The major objectives of this article are to review hemoglobin outcome studies, focusing on the utility of purely observational approaches; the design limitations of hemoglobin target randomized trials; what is known from the trials that have been performed to date; and whether confident recommendations for target ranges can be made. The commonly observed association among lower hemoglobin levels, left ventricular hypertrophy, and higher mortality also has been seen within randomized trials when assigned hemoglobin targets were ignored; critically, however, corresponding relationships were absent when intention-to-treat principles were used, strongly suggesting noncausal associations and the need for randomized designs. This being said, hemoglobin typical target trials often have undesirable features, including inadequate blinding and the use of imbalanced, nonstandardized, nonblinded co-interventions. The trials published to date, spanning hemoglobin levels of approximately 7 to 13 g/dl, suggest that higher treatment targets enhance quality of life but at the price of higher BP, thrombotic events, and reduced dialysis adequacy in hemodialysis patients. To date, there is no convincing evidence that targets that approach the physiologic range (versus intermediate targets) have an effect on left ventricular size or survival. Therefore, depending on the outcome examined, higher hemoglobin levels may have beneficial effects, harmful effects, or no effect, leading to the unsatisfactory situation of having to make opinion-based tradeoff decisions. Whereas the available evidence suggests that 11 g/dl is a reasonable lower bound for the hemoglobin target range, the upper bound remains to be defined and targets above 13 g/dl cannot be routinely recommended.

In clinical practice, determining such an optimum hemoglobin level remains a major challenge. The major objectives of this article are to review hemoglobin outcome studies, focusing on (1) the utility of purely observational studies, (2) the inherent limitations of hemoglobin target trials, (3) what is known from the trials that have been performed to date, and (4) whether the available evidence allows a confident recommendation for target ranges.

Utility of Observational Studies

It would be very convenient to be able to rely purely on observational studies to address the question of optimum hemoglobin targets in patients with CKD. Anemia is consistently associated with LV enlargement, congestive heart failure, hospitalization, and mortality (1–5). Anemia has been shown to come before the outcomes studied, satisfying the temporality requirement, so that several but not all components that are required for causal inference seem to be in place.

Hemoglobin levels in patients with CKD reflect many factors beyond clinical choice. For example, anemia is a typical feature of inflammatory illnesses. An expanding literature suggests that advanced CKD is an inflammatory state, one that is associated with adverse outcomes, as well as with anemia (6–15). Whether inflammation also explains the observation that higher doses of erythropoiesis-stimulating agents (ESA) are associated with higher mortality remains unclear (16). These
observations, showing that anemia–mortality associations may be confounded by measurable variables, must lead one to suspect that confounding by unknown factors is considerable. If the latter hypothesis is correct, then the possibility that the hemoglobin levels that are measured in observational studies represent treatment choices alone seems remote, so purely observational studies that relate hemoglobin level to outcome represent unreliable therapeutic evidence.

Within data sets that are generated from randomized trials, interesting findings have been generated by comparing the outcome associations seen with (1) randomly assigned hemoglobin (irrespective of observed hemoglobin levels) and (2) observed hemoglobin levels (irrespective of randomly assigned hemoglobin levels). In the presence of a causal relationship between hemoglobin levels and outcomes, approaches 1 and 2 should lead to similar conclusions; the evidence, however, suggests that this may not be the case. For example, in the US Normal Hematocrit Study, 1233 hemodialysis patients with overt cardiac disease were randomly assigned to hemoglobin targets of 14 g/dl or to targets of 10 g/dl. The primary end point, a composite of death or first nonfatal myocardial infarction, was equally common in both groups. On intention-to-treat analysis, by random group assignment, more deaths were seen in the higher target hemoglobin group than in the lower. When hemoglobin levels that were achieved during the course of the study were used as an analytical variable, however, mortality rates decreased with increasing hemoglobin values (17). Therefore, in a nonrandomized framework, this study showed the typical association that is seen in most observational settings, namely higher hemoglobin levels and longer survival; within the experimental setting, no such association was seen. Similar findings, in which outcome is more tightly associated with achieved than assigned hemoglobin, have been seen in three other trials in which the study end points were death (18) and LV size (19,20). In sum, although the association between hemoglobin levels and outcome satisfies many criteria, such as biologic plausibility, consistency across studies, and temporality, it seems that identification of appropriate treatment targets will be possible only with experimental designs that are based on random treatment assignment.

Inherent Limitations of Hemoglobin Target Trials

Hemoglobin target trials suffer from all the limitations that are common to randomized trials in general. Definitive trials consume substantial financial and human resources and typically are beset by generalizability issues. For example, patients who are selected rarely are a true random sample of all available patients with the target disorder. Similarly, time must elapse for outcomes to accrue, so generalizability of trial results from historic patients to current patients can be a dilemma without an obvious solution.

Even if one accepts that randomized trials are necessary to define hemoglobin targets, these trials have other limitations that may not be shared by other trials. For example, they differ considerably from typical trials of oral medications, in which patients are assigned either active medication or a placebo of identical appearance, consistency, and taste. Most anemia treatment trials randomly assign hemoglobin targets, as opposed to ESA doses; therefore, ESA therapy is a nonblinded, nonstandardized co-intervention in most trials. Because ESA therapy is almost always required, one embarks on hemoglobin target studies with the full intention to co-intervene and with the full knowledge that this co-intervention will be imbalanced between groups. Similarly, intensive ESA therapy often mandates more intensive co-intervention with intravenous iron and antihypertensive agents, and it is easy to imagine that these co-interventions themselves could affect primary outcome rates directly. Ultimately, then, it may be difficult to prove whether good and bad outcomes reflect the hemoglobin targets per se or the co-interventions needed to reach these targets. Also, because co-intervention may be extreme and applied differently in differing studies, comparison of studies and aggregation of studies may be difficult. For example, some studies have incorporated real-time central monitoring of key clinical parameters into the study design; typically, hemoglobin, ESA dose, BP levels, iron storage variables, and iron dose are communicated to the coordinating center, and treatment advice is communicated back to the study centers (19,21). In contrast, in most other studies, treatment decisions are left entirely to the discretion of the treatment team. When attempted, blinding may be difficult to maintain in studies of long duration, because most studies allow treatment teams to know hemoglobin values during the course of the study.

Finally, another type of trial involves patients who have mild to moderate anemia and whose hemoglobin levels are declining at defined rates; one group is randomly assigned to receive early ESA intervention to maintain current hemoglobin levels, whereas hemoglobin levels are allowed to decline further in the other group, with ESA salvage when anemia becomes severe. Although these trials make intuitive sense, they may not be applicable to patients in whom anemia does not tend to worsen. In addition, it is difficult to describe the nature of the study intervention: Is it time, or is it hemoglobin? In practice, it seems to be an amalgam of both components, as neither a true randomized trial of time (immediate or delayed) nor a true randomized trial of hemoglobin target (high or low) has been performed. In summary, although experimental designs seem to be the only way to determine effects of different hemoglobin target, their limitations are many.

Randomized Trials in CKD

Several randomized trials of hemoglobin targets, or ESA dose, that examined death, cardiovascular events, LV size, quality of life, and rates of loss of renal function have been published, varying in size from 1233 to 11 participants (17–37). Table 1, which summarizes the findings from those with 100 or more participants and at least 6 mo of follow-up, shows that most of the evidence accrued so far has been generated from hemodialysis populations.

When one examines these trials and examines each outcome sequentially, the most striking finding is the lack of uniformity in the direction of hemoglobin target effects. A common clinical approach is to consider hemoglobin values as falling into three
Table 1. Major findings from randomized trials (N ≥ 100, ≥6 mo follow-up) reporting on death, cardiovascular events, left ventricular size, quality of life, or rates of change of renal functiona

<table>
<thead>
<tr>
<th>Study Details</th>
<th>Outcomes, High Hgb, or EPO Dose (versus Low)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, Year, Stage of CKD, N</td>
<td>Intervention Target Hgb or EPO Dose</td>
</tr>
<tr>
<td>Besarab, 1998 (17), HD, N = 1233</td>
<td>14.0 versus 10.0</td>
</tr>
<tr>
<td>Parfrey, 2005 (21), HD, N = 596</td>
<td>13.5 to 14.5 versus 9.5 to 11.5</td>
</tr>
<tr>
<td>Furuland, 2003 (18), c HD, N = 416</td>
<td>13.5 to 16.0 versus 9.0 to 12.0</td>
</tr>
<tr>
<td>Foley, 2000 (19), HD, N = 146</td>
<td>13.0 to 14.0 versus 9.5 to 10.5</td>
</tr>
<tr>
<td>Bahlmann, 1991 (25), HD, N = 129</td>
<td>EPO versus placebo</td>
</tr>
<tr>
<td>CESG, 1990 (26), HD, N = 118</td>
<td>High EPO versus low EPO versus placebo</td>
</tr>
<tr>
<td>Nissenson, 1995 (24), PD, N = 152</td>
<td>ESA versus placebo</td>
</tr>
<tr>
<td>Roger, 2004 (23), ND-CKD, N = 155</td>
<td>Early EPO, 12.0 to 13.0 versus late EPO, 9.0 to 10.0</td>
</tr>
<tr>
<td>Levin, 2005 (20), ND-CKD, N = 152</td>
<td>Early EPO, 12.0 to 14.0 versus late EPO, Hgb 9.0 to 10.5</td>
</tr>
</tbody>
</table>

a ↑, better or higher; ←, equivalent; ↓, worse or lower; CESG, Canadian Erythropoietin Study Group; CKD, chronic kidney disease; CV, cardiovascular; CVA, cerebrovascular accident; CVD, cardiovascular disease; EPO, epoetin; HD, hemodialysis; Hgb, hemoglobin; LVD, left ventricular dilation; LVH, left ventricular hypertrophy; ND-CKD, nondialysis chronic kidney disease; PD, peritoneal dialysis; QoL, quality of life; URR, urea reduction ratio.

b On-study Hgb levels were reported in many ways; reported values are indicative.

c The sample included HD, PD, and ND-CKD patients; 70% were on HD.

d Primary outcome.

The primary outcome was a composite of death or nonfatal myocardial infarction.

e Criterion added during the progress of the trial, in response to the findings of the US Normal Hematocrit Study (13).
categories: Low, intermediate, and high (e.g., <9, 9 to 11.9, and ≥12 g/dl). It is apparent from Table 1 that most of larger trials that have been performed to date have concentrated on comparing intermediate to high targets. The absence of “no treatment” arms must generate intellectual dissatisfaction. For example, showing that intermediate and high target hemoglobin levels have similar effects on survival might lead one to conclude that hemoglobin levels have no survival effect; this assumption could be erroneous, because there is no way to refute the hypothesis that low targets lead to better (or worse) survival than higher targets. The situation is clearer for quality of life, because initial trials included untreated arms, and a pivotal study, the Canadian Erythropoietin Study Group trial, used a three-arm design (placebo, intermediate epoetin dose, and high epoetin dose) (26).

**Mortality and Major Cardiovascular Events**

Most of the studies that have been reported to date have been underpowered to detect differences in major outcomes, such as death and cardiovascular event rates. The US Normal Hematocrit Study is the most obvious exception (17). In the latter study of hemodialysis patients, symptomatic cardiac disease was the main inclusion criterion. Patients who were assigned to a hemoglobin target of 14 g/dl showed an increased rate of the primary, composite outcome of death or nonfatal myocardial infarction, compared with patients who were assigned to target of 10 g/dl, a finding that did not reach statistical significance (hazard ratio 1.3; 95% confidence interval 0.9 to 1.9). In another trial, involving patients without symptomatic heart disease or dilated left ventricles, the 13.5- to 14.5-g/dl group had unexpectedly higher rates of cerebrovascular events, compared with those in 9.5- to 11.5-g/dl target. This latter trial, which was not powered formally for examination of major clinical events, showed no differences in the rates of other cardiovascular events and death (21).

**LV Size**

The evidence that treating anemia ameliorates LV hypertrophy or dilation is scanty. The randomized, controlled evidence that partial treatment is better than no treatment comes from three trials involving a total of 51 patients, which seems an unreliable basic foundation on which to build (30,35,37). As shown in Table 1, most of the evidence has come from trials that compared intermediate or high targets, and the available evidence suggests neutral effects.

**Quality of Life**

Several studies have shown a beneficial effect of higher hemoglobin targets on quality of life, especially within the domains of vitality and fatigue, over a range extending from 6 to 16 g/dl (Table 1). Higher targets also have benefitted depression and physical symptoms in some studies. Findings are somewhat dependent on the quality-of-life instrument used and have been more apparent when the study instruments have been developed in populations with CKD.

**Hemodialysis Access**

Some studies have shown higher rates of vascular access problems with higher hemoglobin targets, and no study has shown a protective effect of higher hemoglobin values (Table 1). In addition, it is worth pointing out that the more recent trials excluded hemodialysis patients with ongoing vascular access problems (19,21). Of the available studies, the US Normal Hematocrit Study had the greatest statistical power to study this issue (17). In the latter study, access thrombosis rates were approximately one third higher in the higher hemoglobin group, and the excess risk was apparent both in patients with synthetic grafts and in patients with natural fistulas.

**BP**

Most of the trials that have been reported to date have shown that higher hemoglobin targets lead to higher BP levels and/or greater requirements for antihypertensive therapy (Table 1). This effect has been present in earlier and in more recent trials; the latter trials have tended to show similar BP levels but greater requirements for antihypertensive therapy, suggesting that awareness of this issue has grown and that rising BP that is associated with higher hemoglobin targets is usually treatable.

**Dialysis Adequacy and Change in GFR**

Compared with intermediate hemoglobin targets, higher hemoglobin targets have been associated with lower measures of intradialytic urea clearance (Table 1). In general, the between-target differences have been small, and the clinical significance of these differences is unclear.

The effect of hemoglobin target on progression of kidney disease is unclear. Although the larger, more recently published studies (Table 1) have not shown an effect in either harmful or beneficial directions, it is likely that the totality of available evidence does not allow definitive conclusions. In practice, it seems reasonable to suggest that aggressive hemoglobin targets should be accompanied by aggressive clinical monitoring, including standard measures of renal function.

**Transfusions**

Avoidance of blood transfusions is very desirable for many reasons, especially in patients in whom renal transplantation is being considered, a fact that often is forgotten when the pros and cons of various approaches to anemia therapy are being debated. The studies that have been performed to date suggest that higher hemoglobin targets are followed by lower transfusion rates (Table 1).

**What Else Do We Need to Know?**

Many would accept that a dominant survival effect might take precedence over other considerations, such as quality of life, LV size, BP, dialysis access integrity, and transfusion requirements. Considering anemic patients with CKD, it easy to imagine four different CKD groupings (late-stage CKD, not on renal replacement; hemodialysis; peritoneal dialysis; and transplant), three hemoglobin groupings (low, intermediate, and high), and two groupings for overt cardiovascular disease (ab-
sent and present). Therefore, a given anemic patient with CKD could belong to one of 4 × 2 × 3, or 24 groups. To date, one definitive survival trial comparing intermediate and high hemoglobin targets in hemodialysis patients with overt cardiovascular disease has been performed. This study showed neutral effects on primary outcomes, as well as clinically relevant adverse effects (17). If evidence from trials with mortality as primary outcome is the main criterion on which to base guidelines, then much of the required evidence remains to be gathered.

Most patients and clinicians also would accept that quality of life is an important consideration in treatment decisions. Although the larger trials consistently have shown quality-of-life benefits with higher hemoglobin targets, most concurrently have shown at least one unexpected harmful clinical effect that would generate concern. No clinical trial that can lead one to reject the hypothesis that higher hemoglobin targets could lead to better quality of life but worse quantity of life has been performed. Until trials prove nonsuperiority of low hemoglobin targets for survival outcomes, the evidence base must be considered less than definitive.

Surrogate outcomes have attractions in controlled trials, especially because they reduce sample size requirements. In this regard, LV size has been used extensively. The use of surrogates such as LV size needs to be validated. Inclusion of sequential surrogates within the setting of definitive hard outcome trials is needed to do this.

To date, the clearest evidence discriminating between hemoglobin targets pertains to quality of life. Many patients might accept a finite degree of risk, with regard to dialysis access, hypertension, dialysis adequacy, and even cardiovascular disease and longevity, in return for quality-of-life benefits. This area of tradeoff analysis remains a neglected but clinically relevant field of investigation.

Most of the recent trials have attempted to normalize hemoglobin levels over 3 to 6 mo in patients who may have been anemic for years. It is conceivable that a given hemoglobin level could be associated with different vascular effects depending on the rate of change of hemoglobin. It is unknown whether protocols that change hemoglobin levels very gradually might be associated with different clinical outcomes from those that are seen in the clinical trials that have been published to date, which typically have attempted to raise hemoglobin levels by at least 0.5 g/dl per wk in the higher target groups.

ESA type and dose have received little attention in recent trials. It is possible that different ESA have different biologic effects that are independent of hemoglobin levels. Similarly, one must wonder whether patients with ESA resistance receive benefit or harm when subjected to treatment protocols that encourage continuous dose escalation in response to inadequate hemoglobin increments.

Worldwide, the vast majority of anemic patients with CKD are treated in clinical and societal environments where the financial implications of different hemoglobin targets cannot be ignored. Although cost–utility analyses of hemoglobin targets have been published, the underlying data have been modeled from available data, as opposed to direct measurement within the framework of a single, large, randomized trial (38). In contrast, formal cost–utility analyses, generated within the setting of ongoing trials, are highly important from a societal perspective. Few if any have been published to date.

Conclusion?

A pragmatic approach to integrating available evidence into clinical management might start with an initial hemoglobin target of ≥11.0 g/dl, a lower bound that has a reasonable evidence base (39,40). Defining an appropriate upper bound for hemoglobin targets remains difficult. Regarding the overall body of evidence, relatively clear positive effects (e.g., better quality of life) coexist with relatively clear negative effects (e.g., risks of arteriovenous dialysis access thrombosis and hypertension); if one further considers what is very unclear (e.g., survival effects), it is difficult to suggest a single target hemoglobin that is appropriate for all patients with anemia and CKD. Although it seems somewhat facile to suggest that hemoglobin targets need to be individualized, the currently available evidence seems to suggest this approach. Patient involvement in setting hemoglobin targets seems to be very important so that an appropriate, patient-specific risk–benefit approach can be identified. At present, defining an appropriate upper bound for hemoglobin targets remains more in the realm of art than of science, and there is insufficient evidence to recommend routinely targets above 13 g/dl. One hopes that the next generation of clinical trials will prove illuminating.

References

8. Gunnell J, Yeun JY, Depner TA, Kaysen GA: Acute-phase response predicts erythropoietin resistance in hemodialy-


