY, USA) provided editorial support, which was funded by AMAG Pharmaceuticals.

Disclosure
WES is an employee of AMAG Pharmaceuticals, Inc. and holds equity in the company. MA has received research funding for data management only from AMAG Pharmaceuticals. MA has consulted for Allergan, AMAG Pharmaceuticals, American Regent/Luitpold, and Pharmacosmos. The authors report no other conflicts of interest in this work.

References


