Treatment of anemia in chronic kidney disease: known, unknown, and both

Robert N Foley
Chronic Disease Research Group, Minneapolis Medical Research Foundation, Minneapolis, MN, USA

Abstract: Erythropoiesis is a rapidly evolving research arena and several mechanistic insights show therapeutic promise. In contrast with the rapid advance of mechanistic science, optimal management of anemia in patients with chronic kidney disease remains a difficult and polarizing issue. Although several large hemoglobin target trials have been performed, optimal treatment targets remain elusive, because none of the large trials to date have unequivocally identified differences in primary outcome rates or death rates, and because other reported outcomes indicate the potential for harm (rates of stroke, early requirement for dialysis, and vascular access thrombosis) and benefit (reductions in transfusion requirements and fatigue).

Keywords: hemoglobin, erythropoietin, oxygen-sensing, target trial, methodology

Introduction

Anemia is a typical feature of chronic kidney disease (CKD). In many countries, CKD is the commonest cause of anemia after that accounted for by iron, vitamin B12, and folate deficiencies, overt hemolysis, and ongoing bleeding. Although loss of renal erythropoietin production has traditionally been advanced as an explanation, there has been a gradual realization that other forces may be contributory. Although this is a nascent field, abnormal oxygen-sensing mechanisms and resistance to erythropoietin may be a predominant issue in a sizeable proportion of patients with advanced CKD. This insight, which, as with many other conditions of relative hormonal deficiency, has been slow to embed itself in clinical consciousness, and has considerable implications. It suggests, for example, that greater efforts are needed to delineate the presence and causes of erythropoietin resistance when CKD and anemia intersect, both within individual patients and as part of the broader agenda for future research initiatives. A practical offshoot of this realization may be that anemia treatment algorithms predicated entirely on ever increasing doses of erythropoiesis-stimulating agents (ESAs) that ignore the issue of ESA resistance may not be optimal. Despite a profusion of mechanistic and therapeutic insights, treatment of anemia may be the singlemost controversial aspect of CKD management.

Somewhat tongue in cheek, as bigger and better randomized trials come and go, the notion that optimal indisputable anemia treatment strategies will become clear seems to recede ever further into the distance. Without wishing to diminish their potential importance, iron management protocols will not be considered in this article for space reasons and because, to date, randomized trials have focused completely on hemoglobin response, without meaningful consideration of differential effects on...
conventional “hard” outcomes, including cardiovascular events, infectious events, and death. Similarly, epoetins that are considered biosimilar to epoetin alfa and beta will not be discussed here. The aspects considered in this review include the basic mechanistic constructs that underpin the bulk of current therapeutic development of ESAs, newer ESAs, hemoglobin target trials in CKD, and finally, some discussion about the implications of resistance to ESAs.

**Erythropoietin: mechanistic constructs**

Circulating erythropoietin consists of 165 amino acids in a single strand with two disulfide bonds and four carbohydrate chains, three of which are N-linked and one of which is O-linked. Plasma levels are inversely related to oxygen content; the half-life is 2–13 hours and metabolism takes place in the kidney, liver, and bone marrow. Ultimately, less than 10% of circulating erythropoietin is excreted by the kidneys.

Although current understanding is incomplete, it is evident that erythropoietin gene (*EPO*) expression is tightly regulated by stimulators, including hypoxia-inducible transcription factors (HIF) and hepatocyte nuclear factor 4α, and by inhibitors, including nuclear factor kappa B and GATA2. HIF1, a key regulator of erythropoietin transcription, also regulates transcription of several other genes induced by hypoxia, including those for glycolytic enzymes, platelet-derived growth factor, and vascular endothelial growth factor. HIF1 binds to a hypoxia-response element residing on the 3′-flanking region of *EPO*. In the presence of hypoxia, HIF1 levels increase and the hypoxia-responsive genes become upregulated. In contrast, in the absence of hypoxia, the oxygen-sensitive HIF1α subunit is hydroxylated by prolyl hydroxylase, and degraded in the proteasome; ultimately, downregulation of *EPO* and other hypoxia-responsive genes occur. In this cascade of events, inhibition of prolyl hydroxylase should result in tonic, hypoxia-independent upregulation of *EPO*, and possibly several other hypoxia-responsive genes.

Following translation of the *EPO* gene, three N-linked and one O-linked carbohydrate chains are added to erythropoietin. These chains normally exhibit heterogeneity in terms of the type of carbohydrate moieties incorporated, chain lengths, and branching configuration. The number of sialic residues also varies between healthy individuals, and up to four residues can be found on each N-linked carbohydrate chain, while up to two residues can be found on the O-linked chain, and variability in sialic acid composition has implications for the overall electrical charge of circulating erythropoietin. In turn, overall electrical charge has effects on the circulating half-life of erythropoietin and interactions with the erythropoietin receptor; in general, greater sialic acid content correlates with longer and greater potency.

While marrow-based red blood cell-producing cells are thought to express the highest density of erythropoietin receptors, erythropoietin receptors have been found in tissues throughout the body, including the kidney, brain, heart, retina, muscle, and vascular endothelium. Ligand-receptor interaction leads to conformational changes in the dimeric transmembrane erythropoietin receptor, conformational changes and phosphorylation of tightly-associated Janus kinase 2 molecules, followed by phosphorylation of tyrosine moieties in the cytoplasmic domain of the erythropoietin receptor. Ultimately, the intracellular portion of the erythropoietin receptor becomes a docking complex for proteins exhibiting Src homology 2 (SH2) domains. After docking, SH2-containing proteins are activated by Janus kinase 2-mediated tyrosine phosphorylation and rendered capable of activating nuclear genes involved in cell growth and differentiation, and prevention of apoptosis.

**Newer erythropoiesis-stimulating agents**

Darbepoetin, with its five amino acid substitutions and two extra carbohydrate chains, can incorporate as many as 22 sialic acid residues, in contrast with native erythropoietin, which can incorporate 14 residues. Because darbepoetin has a longer half-life than erythropoietin, of approximately one day or more, it was anticipated that darbepoetin could be efficacious with longer dose intervals than with epoetins. Multicenter, randomized, controlled trials have shown that darbepoetin once every 1–2 weeks is as effective in terms of attained hemoglobin levels as epoetin 2–3 times per week. While the possibility of extended dose intervals is attractive for a compound requiring parenteral administration, several questions remain. For example, comparisons of costs and savings for optimized dosing schedules of different ESAs would be very helpful for clinical decision-making. In terms of basic mechanistic biology, it is unknown whether the non-erythroid effects of darbepoetin in humans are similar to those of epoetins. Given its non-native chemical composition, there is a surprising lack of clinical information about the immunogenicity of darbepoetin. However, it is noteworthy that pure red cell aplasia has been described with this agent.

In continuous erythropoietin receptor activator (CERA), a methoxypolyethylene glycol polymer chain is added to...
the erythropoietin molecule via amide bonds between the N-terminal amino group of alanine and the σ-amino groups of a lysine at position 45 or 52, through a succinimidyl butanoic acid linker. CERA has a molecular weight twice that of erythropoietin and a half-life of over five days.\textsuperscript{31,32} Randomized controlled trials in patients with anemia and CKD show that CERA, administered every 2–4 weeks, is not inferior to conventional ESA therapy in terms of hemoglobin response.\textsuperscript{33–38} As with darbepoetin, comparative nonerythroid effects and accurate rates of associated pure red cell aplasia have not been quantified in humans.

Several other synthetic proteins, modified erythropoietin derivatives, and peptide-based erythropoietin receptor activators are also being researched. For example, CNTO 528, an antibody fusion protein with a hematopoietic peptide attached to an IgG1 base (a “mimetibody”) has a half-life of approximately six days and shows erythropoietic activity in healthy men.\textsuperscript{39,40} Small erythropoietin-emulating peptides that interact with the erythropoietin receptor are also under active investigation. For example, peginesatide, which consists of two peptide chains linked by a pegylation chain, appears to be distinct from erythropoietin in terms of antigen-antibody cross-reactivity.\textsuperscript{41} In a rat model of anterythropoietin antibody-mediated pure red cell aplasia, peginesatide administration led to correction of anemia.\textsuperscript{42} Similar findings were observed in humans in an open-label, uncontrolled study in which peginesatide was administered subcutaneously to 14 CKD patients with erythropoietin antibody-mediated pure red cell aplasia.\textsuperscript{43} Long-term clinical trials of peginesatide involving approximately 2600 CKD patients have been completed, and detailed study reports are awaited with interest.\textsuperscript{44}

Insights into the biological underpinnings of erythropoiesis have led to the identification of several nonpeptide-based (and potentially orally administered) candidates for the treatment of anemia in patients with CKD and other chronic diseases. As described above, while inhibition of prolyl hydroxylase should lead to enhanced erythropoietin production, nonselective inhibition could have pervasive regulatory effects on many genes, both known and unknown. An interesting proof-of-concept study with the orally active prolyl hydroxylase inhibitor, FG-2216, was recently reported.\textsuperscript{45} A single dose of FG-2216 was administered to six healthy volunteers, six anephric hemodialysis patients receiving conventional ESA therapy, and six hemodialysis patients with kidneys receiving ESAs. Erythropoietin levels increased 12.7-fold, 14.5-fold, and 30.8-fold, respectively, in the three groups. Even though study numbers are small, the findings in the anephric group are especially notable, and suggest that abnormal oxygen sensing may be at least as important in the pathogenesis of renal anemia as inadequate erythropoietin production.

The biological mechanisms described above implicate many gene-based pathways for increasing erythropoietin levels. However, uncontrolled erythropoietin production might be anticipated with many approaches. One notable study in erythropoietin-deficient transgenic mice described an intuitive approach to normalizing hemoglobin levels without incurring polycythemia, ie, use of a vector that adds a hypoxia-response element which, in turn, controls EPO transcription.\textsuperscript{46}

**Hemoglobin targets**

Optimal management of anemia in patients with CKD remains a troublesome issue. This section discusses issues related to methodology and interpretation of the large number of published trials that most heavily influence current treatment. It is not intended to be a systematic review or meta-analysis, in part because several of these have already been published, and in part because of a belief that even the best of the available trials are so intrinsically different that simple aggregation of findings may not be advisable. As shown in Table 1, five large (>500 subjects) clinical trials have been reported to date, ie, US Normal Hematocrit\textsuperscript{47} (USNH), Canadian European Normalization of Hemoglobin\textsuperscript{48} (CENH), Cardiovascular Risk Reduction by Early Anemia Treatment with Epoetin Beta\textsuperscript{49} (CREATE), Correction of Hemoglobin and Outcomes in Renal Insufficiency\textsuperscript{50} (CHOIR), and Trial to Reduce Cardiovascular Events with Aranesp Therapy\textsuperscript{51} (TREAT), targeting comparatively unique subsets of anemic CKD patients with a distinctly heterogeneous array of interventions.

Hemoglobin target trials are nothing if not quirky and problematic. They differ profoundly from placebo-controlled trials using identical-looking pills, and these differences can impede interpretation and comparability with other trials. For example, target allocations are often not concealed (as with CHOIR, CREATE, USNH), a strategy that may reduce logistical complexity, especially when drugs are administered parenterally in nonfixed doses that are adapted to on-treatment biological analytes like hemoglobin. Even trials attempting to hide target allocation tend to become unmasked over time for several potential reasons. Because patients with CKD typically have other serious comorbid illnesses, they often have hemoglobin levels routinely drawn at other specialty clinics. Hence, even the most stringent efforts at
Table 1 Major anemia management trials in patients with chronic kidney disease

<table>
<thead>
<tr>
<th>Trial (n)</th>
<th>CHOIR (1432)55</th>
<th>CREATE (603)49</th>
<th>USNH (1233)47</th>
<th>CENH (596)48,58</th>
<th>TREAT (4038)51</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major enrolment criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On hemodialysis</td>
<td>E</td>
<td>E</td>
<td>I</td>
<td>1 (3–18 months)</td>
<td>E</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73 m²)</td>
<td>15.0–50.0</td>
<td>15.0–35.0</td>
<td>–</td>
<td>–</td>
<td>20.0–60.0</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>I</td>
</tr>
<tr>
<td>CVD, inclusion</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CVD, exclusion</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>LV dimensions</td>
<td>–</td>
<td>Advanced CVD ≤ 3 months</td>
<td>MI, PTCA or CABG ≤ 3 months NYHA Class IV</td>
<td>Symptomatic IHD or CHF, critical angiographic HD</td>
<td>CV event ≤ 12 weeks</td>
</tr>
<tr>
<td>Hemoglobin (g/dL)</td>
<td>&lt; 11.0</td>
<td>11.0–12.5</td>
<td>9.0–11.0</td>
<td>8.0–12.0</td>
<td>≤ 11.0</td>
</tr>
<tr>
<td>Interventions and primary outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Patients</td>
<td>Blinding</td>
</tr>
<tr>
<td>Placebo</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Placebo</td>
</tr>
<tr>
<td>Real-time</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Central monitoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin targets (g/dL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>10.5–11.0</td>
<td>Immediate epoetin</td>
<td>10.0 ± 1</td>
<td>9.5–11.5</td>
<td>13.0</td>
</tr>
<tr>
<td>High</td>
<td>13.0–13.5</td>
<td>13.0–15.0</td>
<td>14.0 ± 1</td>
<td>13.5–14.5</td>
<td>None (rescue if &lt; 9.0)</td>
</tr>
<tr>
<td>Treatment strategies</td>
<td>High and low target groups received 10,000 U epoetin alfa once weekly for three weeks.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thereafter, a “prespecified” dosing algorithm, not reported in primary publication.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dose reviewed every four weeks. If hemoglobin increment &lt; 0.5 g/dL, ↑ by 25% to 50%; if hemoglobin increment &gt; 0.5 g/dL, ↓ by 25%–50%.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>With high target, dose initially ↑ 1.5-fold on study entry; thereafter if 2-week △ hematocrit &lt; 2%, ↑ by 25%; if &gt;2% ↓ 25 U/kg. With low target, adjust by 10–25 U/kg.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Epoetin-naive patients in the higher target initially given 150 IU/kg per week. For hemoglobin deviating from target, Δ 25% or 25 IU/kg, initial dose of darbepoetin or placebo equivalent 0.75 μg/kg every two weeks, could be extended to every four weeks if hemoglobin levels stable, to maximum dose of 600 μg.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary outcome</td>
<td>Time to the first occurrence of death, MI hospitalization for CHF (excluding renal replacement therapy), or stroke.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time to first occurrence of sudden death, MI, acute heart failure, stroke, transient ischemic attack, angina pectoris resulting in hospitalization for 24 hours or more or prolongation of hospitalization, complication of peripheral vascular disease, or cardiac arrhythmia resulting in hospitalization for 24 hours or more</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time to death or a first nonfatal MI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Echocardiographic LVVI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>First primary outcome (spendable α 0.048): time to earliest occurrence of death, nonfatal MI, CHF, stroke, or hospitalization for myocardial ischemia. Secondary primary outcome (spendable α 0.002): time to earliest occurrence of death or end-stage renal disease.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Results

**Success of randomization**
- Hypertension and CABG in high target

**Hemoglobin levels (g/dL)**
- Increment during month 1: ↑ in both arms
  - High, ≈ 13.0
  - Low, ≈ 11.0
- During maintenance phase: High, ≈ 13.3; Low, ≈ 11.4
- Epoetin doses: High, 11,000 IU/week; Low, 5,000 IU/week
- Darbepoetin doses: High, 11,000 iU/week; Low, 5,000 iU/week

**Blood pressure**
- High, ↑ by 0.2 mmHg; Low, ↓ by 0.7 mmHg

**Primary outcome, unadjusted**
- ↑ CV events in high target. Hazards ratio (P = 0.03).

**Primary outcome, adjusted for baseline characteristics**
- Analysis not reported.

**Quality of life**
- Patients not blinded. Similar evolution of LASA, KDQ, and SF-36 scores in both treatment arms, except for the score for the emotional role subscale of the SF-36, which was significantly higher in the low-hemoglobin group. Timing of quality of life assessments not reported.

<table>
<thead>
<tr>
<th>Transfusions</th>
<th>High 5.4%</th>
<th>High 21%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low 8.6%</td>
<td>Low 31%</td>
</tr>
</tbody>
</table>

**Quality of life**
- Patients not blinded. Quality of life with SF-36 better with high hemoglobin target at year 1 (general health, mental health, physical function, physical role, social function, and vitality) and at year 2 (general health, vitality).

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**

**Quality of life was assessed but not formally reported in primary publication.**

**Patients were blinded. SF-36 vitality scores improved more in the higher than in the lower target group. No significance between-group differences were seen in the time course of FACIT Fatigue scores and the KDQ Quality of Social Interaction scores.**

**Patients were blinded. FACT fatigue scores increased more in the darbepoetin group. No differences in SF-36 energy and physical functioning domains.**

**Analysis not reported.**
masking treatment allocation are unlikely to be completely successful. As a result, outcome comparisons that involve subjectivity on the part of patients and site investigators have to be viewed as less than pristine.

Nonrandomly assigned treatments, like iron and additional antihypertensive agents, are usually expected following randomization, and highly imbalanced co-interventions can make it difficult to unravel the mechanisms underlying differences in trial outcomes. Immediate or delayed intervention (CREATE and TREAT) is another trial design that systematically adds nonrandom elements. In the delayed intervention or rescue arms, treatment is determined by nonrandom elements (time or a decrease in hemoglobin below a critical threshold). While the trigger for this intervention may be specified in advance by a well described algorithm, it clearly is not controlled by a notional coin toss.

Hemoglobin targets and principal treatment strategies for the five large trials are shown in Table 1. Regarding the critical issue of masked treatment allocation, only TREAT was placebo-controlled. Of the other trials, CENH incorporated concealment of treatment targets from patients. While concealed treatment allocation is intuitively important for outcomes like quality of life, it may also be important for rating “hard” clinical events. For example, basic physiological tenets teach us that polycythemia can cause vascular thrombosis and profound anemia can cause cardiac decompensation; familiarity with the assigned target hemoglobin may influence site investigators when confronted with common diagnostic challenges in advanced CKD, such as distinguishing nonspecific chest pain from angina pectoris and extracellular fluid volume overload from true heart failure. Unfortunately, blinded event committees have no control over what is written in case record forms at the site level and cannot correct site-level biases.

It has been known from the earliest days of treatment that ESAs can increase blood pressure levels, so active tracking of blood pressure levels, ESA doses, and hemoglobin levels would seem to be a reasonable safety requirement in clinical trials of aggressive hemoglobin management strategies. The CENH trial incorporated centrally controlled real-time monitoring procedures for hemoglobin, iron, and blood pressure levels, with treatment recommendations returned to study sites within days of measurement. In TREAT, a point-of-care device was used for rapid measurement of hemoglobin levels, which were entered into an interactive voice-response system to determine the next dose of darbepoetin.

Table 1, which summarizes the treatment algorithms used to achieve target, shows that heterogeneity between
the studies was large, to the extent that no two studies were strictly comparable. Although CREATE, CHOIR, and TREAT all examined patients with CKD not requiring dialysis, CREATE used a strategy of delayed intervention in one treatment arm, while CHOIR used a typical parallel-group design without a time-varying therapeutic intent, and TREAT used a bit of both, with placebo control and the possibility of delayed intervention if anemia became sufficiently severe. Aggressiveness of ESA dosing varied greatly, even among apparently similar patients. For example, both the active and the control groups in CHOIR received a substantial initial dose of epoetin (10,000 U), and it is doubtful if many practicing clinicians would begin treating epoetin-naïve patients with such a large dose, particularly in the 50% of patients for whom the objective was to maintain current hemoglobin levels.

With regard to primary outcomes, no two studies were identical, although all but CENH share common elements (Table 1). While many trials these days use composite primary outcomes, it is always worthwhile to reflect on the individual components and their implications. In CHOIR, the primary outcome was time to first episode of death, myocardial infarction, heart failure hospitalization (provided dialysis was not employed for treatment), and stroke, with dialysis as a censoring event. With this design, episodes of heart failure that are severe enough to require dialysis are ignored, while less severe episodes are counted; it is certainly unusual that less severe episodes would be more likely to be included in treatment comparisons than more severe episodes. With the composite outcome designs employed in these studies, the first event is counted, and subsequent outcomes are ignored in situations where individuals experience multiple components of the composite outcome. All things being equal, it seems intuitively obvious to posit that patients who develop all the components of a composite outcome might receive more weight in outcome analysis. Composite outcomes can never exclude the possibility that serious but unknown treatment-related effects may exist. As a result, comparisons of death rates, in isolation, need full attention, irrespective of the reported primary outcomes.

Table 1 also summarizes the hemoglobin levels achieved in these five studies. There has long been concern that very rapid increases in hemoglobin may be detrimental; first-month hemoglobin increments in the active treatment arms were notably higher in TREAT and CHOIR. During the maintenance phase, while no study maintained targeted hemoglobin levels, statistical separation was clear. ESA doses differed widely between the studies, with much higher doses in CHOIR than in CREATE, and much higher doses in USNH than in CENH. In the nondialysis CKD studies, these differences may be related to study design, because all patients in CHOIR were treated with epoetin alfa 10,000 U/week for three weeks, which was five times the starting dose used in the early intervention arm of CREATE. Blood pressure levels were higher in the active treatment arms in CHOIR and CREATE. While blood pressure levels were similar in the other studies, more antihypertensive agents were used in the high-target arm of CENH.

On unadjusted analysis, CREATE, TREAT, and USNH showed no differences in the primary outcome rates, whereas primary outcome rates were higher in the high hemoglobin target arm of CHOIR. Not unexpectedly, Table 1 shows that significant differences in baseline characteristics were present in each study. Given that several of these differences could be clinically important, it seems worthwhile to scrutinize the strategies used to deal with this issue. Of the three trials examining cardiovascular outcomes and death, USNH reported a comparison of the primary outcome with adjustment for baseline variables in the original publication, and the findings were very similar with and without covariate adjustment. With CHOIR, while the original publication did not report an adjusted comparison of primary outcomes, a regulatory study report filed at ClinicalTrials.gov shows adjustment for baseline obviated statistical significance, with a P value of 0.111, as opposed to 0.03 in the unadjusted analysis. The contrast between unadjusted significance and adjusted non-significance leads to difficulties with overall interpretation of treatment effects, because it is impossible to reject the hypothesis that imbalanced assignment of factors other than study interventions caused the disparity in primary outcome rates seen on unadjusted analysis. TREAT also exhibited substantial baseline differences. While classical covariate adjustment was not undertaken, hazards ratios were reported that incorporated the strata used to randomize patients in the study–study site, baseline level of proteinuria, and an investigator-designated history of cardiovascular disease.

Quality of life was a secondary outcome in all of these trials, and although findings in this regard would never be considered definitive by purists, the available sample sizes were large. In addition, only TREAT attempted to conceal treatment allocation completely from patients, site investigators, and outcome assessors, and of the others, only CENH concealed treatments from study subjects. While four of these trials formally reported quality of life comparisons, no two trials used identical arrays of instruments. Three of the trials, including those that incorporated patient blinding, reported...
quality of life benefits, predominantly in domains related to fatigue and vitality. In contrast, CHOIR showed a potential loss of quality life with the higher hemoglobin target for the emotional subscale of the Short Form-36 instrument; findings were likely clouded by important imbalances in baseline characteristics and the nonreporting of timing of study assessments and dropout rates. Critically, when it comes to assessing risks and benefits, none of the studies attempted to determine what differences in quality of life scales meant to patients.

Two of the studies examined left ventricular dimensions and found that hemoglobin targets had no effect. Transfusion rates, reported by all studies except CHOIR, were 37%–60% lower with higher targets. While no study showed an irreducible effect on primary study outcomes, several showed differences in important nonprimary event rates or in individual components of the primary composite outcomes. For example, in CREATE, the risk of renal replacement therapy was greater in the immediate intervention arm, in spite of similar rates of change in glomerular filtration rate, vascular access loss was greater with higher targets in USNH, and stroke was greater in CENH. In TREAT, stroke rates were higher in the darbepoetin arm; for outcomes not incorporated in the primary outcome, darbepoetin was associated with higher rates of venous and arterial thromboembolism and higher rates of cancer-related death in subjects with cancer at baseline, without being associated with death attributed to cancer in the overall population and overall cancer rates.

Even though only five trials are examined in detail here, noteworthy differences were present at each level considered, ie, enrollment criteria, interventions, primary outcomes selected, treatment concealment, success of randomization, and study findings. In the presence of such distinct heterogeneity, it is questionable whether attempting to generate aggregate findings is a reasonable approach. It seems apparent that an ideal approach to treating anemia in patients with CKD remains to be identified, and, despite years of research, much more investigation is needed.

Resistance to ESAs
There is a substantial amount of literature relating declining hemoglobin levels to adverse outcomes, including death, in patients with CKD. In addition, resistance to ESA, typically defined by high ratios of ESA dose to hemoglobin levels, is also associated with adverse outcomes. Two hypotheses are compatible with these observations, and have disturbingly different treatment implications, ie, sicker patients need higher doses to achieve a given hemoglobin target and, assuming for a given patient that a given hemoglobin target can be achieved with different doses of ESA, higher ESA doses are harmful. In this regard, a recent subgroup analysis of the TREAT trial provides valuable insights. By design, patients receiving ESAs were excluded and fixed initial darbepoetin doses were employed. This subgroup analysis examined subjects in the active treatment arm, and reported higher rates of both the composite cardiovascular endpoint and death in those with a lesser hemoglobin response. Many, including the author of this report, have tended to interpret these findings as raising concern about current target-based strategies for treating anemia in patients with CKD. It seems difficult to make claims for dose-related harm, given that doses were fixed, and it is plausible that this is a beautiful demonstration of patient-related factors being responsible for a connection between hemoglobin response and harm, at least in this study.

Conclusion
Unraveling the biology of erythropoiesis has identified many potential therapeutic targets. Sadly, the quality and quantity of clinical trial information available to guide ESA use do not permit confident statements about optimally efficacious, safe, and cost-effective therapeutic strategies. With our ever increasing insight into the biology of erythropoietin, it seems natural to question whether different ESAs might have different nonerythroid effects at doses and intervals resulting in similar hemoglobin trajectories. If this is the case, the drug evaluation and approval of ESAs will need to go far beyond the research reported to date.

Disclosure
The author reports associations with Affymax, Amgen, Bayer, Ortho Biotech, and Roche.

References
7. Semenza GL. HIF-1, O(2), and the 3 PHDs: how animal cells signal hypoxia to the nucleus. Cell. 2001;107:1–3.


