Comparative Effectiveness Review

Number 3

Comparative Effectiveness of Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment

This report is based on research conducted by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0026). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

Comparative Effectiveness Review

Number 3

Comparative Effectiveness of Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-02-0026

Prepared by:
Blue Cross and Blue Shield Association Technology Evaluation Center EPC, Chicago, Illinois

Investigators
Jerome Seidenfeld, Ph.D.
Margaret Piper, Ph.D., M.P.H.
Julia Bohlius, M.D., M.Sc.P.H.
Olaf Weingart, M.D.
Sven Trelle, M.D.
Andreas Engert, M.D.
Nicole Skoetz, M.D.
Guido Schwarzer, M.Sc., Ph.D.
Jayne Wilson, B.Sc., M.Sc.
Susan Brunskill, B.Sc., M.Sc.
Chris Hyde, M.B.B.S., M.R.C.P., F.F.P.H.
Claudia Bonnell, R.N., M.L.S.
Kathleen M. Ziegler, Pharm.D.
Naomi Aronson, Ph.D.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested Citation:

Seidenfeld J, Piper M, Bohlius J, Weingart O, Trelle S, Engert A, Skoetz N, Schwarzer G, Wilson J, Brunskill S, Hyde C, Bonnell C, Ziegler KM, Aronson N. Comparative Effectiveness of Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment. Comparative Effectiveness Review No. 3. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-02-0026.) Rockville, MD: Agency for Healthcare Research and Quality. May 2006. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm.

AHRQ expects that Comparative Effectiveness Reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Acknowledgments

The research team would like to acknowledge the efforts of Maxine A. Gere, M.S., for general editorial assistance and program support; Carol Gold-Boyd for administrative support; Tracey Perez, R.N., J.D., for program support; Mark D. Grant, M.D., M.P.H., for methods and editorial input; Gordon Guyatt, M.D., M.Sc., for conceptual consultation regarding quality of life and clinical significance; Steve Goodman, M.D., M.H.S., Ph.D., for statistical consultation; and Elise Berliner, Ph.D., of the Agency for Healthcare Research and Quality, for advice as our Task Order Officer.

AHRQ Contacts

Beth A. Collins-Sharp, Ph.D., R.N.
Director
Evidence-based Practice Center Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Rockville, MD

Elise Berliner, Ph.D.
Task Order Officer
Evidence-based Practice Center Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Rockville, MD

Contents

Tables

Figures

Appendixes

Executive Summary

Background

Anemia (deficiency of red blood cells) occurs in 13-78 percent of patients undergoing treatment for solid tumors and 30-40 percent of patients treated for lymphoma. Tumor type, treatment regimen, and history of prior cancer therapy influence the risk and severity of anemia. For example, among patients with solid tumors, the frequency of anemia severe enough to require red blood cell transfusion is highest for those with lung, gynecologic, and genitourinary tumors. This report focuses on use of epoetin or darbepoetin to manage anemia in patients undergoing cancer treatment with chemotherapy and/or radiation.

Anemia severity is defined by hemoglobin (Hb) concentration. Normal ranges are 12-16 g/dL for women and 14-18 g/dL for men. Mild anemia is defined as Hb from 10 g/dL to the lower limit of normal ranges, while moderate anemia is 8-10 g/dL. Patients are usually transfused if Hb falls to or below 8 g/dL, defined as severe anemia.

Transfusion quickly increases Hb concentration. Serious transfusion-related adverse events are uncommon. For example, in the United States, adverse events due to errors in transfusion are estimated to occur in only 1 in 14,000 units. Risk of hepatitis B infection is estimated to be 1 in 220,000 per unit of blood transfused.

Erythropoietin, a hormone produced primarily in the kidney, participates in regulating red blood cell production (erythropoiesis) and thus Hb concentration. Two erythropoietic stimulants are available commercially in the United States, epoetin alfa (Epogen®, Procrit®) and darbepoetin alfa (Aranesp®), which is a newer and longer acting drug. Epoetin beta, which is pharmacologically and clinically similar to epoetin alfa, is commercially available in Europe and elsewhere. Erythropoietic stimulants are widely used in clinical practice to manage anemia of patients undergoing cancer treatment and to reduce the need for transfusion.

Although it is well established that erythropoietic stimulants improve anemia in patients undergoing cancer treatment, the comparative effectiveness of epoetin and darbepoetin has not been evaluated in a systematic review. Moreover, trials varied substantially in how erythropoietic stimulants have been used, including Hb concentration at start of treatment, doses given, treatment duration, and target Hb concentrations they sought to maintain. A review of these various trials may help maximize benefit, optimize drug usage, and minimize adverse effects from using erythropoietic stimulants to manage anemia in patients undergoing cancer treatment.

The report addresses the following questions:

1. What are the comparative efficacy and safety of epoetin (alfa or beta) and darbepoetin?

2. How do alternative dosing strategies affect the comparative efficacy and safety of epoetin and darbepoetin?

3. How do alternative thresholds for initiating treatment or alternative criteria for discontinuing therapy or duration of therapy affect the efficacy and safety of erythropoietic stimulants?

4. Are any patient characteristics at baseline or early hematologic changes useful to select patients or predict responses to treatment with erythropoietic stimulants?

Conclusions

Comparative efficacy and safety of epoetin and darbepoetin

Three sets of trials were summarized and analyzed: 7 randomized direct comparisons of darbepoetin versus epoetin (pooled N=1,415 patients randomized to epoetin, 1,087 to darbepoetin); 48 randomized controlled trials (RCTs) of epoetin versus control[a] (pooled N=4,518 patients randomized to epoetin, 3,743 to control); and 4 RCTs of darbepoetin versus control[a] (pooled N=598 patients randomized to darbepoetin, 396 to control).

The evidence does not show any clinically significant difference between epoetin and darbepoetin in hemoglobin response, transfusion reduction, and thromboembolic events. (See Table A for details.)

- For hematologic response, five of six trials comparing darbepoetin to epoetin showed no statistically significant difference between these drugs. Pooled results of trials comparing epoetin to control and darbepoetin to control showed no difference; over 50 percent of patients treated with epoetin or darbepoetin had a Hb increase $\geq$2 g/dL, compared with fewer than 20 percent of untreated patients.

- For rates of transfusion, trials comparing darbepoetin to epoetin showed no statistically significant difference between these drugs. Pooled results of trials comparing epoetin or darbepoetin to control showed approximately 30 percent of patients treated with epoetin or darbepoetin were transfused, compared with 50 percent of untreated patients. However, patients varied widely in how likely they were to need a transfusion; the proportion of untreated patients undergoing transfusion ranged from 0 percent to 100 percent in the studies reviewed.

- For thromboembolic events,[b] trials comparing darbepoetin to epoetin showed no statistically significant difference between these drugs. Pooled results of trials comparing epoetin or darbepoetin to control showed that approximately 7 percent of patients treated with epoetin or darbepoetin experienced a thromboembolic event, compared with 4 percent of untreated patients. However, trials varied widely in thromboembolic event rates: 0 percent to 30 percent among treated patients and 0 percent to 23 percent among untreated patients. Several studies sought to maintain Hb levels higher than

[a] Controls received placebo or no erythropoietic stimulant, and each group (treated or control) was transfused as necessary.
[b] Studies usually did not provide a detailed definition of thromboembolic events; those that did included thrombosis and related complications such as thrombophlebitis, transient ischemic attacks, stroke, pulmonary embolism, and myocardial infarctions.

recommended in product labels (≤12 g/dL); however, evidence is insufficient to determine if risk is lower when treatment conforms to Food and Drug Administration (FDA) label recommendations.

- For each of the above outcomes, more evidence is available on epoetin than darbepoetin.

Table A. Summary of Rates of Hematologic Response, Transfusion, and Thromboembolic Events

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control
Hb response rates: Number of studies reporting	6	15	3
Patients analyzed	2,205	3,293	659
Pooled relative risk of Hb increase ≥2 mg/dL (95% CI)	Meta-analysis not done[1]	3.42 (3.03, 3.86)[2]	3.36 (2.48, 4.56)
Pooled event rates (range across studies)	Meta-analysis not done[1]	Epo: 58% (21%−73%) Control: 17% (3%−32%)	Darb: 54% (25%−84%) Control: 17% (9%−18%)
Transfusion rates: Number of studies reporting	6	34	4
Patients analyzed	2,158	5,210	950
Pooled relative risk (95% CI)	1.10 (0.93, 1.29)[2]	0.63 (0.59, 0.67)[2]	0.61 (0.52, 0.72)
Pooled event rates (range across studies)	Darb: 22% (3%−28%) Epo: 20% (12%−43%)	Epo: 30% (0−91%) Control: 47% (0−100%)	Darb: 29% (13%−34%) Control: 51% (25%−67%)
Thromboembolic events: Number of studies reporting	3	30	1
Patients analyzed	1,879	6,092	314
Pooled relative risk (95% CI)	0.86 (0.61, 1.21)	1.69 (1.36, 2.10)	1.44 (0.47, 4.43)[3]
Pooled event rates (range across studies)	Darb: 6% (3%−9%) Epo: 7% (3%−11%)	Epo: 7% (0−30%) Control: 4% (0−23%)	Darb: 5% Control: 3%

[1] Trials defined response differently and initiated and adjusted doses differently; only one randomized controlled trial (n=352) reported significant difference favoring epoetin, but results may be biased since dose was adjusted differently in each arm; five trials (N=1,853) reported no significant differences between arms.
[2] Tests of heterogeneity (I^2) indicated excessive variability among individual study results. Results of this fixed-effects meta-analysis were compared with random-effects meta-analysis; results were not meaningfully different.
[3] Since there was only one trial, this result is a single-study (not pooled) relative risk.
Abbreviations: CI: confidence interval; Hb: hemoglobin.

The evidence is not sufficient for conclusions on effects of either epoetin or darbepoetin on quality of life, tumor response and progression, survival, or adverse outcomes other than thromboembolic events.

- Trials did not completely or consistently report quality of life (QoL) results, so 12 potentially relevant studies were unusable for this analysis, and quantitative analysis could not be performed for the 15 remaining studies. Overall, QoL measures tended to favor treatment with epoetin or darbepoetin. However, the degree of change varied widely across studies and not all positive changes were statistically significant.

  Numeric changes on QoL instrument scales must be empirically evaluated to determine whether the degree of change is perceptible and meaningful to the patient. Currently,

there is not enough evidence to quantify the minimum changes that are clinically meaningful on the most commonly used QoL instrument, Functional Assessment of Cancer Therapy-Anemia (FACT-An) and its subscales. Additional limitations of the evidence are potential bias due to substantial missing data; concerns regarding study validity, including lack of blinding and lack of information on QoL instrument administration; and incomplete reporting of numerical results.

- The limited evidence available (five studies, N=688) does not suggest that erythropoietic stimulants improve solid tumor response to a concurrent course of cancer therapy. Whether erythropoietic stimulants accelerate progression of some cancers, as reported by one study (n=351), is uncertain.

- Of 40 (N=8,249) RCTs reporting on survival, only seven (N=2,188) were actually designed to assess effects on survival (progression free or overall). No studies designed to test survival[c] used epoetin or darbepoetin as currently recommended; rather, all seven trials sought to maintain Hb levels >12 g/dL. Two of the seven trials, one on metastatic breast cancer (n=939) and one on head and neck cancer (n=351), showed poorer overall survival for patients treated with epoetin; this prompted an FDA safety review in May 2004 and revised product labeling to indicate that clinicians should avoid targeting Hb concentrations above 12 g/dL. Of the other five trials, survival appeared poorer with erythropoietic stimulant in three (N=471) and better in two (N=427), but most results were not statistically significant.

  The remaining 33 of the 40 RCTs reporting on survival collected survival data retrospectively from trials designed only to test hematologic and transfusion outcomes. This evidence is not definitive, but might detect a large difference in survival. Analysis of mortality in all 40 trials shows no overall benefit of darbepoetin or epoetin on survival. Neither higher than recommended target Hb nor any other single patient- or treatment-related factor explained why some trials showed a detriment in survival and others did not.

- For other adverse events, reporting is incomplete, representing less than one-third of patients. Studies did not use consistent definitions of events and severity. For epoetin, 15 studies (N=1,949) reported on hypertension, 9 (N=1,422) reported on thrombocytopenia/hemorrhage, 6 (N=522) reported on rash, 3 (N=389) reported on seizures. For darbepoetin, one trial (n=122) comparing darbepoetin to epoetin reported on seizures, and one trial (n=314) comparing darbepoetin to control reported on hypertension. Overall, adverse events were more frequent with epoetin or darbepoetin than control, but pooled results did not show statistically significant differences.

- For each of the above outcomes, more evidence is available on epoetin than darbepoetin.

[c] To test survival, a trial should enroll sufficient numbers of patients with the same tumor (or stratify patients by tumor), and should follow them over an adequate time period.

Alternative dosing strategies

- Twelve trials examined different dosing regimens for epoetin and seven trials examined different dosing regimens for darbepoetin. For each of the following pairs of dosing strategies,[d] one large trial reported no statistically significant difference between strategies: fixed-dose compared to dose based on weight, one trial each for epoetin and darbepoetin; fixed-dose epoetin administered weekly vs. thrice weekly; fixed dose epoetin administered weekly vs. every 3 weeks; and darbepoetin using an initial loading dose versus constant weight-based dosing regimens. The remaining 14 trials were too small to interpret.

Thresholds for initiating treatment or criteria for discontinuing therapy

- Three unblinded randomized trials, not yet published, compared using erythropoietic stimulant therapy soon after mild anemia developed vs. delaying treatment until Hb had fallen below a predefined threshold of moderate anemia. Comparisons were ~11 g/dL vs. 9 g/dL (N=269); ~11 g/dL vs. 10 g/dL (N=204); and ~13 g/dL vs. 10 g/dL (N=216). All patients in the mild anemia arms were treated with an erythropoietic stimulant; of patients in whom treatment was delayed until moderate anemia developed, 19 percent, 63 percent, and 44 percent, respectively, were treated with erythropoietic stimulant. Transfusion was more frequent when treatment was delayed until moderate anemia developed, but the difference was not statistically significant in any study. One trial reported a statistically significant increase in thromboembolic events among patients who were treated for mild anemia compared with those who were treated for moderate anemia.

- No trials compared criteria for discontinuing therapy.

Factors to select patients or predict responses to treatment

- Available evidence does not identify any single factor as clinically useful to guide treatment decisions. Potential predictive factors, measured at baseline (e.g., serum erythropoietin level or observed/predicted ratio [O/P ratio], serum ferritin) or early after starting treatment (e.g., Hb increase, serum ferritin, reticulocyte increase), were evaluated in 26 studies and found to have either weak ability or no ability to discriminate between responders and nonresponders.

- Seven algorithms combining multiple factors, potentially more useful to predict Hb response, are each currently supported only by one study. The largest of these studies do not report sufficient predictive ability for any algorithm to establish clinical utility for selecting treatment.

[d] Rationales for comparing these alternative strategies are: (1) Drug concentrations with fixed-dose strategies may be inadequate for overweight patients and excessive for underweight patients. (2) More frequent dosing schedules are less convenient, but may be more effective to maintain the desired drug concentration range. (3) Front-loading refers to starting at higher dose, then reducing to a maintenance dose, which may increase the proportion of responding patients.

Remaining Issues

- Considerably less evidence exists on darbepoetin than epoetin. Consequently, most conclusions concerning effects of erythropoietic stimulants as a class rest on inferences from the evidence on epoetin.

- More evidence is needed to delineate the effects on survival, tumor progression, and risk of adverse effects when erythropoietic stimulants are administered as currently recommended.

- To interpret changes in anemia-specific quality of life measures, a clear, empirically based definition of the minimum clinically important difference is needed.

Chapter 1. Introduction

This review compares the efficacy and adverse effects of specific erythropoietic stimulants (i.e., epoetin [alfa or beta], darbepoetin alfa) when used to manage anemia in patients undergoing cancer therapy (i.e., chemotherapy and/or radiation).[1] This review also addresses questions relevant to optimizing the use of erythropoietic stimulants as a general class: the outcomes of using alternative thresholds to initiate or discontinue treatment and whether there are early predictors of response to treatment.

Erythropoietin is an endogenous hormone, produced primarily in the kidney, which participates in regulating red blood cell production (erythropoiesis). Two forms of recombinant human erythropoietin—epoetin alfa and epoetin beta (the latter not commercially available in the United States)—have been extensively studied and used clinically for more than a decade to treat various anemias; they have similar clinical efficacy (Halstenson, Macres, Katz, et al., 1991; Storring, Tiplady, Gaines Das, et al., 1998). In a recent review of safety concerns associated with recombinant human erythropoietins, a U.S. Food and Drug Administration (FDA) briefing document noted that "…the biochemical differences between various erythropoietin products are not associated with marked differences in the pharmacodynamic properties of the different products when used at recommended doses, thus effects observed with these non-US-licensed products may also be associated with the U.S. licensed product."

A novel long-acting recombinant erythropoietin--"novel erythropoiesis-stimulating protein" (NESP) or darbepoetin alfa--was developed more recently. Darbepoetin alfa, which produces a similar physiologic response when compared to recombinant human erythropoietin (Joy, 2002), has been tested in prospective clinical trials (Glaspy, Jadeja, Justice, et al., 2003; Hedenus, Hansen, Taylor, et al., 2003; Vansteenkiste, Pirker, Massuti, et al., 2002), and is commercially available in the United States. The epoetins have the same amino acid sequence as endogenous erythropoietin, while darbepoetin alfa has two additional oligosaccharide chains; however, the epoetins and darbepoetin all have pharmacologic actions identical to those of the endogenous hormone (McEvoy, 2005). They increase the number of red blood cells, and thus the blood concentration of hemoglobin, when given to individuals with functioning erythropoiesis.

Anemia, defined as a deficiency in the concentration of hemoglobin-containing red blood cells, is a widely prevalent complication among cancer patients. The National Cancer Institute and others have agreed to use the following classification for anemia based on hemoglobin (Hb) values (National Cancer Institute Cancer Therapy Evaluation Program, 1999):

- Grade 0, within normal limits, hemoglobin values are 12 to 16 g/dL for women and 14 to 18 g/dL for men

[1] This review overlaps somewhat with a critical appraisal of the literature on outcomes of erythropoietin for anemia related to cancer treatment conducted for the National Institute for Health and Clinical Excellence (NICE) in the U.K. (Wilson, Yao, Rafferty, et al., 2005). The evidence base used in the appraisal was an update of the earlier Cochrane review (Bohlius, Langensiepen, Schwarzer, et al., 2005), and included a cost-effectiveness component (Wilson, Yao, Rafferty, et al., 2005). Note that pooled analyses for the appraisal included trials in patients with myelodysplastic syndrome, as well as trials of patients with cancer who were not receiving cancer therapy. These types of trials were excluded from the present analysis, which is limited to patients undergoing cancer treatment. In addition, the search date cutoff for the NICE appraisal was September 2004, whereas the search date cutoff for this analysis was March 2005.

- Grade 1, mild (Hb 10.0 g/dL to less than lower limit of normal)
- Grade 2, moderate (Hb 8 to <10.0 g/dL)
- Grade 3, serious/severe (Hb 6.5 to <8.0 g/dL)
- Grade 4, life threatening (Hb <6.5 g/dL).

Historically, red blood cell transfusion has been the conventional treatment of choice for severe anemia in cancer patients. The literature reports a critical degree of anemia as Hb less than 8 g/dL, while mild-to-moderate anemia (Hb level 8–10 g/dL) usually has been left untreated (Koeller, 1998; Blajchman and Hebert, 2001). Although blood transfusion is the fastest method to alleviate symptoms, short- and long-term risks exist (Engert, 2000). Potential complications associated with blood transfusion include transmitting infectious diseases, transfusion reactions, alloimmunization, and over-transfusion (Goodnough, 2005). However, the risks are quite small. Adverse events due to error in transfusion are estimated to be 1 in 14,000 units in the United States. The risk of transfusion-related acute lung injury is about 1 in 5,000 transfusions. The risk of severe infections is estimated to be to 1 in 220,000 per unit of blood transfused for hepatitis B, 1 in 1,600,000 per unit for hepatitis C, and 1 in 1,800,000 for human immunodeficiency virus (HIV) (Busch, Kleinman, and Nemo, 2003). Emerging bloodborne infections such as the West Nile virus outbreak in 2002 are of concern; screening for West Nile virus was implemented in the U.S. in July 2003 (Pealer, Marfin, Petersen, 2003); that summer, 4,137 cases of West Nile virus infection were reported to the Centers for Disease Control and Prevention (CDC), only 2 of which were known to be transmitted by blood (Goodnough, 2005).

Among cancer patients, the prevalence of anemia varies according to the type of neoplasia (Knight, Wade, and Balducci, 2004). Defining anemia as an Hb range of 9–11 g/dL, one systematic review reported that the prevalence of anemia in solid tumor types (e.g., breast, brain, prostate) varies from 13–78 percent (Knight, Wade, and Balducci, 2004), depending on tumor type. Among patients with solid tumors, the highest frequency of anemia requiring transfusion has been reported for lung, gynecologic (e.g., ovarian), and genitourinary tumors, in part attributable to the use of platinum-based therapies (Groopman and Itri, 1999). Patients with hematologic malignancies frequently experience anemia. At the time of diagnosis, 30 to 40 percent of patients with Hodgkin's or non-Hodgkin's lymphoma and up to 70 percent of patients with multiple myeloma are anemic; the figures are even higher in patients with myelodysplastic syndromes (Garton, Gertz, Witzig, et al., 1995). The type and amount of chemotherapy also influences the extent of anemia. For patients with lymphoma, anemia is present in around 40 percent of patients at diagnosis; however, after 3 to 4 cycles of chemotherapy, up to 70 percent of patients will be anemic (Samol and Littlewood, 2003). Patients with cancer-related anemia not undergoing cancer treatment are a different patient group, with distinct causes of their anemia; they should be analyzed separately from those undergoing treatment for their malignancy, and thus are outside the scope of this report.

Scope and Key Questions

Several evidence-based guidelines have addressed whether recombinant erythropoietin's ability to increase hemoglobin levels reduces the risk for blood transfusions in patients with malignant disease. The most comprehensive guideline is from the American Society of Clinical Oncology (ASCO) and the American Society of Hematology (ASH) (Rizzo, Lichtin, Woolf, et al., 2002). The basis of this guideline is a systematic review commissioned by the Agency for Healthcare Research and Quality (AHRQ) and conducted by the Blue Cross and Blue Shield Association's Technology Evaluation Center Evidence-based Practice Center (BCBSA TEC EPC) (Aronson, Seidenfeld, Piper, et al., 2001; Seidenfeld, Piper, Flamm, et al., 2001). That AHRQ report summarizes and analyzes evidence published through 1999 on the use of epoetin to manage anemia in oncology patients.

In collaboration with authors of the AHRQ report, the Cochrane Haematological Malignancies Group conducted a Cochrane Review on the effects of recombinant erythropoietin in patients with malignant disease (Bohlius, Langensiepen, Schwarzer, et al., 2004). The Cochrane Review included 3,287 patients from 27 studies, published between 1993 and May 2002. Both reviews found that treatment with epoetin statistically significantly reduced the need for red blood cell transfusions. Epoetin-treated patients were more likely to have a hematologic response and less likely to undergo transfusion than untreated patients. The evidence on quality-of-life changes from treatment with epoetin was inconclusive.

A post-hoc analysis amended to the study protocol of Littlewood, Bajetta, Nortier, et al. (2001) generated interest in the effects of epoetin on survival. Some investigators hypothesized that epoetin might improve survival, either by improving tumor oxygenation and thus enhancing cytocidal effects of chemotherapy and/or radiation therapy (Glaspy, 2002), or by some other consequence of reversing anemia shown to predict poor prognosis in patients with malignancy (Caro, Salas, Ward, et al., 2001; Bokemeyer, Oechsle, Hartmann, et al., 2002). Others cautioned that these hypotheses must be tested in randomized controlled trials (Watine and Bouarioua, 2002; Steensma and Loprinzi, 2005), especially given evidence that some malignant cells carry erythropoietin receptors that are able to promote tumor cell proliferation when stimulated (e.g., Westenfelder and Baranowski, 2000; Acs, Zhang, Rebbeck, et al., 2002). Data collected and analyzed for the Cochrane Review also suggested that overall survival of anemic oncology patients receiving epoetin may be greater than among patients receiving only red blood cell transfusion as needed (Bohlius, Langensiepen, Schwarzer, et al., 2004). However, the evidence was only used to generate hypotheses, as the study by Littlewood and co-workers (2001) and other studies included in the Cochrane analysis were not designed to test the effect of epoetin on survival.

To test the effect of erythropoietic stimulants on survival, a trial should have a homogeneous primary tumor type and treatment regimen. Duration of follow-up and number of participants should be sufficient to detect a clinically meaningful difference in overall survival or surrogate outcomes such as tumor response or progression-free survival (Food and Drug Administration Oncologic Drugs Advisory Committee, 2004). Survival data available for the 2004 Cochrane Review were largely from trials designed to test effects of epoetin on hemoglobin response and risk of transfusion. Almost all trials included mixed populations with respect to tumor types and treatment regimens. Data on survival were collected subsequent to these trials' prespecified endpoints, and so do not represent results of the original randomized controlled trial design.

Subsequently, several studies designed to assess overall or progression-free survival have been conducted and published. The evidence thus generated needs to be assessed: two studies demonstrated significantly worse overall survival for patients receiving epoetin (Henke, Laszig, Ruebe, et al., 2003; Leyland-Jones, 2003; Leyland-Jones, Semiglazov, Pawlicki, et al., 2005). Further, other important clinical questions have not yet been resolved, including optimal hemoglobin thresholds to initiate and stop treatment with erythropoietic stimulants, and which patients are most likely to benefit from such treatment. Because both epoetin and darbepoetin alfa are expensive, a systematic review comparing their costs and effectiveness as treatment alternatives also would be useful. In addition, the evidence on darbepoetin alfa has not yet been systematically reviewed.

For further background details on the pathophysiology of cancer-related anemia and a more detailed description of epoetin, readers are referred to the AHRQ evidence report, "Uses of Erythropoietin for Anemia in Oncology" (Aronson, Seidenfeld, Piper, et al., 2001).

Although several types of erythropoiesis-stimulating products currently are approved for use or undergoing active research in other countries --(e.g., other epoetin alfa products [Eprex®, Janssen-Cilag]; epoetin beta [NeoRecormon® and Recormon®, Roche; Epogin®, Chugai]; epoetin omega [Epomax®, Elanex]; epoetin delta [Dynepo®, TKT]; synthetic peptide-based erythropoiesis-stimulating agent [Hematide™, Affymax, Inc., currently in Phase II trials]; continuous erythropoiesis-receptor activator [CERA]) (Deicher and Horl, 2004)-- there are three products commercially available in the U.S. These are Epogen® and Procrit® (both epoetin alfa), and Aranesp® (darbepoetin alfa). Table 1 describes the FDA-labeled indications and dosages for these products. Note, however, that this review includes evidence from trials of epoetin beta (not licensed in the United States) as well as from trials of epoetin alfa and darbepoetin alfa.

The National Comprehensive Cancer Network (NCCN) in its oncology practice guideline on cancer- and treatment-related anemia provides dosing schedules for treatment according to the FDA-approved package inserts (Table 1), as well as "commonly used" regimens for darbepoetin. The first regimen recommends darbepoetin at a dosage of 3 mcg/kg subcutaneously every 2 weeks; in patients without response, the guideline recommends increasing dosage to 5 mcg/kg every 2 weeks. The second common regimen is a fixed-dose regimen of 200 mcg every 2 weeks, with titration to up to 300 mcg every 2 weeks in patients with no or inadequate response (NCCN, 2006).

In 2004, the FDA revised the labeling of erythropoietic stimulants licensed in the Untied States. Studies presented at a May 4, 2004, meeting of the Food and Drug Administration (FDA) Center for Drug Evaluation and Research (CDER) Oncologic Drugs Advisory Committee (ODAC) (Food and Drug Administration Oncologic Drugs Advisory Committee, 2004) showed the potential for increased thromboembolic complications, and possibly worse survival, in patients whose target hemoglobin was higher than 12 g/dL. A subsequent healthcare provider communication highlighted the results of the ODAC meeting and revised the product labeling to include information on thromboembolic complications in the "Warnings" and "Precautions" sections. The revised labeling for all three commercially available products recommends that the target hemoglobin in patients with cancer not exceed 12 g/dL in both men and women.

The following four key questions are addressed in the current review.

1. What are the comparative efficacy and safety of epoetin (alfa or beta) and darbepoetin? Outcomes of interest include hematologic responses, transfusions, tumor response to therapy, overall survival, quality of life, thromboembolic complications, and other adverse events.

2. How do alternative dosing strategies affect the comparative efficacy and safety of epoetin and darbepoetin?

3. How do alternative thresholds for initiating treatment or alternative criteria for discontinuing therapy or duration of therapy affect the efficacy and safety of erythropoietic stimulants?

4. Are any patient characteristics at baseline or early hematologic changes useful to select patients or predict responses to treatment with erythropoietic stimulants? The outcome of interest is limited to hematologic response.

Table 1. Erythropoietic Stimulants Available Commercially in the United States

Drug	Trade name(s)	Half-life	Labeled indications	Initial dose recommendations for anemia of cancer chemotherapy	Recommended dosage adjustments for anemia of cancer chemotherapy
epoetin alfa*	Epogen® (Amgen, Inc., 2005) and Procrit® (Ortho-Biotech, 2005)*	40 hours with a range of 16 to 67 hours	anemia of chronic renal failure anemia in zidovudine-treated HIV-infected patients anemia in cancer patients receiving chemotherapy reduction of allogeneic blood transfusion in surgery patients	150 units/kg SC 3 times weekly or 40,000 units SC weekly (adults)	Reduce dose by 25% when hemoglobin approaches 12 g/dL or hemoglobin increases >1 g/dL in any 2-week period Withhold dose if the hemoglobin exceeds 13 g/dL until the hemoglobin falls to 12 g/dL, then restart dose at 25% below the previous dose Increase dose to 300 units 3 times weekly if response is not satisfactory (i.e., no reduction in transfusion requirements or rise in hemoglobin) after 8 weeks For patients receiving once-weekly therapy, if after 4 weeks of therapy the hemoglobin has not increased by >1 g/dL, in the absence of RBC transfusion, the epoetin alfa dose should be increased to 60,000 units weekly Recommended target hemoglobin: 10 g/dL to 12 g/dL

Table 1. Erythropoietic Stimulants Available Commercially in the United States (continued)

Drug	Trade name(s)	Half-life	Labeled indications	Initial dose recommendations for anemia of cancer chemotherapy	Recommended dosage adjustments for anemia of cancer chemotherapy
darbepoetin alfa	Aranesp® (Amgen, Inc., 2006)	observed half-life after SC administration in chronic renal failure patients: 49 hours (range: 27 to 89 hours) after IV administration, there is a distribution half-life of ~1.4 hours and a mean terminal half-life of 21 hours after IV administration, the terminal half-life of darbepoetin alfa is approximately 3-fold longer than epoetin alfa	anemia of chronic renal failure anemia in cancer patients receiving chemotherapy	2.25 mcg/kg SC weekly or 500 mcg SC every 3 weeks	If the hemoglobin exceeds 13 g/dL, doses should be temporarily withheld until the hemoglobin falls to 12 g/dL. At this point, therapy should be reinitiated at a dose approximately 25% below the previous dose If hemoglobin increases by more than 1.0 g/dL in a 2-week period or if the hemoglobin exceeds 12 g/dL, the dose should be reduced by approximately 25% If there is less than a 1.0 g/dL increase in hemoglobin after 6 weeks of therapy, the dose should be increased up to 4.5 mcg/kg Target hemoglobin should not exceed 12 g/dL in men or women

Abbreviations: IV, intravenously; SC, subcutaneously
*Epoetin alfa preparations are derived from the same source and are identical in composition (McEvoy, 2005)

Key Questions 1–3 address questions of therapeutic outcome, for which we required evidence from randomized controlled trials. Key Question 4 addresses predicting responses to erythropoietic stimulants, to which we applied an approach used to evaluate diagnostic tests.

Two reviewers screened all article titles and abstracts identified by the search strategy (see Search Strategy; Appendix A). If eligibility could not be assessed satisfactorily from the title and abstract, we retrieved the article in full text.

Types of participants

- All trials included patients diagnosed with malignant disease and undergoing treatment with chemotherapy or radiotherapy. Other reasons for anemia, such as hemolysis, iron deficiency, and occult bleeding, should have been ruled out.

- Trials were excluded if (a) patients were not undergoing treatment for cancer, or (b) treatment was high-dose myeloablative therapy with stem-cell transplant, or (c) patients had myelodysplastic syndrome.

- Also excluded were trials using epoetin for short-term preoperative treatment to correct anemia or to support collection of autologous blood prior to cancer surgery.

Types of interventions

- Trials were included for Key Question 1 if they directly compared epoetin and darbepoetin in patients undergoing cancer treatment. Also included were studies comparing epoetin or darbepoetin versus observation (alone or with placebo) until red blood cell transfusions were necessary.

- If epoetin (alfa or beta) was not administered subcutaneously or intravenously at doses of at least 300 U/kg body weight per week for at least four weeks, trials or study arms were excluded for Key Question 1 (e.g., arms a and b from Cazzola, Messinger, Battistel, et al., 1995). Data were abstracted on all darbepoetin doses for which outcomes were reported separately by study arm/dose level.

- For Key Questions 2 and 3, trials were included if they directly compared two different methods for using epoetin or darbepoetin to manage anemia in patients undergoing cancer treatment:

    - Alternative dosages or treatment schedules are relevant interventions for Key Question 2.

    - Alternative thresholds to initiate therapy; criteria to discontinue therapy; or durations of therapy are relevant interventions for Key Question 3.

- No minimal epoetin (alfa or beta) or darbepoetin dose was required for trials comparing alternative dosing schemes or treatment schedules (Key Question 2).

- Interventions relevant to Key Question 4 were laboratory measures for hematologic parameters at baseline or in the first 4 weeks of treatment that might be used to predict responses to epoetin or darbepoetin.

- Adjusting epoetin or darbepoetin dose based on hematologic response was allowed for all Key Questions.

- Concomitant supportive treatments, e.g., granulocyte colony-stimulating factors (G-CSF) or iron supplementation, and cancer therapies had to be given equally in all study arms.

Types of outcome measures

- **Hematologic response**. Proportion of patients with an increase in hemoglobin level of 2 g/dL or more by end of study or an increase in hematocrit of 6 points or more by end of study, independent of blood transfusions. Of studies that reported hematologic responses, 2 g/dL or more was the most consistently used definition. It was also a robust response,

not easily achieved in those receiving placebo or no treatment. Data from studies using other definitions were abstracted and summarized in the report, but were not pooled for meta-analysis with data conforming to this definition. Note that study lengths were 6–16 weeks in duration. Thus, this aggregate outcome measure does not conflict with the FDA labeling, which states that dosage should be reduced if Hb increases more than 1 g/dL in any 2-week period, as in any study, individual patients may experience Hb increases that require dose reduction or temporary discontinuation.

- **Transfusion**. Proportion of patients receiving red blood cell transfusions.

- **Quality of life (QoL)**. Preferred measures were validated instruments, such as SF-36; EORTC Quality of life Questionnaire (QLQ-C30); Functional Assessment of Cancer Therapy (FACT, including G-General; F-Fatigue; An-Anemia). Visual analog scales (VAS) (including versions named linear analog self-assessment [LASA] and cancer linear analog scale [CLAS]), although initially excluded, were also abstracted. Sample size and amount of missing data for QoL measures were extracted.

- **Tumor response**. Tumor response was only evaluated from studies prospectively designed to assess tumor response. These were trials with a homogeneous patient population undergoing a predefined cancer therapy.

- **Overall survival.** For some studies that did not report survival, unpublished survival data were obtained from investigators by the Cochrane Hematologic Malignancies Review Group, who made the data available for this review.

- **Adverse effects.** Included thromboembolic events, hypertension, thrombocytopenia and/or hemorrhage, rash and similar symptoms, and seizures. Additionally, we abstracted data on development of antibodies to epoetin or darbepoetin, since such antibodies might also bind to and neutralize endogenous erythropoietin, thus impairing normal erythropoiesis.

Key Questions 1–3 assessed all outcomes cited here except for tumor response, which was assessed in Key Question 1 only. For Key Question 4, hematologic response was the only outcome assessed.

Types of studies

- All studies included for Key Questions 1–3 were randomized controlled trials, with at least 10 participants per study arm, published in any language. Ongoing studies and interim analyses were excluded.

- For Key Question 1, trials compared (a) epoetin to darbepoetin, or (b) epoetin to no epoetin, or (c) darbepoetin to no darbepoetin.

- For Key Question 2, trials directly compared at least two alternative dosing schemes or treatment schedules.

- For Key Question 3, included studies directly compared (a) at least two different thresholds to initiate treatment, or (b) at least two alternative criteria to discontinue treatment, or (c) at least two durations of treatment.

- For Key Question 4, non-randomized controlled clinical trials and prospective cohort studies were included in addition to randomized controlled clinical trials.

Studies included in Key Question 4 were designed to prospectively test predictive factors for hematologic response in patients responding and not responding to treatment with erythropoietic stimulants. Predictive factors were patient characteristics at baseline or early hematologic changes in the first four weeks after initiating treatment.

Chapter 2. Methods

Technical Expert Panel

A technical expert panel (TEP) provided consultation for the systematic review (see Appendix E for a list of panel members). Specifically, they helped develop the final key questions, systematic review protocol, and commented on an early draft of the review.

Literature Search

The following databases were searched electronically.

- MEDLINE (January 1999 to March 2005),
- EMBASE (January 1999 to March 2005), and
- Cochrane Central Register of Controlled Trials Register (CENTRAL, January 1999 to March 2005).

The full search strategy is displayed in Appendix A. Literature search databases included fields for errata and other trial-related publications such as letters and special reports; the contents of all documents and publications related to included trials were screened.

Data previously abstracted from studies reviewed for the first Cochrane Review (Erythropoietin for patients with malignant disease; Bohlius, Langensiepen, Schwarzer, et al., 2005) or AHRQ report (Seidenfeld, Aronson, Piper, et al., 2001) were updated if necessary and included in the present report.

We sought additional studies by searching reference lists of included studies, relevant review articles, and relevant clinical practice guidelines.[2]

The following conference proceedings were searched electronically or by hand if they were unavailable in electronic format:

- American Society of Clinical Oncology (January 1999–May 2005),
- American Society of Hematology (January 1999–March 2005),

[2] Guidelines searched included those of the American Society of Hematology/American Society of Clinical Oncology (ASH/ASCO; Rizzo, Lichtin, Woolf, et al., 2002), Cancer Care Ontario Practice Guidelines (CCOPG; Quirt, Bramwell, Charette, et al., 2005), European Organization for Research and Treatment of Cancer (EORTC; Bokemeyer, Aapro, Courdi et al., 2004), Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC; Federation Nationale des Centres de Lutte Contre le Cancer, 2003), and the National Comprehensive Cancer Network (NCCN; National Comprehensive Cancer Network, 2004)

- European Society of Medical Oncology (January 1999–March 2005).

Abstracts selected from conference proceedings were traced for full-text publications.

Finally, from the Food and Drug Administration (FDA) web site, we identified one briefing document from a May 2004 meeting of the Oncologic Drugs Advisory Committee (ODAC) plus an additional Microsoft® PowerPoint® presentation prepared by medical reviewers of the FDA, and three documents, plus additional PowerPoint® presentations prepared by the companies Roche, Johnson & Johnson, and Amgen. All of these documents are publicly available through the FDA briefing document at http://www.fda.gov/ohrms/dockets/ac/04/slides/4037s2.htm (slides) and http://www.fda.gov/ohrms/dockets/ac/04/briefing/4037b2.htm (briefing documents).

Study Selection

We assessed titles and/or abstracts of citations identified from literature searches for inclusion, using the criteria described in the "Introduction and Scope" section. Full-text articles of potentially relevant abstracts were retrieved and a second review for inclusion was conducted by reapplying the inclusion criteria. Results published only in abstract form lack adequate information to assess the validity of the data. Nevertheless, in an effort to include the most recent data possible, we included abstracts from the conference proceedings listed if full-text articles were not published subsequently. The QUOROM diagrams (Figures 1 and 2) outline the selection of articles for inclusion in the review. Table 2 provides the included citations and table/figure designations for the Key Questions.

Data Extraction

A standardized data extraction form was used (Appendix B). Data extraction from randomized, controlled trials (RCTs) on epoetin or darbepoetin versus control for Key Question 1 was independently performed by two reviewers. In addition, plots and tables were fact-checked by a third reviewer. For all other studies and questions, data were extracted by one reviewer then checked by a second reviewer. Disagreements arising at any stage were resolved by discussion and consensus.

Handling of discrepant data. For studies published in multiple articles, reports or presentations, we extracted the most recent or most comprehensive data. The data of any study taken from different sources were compared. If data from different sources were discrepant, data were selected for analysis using the following rules:

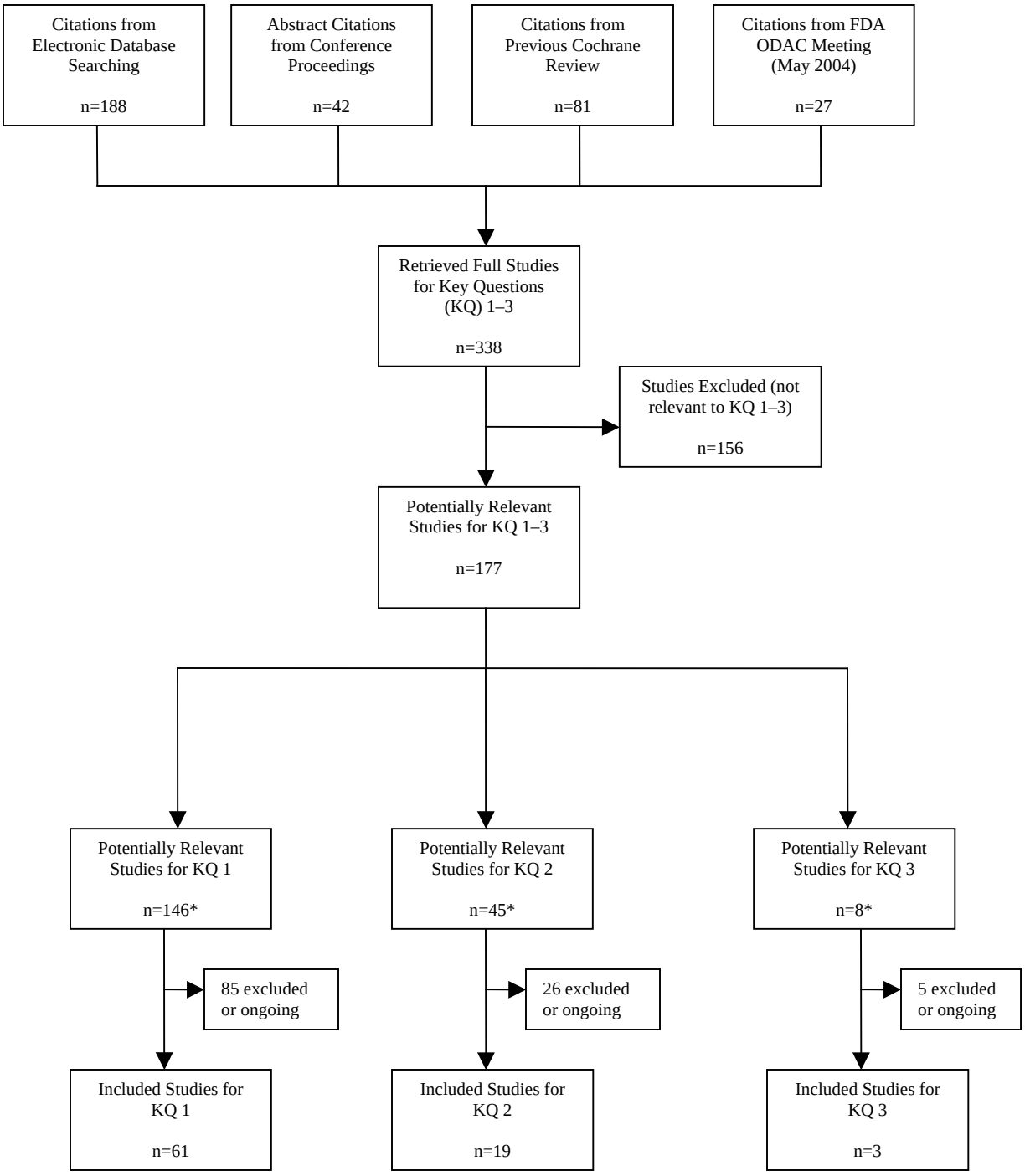
- For survival, data with longest follow up or highest number of deaths were used for analysis.

- For other outcomes, the most complete data sets were used (i.e., those with the largest sample size), or with consistently defined outcomes across trials.

- If different results were available from the same study (e.g., adjusted and unadjusted) we used the unadjusted data for a base-case analysis, then explored the influence of alternative results in sensitivity analyses.

Handling of incompletely reported numbers. If a study only reported the overall number of randomized patients but failed to report the number of patients per study arm we assigned 50 percent of the study patients to each of the study arms. In some cases, this reflected a reported 1:1 assignment; in other cases it was assumed as the most common trial design. This occurred in 10 out of 46 studies of epoetin vs. control, no studies of darbepoetin vs. control, and two studies of epoetin vs. darbepoetin.

Figure 1. QUOROM Diagram, Key Questions 1–3



*Note: There is some study overlap.

Figure 2. QUOROM Diagram, Key Question 4

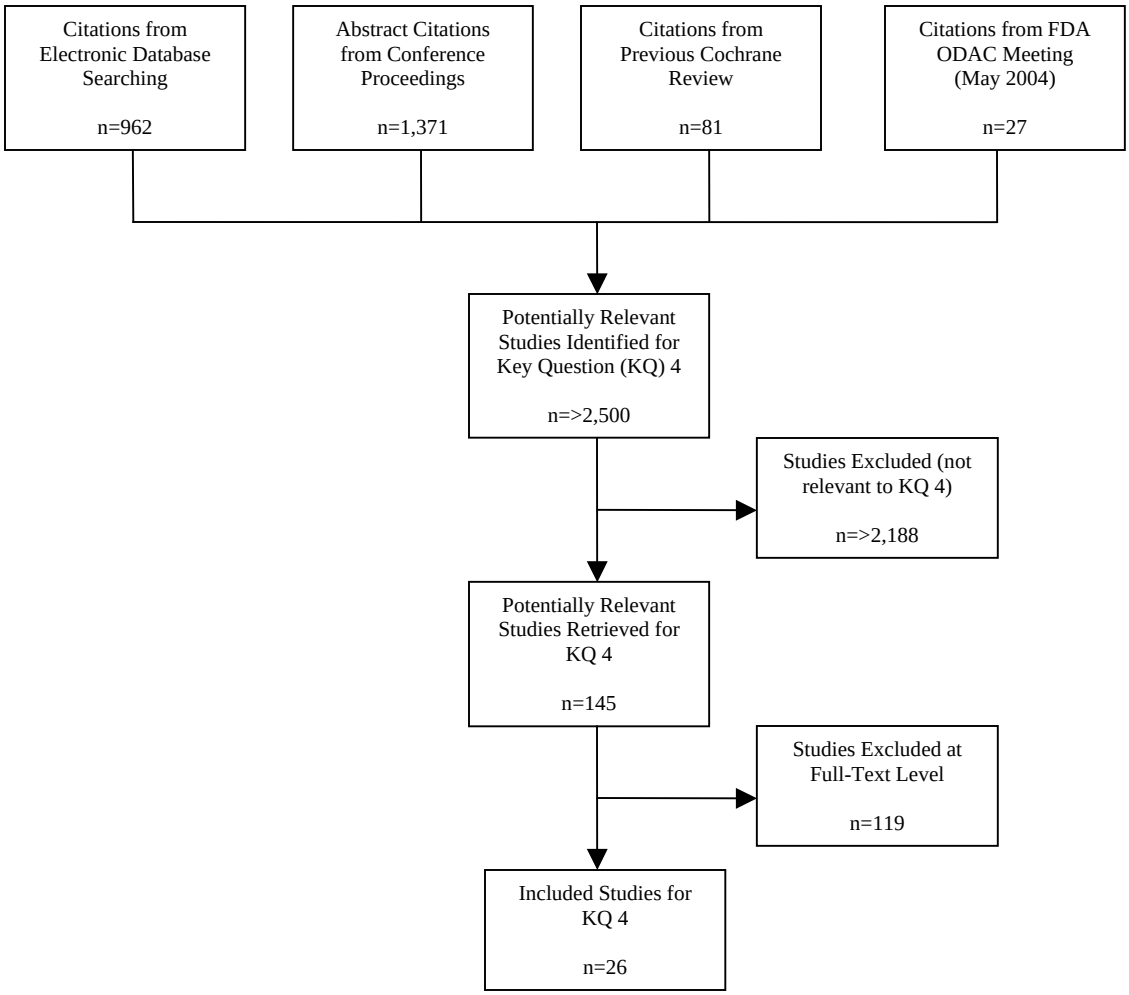


Table 2. Included Studies and Figure/Table Designations

A. Included Studies, Key Question 1

Study	Publication type(s)	Figure/table designation
Darbepoetin versus Epoetin		
Glaspy and Tchekmedyian, 2002B	Full Text	Glaspy 2002 PB a Glaspy 2002 PB b Glaspy 2002 PB c Glaspy 2002 PB d
Glaspy, Jadeja, Justice, et al., 2003	Full Text	Glaspy 2003a Glaspy 2003b Glaspy 2003c
Waltzman, Croot, Williams, 2005[1]	Abstract	Waltzman 2005
Schwartzberg, Yee, Senecal, et al., 2004	Full Text	Schwartzberg 2004a Schwartzberg 2004b Schwartzberg 2004c
Alexopolous and Kotsori, 2004	Abstract	Alexopolous 2004
Glaspy, Berg, Tomita, et al., 2005	Abstract	Glaspy 2005
Glaspy, Jadeja, Justice, et al., 2002A	Full Text	Glaspy 2002 PA c Glaspy 2002 PA d Glaspy 2002 PA e
Epoetin versus Control		
Aravantinos, Linardou, Makridaki, et al., 2003	Full Text	Aravantinos 2003
Bamias, Aravantinos, Kalofonos, et al., 2003	Full Text	Bamias 2003
Boogaerts, Coiffier, Kainz, 2003/Coiffier and Boogaerts, 2001	Full Text, Abstract, Unpublished Data, FDA Documents	Boogaerts 2003 Coiffier 2001
Carabantes, Benavides, Trujillo, et al., 1999	Abstract	Carabantes 1999
Cascinu, Fedeli, Del Ferro, et al., 1994	Full Text, Unpublished Data	Cascinu 1994
Case, Bukowski, Carey, et al., 1993	Full Text, Unpublished Data, FDA Documents	Case 1993
Cazzola, Messinger, Battistel, et al, 1995	Full Text, Unpublished Data, FDA Documents	Cazzola 1995c Cazzola 1995d
Chang, Couture, Young, et al., 2005	Full Text	Chang 2005
Dammacco, Castoldi, Rodjer, et al., 2001	Full Text, Unpublished Data, FDA Documents	Dammacco 2001
Del Mastro, Venturini, Lionetto, et al., 1997	Full Text, Unpublished Data	Del Mastro 1997
Dunphy, Harrison, Dunleavy, et al., 1999	Full Text	Dunphy 1999
EPO-CAN-15	FDA Documents	EPO-CAN-15
EPO-CAN-20	FDA Documents	EPO-CAN-20
EPO-GBR-07	FDA Documents	EPO-GBR-07
GOG-191	FDA Documents	GOG-191
Henke, Laszig, Ruebe, et al., 2003	Full Text, FDA Documents	Henke 2003
Henry, Brooks, Case, et al., 1995	Full Text, Unpublished Data, FDA Documents	Henry 1995
Henze, Michon, Morland, et al., 2002	Abstract	Henze 2002
Huddart, Welch, Chan, et al., 2002	Abstract	Huddart 2002
Iconomou, Koutras, Rigopoulos, et al., 2003	Full Text	Iconomou 2003
INT-1	FDA Documents	INT-1
INT-3	FDA Documents	INT-3

[1] As this report went to press, a full-text version of this trial was published (Waltzman, Croot, Justice, et al., 2005).

Table 2. Included Studies and Figure/Table Designations

A. Included Studies, Key Question 1 (continued)

Study	Publication type(s)	Figure/table designation
Epoetin versus Control (continued)		
Janinis, Dafni, Aravantinos, et al., 2003	Abstract	Janinis 2003
Kunikane, Watanabe, Fukuoka, et al., 2001	Full Text	Kunikane 2001a Kunikane 2001b
Kurz, Marth, Windbichler, et al., 1997	Full Text, Unpublished Data	Kurz 1997
Leyland-Jones, 2003[1]	Full Text, FDA Documents	Leyland-Jones, 2003
Littlewood, Bajetta, Nortier, et al., 2001	Full Text, Unpublished Data, FDA Documents	Littlewood 2001
Machtay, Pajak, Suntharalingam, et al., 2004	Abstract, FDA Documents	Machtay 2004
N93 004[2]	FDA Documents	N93 004
Oberhoff, Neri, Amadori, et al., 1998	Full Text, Unpublished Data, FDA Documents	Oberhoff 1998
O'Shaughnessy, Vukelja, Holmes, et al., 2005	Full Text	O'Shaughnessy, 2005
Osterborg, Boogaerts, Cimino, et al., 1996	Full Text, Unpublished Data, FDA Documents	Osterborg 1996a Osterborg 1996b
Osterborg, Brandberg, Molostova, et al., 2002/Osterborg, Brandberg, Hedenus, 2005	Full Text, Unpublished Data, FDA Documents	Osterborg 2002 Osterborg 2005
P-174, 2004	FDA Documents	P-174
Quirt, Micucci, Moran, et al., 1996	Abstract	Quirt 1996
Razzouk, Hockenberry, Hinds, et al., 2004	Abstract	Razzouk 2004
Rose, Rai, Revicki, et al., 1994	Abstract, Unpublished Data, FDA Documents	Rose 1994
Rosenzweig, Bender, Lucke, et al., 2004	Full Text, FDA Documents	Rosenzweig, 2004
Savonije, Van Groeningen, Van Bochove, et al., 2004[3]	Abstract	Savonije 2004
Silvestris, Romito, Fanelli, et al., 1995	Full Text	Silvestris 1995
ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998	Full Text, Unpublished Data, FDA Documents	ten Bokkel 1998a ten Bokkel 1998b
Thatcher, De Campos, Bell, et al., 1999	Full Text, Unpublished Data, FDA Documents	Thatcher 1999a Thatcher 1999b
Thomas, McAdam, Thomas, et al., 2002	Abstract	Thomas 2002
Throuvalas, Antonadou, Boufi, et al., 2000	Abstract, Unpublished Data	Throuvalas 2000
Vadhan-Raj, Skibber, Crane, et al., 2004	Abstract, FDA Documents	Vadhan-Raj 2004
Welch, James, Wilkinson, 1995	Full Text	Welch 1995
Witzig, Silberstein, Loprinzi, et al., 2005	Full Text, FDA Documents	Witzig 2005
Wurnig, Windhager, Schwameis, et al., 1996	Full Text	Wurnig 1996

[1] As this report went to press, a full-text version of this trial was published (Leyland-Jones, Semiglazov, Pawlicki, et al., 2005)
[2] As this report went to press, a full-text version of this trial was published (Grote, Yeilding, Castillo, et al., 2005).
[3] As this report went to press, a full-text version of this trial was published (Savonije, van Groeningen, van Bochove, et al., 2005)

Table 2. Included Studies and Figure/Table Designations

A. Included Studies, Key Question 1 (continued)

Study	Publication type(s)	Figure/table designation
Darbepoetin versus Control		
Hedenus, Hansen, Taylor, et al., 2002	Full Text	Hedenus 2002a Hedenus 2002b Hedenus 2002c
Hedenus, Adriansson, San Miguel, et al., 2003[1]	Full Text	Hedenus 2003
Kotasek, Steger, Faught, et al., 2003	Full Text	Kotasek 2003a Kotasek 2003b Kotasek 2003c Kotasek 2003d Kotasek 2003e Kotasek 2003f
Vansteenkiste, Pirker, Massuti, et al., 2002	Full Text	Vansteenkiste 2002

[1]As this report went to press, an additional analysis of quality of life data from this trial was published (Littlewood, Kallich, San Miguel, et al., 2006)

B. Included Studies, Key Question 2

Study	Publication type(s)	Figure/table designation
Cazzola, Beguin, Kloczko, et al., 2003	Full Text	Cazzola 2003
Cazzola, Messinger, Battistel, et al, 1995	Full Text, Unpublished Data, FDA Documents	Cazzola 1995
Glaspy and Tchekmedyian, 2002B	Full Text	Glaspy 2002 part B
Glaspy, Jadeja, Justice, et al., 2003	Full Text	Glaspy 2003
Glimelius, Linne, Hoffman, et al., 1998	Full Text	Glimelius 1998
Granetto, Ricci, Martoni, et al., 2003	Full Text	Granetto 2003
Hedenus, Hansen, Taylor et al. 2002	Full Text	Hedenus 2002
Hesketh, Arena, Patel, et al., 2004	Full Text	Hesketh 2004
Johansson, Wersall, Brandberg, et al., 2001	Full Text	Johansson 2001
Justice, Kessler, Jadeja, et al., 2005	Full Text	Justice 2005
Kotasek, Canon, San Miguel, et al., 2004	Abstract	Kotasek 2004
Kotasek, Steger, Faught et al. 2003	Full Text	Kotasek 2003
Kunikane, Watanabe, Fukuoka, et al., 2001	Full Text	Kunikane 2001
Olsson, Svensson, Sundstrom, et al., 2002	Full Text	Olsson 2002
Osterborg, Boogaerts, Cimino, et al., 1996	Full Text, Unpublished Data, FDA Documents	Osterborg 1996
Sakai, Ohashi, Hirashima, et al., 2004	Abstract	Sakai 2004
Steensma, Molina, Sloan, et al., 2005[1]	Abstract	Steensma 2005
ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998	Full Text	ten Bokkel 1998
Thatcher, De Campos, Bell, et al., 1999	Full Text	Thatcher 1999

[1]As this report went to press, a full-text version of this trial was published (Steensma, Molina, Sloan et al., 2006).

C. Included Studies, Key Question 3

Study	Publication type(s)	Figure/table designation
Rearden, Charu, Saidman, et al., 2004	Abstract	Rearden 2004
Straus, Testa, Riggs, et al., 2003	Abstract	Straus 2003
Crawford, Robert, Perry, et al., 2003	Abstract	Crawford 2003

Table 2. Included Studies and Figure/Table Designations

D. Included Studies, Key Question 4

Study	Publication type(s)	Figure/table designation
Boogaerts, Coiffier, Kainz, 2003	Full Text	Boogaerts 2003
Cascinu, Fedeli, Del Ferro, et al., 1994	Full Text	Cascinu 1994
Case, Bukowski, Carey, et al., 1993	Full Text	Case 1993
Cazzola, Beguin, Kloczko, et al., 2003	Full Text	Cazzola 2003
Cazzola, Messinger, Battistel, et al, 1995	Full Text	Cazzola 1995
Chang, Couture, Young, et al., 2005	Full Text	Chang 2005
Demetri, Kris, Wade, et al. 1998	Full Text	Demetri 1998
Fjornes, Wiedemann, Sack, et al., 1998	Full Text	Fjornes 1998
Garton, Gertz, Witzig, et al., 1995	Full Text	Garton 1995
Glaspy, Bukowski, Steinberg, et al., 1997	Full Text	Glaspy 1997
Glimelius, Linne, Hoffman, et al., 1998	Full Text	Glimelius 1998
Gonzalez, Ordonez, Jua, et al., 1999	Abstract	Gonzalez 1999
Gonzalez-Baron, Ordonez, Franquesa, et al., 2002	Full Text	Gonzalez-Baron 2002
Hedenus, Hansen, Taylor, et al. 2002	Full Text	Hedenus 2002
Henry, Abels, and Larholt	Letter	Henry 1995a
Kasper, Terhaar, Fossa, et al., 1997	Full Text	Kasper 1997
Katodritou, Speletas, Kapetanos, et al., 2004	Abstract	Katodritou 2004
Littlewood, Zagari, Pallister, et al., 2003	Full Text	Littlewood 2003
Ludwig, Fritz, Leitgeb, et al., 1994	Full Text	Ludwig 1994
Ludwig, Sundal, Pecherstorfer, et al. 1995	Full Text	Ludwig 1995
McKenzie, Lefebvre, Rosberg, et al., 2004	Abstract	McKenzie 2004
Miller, Platanias, Mills, et al., 1992	Full Text	Miller 1992
Musto, Falcone, D'Arena, et al., 1997	Full Text	Musto 1997
Oberhoff, Neri, Amadori, et al., 1998	Full Text	Oberhoff 1998
Osterborg, Boogaerts, Cimino, et al., 1996	Full Text	Osterborg 1996
Witzig, Silberstein, Loprinzi, et al., 2004	Full Text	Witzig 2004

If percentages but not absolute numbers were reported for any outcome, we calculated absolute numbers based on the reported percentage and sample size per arm.

Some studies reported Kaplan-Meier estimates but not absolute numbers. In these cases, we used the Kaplan-Meier estimates as percentages and recorded the Kaplan-Meier estimates in the relevant evidence tables.

Allocation of treatment arms for Key Questions 2 and 3. To compare different active study arms, we allocated them to "intervention" and "control" arms as displayed in Table 3.

Table 3. Allocation of Study Arms

Type of Intervention	Arm Assigned to "Intervention"	Arm Assigned to "Control"
Dose escalation	Higher (single) dose	Lower (single) dose
Weight-based/fixed	Fixed dose	Weight-adjusted dose
Frequency of administration	Lower frequency	Higher frequency
Front-loading/titration schedules	Group with changing dose	Group with constant dose
Initiating treatment	Early or "immediate" therapy	Late or "delayed" therapy

Quality Assessment

Key Questions 1–3

Study quality characteristics were abstracted. Items abstracted included whether allocation was random, whether treatment allocation was concealed; blinding of participants and clinicians to treatment received; whether loss of patients was similar across study arms; whether analysis was intention-to-treat (ITT); whether participant characteristics were similar at baseline across study arms. These categories were used only for descriptive purposes.

For the subgroup analysis, studies that met all three criteria below were defined as higher-quality trials.

- The study was a randomized controlled trial.
- The study was double-blind.
- At least one of the following conditions was true:
    - less than 10 percent of subjects within each study arm were excluded from the analysis AND the percentage of subjects excluded from analysis in each arm was less than 2:1; OR
    - less than 5 percent of subjects were excluded in each study arm.

One reviewer performed the quality assessment, and a second reviewer checked the results. Discordance was resolved by consensus.

In the original Cochrane Review (including studies for Key Question 1 published before May 2002) all first authors or sponsoring pharmaceutical companies of the included trials were contacted to obtain information on the study design. This was not done with any other studies.

Key Question 4

Included studies were first classified in a manner analogous to the different phases of clinical trials evaluating interventions (phase I–IV). Possible classification systems for predictive factor studies have been developed (Boracchi and Biganzoli, 2003; Infante-Rivard, Villeneuve, Esnaola, 1989; McGuire, 1991; Pepe, 2003; Schumacher, Hollander, Schwarzer, et al., 2001; Simon and Altman, 1994), but agreement on a standard system is lacking. Therefore, a 3-level classification system was developed for this review and is summarized in Table 4.

In addition to study classification, studies included for Key Question 4 were assessed for specific quality criteria. Although specific assessment tools for predictive factor studies were not found, studies of predictive factors are related to diagnostic and prognostic factor studies. Several authors have formulated minimum criteria for these kinds of studies or statistical methods employed (Boracchi and Biganzoli, 2003; Infante-Rivard, Villeneuve, Esnaola, 1989; McGuire, 1991; Pepe, 2003; Hollander, Schwarzer, et al., 2001; Simon and Altman, 1994;

Altman, 2001; Concato, Feinstein, Holford, 1993; Justice, Covinsky, Berlin, 1999; Altman and Royston, 2000; Hollander and Schumacher, 2001; Bossuyt, Reitsma, Bruns, et al., 2003). From these guidelines, a list of 19 quality assessment criteria was developed (Table 5).

Table 4. Classification System for Predictive Factor Studies

Classification	Description	Utility
I	Exploratory study, i.e., no clear statement if possible predictive factors had been defined before the study and/or analysis started, no refutable hypotheses	Hypothesis-generating
II	Study prospectively evaluating/testing possible predictive factors, i.e., a restricted set of factors had been defined before the study started, refutable hypotheses	Hypothesis-testing
III	Study fulfilling the criteria as defined by Simon & Altman 2001 (e.g., prospective study, prespecified hypotheses, study specifically designed to evaluate predictive factors, prospective power calculation) or a randomized controlled trial employing a predictive factor/model in one arm and standard treatment in the other arm	Results may be used to guide clinical practice

Table 5. Quality Criteria Assessed for Studies Included in Key Question 4

Assessed for all studies:	
1	Study classification (see above)
2	Refutable hypothesis reported (Authors should state minimum requirements of performance measures or other requirements that a predictive factor is satisfying.)
3	Objective prospectively defined
4	Inclusion criteria defined for predictive factors study (Yes if inclusion explicitly stated [e.g., all patients were included for which baseline erythropoietin levels and data for response status were available]); Unclear if inclusion criteria were not explicitly stated but reasonable to assume that all patients treated with Epo/evaluated for Hb response were included; No for all other studies)
5	Sample size calculation and method used if applicable
6	Number and characteristics of excluded patients reported (Yes/Partially if explicitly stated; Unclear if not explicitly stated but reasonable to assume that all patients treated/evaluated for Hb response were included; No for all other studies)
7	Missing data handling reported, including losses to follow-up reported
8	Internal validation of discovered predictive factors and method used if applicable (e.g., splitting sample in training and validation set)
9	Follow-up of patients at least 4 weeks
10	Selection process of possible predictive factors explained and adequate (e.g., based on previous studies, biological hypotheses)
11	Cut-off values for continuous variables explained and adequate (Yes if based on statistical tests for example; Partially if method unsatisfactory [e.g., arbitrarily chosen or medians used]; No for all other studies)
12	Performance measures reported (e.g., sensitivity, specificity)
13	Method of statistical analysis (just descriptive no assessment of adequacy)
14	Prognostic variables fully defined (This is mostly relevant for non-standard laboratory values but may also apply to factors not clearly described)
Assessed if multivariate methods were used:	
15	Statistical package used (just descriptive no assessment of adequacy)
16	Coding of variables reported (relevant for a continuous variable coded as ranked variable)
17	Problem with overfitting (A cut-off of 10 events per tested variable was chosen for the label "probable")
18	Conformity of linearity for ranked variables reported
19	Tests of interaction performed

Data Synthesis

Where data allowed, quantitative methods were used to summarize outcomes of epoetin or darbepoetin treatment. Known clinical heterogeneity, and discovered statistical heterogeneity in some cases warranted exploration of patient subgroups. For a discussion of heterogeneity, impact on meta-analysis, and methods of evaluation, see Appendix F.

Procedure. Most analyses were performed using Review Manager (RevMan), 4.2.5; the statistical software package R (Ihaka and Gentleman, 1996) was used for additional analyses (e.g., meta-regression) that cannot be done with RevMan 4.2.5.

A fixed-effects model was initially assumed for all meta-analyses. For binary data, the relative risk was used as a measure of treatment effect and we used the Mantel-Haenszel method for pooling in RevMan. The p-value of the homogeneity test and the I^2 statistic were used to describe the extent of heterogeneity inherent in a meta-analysis. When the value of I^2 was greater than 25 percent, a random-effects analysis (RevMan) was also conducted. For primary outcome measures potential causes of heterogeneity were explored by performing sensitivity and subgroup analyses. The statistical significance of differences in effect among subgroups was calculated by the inverse variance method. The resulting p-value of subgroup differences is based on the partitioning of heterogeneity: Chi^2(between groups) = Chi^2(all) – Chi^2(within groups).

The estimated overall relative risk and a range of plausible values for the baseline-risk were used to estimate numbers needed to treat (NNT) and numbers needed to harm (NNH) for selected outcomes. Where there was significant statistical heterogeneity across studies, a L'Abbe plot (L'Abbe, Detsky, O'Rourke, 1987) was utilized to assess the constancy of the pooled treatment effect prior to calculating NNT or NNH.

Time-to-event data, i.e., overall survival, were calculated as hazard ratios (HR) based on individual patient data (IPD). If IPD were not available the HR was calculated (i) from published reports, using methods described in Parmar, Torri, and Stewart (1998), or (ii) from binary mortality data. For the latter method, numbers of deaths and sample sizes were imputed in the corresponding section in RevMan and processed with "calculate."

In addition to subgroup analyses, a fixed-effects meta-regression, i.e., method "1" in Thompson and Sharp (1999), was conducted for the outcome "proportion of participants transfused." For this analysis, data from RCTs comparing epoetin or darbepoetin versus control were pooled together. All covariates showing a significant effect ($p < 0.05$) in univariate analyses were included in the regression. For model selection, the data set was restricted to studies that provided information on all variables found statistically significant in univariate analyses. Next, a back-wise selection method was used; the covariate with the largest p-value was removed consecutively until the only remaining covariates were significant according to the Akaike Information Criterion (Akaike, 1969). For a more detailed description of the meta-regression see the subsection on "meta-regression" in the section on transfusion for Key Question 1.

Several studies compared different epoetin or darbepoetin dosages, routes, or schedules of administration versus one control group. For each of these studies, we artificially divided and randomly assigned control patients to the corresponding number of separate control groups for entry into RevMan (base model). As this might influence study weighting and thus pooled results, we merged the two (or more) active arms of any such study into one experimental arm

and compared it to that study's full control group. Results of these alternative analytic approaches were compared and described for each outcome.

Sensitivity Analysis and Subgroup Analysis

Subgroup analysis. We extracted data on the following patient, trial, publication, and quality characteristics, which were used for subgroup analyses when appropriate (Figure 3). However, formal subgroup analyses were performed only for Key Question 1. For Key Questions 2, and 3, insufficient numbers of trials addressing the same question were available to permit formal subgroup analysis.

- Patient baseline characteristics
    - Hemoglobin (Hb) at study entry (Hb $\leq$10 g/dL versus >10 but <12 g/dL versus $\geq$12 g/dL vs. unclear). Categorizations were based on the aggregated mean or median Hb at baseline. If hematocrit (Hct) was reported instead of Hb, we documented the Hct and converted it into Hb for categorization. If the baseline Hb or Hct was not reported, the study was categorized as "unclear."
    - Solid tumors versus hematologic malignancies versus mixed (including both solid and hematological malignancies) vs. unclear. Studies including solid tumors only were categorized as "solid tumors." Studies including hematological malignancies only were categorized as "hematological malignancies." Studies including both hematological and solid tumors were categorized as "mixed." Studies with imprecise information on the population evaluated, e.g., "cancer patients," were categorized as "unclear."
    - Age (elderly [aged >65 years] versus non-elderly adults versus children [$\leq$18 years]). Studies were categorized as "adults" if the majority of the population were adults. If a study was restricted to children ($\leq$18 years), the study was categorized as "children." If the study was restricted to elderly patients (e.g., age >65 years), the study was categorized as "elderly patients" (however, no included studies met the latter criterion).
    - Ethnicity. Not applied, as data were not available.
    - Gender (female versus male patients). Not applied, as data were not available.
- Treatment protocols
    - Type of treatment given. All studies were assigned to the following five different study groups:
        - Platinum-based chemotherapy: More than 70% of the study population received platinum-based chemotherapy.

– Some patients receiving platinum-based chemotherapy: Less than 70% of the patients received platinum-based chemotherapy.

– Chemotherapy without platinum: Studies with all patients receiving platinum-free chemotherapy

Figure 3. Patient, Study, and Reporting Variables Prespecified for Subgroup Analysis

- **Baseline characteristics of study populations**
  - average baseline Hb concentration
    - <10; 10–12; or >12 g/dL, unclear
  - type of malignancy
    - only solid tumors; only hematologic malignancies; mixed populations, unclear
  - age range
    - only adult patients; only pediatric patients
- **Treatment protocols**
  - therapies for malignancy
    - platinum for all; platinum for some; platinum for none; radiation ± chemotherapy, unclear
  - iron supplementation
    - fixed dose; if stores inadequate; not specified/no iron
  - study and treatment duration
    - 6–9 weeks; 12–16 weeks; >20 weeks, unclear
  - epoetin regimen
    - weight-based versus fixed-dose; thrice versus once weekly; dose adjustments
- **Publication type, quality ratings, and methods**
  - publication type
    - full-text; abstract only; unpublished; reported to FDA ODAC
  - overall quality rating
    - high-quality study; low quality study (based on next three factors)
      - randomization
        - randomized, controlled trial (excluded if not randomized)
      - double-blinding
        - investigators explicitly described trial as double-blinded
      - minimal loss to follow-up and analysis
        - intent-to-treat (ITT) analysis, or <10% loss with <2:1 ratio of loss per arm, or <5% loss per arm
  - other methodologic differences (see Study Characteristics tables, Appendix C)
    - placebo use
      - controls given placebo, controls untreated
    - allocation concealment
      - adequate, inadequate
    - trial arms well-balanced at baseline
      - groups well-balanced, important differences at baseline, inadequate information to assess balance
    - transfusion decisionmaking
      - at specified trigger; at physician discretion; not specified

- Radiotherapy/chemoradiotherapy: Patients receiving an anticancer regimen mainly based on radiotherapy. Whether chemotherapy was concomitantly administered was not evaluated in this analysis.

- Unclear: Some studies failed to report the anticancer treatment given. If insufficient information on the therapy or the cancer entity was reported, the study was categorized as "unclear."

- Iron supplementation (fixed vs. as necessary vs. unclear). Studies using a fixed dose and schedule of iron supplementation for all patients were categorized as "fixed." Studies supplementing patients with iron as necessary, i.e., if iron stores were measured and found deficient, were categorized as "as necessary." Studies either not using iron or not reporting on iron usage were categorized as "unclear."

- Duration of epoetin or darbepoetin treatment. Duration of treatment with epoetin or darbepoetin was categorized into the following subgroups: 6 to 9 weeks, 12 to 16 weeks, more than 20 weeks and unclear if the reporting was insufficient.[3] The following assumptions were made. If, for example, a study reported that epoetin was given for three chemotherapy cycles with a cycle length of three weeks, the duration of epoetin treatment was calculated to be 3 X 3 weeks = 9 weeks.

- For overall survival additionally: duration of follow up. The duration of followup was split into studies with follow up less than 1 year and studies with duration of followup greater than 1 year. If the duration of follow up was not reported or was not estimable from the available information the study was categorized as "unclear."

- Reporting and quality

- Study quality (high- versus low-quality studies). Studies were grouped into "higher" and "lower" quality studies. Higher-quality studies were randomized controlled trials; were double-blinded; and either, a) less than 10% of subjects within each study arm were excluded from the analysis AND the percentage of subjects excluded from analysis in each arm was less than 2:1; OR b) less than 5% of subjects were excluded in each study arm.

- Source of data (full-text publications versus abstract publications versus unreported data versus documents presented at FDA hearing). Data taken from full-text reports were categorized as "full-text publications." Data taken from abstract publications were categorized as "abstract publications." Unreported data of published studies that were submitted by the investigators for the first Cochrane Review were categorized as "unpublished data." Data of either unpublished or published studies that were reported and taken from one of the FDA documents were categorized as "FDA documents."

[3] Although discontinuous, these categories include all studies, i.e., no trials had treatment durations of 9–12 or 16–20 weeks.

Sensitivity analysis. With sensitivity analysis we explored the influence of single large studies in a meta-analysis and the use of different data sets, e.g., adjusted vs. unadjusted data.

For Key Question 4 ("Factors Predicting Response"), to allow an assessment of the power of different predictive factors, performance measures, i.e., specificity, sensitivity, predictive values were calculated whenever possible. Specificity and sensitivity depend on the study definitions of a positive test and of Hb outcome. Predictive values depend on prevalence, which for purposes of comparison across studies was assumed to be a number similar to the pooled result for hematologic outcomes in Key Question 1 of this review.

Peer Review

We requested peer review of the draft of this report from content or methodology experts and professional or patient advocacy organizations. The draft report was reviewed by external reviewers, including members of the technical expert panel, other invited technical experts, and stakeholders (see Appendix E). Revisions were made to the draft report based on reviewers' comments.

Chapter 3. Results

Key Question 1. What are the comparative efficacy and safety of epoetin (alfa or beta) and darbepoetin?

Overview of Evidence and Findings for KQ1

Three sets of relevant trials were summarized and analyzed for Key Question 1 (for full study details, please refer to Appendix C, KQ1 Appendix Tables C1–C42). Seven studies directly compared epoetin versus darbepoetin (pooled N=1,415 randomized to epoetin, 1,087 to darbepoetin); 48 RCTs tested epoetin versus control (pooled N=4,518 to epoetin, 3,743 to control); and four RCTs tested darbepoetin versus control (pooled N=598 to darbepoetin, 396 to control). Trials within each set differed with respect to outcomes reported, and variables prespecified for subgroup analysis on: study samples' baseline characteristics; treatment protocols; and publication type, quality ratings, and methods. Effects of baseline hemoglobin concentration on outcomes are also relevant to Key Question 3, which examines alternative thresholds for initiating treatment. To avoid duplication of Forest plots, those shown for this Key Question have trials grouped by mean (or median) baseline hemoglobin.

No trials reported outcomes separately by elderly vs. non-elderly adults, ethnicity, or gender. Only two trials studied pediatric populations (Razzouk, Hockenberry, Hinds, et al., 2004; Henze, Michon, Morland, et al., 2002); each compared epoetin versus control (N=456; 228 each to epoetin and control).

Major findings are summarized in Tables 6–12.

Table 6. Overview: Hematologic Response

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Number of studies	6	15	3	no combined analysis for this intermediate (surrogate) outcome; transfusion risk is the relevant primary outcome
Patients analyzed	2,205	3,293	659	
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2	not amenable to meta-analysis: trials defined response, initiated and adjusted doses, differently; only one RCT (N=352) found significant difference favoring epoetin, but may be biased: dose adjusted differently in each arm; five trials (N=1,853) reported no significant differences	3.42[1] (3.03, 3.86)[1] <0.00001 66%[1]	3.36 (2.48, 4.56) <0.00001 0	

[1] Since I^2>25%, compared fixed-effects analysis with random-effects analysis showing RR=3.73; (95% CI: 2.94, 4.74); p<0.00001
CI: confidence interval; RCT: randomized, controlled trial; RR: relative risk

Table 7. Overview: Transfusion Rates

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Number of studies	6	34	4	32[3]
Patients analyzed	2,158	5,210	950	5,063[3]
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2	1.10[1] (0.93, 1.29)[1] 0.27 42.8%[1]	0.63[2] (0.59, 0.67)[2] <0.00001 62.9%[2]	0.61 (0.52, 0.72) <0.00001 0	no significant difference, darbepoetin versus epoetin (p=0.35) by univariate analysis of all 38 trials; fixed-effects meta-regression[3] shows risk reduced more with solid tumors, shorter studies, and unpublished data
Number needed to treat (95% CI) by baseline risk : 30% 50% 70%		 9 (8, 10) 5 (5, 6) 4 (3, 4)	 9 (7, 12) 5 (4, 7) 4 (3, 5)	

[1] Since I^2>25%, compared fixed-effects analysis with random-effects analysis showing RR=0.87; 95% CI: 0.63, 1.20; p=0.40
[2] Since I^2>25%, compared fixed-effects analysis with random-effects analysis showing RR=0.60; 95% CI: 0.53, 0.67), p<0.00001
[3] Six trials (N=1,097) lacking information on one or more meta-regression variables were omitted from this analysis. Fixed-effect meta-regression analysis compared darbepoetin with epoetin indirectly, and explored causes of heterogeneity.

Table 8. Overview: Quality of Life

Parameter	Darbepoetin vs. epoetin	(Epoetin or darbepoetin) vs. control[1]
Number of studies	1	15 (13 of epoetin; 2 of darbepoetin)
Patients analyzed	731[2]	3, 610 randomized[3] (2,947 to epoetin; 663 to darbepoetin)
QoL instruments	FACT-An, FACT-fatigue subscale	FACT-An, FACT-G, FACT-fatigue and FACT-anemia non-fatigue subscales; various general measures; 3-item VAS
Results	No statistically significant differences between study arms in changes from baseline to 16 weeks	QoL results could not be combined quantitatively because of incomplete reporting; therefore we evaluated patterns of tabulated results. Main findings are: o no results significantly favored control for any QoL measure; o for each FACT measure, balance among results significantly favoring treatment, not significantly different, and significantly favoring control, favors treatment; o results from general measures were inconclusive due to heterogeneity of measures and few studies reporting any one measure; o for each VAS item, balance among results significantly favoring treatment, not significantly different, and significantly favoring control, favors treatment. Analysis of study quality detected threats to validity in most studies, including lack of blinding, unclear allocation concealment, missing data, and insufficient detail on methods of QoL instrument administration. The clinical significance of study results is uncertain.

[1]Studies of epoetin vs. control and darbepoetin vs. control were analyzed together as a class.
[2]40% of randomized patients not evaluable for QoL; study available only as abstract/poster.
[3]Proportion of enrolled patients not evaluable for QoL varied by study and by instrument, ranging from 0 to 63% and averaging close to 20%.
An: anemia; FACT: Functional Assessment of Cancer Therapy; G: general; QoL: quality of life; VAS: visual analog scale

Table 9. Overview: Survival

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Number of studies[1]	1	35	4	39
Patients analyzed	358	6,918	973	7,891
Pooled hazard ratio	Single Trial HR = 1.25[2]	1.11	0.96	1.08
(95% CI)	(0.76, 2.07)	(1.00, 1.22)	(0.78, 1.17)	(0.98, 1.18)
p-value for test of overall effect	0.4	0.05	0.66	0.11
Test for heterogeneity I^2		0%	72.2%	13.4%
HR (95% CI) for subgroups[3]:				
Labeled use		0.91 (0.47, 1.78)		0.91 (0.47, 1.78)
Unlabeled use		1.12 (1.01, 1.24)		1.09 (0.99, 1.19)
HR (95% CI) for subgroups[3]:				
Max Hb target: 13 g/dL		0.91 (0.47, 1.78)		0.91 (0.47, 1.78)
14 g/dL		1.16 (1.00, 1.35)		1.16 (1.00, 1.35)
15 g/dL		1.03 (0.90, 1.19)		1.01 (0.90, 1.13)
16 g/dL		1.67 (1.13, 2.48)		1.67 (1.13, 2.48)
Trend analysis		p=0.67		

[1] Only 7 (N=2,188) studies were prospectively designed to evaluate survival. Other studies may have collected retrospective data after study closure, so that patient management was no longer protocol-directed.
[2] Darbepoetin compared to epoetin
[3] Subgroup analyses (two shown here) failed to distinguish adverse studies (i.e. poorer survival with epoetin) from others.
CI: confidence interval; Hb: hemoglobin; HR: hazard ratio

Table 10. Overview: Tumor Response and Progression

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Tumor Response				
Number of studies	none	5[1]	none	not applicable
Patients analyzed		688		
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2		1.00 (0.92, 1.10) 0.91 0		
Tumor Progression		one trial (n=351) reported decreased progression-free survival with epoetin; four smaller trials (total N=585) reported no significant effect, but three of four closed prematurely and all likely were underpowered	one trial (n=314) reported progression-free survival did not differ significantly between arms over 24 months followup	

[1] Studies reported on solid tumors only; none reported on hematologic malignancies.
CI: confidence interval;

Table 11. Overview: Thromboembolic Events

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Number of studies	3	30	1	31
Patients analyzed	1,879	6,092	314	6,406
Pooled risk ratio[1] (95% CI) p-value for test of overall effect Test for heterogeneity I^2	0.86 (0.61, 1.21) 0.40 0%	1.69 (1.36, 2.10) <0.00001 0%	Single trial RR = 1.44 (0.47, 4.43)	1.68 (1.36, 2.08) <0.00001 0%
Number needed to harm (95% CI) by baseline risk: 2.5% 5% 10% 20%		 58 (36, 111) 29 (18, 56) 15 (9, 28) 7 (5, 14)		
RR (95% CI) for subgroups[2]: Labeled use (6.4% of patients) Unlabeled use (93.6% of patients)		 0.70 1.75 (p=0.046)		
RR (95% CI) for subgroups[2]: Max Hb target: 13 g/dL 14 g/dL 15 g/dL 16 g/dL Trend analysis		 0.70 (0.29, 1.67) 1.71 (1.23, 2.40) 1.92 (1.22, 3.02) 1.66 (1.08, 2.54) p=0.74		

[1] Unless otherwise noted

[2] Subgroup analyses are consistent with the explanation that Hb target >13 g/dL increases thromboembolic event risk; but may be confounded by small numbers in the ≤13 g/dL category and by other factors. There is no clear relationship between incremental increases in target Hb > 13 g/dL and RR for thromboembolic events; the trend is not statistically significant (p=0.742).

CI: confidence interval; Hb: hemoglobin; RR: relative risk;

Table 12. Overview: Other Adverse Events

Parameter	Darbepoetin vs. epoetin	Epoetin vs. control	Darbepoetin vs. control	(Epoetin or darbepoetin) vs. control
Hypertension[1]				
Number of studies	none	15	1	not done
Patients analyzed		1,949	314	
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2		1.22 (0.98, 1.52) 0.07 8.2%	1.54 (0.56, 4.22) 0.40 not applicable	
Thrombocytopenia/Hemorrhage				
Number of studies	none	9	none	not applicable
Patients analyzed		1,422		
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2		1.08 (0.76, 1.53) 0.66 0		
Rash				
Number of studies	none	6	none	not applicable
Patients analyzed		522		
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2		1.77 (0.82, 3.81) 0.14 0		
Seizures				
Number of studies	1	3	none	not applicable
Patients analyzed	122	389		
Pooled relative risk (95% CI) p-value for test of overall effect Test for heterogeneity I^2	no seizures in either study arm	1.19 (0.33, 4.35) 0.79 0		

[1] definition of hypertension not consistently reported

CI: confidence interval

Detailed Analysis

KQ1 Outcome I. Hematologic Response

This analysis excludes trials with mean or median baseline Hb >12 g/dL, and defines hematologic response as proportion of patients with hemoglobin (Hb) concentration increased from baseline by ≥2 g/dL, or hematocrit (Hct) by six percent, before end of study (see "Introduction" for rationale). Data were abstracted and summarized from trials that defined hematologic response differently, and are reported here qualitatively, but were not included for meta-analyses.

Darbepoetin versus Epoetin

Six trials (Appendix C Tables C6, C7, C10, and C13), compared hematologic response rates of patients randomized to darbepoetin versus epoetin (Glaspy and Tchekmedyian, 2002B; Glaspy, Jadeja, Justice, et al., 2003; Waltzman, Croot, Williams, 2005; Schwartzberg, Yee, Senecal, et al., 2004; Alexopoulos and Kotsori 2004; Glaspy, Berg, Tomita, et al., 2005). All were rated as poor study quality, since each was unblinded and described randomization methods inadequately. Results of these trials were not amenable to meta-analysis due to differences in definition of hematologic response, differences in initial doses, and, in three studies (Glaspy and Tchekmedyian, 2002B; Glaspy, Jadeja, Justice, et al., 2003; Waltzman, Croot, Williams, 2005), differences in dose adjustments between epoetin and darbepoetin arms. Three studies compared a darbepoetin dose used commonly in U.S. practices (200 mcg every two weeks; NCCN 2005) with a labeled epoetin dose (40,000 IU/week) (Glaspy, Berg, Tomita, et al., 2005; Schwartzberg, Yee, Senecal, et al., 2004; Waltzman, Croot, Williams, 2005). Study characteristics and results are summarized in Table 13.

Results. In all but one study, differences in hematologic response rates were not statistically significant, whether measured as defined for this review (proportion with Hb increased by ≥2g/dL from baseline by end of study), or otherwise. The exception was Waltzman, Croot, Williams, (2005), which reported a statistically significant difference in responses by week 17 that favored epoetin. However, this study adjusted dose for inadequate initial response at different times in the two arms (Table 13), potentially biasing the results. Patients with <1 g/dL rise in Hb had the dose increased 1.5-fold at week 6 if randomized to darbepoetin (from 200 to 300 mcg every 2 weeks), but at week 4 if randomized to epoetin (from 40,000 to 60,000 IU/week).

Taken together, trials directly comparing darbepoetin versus epoetin did not demonstrate that one drug achieves hematologic response in a larger proportion of patients than the other. However, conclusions from direct comparisons were limited since trials defined hematologic response, and initiated and adjusted doses, differently. Therefore, we also examined indirect evidence from trials comparing epoetin versus control or darbepoetin versus control.

Table 13. Study Characteristics and Results of RCTs Directly Comparing Hb Response Rates for Darbepoetin versus Epoetin

Trial	N		Response: ≥2 g/dL ?	Hb Response Rates		comment
	Darb	Epo		Darb	Epo	
Darbepoetin 200 mcg once per 2 weeks versus Epoetin 40,000 IU once weekly						
Waltzman 2005	177	175	yes	41.8%	57.7%	arms differed in dose adjustment for inadequate response[1]
				RR=0.72 (95% CI: 0.58, 0.90) p=0.004		
Glaspy 2005	606	603	no[2]	90.3% (95% CI: 87.5%, 93.1%)	95.5% (95% CI: 93.6%, 97.4%)	
Schwartzberg 2004	157	155	no[3]	68.8%	72.3%	
				no significant difference		
Other Doses						
Alexopoulos 2004[4]	25	25	no[4]	44%	44%	
				no significant difference		
Glaspy 2002 Part B[5]	31-33, each of 4 arms	32	yes	56% to 81%	59%	dose-finding study; dose adjusted only in Epo arm[5]
				no significant difference, lowest two darb doses versus epo		
Glaspy 2003[6]	30-32, each of 3 arms	30	yes	57% to 67%	50%	dose-finding study of front-loaded darb, not increased for inadequate response[6]
				no significant difference, any darb arm versus epo arm		

[1] Waltzman 2005 patients with <1 g/dL Hb rise from baseline had 1.5-fold dose increase at week 6 if randomized to darbepoetin (from 200 to 300 mcg Q2W), but at week 4 if randomized to epoetin (from 40,000 to 60,000 IU/week).

[2] Glaspy 2005 defined response as reaching Hb ≥11 g/dL and remaining between 11 and 13 g/dL.

[3] Schwartzberg 2004 defined response as reaching Hb ≥12 g/dL or increasing by 2 g/dL from baseline to end of study.

[4] Alexopoulos 2004 compared 150 mcg darbepoetin once weekly versus 10,000 IU epoetin thrice weekly, and defined Hb response as increasing by ≥1.5 g/dL over baseline by end of study.

[5] Glaspy 2002 Part B compared arms given 3, 5, 7, or 9 mcg/kg darbepoetin Q2W versus epoetin 40,000 IU QW; dose increase for inadequate Hb response only permitted for epoetin arm.

[6] Glaspy 2003 compared three arms given different front-loaded darbepoetin regimens versus epoetin 40,000 IU QW; dose increase for inadequate Hb response only permitted for epoetin arm.

Epoetin versus Control. Characteristics of reporting studies are enumerated in Table 14.

Fifteen trials (N=3,293; 1,844 to epoetin, 1,449 to control) reported hematologic response rates as defined for this review (Bamias, Aravantinos, Kalofonos, et al., 2003; Boogaerts, Coiffier, Kainz, 2003; Case, Bukowski, Carey, et al., 1993; Cazzola, Messinger, Battistel, et al, 1995; Chang, Couture, Young, et al., 2005; Dammacco, Castoldi, Rodjer, et al., 2001; Henry, Brooks, Case, et al., 1995; Iconomou, Koutras, Rigopoulos, et al., 2003; Littlewood, Bajetta, Nortier, et al., 2001; Oberhoff, Neri, Amadori, et al., 1998; Osterborg, Boogaerts, Cimino, et al., 1996; Osterborg, Brandberg, Molostova, et al., 2002; Rose, Rai, Revicki, et al., 1994; Savonije, Van Groeningen, Van Bochove, et al., 2004; Witzig, Silberstein, Loprinzi, et al., 2005). Two of the 15 studies (Cazzola, Messinger, Battistel et al., 1995; Osterborg, Boogaerts, Cimino et al., 1996) tested two different epoetin doses and were evaluated as two trials each.

Eight others (Carabantes, Benavides, Trujillo, et al., 1999; Cascinu, Fedeli, Del Ferro, et al., 1994; Del Mastro, Venturini, Lionetto, et al., 1997; Henke, Guttenberger, Barke, et al., 1999; Henke, Laszig, Ruebe, et al., 2003; Huddart, Welch, Chan, et al., 2002; Kurz, Marth, Windbichler, et al., 1997; Silvestris, Romito, Fanelli, et al., 1995) used different definitions or did not report separately by study arm.

Table 14. Study Characteristics and Subgroup Analyses of RCTs Reporting Hematologic Responses (as defined in Scope and Key Questions)

Outcome Subgroup	Epoetin versus Control					Darbepoetin versus Control				
	# Studies	# Total Patients	#Epo/#Ctl Patients	RR	95% CI (p-value)	# Studies	# Total Patients	#Darb/#Ctl Patients	RR	95% CI (p-value)
Hb Response	15	3,293	1844/1449	3.42	3.03; 3.86	3	659	427/232	3.36	2.48; 4.56
(Heterogeneity)					**(<0.0001)**					(0.98)
Subgroup Analyses: Patient Baseline Characteristics										
Baseline Hb <10	11	2,372	1,329/1,043	3.24	2.82; 3.73	(all)	659	427/232		
Baseline Hb 10-12	4	921	515/406	3.98	3.11; 5.10					
Baseline Hb >12										
Baseline Hb ?										
(Group difference[1])					(0.563)					
Solid tumors	7*	1,660	925/735	3.30	2.80; 3.88	1	249	198/51	3.51	1.74; 7.08
Hematologic	6*	1,093	643/450	3.30	2.68; 4.06	2	410	229/181	3.31	2.37; 4.63
Mixed	3	450	276/264	4.32	3.04; 6.13					
(Group difference[1])					(0.136)					(0.9715)
Children										
Adults	(all)	3,293	1844/1449			(all)	659	427/232		
(Group difference[1])										
Subgroup Analyses: Treatment Protocols										
Chemo, all plat	3	584	347/237	2.89	2.18; 3.84					
Chemo, some plat	5	1,053	535/518	3.12	2.56; 3.81					
Chemo, no plat	7	1,656	962/694	3.84	3.21; 4.58					
Chemo, plat ?										
Chemo+RT or RT										
Unknown										
(Group difference[1])					(0.212)					
Iron, fixed	2	441	222/219	2.43	1.92; 3.07					
Iron, as needed	10	2,249	1,244/1,005	4.13	3.51; 4.85					
Iron ?	3	603	378/225	2.25	1.94; 3.35					
(Group difference[1])					**(0.002)**					
Epo tx 6-9 weeks	1	86	57/29	8.91	2.30; 34.50					
Epo tx 12-16 weeks	11	2,560	1,376/1,184	3.31	2.91; 3.77	(all)	659	427/232		
Epo tx >20 weeks	4	647	411/236	3.65	2.62; 5.05					
Epo tx ? weeks										
(Group difference[1])					(0.1509)					

[1] p value for differences among subgroup categories calculated by inverse variance method (see Methods/Data Extraction and Analysis/Statistical Data Analysis)

* Note: Littlewood 2001 was split into two subsets for malignancies: solid and hematologic malignancies since Hb responses were reported separately

CI: confidence interval; Ctl: control; darb: darbepoetin; epo: epoetin; Hb: hemoglobin; plat: platinum; RT: radiotherapy; tx: treatment

Table 14. Study Characteristics and Subgroup Analyses of RCTs Reporting Hematologic Responses (as defined in Scope and Key Questions), continued

Outcome Subgroup	Epoetin versus Control					Darbepoetin versus Control				
	# Studies	# Total Patients	#Epo/#Ctl Patients	RR	95% CI (p-value)	# Studies	# Total Patients	#Darb/#Ctl Patients	RR	95% CI (p-value)
Subgroup Analyses: Reporting and Study Quality										
High quality	6	1,530	864/666	2.94	2.53; 3.43	(all)	659	427/232		
Low quality	9	1,763	980/783	4.13	3.40; 5.01					
(Group difference[1])					**(0.0414)**					
Data from full text	9	1966	1,055/911	4	3.39; 4.71	(all)	659	427/232		
Data from abstract	1	314	211/104	2.25	1.66; 3.04					
Data unpublished	5	1012	578/434	3.05	2.45; 3.80					
Data from FDA										
(Group difference[1])					**(0.0416)**					

[1] p value for differences among subgroup categories calculated by inverse variance method (see Methods/Data Extraction and Analysis/Statistical Data Analysis)
CI: confidence interval; Ctl: control; darb: darbepoetin; epo: epoetin; Hb: hemoglobin; plat: platinum; RT: radiotherapy; tx: treatment

Trials that defined hematologic response rates as in this review differed with respect to several variables prespecified for subgroup analysis (Figure 3, Table 14). Baseline characteristics of study populations differed by average baseline Hb concentration and type of malignancy. Treatment protocols differed by therapies for malignancy, iron supplementation, and duration of epoetin treatment. Trials also varied with respect to publication type and overall quality rating.

Results. Each trial reported significantly more hematologic responses among patients randomized to epoetin than among patients randomized to controls. Trials that used the most common definition of hematologic response were pooled for meta-analysis. A test for heterogeneity across these 15 trials was strongly significant ($p<0.0001$, $I^2 = 66.0$ percent). Therefore, both fixed- and random-effects meta-analyses were conducted and showed no substantive difference in the results.

Meta-analysis of data from all 15 trials[1] (Figure 4) yielded:

- Fixed-effects: relative risk (RR) = 3.42 (95 percent CI: 3.03, 3.86), $p<0.00001$
- Random-effects: RR = 3.73 (95 percent CI: 2.94, 4.74), $p<0.00001$
- Pooled response (event) rates (range across trials) were 58 percent (20.8 percent to 72.7 percent) for epoetin treatment arms and 16.5 percent (2.8 percent to 31.7 percent) for control arms.

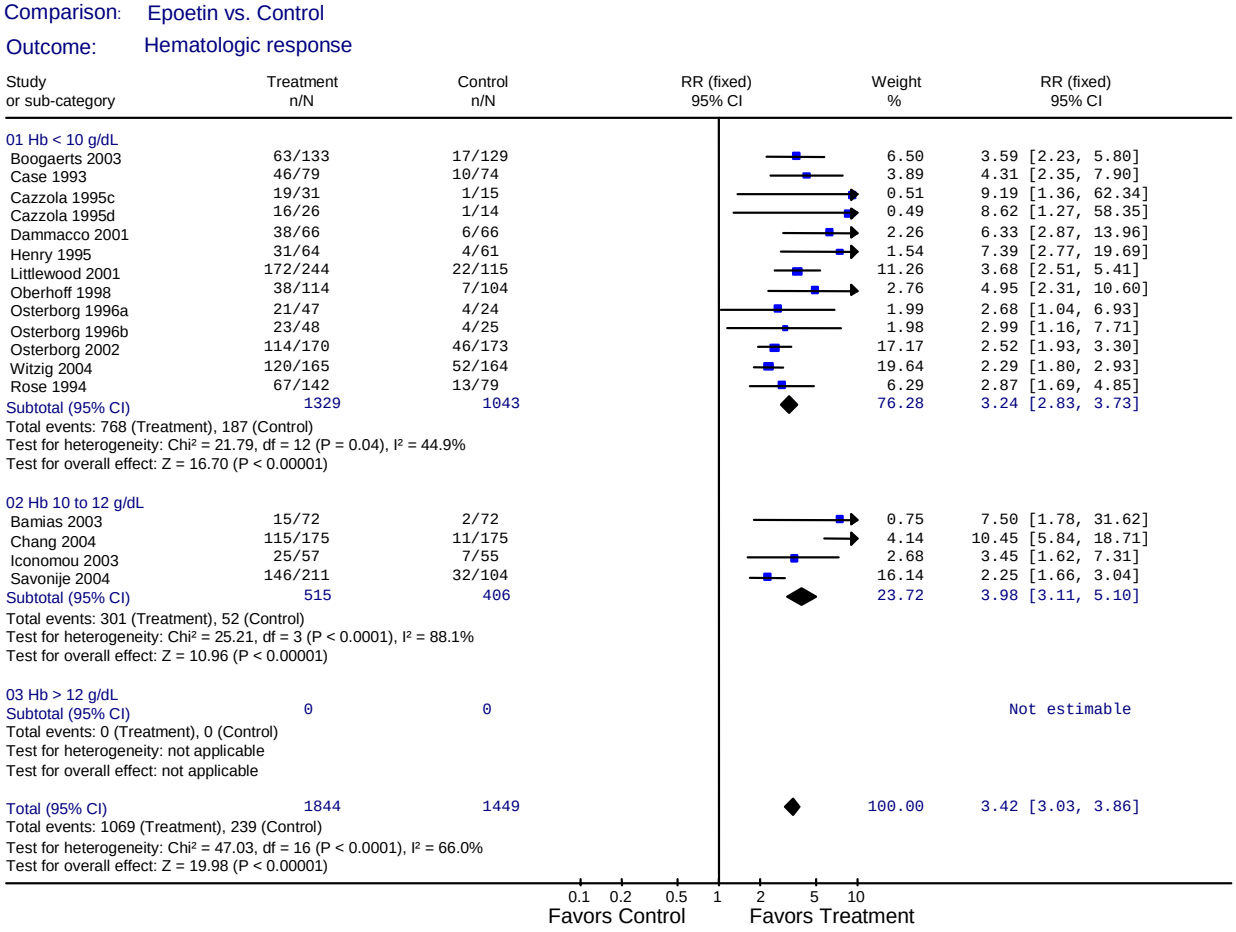
[1] In the Cazzola and Osterborg studies, two different epoetin dosages were compared with one control group. For the meta-analysis, each trial's control group was split artificially into two groups. Given the low total weight for these two studies (4.98%), it is unlikely that splitting the controls influenced the meta-analytic results.

- RRs ranged from 2.25 (95 percent CI: 1.66, 3.04; Savonije 2004) to 10.45 (95% CI: 5.84, 18.71; Chang 2004).

Univariate subgroup analyses found three statistically significant associations (Table 14). All subgroup differences were in magnitude rather than direction of effect: hematologic responses were consistently more frequent in epoetin arms than in controls for all subgroups. Variables significantly associated with increased likelihood (larger RR) of hematologic response were: iron supplementation as needed (vs. fixed iron or iron unknown); lower quality studies (vs. higher quality studies); and full-text publication (vs. abstract only or unpublished data).

However, availability of only two trials (N=441) in the "fixed iron" subgroup (Table 14) limits the analysis on effects of iron. These data were compared with 10 trials (N=2,299) that gave iron supplementation as necessary and three (N=603) that did not report on iron supplementation. The significant difference found in univariate analysis might be confounded by other factors.

Figure 4. Fixed-Effects Meta-Analysis of Data on Hematologic Response Rates from 15 RCTs of Epoetin versus Control



Five trials that defined hematologic response differently from those in the pooled analysis also reported greater response rates in the arms randomized to epoetin than in control arms (Appendix C Table C14). Definitions included reaching and maintaining Hb >10 g/dL (Cascinu, 1994; Del Mastro, 1997), reaching Hb>14 for women or >15 for men (Henke 2003), a 2 g/dL increase or reaching Hb >12 g/dL (Kurz 1997), and a 2 g/dL increase or an increase in reticulocyte counts $>40x10^9$ (Huddart 2002).

Darbepoetin versus Control. Characteristics of reporting studies are enumerated in Table 14.

Three of four trials comparing darbepoetin versus control (Hedenus, Hansen, Taylor, et al., 2002; Hedenus, Adriansson, San Miguel, et al., 2003; Kotasek, Steger, Faught, et al., 2003) reported the proportion of hematologic responders as defined for this review (N=659; 427 to darbepoetin, 232 to control). Two of these studies (Hedenus, Hansen, Taylor, et al., 2002; Kotasek, Steger, Faught, et al., 2003) tested several doses and were evaluated as three and six trials, respectively. The fourth trial used a different definition of response and was not included in the meta-analysis (Vansteenkiste, Pirker, Massuti, et al., 2002).

Trials that reported Hb response rates as defined for this review differed with respect to several variables prespecified for subgroup analysis (Figure 3, Table 14). Patient groups differed only by

type of malignancy. Treatment protocols differed by therapies for malignancy and use of iron supplementation.

Results. As with the epoetin versus control trials, each trial reported more frequent hematologic responses among patients treated with darbepoetin than among controls (see Figure 5). Results were not statistically significant for any arm from the two dose-finding studies (Hedenus 2002; Kotasek 2003), but were significant for the third trial (Hedenus 2003). A test for heterogeneity across trials included for Hb response was not statistically significant ($p=0.98$, $I^2 = 0$ percent). An I^2 value of zero percent indicates no observed statistical heterogeneity, thus only a fixed-effects meta-analysis was done.

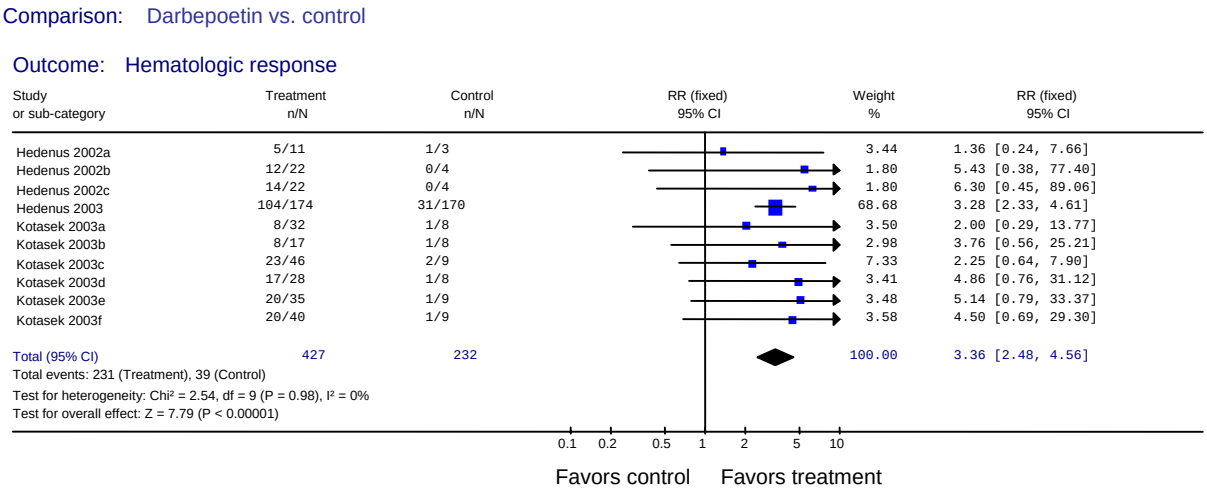
Fixed-effects meta-analysis[1] (Figure 5) yielded:

- RR = 3.36 (95% CI: 2.48, 4.56), $p<0.00001$
- pooled response rates (range by trial arms): darbepoetin arms 54.1% (25% to 84 percent); control arms: 16.9% (9% to 18.2%)
- RR (likelihood) to achieve response across the trials' darbepoetin dose arms ranged from 1.36 to 6.30 (Hedenus 2002a, 95% CI: 0.24, 7.66; Hedenus 2002c, 95% CI: 0.45, 89.06).

Univariate subgroup analyses found no statistically significant differences.

[1] In two studies, three (Hedenus 2002) or six (Kotasek 2003) different darbepoetin dosages were compared with one control group. For the meta-analysis the control group was split artificially into the same number of dose groups. As this might influence weighting of the studies, the analysis was repeated with the all relevant dose arms of each study merged into a single experimental arm compared to the entire control group. The overall result (RR 3.45 (95% CI: 2.53, 4.71) was similar to the base model. Additionally, a meta-analysis was performed using FastPro, which allows multi-dose entries with a single control arm, and combination using an empirical Bayes method. Setting 2.25 mcg/kg per week as the standard dose, the results were again similar: RR 3.50 (95% CI: 2.03, 6.04).

Figure 5. Fixed-Effects Meta-Analysis of Data on Hematologic Response Rates from Three RCTs of Darbepoetin versus Control



KQ1 Outcome II. Transfusion Rates

For purposes of this report, transfusion rate is defined as the proportion of patients transfused with red blood cells (or whole blood) at least once during the study.

Evidence for Comparative Effectiveness

Darbepoetin versus Epoetin. Characteristics of reporting studies are enumerated in Table 15.

Six RCTs (N=2,158; 1,169 to darbepoetin, 989 to epoetin) compared darbepoetin versus epoetin for their effects on transfusion rates (Appendix C Table C21; Glaspy, Jadeja, Justice, et al., 2002A; Waltzman, Croot, Williams, 2005; Schwartzberg, Yee, Senecal, et al., 2004; Alexopolous and Kotsori, 2004; Glaspy, Berg, Tomita, et al., 2005; Glaspy and Tchekmedyian, 2002B). All were judged to be of poor quality, since each was unblinded and described randomization methods inadequately. Another trial monitored, but did not report, transfusion rates (Glaspy, Jadeja, Justice, et al., 2003). Available studies defined transfusion rate consistently, permitting pooled analysis of data from trials comparing adequate doses of the two drugs. One study reported separately on three patient groups, each with a different malignancy (Schwartzberg 2004 arms a-c). Two studies compared different doses of darbepoetin versus a single dose of epoetin (Glaspy 2002A and B; Figure 6). The meta-analysis evaluated darbepoetin doses of 1.5, 2.25, and 4.5 mcg/kg weekly from one trial (Glaspy 2002A arms c-e), and all doses (3, 5, 7, and 9 mcg/kg biweekly) from the other (Glaspy 2002B arms a-d) as three and four trials, respectively. Thus, the meta-analysis included a total of 13 comparisons.

Trials that reported transfusion rates differed with respect to several variables prespecified for subgroup analysis (Figure 3, Table 15). Patient groups varied by average baseline Hb concentration, but univariate subgroup analysis was not done since the variation was minimal (Appendix C Table C7). Treatment protocols differed by therapies for malignancy and epoetin/darbepoetin treatment duration. The trials also varied with respect to publication type.

Results. Seven of 13 comparisons for relative risk of transfusion (RR, darbepoetin to epoetin, Figure 6) favored darbepoetin. RR ranged from 0.12 to 0.62 (Glaspy 2002 PB a, 95% CI: 0.01, 1.11; Glaspy 2002 PB b, 95% CI: 0.21, 1.88). The other six comparisons (darbepoetin to epoetin) favored epoetin, and RR ranged from 1.16 to 1.56 (Glaspy 2002 PA c, 95% CI: 0.41, 3.25; Schwartzberg 2004b, 95% CI: 0.74, 3.27). However, no single comparison was statistically significant: each RR had 95% CI limits that included 1.0. A test for heterogeneity across studies just reached statistical significance ($p=0.05$); an I^2 value of 42.8% suggested moderate heterogeneity (Higgins, Thompson, Deeks et al., 2005). However, fixed- and random-effects meta-analyses showed no meaningful difference in the results; although point estimates for the two types of meta-analysis were on opposite sides of 1.0, confidence intervals for both included 1.0, overlapped considerably, and were not statistically significantly different.

Table 15. Study Characteristics and Subgroup Analyses of RCTs Reporting Transfusion Responses

Outcome Subgroup	Darbepoetin versus Epoetin					Epoetin versus Control					Darbepoetin versus Control				
	# Studies	#Total Patients	#Darb/#Epo Patients	Relative Risk	95% CI (p-value)	# Studies	#Total Patients	#Epo/#Ctl Patients	Relative Risk	95% CI (p-value)	# Studies	#Total Patients	#Darb/#Ctl Patients	Relative Risk	95% CI (p-value)
Transfusion	6	2,158	1,169/989	1.10	0.93; 1.29	34	5,210	2,859/2,351	0.63	0.59; 0.67	4	950	566/384	0.61	0.52; 0.72
(Heterogeneity)					(0.27)					**(<0.00001)**					(1.00)
Subgroup Analyses: Patient Baseline Characteristics															
Bsln Hb <10	2	199	144/55	0.55	0.31; 0.96	15	2,805	1,547/1,258	0.70	0.64; 0.76	3	636	410/226	0.61	0.49; 0.76
Bsln Hb 10-12	4	1,959	1,025/934	1.16	0.97; 1.37	12	1,781	972/809	0.42	0.36; 0.50	1	314	156/158	0.60	0.47, 0.78
Bsln Hb >12						5	302	179/123	0.56	0.40; 0.80					
Bsln Hb ?						2	322	161/161	0.80	0.68; 0.95					
(Group difference[1])										**(<0.0001)**					(0.967)
Solid tumors	(all)					22	2,924	1,620/1,304	0.5	0.45; 0.56	2	552	344/208	0.59	0.48; 0.73
Hematologic						6	1,111	647/464	0.74	0.66; 0.84	2	398	222/176	0.64	0.49; 0.83
Mixed/unknown[2]						7	1,175	592/583	0.74	0.67; 0.83					
(Group difference[1])										**(<0.0001)**					(0.6984)
Children						2	454	227/227	0.87	0.77; 0.99					
Adults	(all)					32	4,756	2,632/2,124	0.59	0.55; 0.64	(all)				
(Group difference[1])										**(0.0001)**					
Subgroup Analyses: Treatment Protocols															
Chemo, all plat						13	1,251	744/507	0.51	0.45; 0.58	1	314	156/158	0.60	0.47; 0.78
Chemo, some plat	2	1,471	745/726	1.24	1.03; 1.41	7	1,478	744/734	0.59	0.50; 0.68	1	238	188/50	0.56	0.38; 0.83
Chemo, no plat						8	1,733	999/734	0.72	0.64; 0.80	2	398	222/176	0.64	0.49; 0.83
Chemo, plat unknown															
Chemo+RT						2	113	56/57	0.31	0.13; 0.71					
Unknown	4	687	424/263	0.75	0.54; 1.04	4	635	316/319	0.76	0.67; 0.87					
(Group difference[1])										**(<0.0001)**					(0.8824)
Iron, fixed						5	898	450/448	0.51	0.41; 0.65					
Iron, as needed						18	3,030	1,684/1,346	0.65	0.59; 0.71	1	332	167/165	0.65	0.49; 0.86
Iron unknown	(all)					11	1,282	725/557	0.64	0.58; 0.72	3	618	399/219	0.59	0.47; 0.72
(Group difference[1])										**(0.0195)**					(0.5269)

Table 15. Study Characteristics and Subgroup Analyses of RCTs Reporting Transfusion Responses (continued)

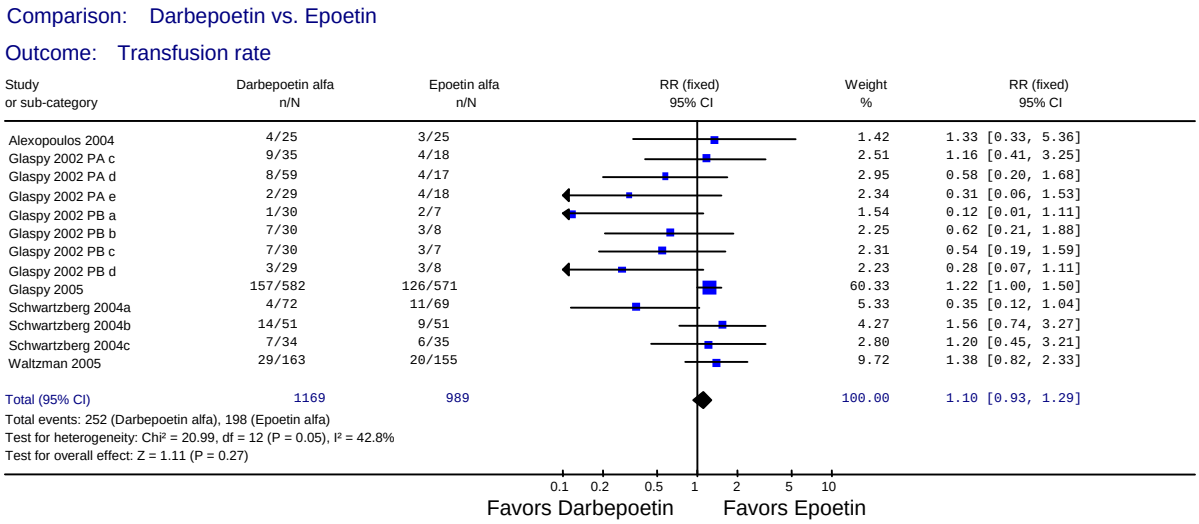
Outcome Subgroup	Darbepoetin versus Epoetin					Epoetin versus Control					Darbepoetin versus Control				
	# Studies	#Total Patients	#Darb/#Epo Patients	Relative Risk	95% CI (p-value)	# Studies	#Total Patients	#Epo/#Ctl Patients	Relative Risk	95% CI (p-value)	# Studies	#Total Patients	#Darb/#Ctl Patients	Relative Risk	95% CI (p-value)
Subgroup Analyses: Treatment Protocols (continued)															
Epo tx 6-9 weeks						5	320	182/138	0.43	0.28; 0.65					
Epo tx 12-16 weeks	(all)					18	3,189	1,689/1,500	0.64	0.59; 0.69	(all)				
Epo tx >20 weeks						10	1,329	802/527	0.67	0.60; 0.75					
Epo tx ? Weeks						1	372	186/186	0.4	0.23; 0.67					
(Group difference[1])										**(0.0062)**					
Subgroup Analyses: Reporting and Quality															
High quality						13	2,190	1,194/996	0.69	0.63; 0.76	(all)				
Low quality	(all)					21	3,020	1,665/1,355	0.58	0.52; 0.63					
(Group difference[1])										(0.2342)					
Data from full text	3	637	399/238	0.72	0.52; 1.01	18	2,472	1,376/1,096	0.56	0.50; 0.63	3	636	410/226	0.61	0.49; 0.76
Data from abstract	3	1,521	770/751	1.25	1.03; 1.5	10	1,560	834/726	0.62	0.55; 0.69					
Data unpublished						6	1,178	649/529	0.75	0.66; 0.84					
Data from FDA											1	314	156/158	0.6	0.47; 0.78
(Group difference[1])										**(0.0003)**					(0.967)

[1] p value for differences among subgroup categories calculated by inverse variance method (see Methods/Data Extraction and Analysis/Statistical Data Analysis)
[2] The Littlewood 2001study was split into two separate studies for this analysis (solid tumors and hematological malignancies), therefore the overall number of studies in this subgroup analysis appears to be 35 instead of 34. The Thomas 2002 (n=127) study did not report type of malignancy investigated and was classified in the 'mixed' category.

Meta-analysis[6] showed (Figure 6):

- Fixed-effects RR = 1.10 (darbepoetin to epoetin; 95% CI: 0.93,1.29), p=0.27
- Random-effects RR= 0.87 (darbepoetin to epoetin; 95% CI: 0.63, 1.20), p=0.40
- pooled transfusion rates (ranges across trials and dose arms): darbepoetin arms, 21.6% (3.3% to 27.5%); epoetin arms, 20% (12% to 42.9%)
- subgroup analyses were not done since the few differences between trials were either minimal (baseline Hb) or lacked adequate information (therapies for malignancy).

Figure 6. Fixed-Effects Meta-Analysis of Data on Transfusion Rates from Six RCTs of Darbepoetin versus Epoetin



With respect to effects on transfusion rates, the fixed-effects and random-effects meta-analyses support neither superiority nor inferiority for darbepoetin compared with epoetin. The fixed-effects point estimate favors epoetin, while the random-effects point estimate favors darbepoetin; however, the confidence intervals overlap, and each includes 1.0 (no difference). The analyses do not exclude the possibility that a larger and more homogeneous data set might show superiority (within these confidence limits) for one of the drugs. We evaluated indirect evidence (studies of epoetin versus control and darbepoetin versus control) to further compare effects on transfusion rates.

[6] In two of the six included studies, three (Glaspy 2002 Part A) and four (Glaspy 2002 Part B) different darbepoetin doses were compared with one control group each. For meta-analysis, the control groups were split artificially into the corresponding number of groups. As this might influence weighting of studies, the analysis was repeated with all dose arms of each study merged into one experimental arm, then compared to the trial's full control group. The overall result (RR=1.10, 95% CI: 0.93, 1.29) was almost identical to the base model. Additionally, a second meta-analysis used FastPro, which allows multi-dose entries with a single control arm, and combines results using an empirical Bayesian method. With standard dose set as 2.25 μg/kg weekly, relative risk was 0.99, 95% CI: 0.70, 1.39.

Epoetin versus Control. Characteristics of reporting studies are enumerated in Table 15. Thirty-four RCTs (N=5,210; 2,859 to epoetin, 2,351 to control) reported transfusion rates as defined for this review (Appendix C Tables C2, C3, and C8; Aravantinos, Linardou, Makridaki, et al., 2003; Bamias, Aravantinos, Kalofonos, et al., 2003; Boogaerts, Coiffier, Kainz, 2003; Carabantes, Benavides, Trujillo, et al., 1999; Cascinu, Fedeli, Del Ferro, et al., 1994; Case, Bukowski, Carey, et al., 1993; Cazzola, Messinger, Battistel, et al, 1995; Chang, Couture, Young, et al., 2005; Dammacco, Castoldi, Rodjer, et al., 2001; Del Mastro, Venturini, Lionetto, et al., 1997; Dunphy, Harrison, Dunleavy, et al., 1999; Henry, Brooks, Case, et al., 1995; Henze, Michon, Morland, et al., 2002; Huddart, Welch, Chan, et al., 2002; Iconomou, Koutras, Rigopoulos, et al., 2003; Janinis, Dafni, Aravantinos, et al., 2003; Kunikane, Watanabe, Fukuoka, et al., 2001; Kurz, Marth, Windbichler, et al., 1997; Littlewood, Bajetta, Nortier, et al., 2001; Oberhoff, Neri, Amadori, et al., 1998; Osterborg, Boogaerts, Cimino, et al., 1996; Osterborg, Brandberg, Molostova, et al., 2002; Quirt, Micucci, Moran, et al., 1996; Razzouk, Hockenberry, Hinds, et al., 2004; Rose, Rai, Revicki, et al., 1994; Savonije, Van Groeningen, Van Bochove, et al., 2004; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Thatcher, De Campos, Bell, et al., 1999; Thomas, McAdam, Thomas, et al., 2002; Throuvalas, Antonadou, Boufi, et al., 2000; Vadhan-Raj, Skibber, Crane, et al., 2004; Welch, James, Wilkinson, 1995; Witzig, Silberstein, Loprinzi, et al., 2005; Wurnig, Windhager, Schwameis, et al., 1996). Five trials (Cazzola 1995c-d; ten Bokkel Huinink 1998a-b; Kunikane 2001a-b; Thatcher 1999a-b; Osterborg 1996a-b) each tested two different doses or methods of titrating dose; each study was evaluated as two trials.

One other trial focusing on QoL outcomes was excluded (Appendix C Table C24: O'Shaughnessy, Vukelja, Holmes, et al., 2005), since patients were removed from either arm of this double-blind study if 1) Hb fell below 8 g/dL; 2) they were transfused for another clinical indication; or 3) they received non-study ("commercial") epoetin based on clinical necessity.

Trials that reported transfusion rates differed with respect to several variables prespecified for subgroup analysis (Figure 3, Table 15). Baseline characteristics of study populations differed by average baseline Hb concentration, type of malignancy, and age range. Treatment protocols differed by therapies for malignancy, iron supplementation, and epoetin treatment duration. Trials also varied with respect to publication type and overall quality rating.

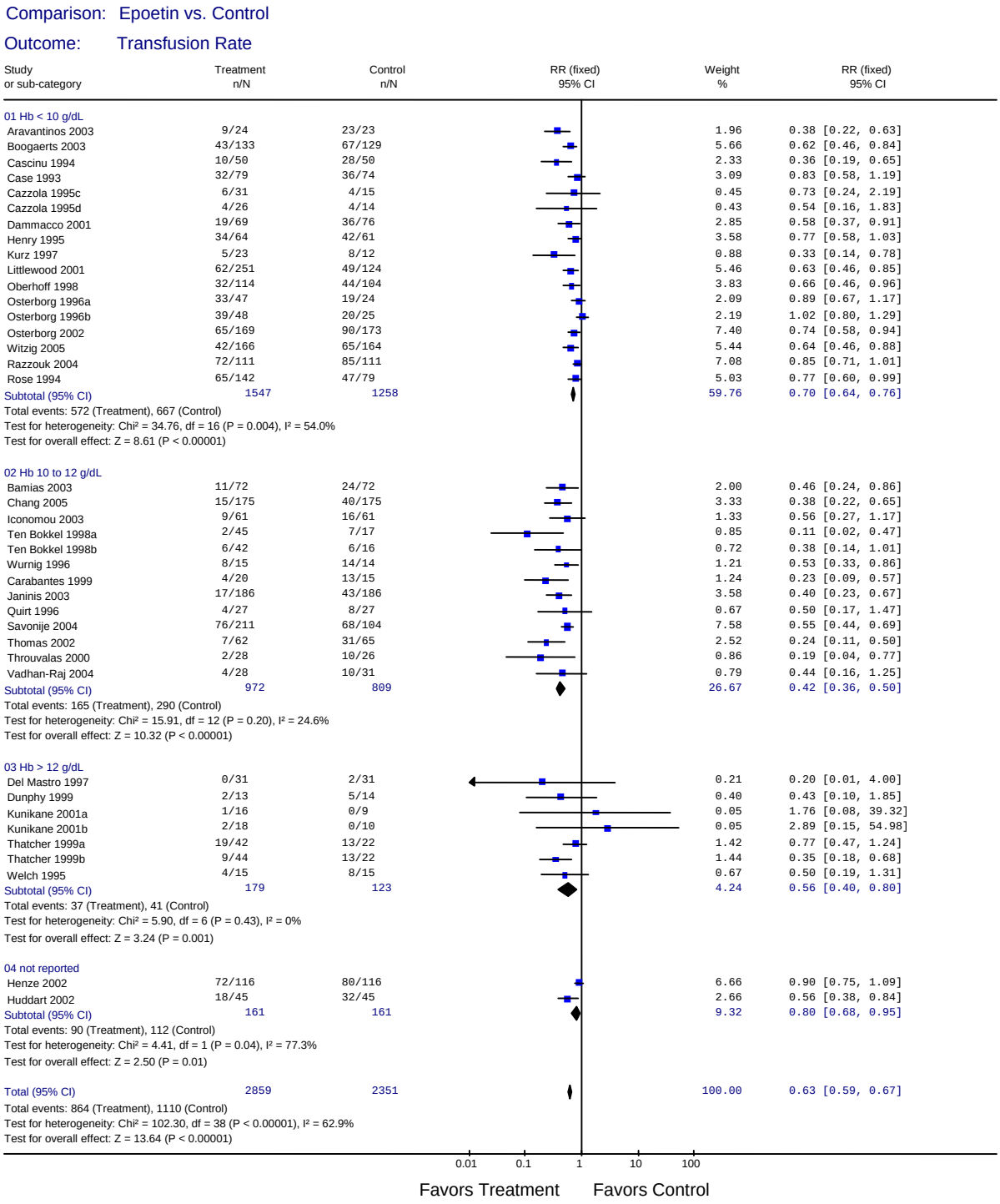
Results. The overwhelming majority of trials reported fewer transfusions among those randomized to epoetin than among those randomized to control. However, differences between epoetin and control arms (or reductions in risk of transfusion) were not always statistically significant (see Figure 7). A test for heterogeneity across trials included for transfusion was strongly significant ($p<0.00001$, $I^2= 62.9$ percent). Fixed- and random-effects meta-analyses showed no substantive difference in results.

Meta-analysis of data from all 34 RCTs[7] (Figure 7) yielded:

- Fixed-effects RR = 0.63 (95 percent CI: 0.59; 0.67), $p<0.00001$
- Random-effects RR = 0.60 (95 percent CI: 0.53; 0.67), $p<0.00001$

[7] In two of the six included studies, three (Glaspy 2002 Part A) and four (Glaspy 2002 Part B) different darbepoetin doses were compared with one control group each. For meta-analysis, the control groups were split artificially into the corresponding number of groups. As this might influence weighting of studies, the analysis was repeated with all dose arms of each study merged into one experimental arm, then compared to the trial's full control group. The overall result (RR=1.10, 95% CI: 0.93, 1.29) was almost identical to the base model. Additionally, a second meta-analysis used FastPro, which allows multi-dose entries with a single control arm, and combines results using an empirical Bayesian method. With standard dose set as 2.25 μg/kg weekly, relative risk was 0.99, 95% CI: 0.70, 1.39.

Figure 7. Meta-Analysis of Data on Relative Risk of Transfusion from 34 RCTs of Epoetin versus Control



- pooled transfusion rates (range across trials): epoetin arm, 30.2 percent (0 percent to 91.4 percent); control arms, 47.2 percent (0 percent to 100 percent)

- RR ranged from 0.11 to 2.89 (ten Bokkel 1998a, 95 percent CI: 0.02, 0.47; Kunikane 2001b, 95 percent CI: 0.15, 54.98).

Epoetin consistently reduced transfusion risk in all subgroups analyzed (see Table 15). Seven variables were statistically significant predictors (by univariate analysis; see Methods/Data Extraction and Analysis/Statistical Data Analysis) for subgroups with a smaller relative risk of transfusion in epoetin arms compared with controls (p values in bold font, Table 15). Univariate analysis also suggested transfusion risk may have been reduced to a greater extent in trials whose participants had mean baseline Hb from 10 to 12 g/dL (RR=0.42; 95 percent CI: 0.36, 0.50) than in trials with baseline Hb <10 g/dL (RR=0.70; 95 percent CI: 0.64, 0.76) or >12 g/dL (RR=0.56; 95 percent CI: 0.40, 0.80). However, subgroup differences for other patient and study variables may have confounded this result.

Seeking better insight into potentially important subgroup differences identified in univariate analyses, we used meta-regression to explore independent sources of heterogeneity across included trials (follows next two sections).

Darbepoetin versus Control. Characteristics of reporting studies are enumerated in Table 15.

Four trials (N=950; 566 to darbepoetin, 384 to control) reported effects of darbepoetin on transfusion rates (Appendix C Tables C4, C5, and C9; Hedenus, Hansen, Taylor, et al., 2002; Hedenus, Adriansson, San Miguel, et al., 2003; Kotasek, Steger, Faught, et al., 2003; Vansteenkiste, Pirker, Massuti, et al., 2002). Two trials (Hedenus 2002a-c; Kotasek 2003a-f) tested different doses of darbepoetin (three and six, respectively) versus single control groups; therefore, each dosage arm was analyzed as a separate study.

Trials that reported transfusion rates differed with respect to several variables prespecified for subgroup analysis (Figure 3, Table 15). Patient groups differed by average baseline Hb concentration and malignancy type. Treatment protocols differed by therapies for malignancy and iron supplementation. The trials also varied with respect to type of publication.

Results. Each trial comparing darbepoetin versus control reported proportionally fewer transfusions in darbepoetin arms than in controls. However, risk reduction was not statistically significant in any individual dose arm from multi-dose trials (Figure 8), most likely because each trial's single control arm was artificially split into smaller groups for the analysis. A test for heterogeneity across trials and dose arms included for transfusion was not statistically significant (p=1.00, I^2 =0 percent). An I^2 value of zero percent indicates absence of statistical heterogeneity, thus only a fixed-effects meta-analysis was done.

Meta-analysis of data from the four RCTs[8] (Figure 8) yielded:

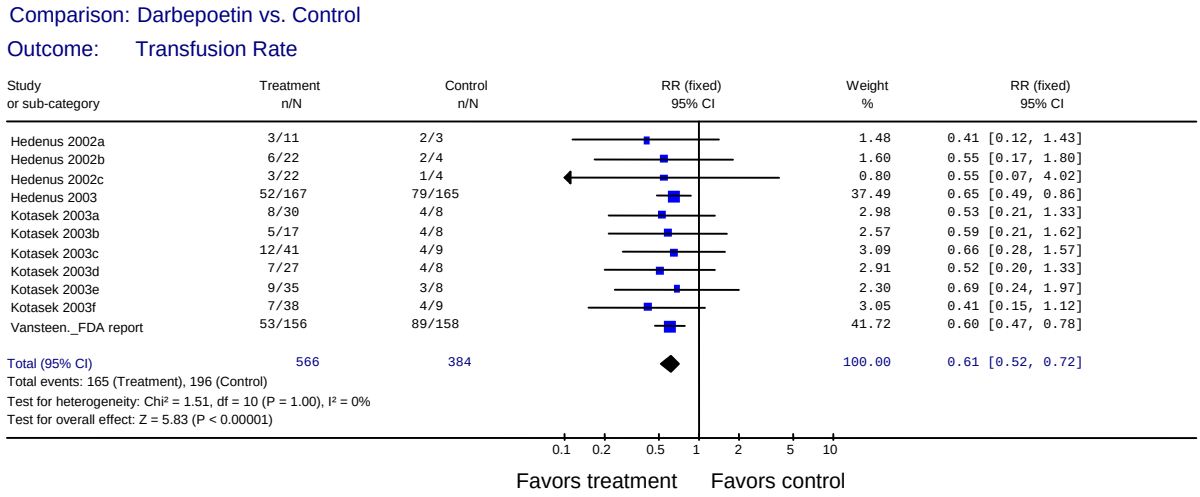
[8] In two studies, 3 (Hedenus 2002a-c) and 6 (Kotasek 2003a-f) different darbepoetin dose arms were compared with one control group each. For meta-analysis, the control groups were split artificially into the corresponding number of groups. As this might influence weighting of studies, the analysis was repeated with all dose arms of each study merged into one experimental arm, then compared to the trial's full control group. The overall result (RR 0.61; 95% CI: 0.51, 0.72) was almost identical to the base model. Additionally, a second meta-analysis used FastPro, which allows multi-dose entries with a single control arm, and combines results using an empirical Bayesian method. With standard dose set as 2.25 mcg/kg weekly, relative risk was slightly more favorable: 0.51; 95 % CI: 0.40, 0.64.

- Fixed-effects RR = 0.61 (95% CI 0.52; 0.72), p<0.00001

- pooled transfusion rates (ranges across trials and dose arms): darbepoetin arms, 29.2% (13.6% to 34.0%); control arms, 51% (25% to 66.7%)

- RR ranged from 0.41 to 0.69 (Hedenus 2002a, 95% CI: 0.12, 1.43; Kotasek 2003e, 95% CI: 0.24, 1.97).

Figure 8. Meta-Analysis of Data on Transfusion Rates from Four RCTs of Darbepoetin versus Control



For each variable tested by univariate analysis, there were no statistically significant differences among subgroups.

Indirect Comparison. Thirty four trials (N=5,210) compared epoetin versus control. Pooled RR of transfusion for epoetin treated patients compared to control was 0.63 (95 percent CI: 0.59, 0.67; p<0.00001). Four trials (N=950) compared darbepoetin versus control. Pooled RR of transfusion for darbepoetin treated patients compared to control was 0.61 (95 percent CI: 0.52, 0.72; p<0.00001). Pooled transfusion rates also showed similar results for epoetin or darbepoetin vs. control: epoetin 30 percent versus control 47 percent darbepoetin 29 percent versus control 51 percent

The actual benefit of treatment with an erythropoietic stimulant depends on the patient's underlying risk of transfusion. Trials ranged widely with respect to the percent of control arm patients who underwent transfusion: 0-100 percent in epoetin trials and 25 percent to 66.7 percent in darbepoetin trials. To illustrate, we calculated overall number-needed-to-treat (NNT) with epoetin[9] or darbepoetin to spare one patient from transfusion, at representative baseline transfusion risks (Table 16).

[9] Since the epoetin-versus-control meta-analysis showed statistically significant heterogeneity across trials, we used a L'Abbe plot (transfusion rate in the epoetin arm as a function of transfusion rate in the corresponding control arm; see Methods and Appendix F) to confirm that RR was relatively constant across the range of baseline risks, justifying calculation of an overall number-needed-to-treat (NNT) to spare one patient from transfusion.

Table 16. Calculated Numbers Needed to Treat (NNT) to Spare One Patient from Transfusion, at Representative Baseline Risks of Transfusion

Baseline Risk	Epoetin			Darbepoetin		
	NNT	95% confidence interval		NNT	95% confidence interval	
		lower limit	upper limit		lower limit	upper limit
30%	9.01	8.13	10.10	8.55	6.94	11.90
50%	5.41	4.88	6.06	5.13	4.17	7.14
70%	3.86	3.48	4.33	3.66	2.98	5.10

At each level of baseline risk, the NNT (rounded to the nearest whole number) to spare one patient from transfusion is the same for darbepoetin as for epoetin, except that confidence intervals are slightly wider.

Meta-regression of RCTs that Compared Epoetin or Darbepoetin versus Control. To better compare darbepoetin with epoetin indirectly for their transfusion-sparing effects, and also to explore causes of heterogeneity in meta-analysis on red blood cell transfusion rates, we used a fixed-effect meta-regression analysis.

Pooling studies. Because epoetin (RR=0.63; 95% CI: 0.59, 0.67) and darbepoetin (RR=0.61; 95% CI: 0.52, 0.72) appeared similar in their ability to reduce transfusion risk, we pooled studies comparing epoetin versus placebo/no treatment and studies comparing darbepoetin versus placebo/no treatment, to increase statistical power. After pooling, the following categorical variables (subgroups defined in Table 15) were statistically significant in univariate analyses (p values calculated by inverse variance method; see Methods/Data Extraction and Analysis/ Statistical Data Analysis):

- Hemoglobin at study entry ($p<0.0001$)
- Type of malignancy ($p<0.0001$)
- Type of treatment ($p<0.0001$)
- Iron supplementation ($p=0.041$)
- Duration of epoetin or darbepoetin treatment ($p=0.0042$)
- Type of publication ($p=0.0008$)
- Age range (adults versus children) ($p<0.0001$)

Univariate analyses showed that neither study quality nor type of erythropoietic stimulant (epoetin or darbepoetin; $p=0.35$) were statistically significant predictors for RR of transfusion. Consequently, both variables were omitted from the meta-regression.

Evidence Regarding the Class of Erythropoietic Stimulants

The meta-regression explored whether variables found statistically significant in univariate analyses contributed independently to heterogeneity in meta-analysis of transfusion risk reduction by erythropoietic stimulants.

Adjustments for inadequate information. For iron supplementation, the "unclear" subgroup (i.e., studies that did not report on iron use) included 14 trials with 1,900 patients (34.6 weight percent of the overall analysis), which hampered meaningful analysis. Consequently, we omitted the iron supplementation variable from further univariate or multivariate analyses, despite current uncertainties concerning optimal adjunctive iron therapy. All other factors significant in initial univariate analyses remained significant after omitting the iron supplementation variable.

To define a data set limited to trials with unambiguous information on each significant variable, we also omitted six trials with 1,097 patients (19.7 weight percent of the overall analysis) that were classified as "unclear" for one or more variable(s) (Quirt, Micucci, Moran, et al., 1996; Henze, Michon, Morland, et al., 2002; Huddart, Welch, Chan, et al., 2002; , McAdam, Thomas, et al., 2002; Janinis, Dafni, Aravantinos, et al., 2003; Razzouk, Hockenberry, Hinds, et al., 2004). Since Razzouk 2004 and Henze 2002 were the only trials on pediatric patients, the age range variable was deleted, and five variables were included in the meta-regression analysis.

Meta-regression. We used a back-wise selection method to derive the model; the covariate with the largest p value was consecutively removed until all remaining covariates were significant according to the Akaike Information Criterion (Akaike, 1969; see Methods). Variables included in the final model were "type of malignancy," "duration of treatment," and "source of data" (see Table 17). For each combination of variable subgroups, the relative risk can be calculated from Table 17. The following examples approximate the range of possible risk ratios:

- ln RR = -0.52-0.19-0.24-0.08 = -1.03; RR = $e^{-1.03}$ = 0.36 (for patients with solid tumors, treated/followed 6-9 weeks, with results in full text publications).

- ln RR = -0.52+0.08+0.15+0.14 = -0.15; RR = $e^{-0.15}$ = 0.86 (for patients with hematologic malignancies, treated/followed more than 20 weeks, with unpublished results).

For each statistically significant variable, meta-regression results of Table 17 suggest the following subgroup differences in magnitude of treatment benefit from an erythropoietic stimulant (hypotheses that may explain these differences are suggested):

- RR appears smaller (suggesting larger benefit) in trials on patients with solid tumors than in those on patients with hematologic malignancies. (Hypothesis: some patients with a hematologic malignancy may be less able to respond due to bone marrow involvement.)

- RR appears smaller (suggesting larger benefit) in trials that treated/followed patients for shorter durations, relative to those treated/followed for longer durations. (Hypothesis: risk reduction may be greatest soon after the first few weeks of treatment.)

- RR appears larger (suggesting smaller benefit) in trials that provided unpublished results, relative to trials that reported data in full text or abstract publications. (Hypothesis: treatment benefit may be estimated more conservatively when investigators provide fuller, more complete access to primary data.)

Table 17. Meta-Regression Analysis for Red Blood Cell Transfusion

Variable	Effect	S.E. (effect)	95% CI	p value
(Intercept)	-0.52	0.09	-0.69; -0.35	<0.0001
hematological malignancies	0.08	0.06	-0.03; 0.19	0.1368
mixed	0.10	0.06	-0.02; 0.22	0.1061
solid tumors	-0.19	0.05	-0.29; -0.08	0.0004
6 to 9 weeks	-0.24	0.15	-0.53; 0.05	0.1097
12 to 16 weeks	0.09	0.08	-0.07; 0.25	0.2811
>20 weeks	0.15	0.09	-0.02; 0.32	0.0849
abstract publication	-0.06	0.08	-0.22; 0.11	0.4949
full text publication	-0.08	0.06	-0.19; 0.03	0.1371
unpublished data	0.14	0.06	0.02; 0.26	0.0216

CI: confidence interval; SE: standard error

KQ1 Outcome III. Quality of Life

For purposes of this report, we required health-related quality of life (QoL) to be measured as change from baseline to final followup and change in treatment arm(s) compared to that in the control arm. Ideally, studies would also report the percentage of patients in each study arm that achieved a prespecified minimum amount of improvement known from prior studies to be clinically significant. However, only two studies reported results in this format (Vansteenkiste, Pirker, Massuti, et al., 2002; Witzig, Silberstein, Loprinzi, et al., 2005), and used different thresholds for improvement without documenting the validity of these thresholds.

We required the use of a validated instrument, such as the SF-36; European Organisation for Research and Treatment of Cancer (EORTC) Quality of life Questionnaire (QLQ-C30) or the Functional Assessment of Cancer Therapy (FACT; Table 18). Some QoL scales are general (also referred to as "generic" or "global" in some publications) measures of QoL in cancer patients (e.g. FACT-G; EORTC QLQ-C30) or in patients with any type of condition (e.g. SF-36), while others are targeted to specific cancers or symptoms. For example, the FACT-fatigue subscale of the FACT-Anemia symptom-specific instrument is sensitive to different, aspects of fatigue as a consequence of anemia. Thus, improvement in the FACT-fatigue subscale indicates improvement in that domain of QoL-related symptoms, but not necessarily in general or overall QoL. To demonstrate improvement in overall QoL there should be improvement in general QoL measures in addition to symptom-specific measures.

Study qualities not required for this review but nevertheless desirable include blinding and a plan for minimizing the effect of QoL instrument administration on results (see Introduction of Aronson, Seidenfeld, Piper, et al., 2001). In addition, while missing data tend to be unavoidable in QoL evaluations of cancer patients, the best methodologic practice is to prespecify how missing data will be handled to avoid significant bias in results (Donaldson and Moinpour, 2005).

We also abstracted QoL data from unidimensional visual analog scales (VAS) that in trials of epoetin or darbepoetin typically evaluate 3 items: energy, daily activities, and overall QoL. However, while some VAS QoL scales have been formally validated, validation of the 3-item

VAS scale used in these studies has not been well documented (Introduction of Aronson, Seidenfeld, Piper, et al., 2001). Thus, we give less weight to the results of VAS results in these studies.

Table 18. Description of the FACT Scales and Subscales Evaluated in this Review

FACT instrument or subscale	Type of instrument	Domains addressed (#questions)	Range of scale
FACT-G(eneral)	General	Physical well-being (7) Social/family well-being (7) Emotional well-being (6) Functional well-being (7)	0-108
FACT-An(emia)	Symptom-specific	Includes FACT-G, all domains (27)[1]	0-188
		Anemia-specific symptoms (20)	
FACT-fatigue subscale	Symptom-specific	Fatigue-specific questions from the anemia-specific questions of FACT-An (13)	0-52
FACT non-fatigue anemia subscale	Symptom-specific	Questions from the anemia-specific questions of FACT-An that are not part of the FACT-fatigue subscale	0-28

[1]While FACT-Anemia incorporates FACT-G, it was not classified as a general instrument since the results could be dominated by either the general FACT-G or the symptom-specific subscales.

Evidence for Comparative Effectiveness

Darbepoetin versus Epoetin. Characteristics of reporting studies are enumerated in Table 13.

Six trials directly compared darbepoetin to epoetin and measured QoL outcomes. Of those, one trial (Glaspy, Berg, Tomita, et al., 2005) used standard darbepoetin doses and administration schedules; used a validated instrument; and reported results separately for darbepoetin and epoetin study arms. Of the total number of patients randomized in Glaspy, Berg, Tomita, et al. (2005; N=1,220), 60 percent (N=731) were evaluable for QoL. We excluded the other 5 studies because they used nonstandard darbepoetin doses and administration schedules (Glaspy, Jadeja, Justice, et al., 2003); reported QoL results by Hb change rather than by study arm (Glaspy, Jadeja, Justice, et al., 2002A; Glaspy and Tchekmedyian, 2002B); did not include QoL results in an abstract-only publication (Alexopolous and Kotsori, 2004); or were intended to validate an anemia questionnaire (Schwartzberg, Yee, Senecal, et al., 2004).

Results. In this open-label study, Glaspy, Berg, Tomita, et al. (2005) randomized patients with a variety of solid or hematological malignancies and receiving different treatment regimens to either epoetin or darbepoetin, and evaluated the FACT-An and FACT-fatigue subscale at 17 weeks. There were positive changes in both scores in the darbepoetin arms and in the epoetin arms (Table 19). The differences between changes from baseline to 17 weeks in the darbepoetin and epoetin arms were not statistically significant, suggesting no difference in treatment effect on anemia-related symptoms. While the primary endpoint of the trial was transfusion incidence and no power calculations were reported for QoL, the trial was large (N=1,220 randomized) and thus likely to detect meaningful differences between treatment arms on QoL scales.

The availability of only one study limited the evaluation of the impact of darbepoetin versus epoetin on anemia-related fatigue. In addition, 40% of randomized patients were not evaluable for QoL outcomes, and the authors did not present a prespecified plan for avoiding bias due to missing data in the abstract or poster for this otherwise unpublished study. Because this study

was not available as a full publication, trial design and methods related to QoL could not be fully evaluated.

Table 19. Results for Functional Assessment of Cancer Treatment (FACT) Quality of Life Scales for Studies Comparing Darbepoetin to Epoetin

Study	Study Arm	Dose	Follow-up (wks)	N Rand	N Evaluable for QoL	% Not eval-uable	Blinded?	FACT-Anemia			FACT Fatigue subscale		
								Change in score, from baseline	Diff, Darb - Epo	95% CI of Diff	Change in score, from baseline	Diff, Darb – Epo	95% CI of Diff
Glaspy 2005	Darb	200 mg every 2 weeks	17[1]	613	374[2]	39	No	7.1	0.60	-3.1; 4.3	4.2[3]	0.7	-0.8; 2.2
	Epo	40,000 IU weekly		607	357[2]	41		6.5			3.5[3]		

[1]Final assessment, assumed to be week 17, 1 week after completion of 16 weeks of therapy

[2]For the FACT-fatigue subscale, the numbers of evaluable patients were 373 for the Darb arm and 356 for the Epo arm

[3] Analysis of covariance adjusting for stratification variables (variables not reported)

rand = randomized; Con = control; Epo = erythropoietin alfa or beta; Darb = darbepoetin; Hb = hemoglobin; wks = weeks; N, number of patients

Epoetin vs. Control. Of twenty-four potentially relevant RCTs, thirteen were included for this review (N=2,947 randomized) (Bamias, Aravantinos, Kalofonos, et al., 2003; Boogaerts, Coiffier, Kainz, 2003; Dammacco, Castoldi, Rodjer, et al., 2001; Littlewood, Bajetta, Nortier, et al., 2001; Witzig, Silberstein, Loprinzi, et al., 2005; Del Mastro, Venturini, Lionetto, et al., 1997; Razzouk, Hockenberry, Hinds, et al., 2004; Iconomou, Koutras, Rigopoulos, et al., 2003; Chang, Couture, Young, et al., 2005; Osterborg, Brandberg, Molostova, et al., 2002; Abels 1992;[10] Kurz, Marth, Windbichler, et al., 1997; Thatcher, De Campos, Bell, et al., 1999). Of the total number of patients randomized, 81 percent (range across studies, 41–100 percent) were evaluable for QoL (N=2,374). More than 20 percent of data were missing (range: 22–59 percent) from 5 studies. Characteristics of reporting studies are enumerated in Tables 10, 11, and 12.

Of the twenty-five potentially relevant RCTs, the 11 studies we excluded accounted for 35% of the patients randomized (N=1,619); thus, data on over one-third of the randomized patients could not be included because of reporting problems. We excluded eight studies because authors did not state the numbers of participants evaluated for QoL results (Huddart, Welch, Chan, et al., 2002; Janinis, Dafni, Aravantinos, et al., 2003; Leyland-Jones, Semiglazov, Pawlicki 2005; O'Shaughnessy, 2002; Quirt, Micucci, Moran, et al., 1996; Rose, Rai, Revicki, et al., 1994; Savonije, Van Groeningen, Van Bochove, et al., 2005; Thomas, 2004; Welch, James, Wilkinson, 1995); one study because the authors did not report the number of patients evaluated for QoL separately for the epoetin and control groups (Carabantes, Benavides, Trujillo, et al., 1999); and two studies (Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995) because they duplicated the QoL results reported in another included study (Abels, 1992).

Darbepoetin vs. Control. Two studies (N=663 randomized) compared darbepoetin treatment vs. control and reported QoL outcomes (Vansteenkiste, Pirker, Massuti, et al., 2002; Hedenus, Adriansson, San Miguel, et al., 2003). There were no excluded studies. Characteristics of reporting studies and results are enumerated in Tables 20–23.

Evidence Regarding the Class of Erythropoietic Stimulants

Erythropoietic stimulants are considered to have similar pharmacodynamic properties (Food and Drug Administration Oncologic Drugs Advisory Committee Meeting Briefing Information, 2004). Moreover, erythropoietic stimulants raise Hb levels and ameliorate anemia and its consequences. Studies directly comparing epoetin and darbepoetin show similar ability to elicit Hb response, and based on one large study do not appear to differ in effects on QoL related to the symptoms of anemia. Therefore, we analyzed epoetin and darbepoetin vs. control trials that reporting QoL outcomes together for more robust results.

[10] Abels 1992 pooled three studies: Abels 1992 (n=124); Case, Bukowski, Carey, et al., 1993 (n=157); and Henry, Brooks, Case, et al., 1995 (n=132). These studies had slightly different protocols and in the Abels 1992 study patients did not receive chemotherapy. Thus, the Abels 1992 report includes some patients who do not exactly fulfill inclusion criteria for this review.

Table 20. Results for Symptom-Specific FACT Quality of Life Scales from Studies Comparing Epoetin or Darbepoetin to Placebo or No Treatment

Study	Study arm	Follow-up (wks)	N rand	N evaluable for QoL	% not evalu-able	Blinded ?	FACT-An		FACT-G		FACT anemia subscale		FACT fatigue subscale		FACT non-fatigue anemia subscale	
							% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]
MEAN/MEDIAN BASELINE Hb <10 g/dl																
Boogaerts 2003	Ctl	12	129	109	16	No							Not given		Not given	
	Epo		133	104	22								11	<0.05	5	0.076, NS
Littlewood 2001	Ctl	4-24	124	88	29	Yes			-5							
	Epo		251	194	23				3	0.004						
Littlewood 2001	Ctl	4-24	124	90	27								-8			
	Epo		251	200	20								10	0.004		
Osterborg 2002	Ctl	16	173	101	42	Yes	8									
	Epo		170	105	38		13	<0.05								
Osterborg 2002	Ctl	16	173	103	40				5							
	Epo		170	106	38				9	<0.05						
Osterborg 2002	Ctl	16	173	130	25								10		10	
	Epo		170	133	22								18	>0.05, NS	12	>0.05, NS
Witzig 2005	Ctl	16	170	139	18	Yes	0									
	Epo		174	148	15		0	0.4, NS								
Witzig 2005	Ctl	16	170	140	18				0							
	Epo		174	148	15				-2	0.6, NS						
Witzig 2005	Ctl	16	170	148	13								1			
	Epo		174	151	13								6	0.18, NS		
Hedenus 2003	Ctl	12	173	151	13	Yes							2			
	Darb		176	152	14								9	NS		

[1]% change calculated as (end-baseline)/baseline; positive values indicate improved quality of life.

[2]comparing %change in treatment arm to %change in control arm.

[3]%change not reported; 56% of darbepoetin-treated patients and 44% of placebo-treated patients had an improved FACT-fatigue score (p=0.052); 32% of darbepoetin-treated patients and 19% of placebo-treated patients showed ≥25% improvement in FACT-fatigue score (p=0.019).

Table 20. Results for Symptom-Specific FACT Quality of Life Scales from Studies Comparing Epoetin or Darbepoetin to Placebo or No Treatment (continued)

Study	Study arm	Follow-up (wks)	N rand	N evaluable for QoL	% not evalu-able	Blinded ?	FACT-An		FACT-G		FACT anemia subscale		FACT fatigue subscale		FACT non-fatigue anemia subscale	
							% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]	% Change[1]	p-value[2]
MEAN/MEDIAN BASELINE Hb 10 to 12 g/dl																
Chang 2005	Ctl	12	175	170	3	No					-8		-11		-4	
	Epo		175	168	4						4	<0.0001	5	<0.0001	1	<0.001
Iconomou 2003	Ctl	12	61	55	10	No							-4			
	Epo		61	57	7								21	0.022		
Vansteenkiste 2002	Ctl	12 wks	158	128	19	Yes							Not given			
	Darb		156	127	19								Not given	0.019[3]		
MEAN/MEDIAN BASELINE Hb >12 g/dl																
No trials																

[1]% change calculated as (end-baseline)/baseline; positive values indicate improved quality of life.
[2]comparing %change in treatment arm to %change in control arm.
[3]%change not reported; 56% of darbepoetin-treated patients and 44% of placebo-treated patients had an improved FACT-fatigue score (p=0.052); 32% of darbepoetin-treated patients and 19% of placebo-treated patients showed ≥25% improvement in FACT-fatigue score (p=0.019).

Table 21. Results for FACT Submeasures Categorized as Significantly Pro-treatment, Not Significant, or Significantly Pro-control

Study	N evaluable for QoL	% of enrolled not evaluable	Blinded?	FACT Scale	significantly pro-tx	not significant	significantly pro-control
Littlewood 2001	282	25	**Yes**	**FACT-G**	**X**		
Osterborg 2002	209	39	**Yes**	**FACT-G**	**X**		
Witzig 2005	288	16	**Yes**	**FACT-G**		**X**	
Osterborg 2002	206	40	**Yes**	**FACT-An**	**X**		
Witzig 2005	287	16	**Yes**	**FACT-An**		**X**	
Chang 2005	338	4	**No**	**FACT anemia subscale***	**X**		
Boogaerts 2003	213	19	**No**	**FACT fatigue subscale**	**X**		
Littlewood 2001	290	22	**Yes**	**FACT fatigue subscale**	**X**		
Osterborg 2002	263	24	**Yes**	**FACT fatigue subscale**		**X**	
Witzig 2005	299	13	**Yes**	**FACT fatigue subscale**		**X**	
Hedenus 2003	303	14	**Yes**	**FACT fatigue subscale**		**X**	
Chang 2005	338	4	**No**	**FACT fatigue subscale**	**X**		
Iconomou 2003	112	8	**No**	**FACT fatigue subscale**	**X**		
Vansteenkiste 2002	255	19	**Yes**	**FACT fatigue subscale**	**X**		
Boogaerts 2003	213	19	**No**	**FACT non-fatigue anemia subscale**		**X**	
Osterborg 2002	263	24	**Yes**	**FACT non-fatigue anemia subscale**		**X**	
Chang 2005	338	4	**No**	**FACT non-fatigue anemia subscale**	**X**		

*Includes the FACT-fatigue subscale

Table 22. Results for General QoL Scales from Studies Comparing Epoetin or Darbepoetin to Placebo or No Treatment

Study	Study arm	Time (wks)	N randomized	N evaluable for QoL	% not evalu-able	Blinded?	QoL measure	% change[1]	p-value[2]
MEAN/MEDIAN BASELINE Hb < 10 g/dl									
Bamias 2003	Ctl	≤24	72	27	63	No	EORTC QLQ-C30	Not given	
	Epo		72	32	56			Not given	NS
Boogaerts 2003	Ctl	12	129	109	16	No	SF-36 PCS	Not given	
	Epo		133	104	22			9	<0.05
Dammacco 2001	Ctl	12	76	72	5	Yes	NHP (each of 6 domains)	Not given	
	Epo		69	66	4			Not given	NS
Littlewood 2001	Ctl	4-24	124	90	27	Yes	SF-36 PCS	-1	
	Epo		251	200	20			5	0.0512, NS
Littlewood 2001	Ctl	4-24	124	90	27		SF-36 MCS	-1	
	Epo		251	200	20			5	0.0952, NS
Witzig 2005	Ctl	16	170	147	14	Yes	SDS	3	
	Epo		174	151	13			1	0.39, NS
MEAN/MEDIAN BASELINE Hb 10 to 12 g/dl									
No trials									
MEAN/MEDIAN BASELINE Hb >12 g/dl									
Del Mastro 1997	Ctl	26	31	26	16	No	PDI	3	
	Epo		31	27	13			0	0.4, NS
MEAN/MEDIAN BASELINE Hb UNCLEAR									
Razzouk 2004A	Ctl	16	111	86	23	Yes	Patient reported PedQL-I total score	8	
	Epo		111	94	15			10	NS

[1]% change calculated as (end-baseline)/baseline; positive values indicate improved quality of life.
[2]comparing %change in treatment arm to %change in control arm.

EORTC QLQ-C3, European Organization for Research & Treatment of Cancer quality of life questionnaire; SF-36 PCS & MCS, Medical Outcomes Study, Short Form 36, Physical/Mental Component Score; NHP, Nottingham Health profile; SDS, Symptom Distress Scale; PDI, Psychological Distress Inventory; PedQL-I, Pediatrics Quality of Life Inventory; Hb, hemoglobin; Con, control; Epo, epoetin; QoL, quality of life; wks, weeks; NS, not significant; N, number of patients

Table 23. Results for Visual Analog Scales (VAS) from Studies Comparing Epoetin or Darbepoetin to Placebo or No Treatment

Study	Study arm	Time (wks)	N rand	N evaluable for QoL	% not evalu-able	Blinded?	% change: energy[1]	p value: energy[2]	% change: daily activities[1]	p value: daily activities[2]	% change: overall QoL[1]	p value: overall QoL[2]
MEAN/MEDIAN BASELINE Hb < 10 g/dl												
Abels 1992	Ctl	12	207	143	31	Yes	Not given		Not given		-5	
	Epo		206	159	23		Not given	>0.05, NS	Not given	>0.05, NS	9	<0.05
Boogaerts 2003	Ctl	12	129	112	13	No					2	
	Epo		133	111	17						18	0.004
Dammacco 2001	Ctl	12	76	72	5	Yes	Not given		Not given		Not given	
	Epo		69	66	4		Not given	NS	Not given	NS	Not given	NS
Kurz 1997	Ctl	12	12	12	0	Yes	Not given		Not given		Not given	
	Epo		23	23	0		Not given	0.71, NS	Not given	0.53, NS	Not given	0.77, NS
Littlewood 2001	Ctl	4-24	124	108	13	Yes	-13		-13			
	Epo		251	228	9		18	0.0007	16	0.0018		
Littlewood 2001	Ctl	4-24	124	107	14						-12	
	Epo		251	228	9						9	0.0048
Witzig 2005	Ctl	16	170	147	14	Yes					-6	
	Epo		174	150	14						-9	0.58, NS
MEAN/MEDIAN BASELINE Hb 10 to 12 g/dl												
Chang 2005	Ctl		175	169	3	No	-9		-8		-10	
	Epo		175	166	5		6	<0.014	7	<0.01	6	<0.001
Iconomou 2003	Ctl	12	61	55	10	No	-3		-3		-2	
	Epo		61	57	7		14	0.022	19	0.003	15	0.03
MEAN/MEDIAN BASELINE Hb >12 g/dl												
Thatcher 1999	Ctl	16 to 24	44	27	39	No	3		26		16	
	Epo		42	33	21		-4	NS	6	NS	24	NS
	Ctl	16 to 24	44	27	39		3		26		16	
	Epo (higher dose)		44	32	27		6	NS	10	NS	11	NS

[1]% change calculated as (end-baseline)/baseline; positive values indicate improved quality of life.
[2]comparing %change in treatment arm to %change in control arm.
rand = randomized; Con = control; Epo = erythropoietin alfa or beta; Hb = hemoglobin; wks = weeks; NS = not significant (p>0.05); N, number of patients; QoL, quality of life

Analysis of Epoetin versus Control and Darbepoetin versus Control

Fifteen controlled studies randomized a total of 3,610 patients to treatment; 81 percent of randomized patients were evaluable for QoL (N=2,932).

Analysis of study quality detected threats to validity in most included studies. Six of 15 studies were not blinded (see Tables 20-23). In 6 studies, allocation concealment was unclear (Bamias, Aravantinos, Kalofonos, et al., 2003; Chang, Couture, Young, et al., 2005; Dammacco, Castoldi, Rodjer, et al., 2001; Iconomou, Koutras, Rigopoulos, et al., 2003; Razzouk, Hockenberry, Hinds, et al., 2004; Witzig, Silberstein, Loprinzi, et al., 2005). Several comparisons that significantly favored treatment were in relatively large studies for which only approximately 60–75 percent of randomized patients were evaluable for QoL (e.g., Osterborg, Brandberg, Molostova, et al., 2002; Littlewood, Bajetta, Nortier, et al., 2001). Most studies with adequate followup used a last observation carried forward approach to impute missing data, which may distort results in either direction but particularly in favor of treatment if subjects whose outcomes are worsening are lost early (Streiner, 2002). Finally, many studies provided limited details about the timing and circumstances under which the QoL measures were administered.

Results. We preferred to analyze QoL results by quantitative techniques, but could not because of incomplete reporting in several trials of both the numerical results (e.g., some study publications reported only percentage change without baseline value) and measures of their dispersion. Meta-analysis of subsets of trials reporting sufficient data may not be representative of all trials reporting QoL results and could risk significant bias. For example, of 8 studies that administered a FACT QoL instrument, all reported results for the FACT-fatigue subscale. However, only 3 reports included information on result variance. In the absence of sufficient data for a representative meta-analysis, we evaluated the results of all included studies reporting QoL results similarly, based on patterns in the tabulated results ("vote-counting"). We stratified results according to the specific measurement tool employed, distinguishing QoL measured by condition-specific FACT subscales from that measured by global instruments or by VAS.

Results from FACT scales. Eight studies contributed data on QoL assessed by FACT modules (Table 20; for details on FACT scales, refer to Table 18). Of the total number of patients randomized in these studies (N=2,459), 84 percent (N=2,073) were evaluable for QoL (range evaluable across studies, 60–86 percent). The relatively large studies by Littlewood, Bajetta, Nortier, et al. (2001) and Osterborg, Brandberg, Molostova, et al. (2002) reported on QoL for only 75–78 percent and 60–77 percent of 375 and 343 randomized patients, respectively, depending on FACT measure.

When authors clearly presented numerical results, results generally favored treatment (though were not necessarily statistically significant), with only 2 comparisons (1 each in FACT-An and FACT-G; note that FACT-G is a general measure, and FACT-An includes the FACT-G as well as the symptom-specific modules FACT-fatigue subscale and FACT-anemia non-fatigue subscale.) showing no difference or slightly favoring control (Table 20, shaded). All symptom-specific instrument comparisons favored treatment, although not all were statistically significant. For each FACT measure, the balance among study results significantly in favor of treatment, not significantly different, and significantly in favor of control, favors active treatment (Table 21). The FACT-fatigue subscale was used most often; it significantly favored treatment in 5 trials and

favored treatment, but not significantly, in three other trials. However, for each FACT measure used in more than 1 trial, both significant and nonsignificant results were reported. Five of 11 results (45 percent) reported by blinded trials were significant, while 5 of 6 results (83 percent) reported by unblinded trials were significant (Table 21).

Six of 10 significantly favorable results were in studies that had 19 percent or more missing data. Only 2 studies compared darbepoetin to control; results are qualitatively similar to those comparing epoetin to control. Although there is consistency in the direction of effect on FACT-based measures, there is marked variation in size of effect. Without complete data allowing quantitative meta-analysis, it is not possible to determine the size or the statistical significance of the effect.

Results from general instruments. Seven studies contributed data on QoL assessed by validated general instruments other than FACT-G (Table 22). Of the total number of patients randomized (N=1,554), 79% (N=1,231) were evaluable for QoL (range evaluable across studies, 41–96 percent). Bamias, Aravantinos, Kalofonos, et al. (2003) is a relatively small study, with more than 50 percent of patients in each arm not evaluable for QoL results. Only one result significantly favored treatment and none significantly favored control. The rest of the study findings were not significantly different and where numerical results were given, 3 slightly favored treatment and 2 slightly favored control. The heterogeneity of measures and small number of studies reporting any one measure makes an overall pattern difficult to identify with confidence.

Only four studies administered both symptom-specific and global scales (Osterborg, Brandberg, Molostova, et al., 2002; Witzig, Silberstein, Loprinzi, et al., 2005; Littlewood, Bajetta, Nortier, et al., 2001; Boogaerts, Coiffier, Kainz, 2003). The pattern of results is different for each study (Table 24); thus, there are insufficient data to determine whether the positive results from symptom-specific scales are routinely detectable on general QoL scales.

Table 24. Significant vs. Nonsignificant QoL Results for Studies Reporting both Symptom-Specific and General Scale Results

Study	Symptom-specific measure		General measure	
	FACT-fatigue subscale	FACT-anemia non-fatigue subscale	FACT-G	Other general instrument
Littlewood 2001	p<0.05		p<0.05	NS
Osterborg 2002	NS	NS	p<0.05	
Witzig 2005	NS		NS	NS
Boogaerts 2003	p<0.05	NS		p<0.05

Results from VAS instruments. Nine studies contributed VAS data on the impact of epoetin on QOL (Table 23). Thatcher 1999 tested 2 different epoetin doses and is counted as two trials. Of the total number of patients randomized (N=2,176), 86 percent (N=1,865) were evaluable for QoL (range evaluable across studies, 71–100%). The balance among studies reporting significantly in favor of treatment, no significant difference, and significantly in favor of control is: 3 studies vs. 5 studies vs. 0 studies for VAS-energy; 3 vs. 5 vs. 0 for VAS-abilities; and 5 vs. 5 vs. 0 for VAS-overall (all respectively). Because several studies did not report numerical data, it could not be determined whether, without regard to statistical significance, treatment or control was more often favored.

KQ1 Outcome IV. Survival

We abstracted death events from included studies reporting this outcome. Hazard ratios (HR) for death were calculated as reported in Methods, Statistical Data Analysis. While all studies reporting survival are included, only 7 studies of either epoetin or darbepoetin (Henke, Laszig, Ruebe, et al., 2003; Leyland-Jones 2003; Machtay, Pajak, Suntharalingam, et al., 2004; GOG-191, 2004; N93 004, 2004, Vansteenkiste, Pirker, Massuti, et al., 2002; EPO-CAN-15) were actually designed to evaluate either overall or progression-free survival. Other studies were neither designed nor powered for this outcome, and in some studies evaluated retrospective data, collected after the study closed and when patient management was no longer directed by the study protocol. Additionally, studies differed in tumor type studied (e.g. solid vs. solid and hematologic), and in homogeneity of tumor type (e.g. one type of solid tumor vs. many types of solid tumors). Disease progression patterns of different types of malignancies can significantly influence survival outcomes. Moreover, the underlying mortality in each patient population studied interferes with the observation of specific effects on overall mortality, and cause-specific mortality data were not available. Studies also differed in length of reported followup, and seldom reported survival at several different time points.

We pooled all-cause survival data in a meta-analysis to update and test the hypothesis of improved survival with epoetin treatment suggested by an earlier analysis (Bohlius, Langensiepen, Schwarzer et al., 2005). While we would have preferred to utilize data on cause-specific mortality (e.g. from tumor progression, thrombosis, CVD), these data were not available. Given the limitations of the data as described, quantitative pooling of all-cause mortality is necessarily problematic, and the use of these data is largely hypothesis-generating, rather than conclusive.

Evidence for Comparative Effectiveness

Darbepoetin versus Epoetin. Only one of the included studies comparing darbepoetin to epoetin assessed overall survival (Waltzman, Croot, Williams, 2005); survival was a secondary outcome. In this study of 358 randomized patients undergoing chemotherapy for solid tumors, the authors reported that 16% of patients in the darbepoetin arm and 13% of those in the epoetin arm died “on study”. In absolute numbers, 29 of 180 patients in the darbepoetin arm died, as did 23 of 178 in the epoetin arm, not significantly different at p=0.4.

This single trial directly comparing commonly-used doses of darbepoetin and epoetin found no difference in survival. A limitation of the trial is that it was not powered for survival outcomes; the primary outcome was comparison of hematologic response rates. Additional limitations are the short followup time (17 weeks) and variety in tumor types and chemotherapy regimens. Given limited direct evidence from only one trial, we evaluated indirect evidence (epoetin vs. control, darbepoetin vs. control) for effect on survival outcomes.

Epoetin versus Control. Characteristics of reporting studies are shown in Table 25. Thirty-five trials (N=6,918; 3,825 randomized to epoetin, 3,093 to control) reported survival outcomes (Bamias, Aravantinos, Kalofonos, et al., 2003; Cascinu, Fedeli, Del Ferro, et al., 1994; Case, Bukowski, Carey, et al., 1993; Cazzola, Messinger, Battistel, et al, 1995; Chang, Couture, Young,

Table 25. RCTs Reporting Survival: Overall and Subgroup Analyses of Hazard Ratios for Death, Epoetin Compared to Control

Outcome Subgroup	Epoetin versus Control					Darbepoetin versus Control				
	# Studies	#Total Patients	#Epo/#Ctl Patients	Hazard Ratio for death	95% CI (p-value)	# Studies	#Total Patients	#Darbepoetin/ #Ctl Patients	Hazard Ratio for death	95% CI (p-value)
Overall Survival	35	6,918	3,825/3,093	1.11	1.00; 1.22	4	973	583/390	0.96	0.78; 1.17
(Heterogeneity)					(0.48)					**(0.03)**
Subgroup Analyses: Patient Baseline Characteristics										
Bsln Hb <10	14	2,830	1,590/1,240	0.96	0.83; 1.10	3	659	428/231	1.31	0.95; 1.81
Bsln Hb 10-12	7	1,398	782/616	1.17	0.93; 1.49	1	314	155/159	0.78	0.60, 1.01
Bsln Hb >12[1]	7	1,696	870/826	1.27	1.05; 1.54					
Bsln Hb unclear[1]	7	994	583/411	1.63	1.07; 2.49					
(Group difference[2])					**(0.025)**					**(0.015)**
Solid tumors	23	4,526	2,420/2,106	1.22	1.07; 1.38	2	563	353/210	0.77	0.60; 1.00
Hematologic	6	1,044	626/418	1.02	0.81; 1.29	2	410	230/180	1.36	0.98; 1.89
Mixed	6	1,348	779/569	0.86	0.68; 1.08					
(Group difference[2])					**(0.027)**					**(0.008)**
Subgroup Analyses: Treatment Protocols										
Chemo, any	10	1,474	884/590	1.14	0.74; 1.7	1	314	155/159	0.78	0.60; 1.01
Chemo, some plat	4	955	482/473	1.01	0.79; 1.30	1	249	198/51	0.55	0.11; 2.61
Chemo, no plat	13	3,302	1,859/1,443	1.06	0.92; 1.21	2	410	230/180	1.36	0.98; 1.89
Chemo+RT or RT	8	1,187	600/587	1.27	1.05; 1.55					
(Group difference[2])					(0.4134)					**(0.027)**
Iron, fixed	2	360	181/179	1.08	0.82; 1.42					
Iron, as needed	19	3,522	1,964/1,558	0.99	0.86; 1.13	1	344	175/169	1.36	0.98; 1.89
Iron unknown	13	3,036	1,680/1,356	1.32	1.11; 1.55	3	629	408/221	0.77	0.60; 1.00
(Group difference[2])					**(0.033)**					**(0.008)**

Outcome Subgroup	Epoetin versus Control					Darbepoetin versus Control				
	# Studies	#Total Patients	#Epo/#Ctl Patients	Hazard Ratio for death	95% CI (p-value)	# Studies	#Total Patients	#Darbepoetin/ #Ctl Patients	Hazard Ratio for death	95% CI (p-value)
Subgroup Analyses: Treatment Protocols (continued)										
Epo tx 6-9 weeks	6	823	461/362	1.25	0.97; 1.59					
Epo tx 12-16 weeks	19	3,679	2,009/1,670	1.05	0.90; 1.23	(all)				
Epo tx >20 weeks	7	1,958	1,113/845	1.13	0.95; 1.33					
Epo tx ? Weeks	3	458	242/216	1.02	0.71; 1.46					
(Group difference[2])					(0.68)					

Table 25. RCTs Reporting Survival: Overall and Subgroup Analyses of Hazard Ratios for Death, Epoetin Compared to Control (continued)

Outcome Subgroup	Epoetin versus Control					Darbepoetin versus Control				
	# Studies	#Total Patients	#Epo/#Ctl Patients	Hazard Ratio for death	95% CI (p-value)	# Studies	#Total Patients	#Darbepoetin/#Ctl Patients	Hazard Ratio for death	95% CI (p-value)
Subgroup Analyses: Reporting and Quality										
High quality	20	4,384	2,380/2,004	1.14	1.02; 1.27	(all)				
Low quality	15	2,534	1,445/1,089	1	0.81; 1.24					
(Group difference[2])					(0.3087)					
Full Text	8	1,800	983/817	0.98	0.84; 1.13	2	315	253/62	0.55	0.11; 2.61
Abstract	3	678	394/284	1.27	0.79; 2.06					
Unpublished data	5	384	199/185	0.6	0.25; 1.41					
FDA documents	19	4,056	2,249/1,807	1.25	1.08; 1.44	2	658	330/328	0.96	0.79; 1.18
(Group difference[2])					(0.17)					(0.48)
Followup <1 year	24	3,393	1,998/1,395	1.00	0.77; 1.31	2	315	253/62	0.55	0.11; 2.61
Followup >1 year	11	3,525	1,827/1,698	1.12	1.01; 1.25	2	658	330/328	0.96	0.79; 1.18
(Group difference[2])					(0.43)					(0.48)

[1]The N93-004 epoetin trial was published in full in December, 2005 (Grote, Yeilding, Castillo, et al., 2005) and included information on baseline Hb which classified it into subgroup Hb >12. A re-categorized analysis of epoetin vs. control trials resulted in subgroup Hb>12 RR 1.28 (95% CI, 1.06, 1.54), an insignificant change. Because this did not alter the interpretation of results, we did not alter our presentation of the overall analysis.

[2]p-value for differences among subgroup categories calculated by inverse variance method (see Methods/Data Extraction and Analysis/Statistical Data Analysis).

et al., 2005; Coiffier and Boogaerts, 2001; Dammacco, Castoldi, Rodjer, et al., 2001; Del Mastro, Venturini, Lionetto, et al., 1997; Dunphy, Harrison, Dunleavy, et al., 1999; Henke, Laszig, Ruebe, et al., 2003; Henry, Brooks, Case, et al., 1995; Kurz, Marth, Windbichler, et al., 1997; Leyland-Jones, 2003; Littlewood, Bajetta, Nortier, et al., 2001; Machtay, Pajak, Suntharalingam, et al., 2004; Oberhoff, Neri, Amadori, et al., 1998; O'Shaughnessy, Vukelja, Holmes, et al., 2005; Osterborg, Brandberg, Hedenus, 2005; Osterborg, Boogaerts, Cimino, et al., 1996; Razzouk, Hockenberry, Hinds, et al., 2004; Rose, Rai, Revicki, et al., 1994; Savonije, Van Groeningen, Van Bochove, et al., 2004; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Thatcher, De Campos, Bell, et al., 1999; Throuvalas, Antonadou, Boufi, et al., 2000; Vadhan-Raj, Skibber, Crane, et al., 2004; Witzig, Silberstein, Loprinzi, et al., 2005; EPO-CAN-15, 2004; EPO-CAN-20, 2004; EPO GBR-07, 2004; GOG-191, 2004; N93 004, 2004; INT-1, 2004; INT-3, 2004; P-174, 2004).

Trials that reported survival differed with respect to several variables. Baseline characteristics of study populations differed by average baseline Hb concentration, type of malignancy, treatment for malignancy, and age. One study included pediatric patients (Razzouk, Hockenberry, Hinds, et al., 2004). Treatment protocols also differed by therapies for iron supplementation, and duration of epoetin treatment. Trials varied with respect to publication type, overall quality rating, and duration of followup. In addition, several trials (Cazzola, Messinger, Battistel, et al, 1995; INT-1, 2004; Henke, Laszig, Ruebe, et al., 2003; Osterborg, Boogaerts, Cimino, et al., 1996; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Throuvalas, Antonadou, Boufi, et al., 2000) tested 2 or more doses, such that one or more treatment arms received doses that are 50 percent higher than currently recommended. However, for these studies survival data were available only for pooled treatment arms. Thatcher, De Campos, Bell, et al. (1999) also tested 2 different doses but reported death events by treatment arm; treatment arm b tested a higher than usual dose.

Results. A test for heterogeneity across trials included for survival outcomes was not statistically significant (p=0.48, I^2=0.0%). An I^2 value of 0% indicates no observed statistical heterogeneity.

Meta-analysis of data from all 35 trials (Figure 9) yielded:

- Fixed-effects HR=1.11 (95% CI 1.00; 1.22)[11,12,13] , p=0.05

[11] Thatcher, De Campos, Bell, et al. (1999) compared 2 different epoetin doses to one control group; for this study, survival data were available for each epoetin arm and the control arm was randomly divided into 2 separate control arms for meta-analysis. As this study contributed only 0.47% weight to the meta-analysis, the influence of changes in control arm weight was judged to be negligible.

[12] Two studies (Littlewood, Bajetta, Nortier, et al., 2001; Henke, Laszig, Ruebe, et al., 2003) reported both adjusted (for potential prognostic factors) and unadjusted survival data. We used unadjusted data for our base model, but the overall result was similar when adjusted data were used (see below). If the analysis used either the best-case or worst-case results from these two studies, the results varied only minimally.

Unadjusted data	HR 1.11 (95% CI 1.00; 1.23) (base model)
Adjusted data	HR 1.11 (95% CI 1.01; 1.23)
Best case scenario	HR 1.10 (95% CI 0.99; 1.21)
Worst case scenario	HR 1.12 (95% CI 1.01; 1.23)

[13] For the Leyland-Jones study there is a discrepancy between the numbers of death events used in this review (148) and those reported by NICE (141). The major difference between this review and the NICE report is that we retrieved the survival data from the briefing document for the FDA-ODAC hearing in May 2004. The FDA briefing document notes that the reported percentages of patients who survived or died are based on Kaplan-Meier estimates. The obtained hazard ratio (1.37), derived from the data reported in the FDA document (see Figure 9) is the same as that reported in the Leyland-Jones et al., 2005

- Total event rates were 26.4% for epoetin treatment arms and 26.8% for control arms.

- HRs ranged from 0.13 to 2.7 with one extreme value of 7.39.

Meta-analysis of all studies included for survival outcomes does not show improved survival with epoetin treatment. The point estimate of the HR for death is greater than 1 but not significant; the 95% confidence interval of the estimate indicates either no effect or a slight detrimental effect of epoetin.

Since many trials lacked information on baseline Hb and iron supplementation, subgroup analysis was not informative. While there appeared to be a statistically significant increase in HR for death for the subgroup of patients with solid tumors receiving epoetin, the effect was small and the confidence interval overlapped substantially with that for hematologic malignancies, which did not differ significantly from 1.0. No subgroups had HR point estimates significantly less than 1.0.

Most studies only provided qualitative information on followup duration (e.g. duration of chemotherapy plus 28 days). One study (Leyland-Jones, 2003) reported an increase in deaths in the epoetin arm, compared to control, within the first 4 months of followup. Thus, an analysis of early events across trials might be informative but for most studies data were unavailable to differentiate early (e.g., <4 months) from late events, or analyze survival at specific times across studies.

Additional Analyses. After initial review of results, we conducted additional analyses not anticipated in the original protocol in order to answer specific questions or explore new hypotheses. We used an influence analysis to identify those studies that most strongly influenced the pooled HR for death. We conducted a subgroup analysis of those studies that administered epoetin according to current labeled criteria vs. those studies that used criteria exceeding the labeled limits of initial dose or target Hb value. We also compared HR for death among subgroups defined by 1 g/dL increments in maximum Hb target value. Finally, because few studies were actually designed to prospectively evaluate survival outcomes during the followup period specified by the original study protocol, we evaluated subgroups according to whether or not trials had key design characteristics necessary for reliable survival outcomes.

Results of the influence analysis, which omits each study, one at a time, and pools the remaining studies, are shown in Figure 10. Three studies that most strongly change the results are Henke, Laszig, Ruebe, et al., 2003; Leyland-Jones, 2003; and Littlewood, Bajetta, Nortier, et al., 2001 (Table 26). Omitting Leyland-Jones (2003) or Henke, Laszig, Ruebe, et al. (2003) reduces the HR point estimate such that the result is clearly nonsignificant. Omitting Littlewood, Bajetta, Nortier, et al. (2001) increases the HR point estimate and the result is statistically significant for decreased survival with epoetin treatment.

---

publication. The NICE report used the information from the Leyland-Jones 2003 paper, in which the absolute number of deaths was not specified, for the estimation of death events. We assume that the NICE team used the reported 70% survival rate in the epoetin arm to estimate the absolute number of deaths (i.e. 30% of 470 = 141).

Figure 9. Meta-Analysis of Data on Survival from 35 RCTs of Epoetin versus Control

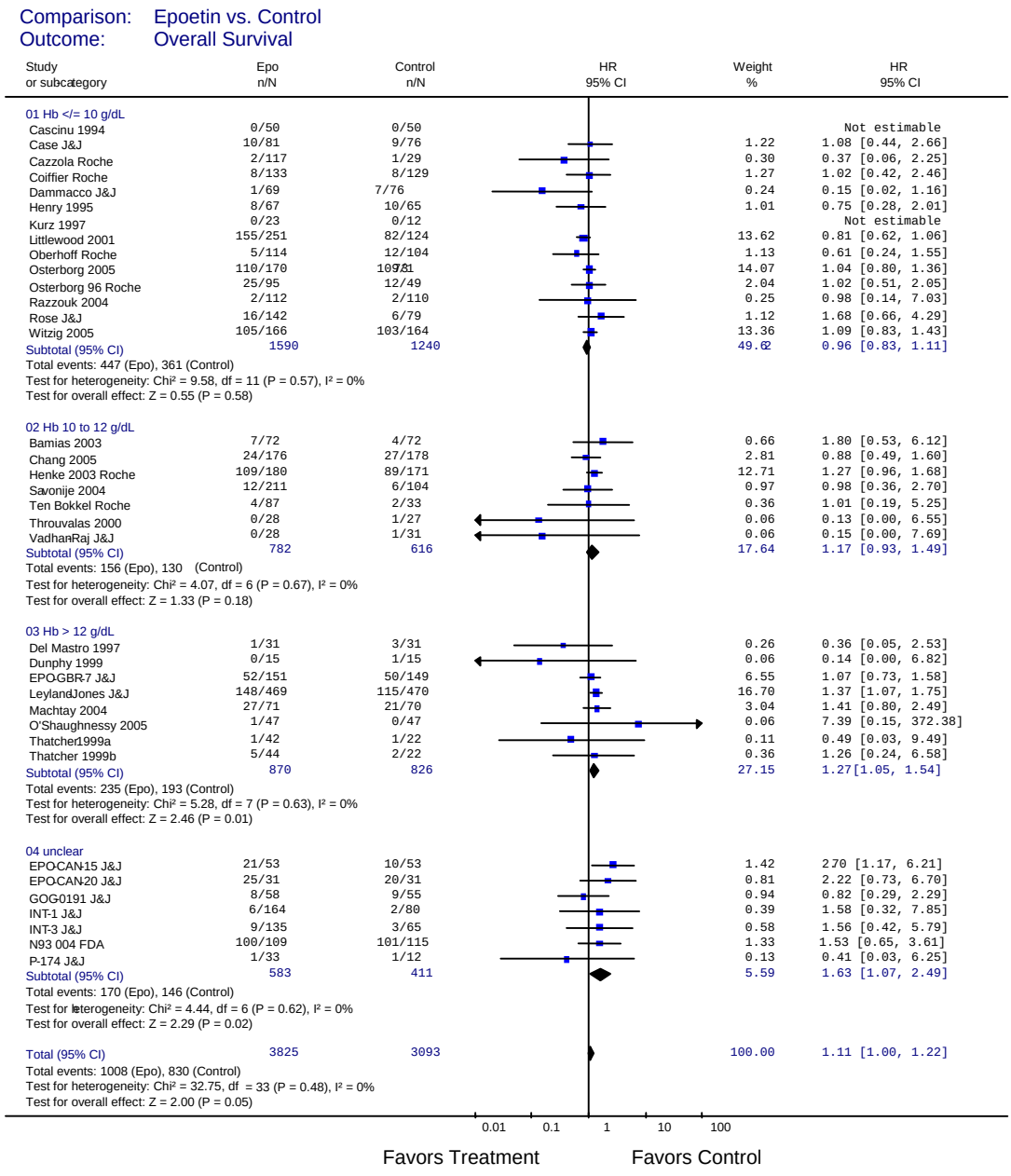


Figure 10. Influence Analysis: Hazard Ratios for Death Recalculated after Omission of One Study at a Time; Point Estimates (Squares) and 95% Confidence Intervals (Lines)

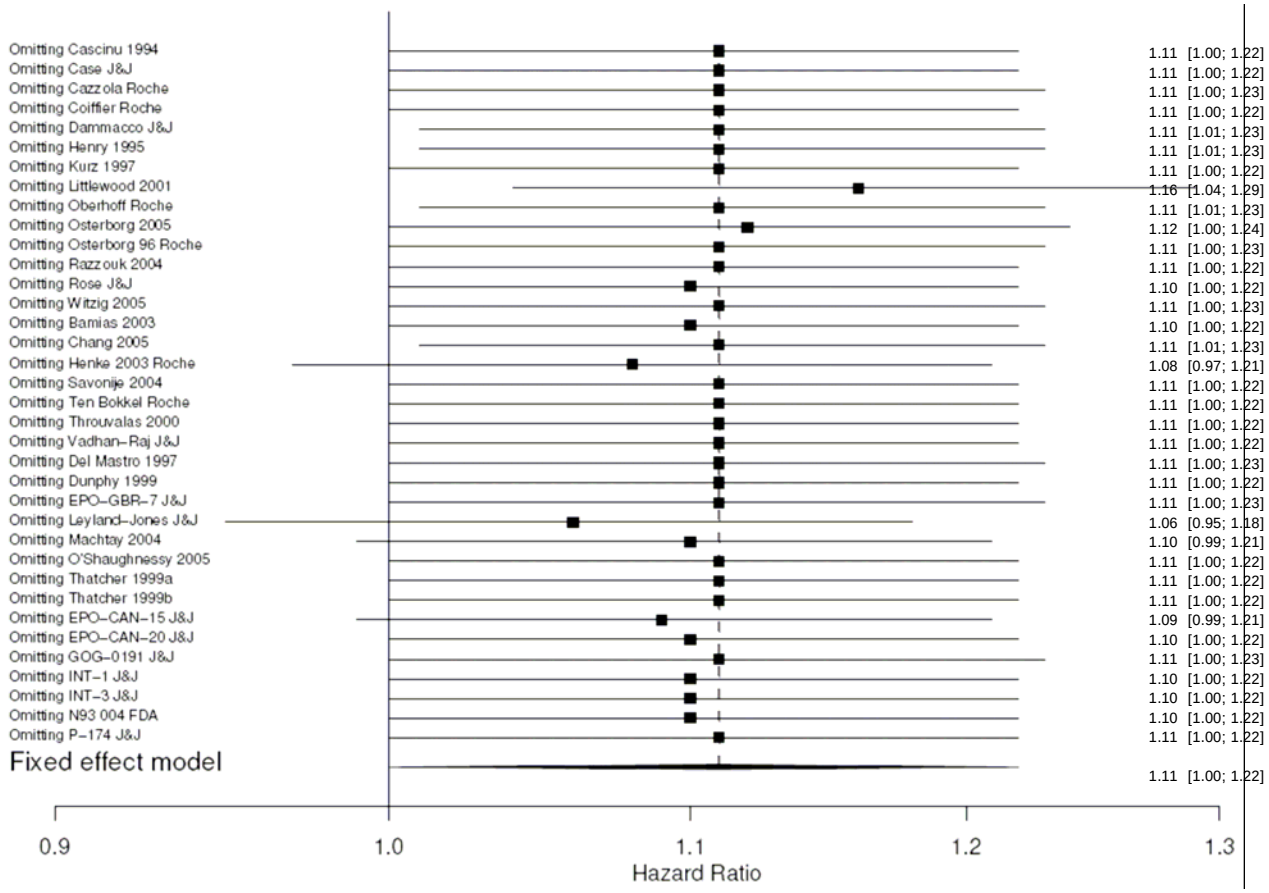


Table 26. Key Characteristics of Studies that Strongly Influence Survival Meta-Analysis

Study omitted	Starting epoetin dose	Baseline Hb category	Hb target, upper limit	Pooled HR for death after study omitted (95% CI)
Leyland-Jones 2003	1x40,000 IU/wk	12	14	1.06 (0.95; 1.18)
Henke 2003	3x300 IU/kg/wk	10-12	15	1.08 (0.97; 1.21)
(none)				1.11 (1.00; 1.22)
Littlewood 2001	3x150 IU/kg/wk	10	15	1.16 (1.04; 1.29)

Over time, clinical trials of epoetin have recruited patients with higher baseline Hb and/or administered higher doses to raise Hb to higher target values, beyond that specified in the labeled drug administration criteria. Differences in these variables among the three studies, however, do not clearly explain the contrasting results of Henke, Laszig, Ruebe, et al. (2003) and Leyland-Jones (2003) vs. Littlewood, Bajetta, Nortier, et al. (2001). While Henke, Laszig, Ruebe, et al. (2003) used twice the labeled dose of epoetin, Leyland-Jones (2003) administered a standard dose as in the Littlewood study (Appendix C Table C28); all three studies targeted a final Hb value well above the current labeled limit of 13 g/dL. The studies each enrolled patients in different baseline Hb categories, which were <10, 10-12, and >12 for Littlewood, Bajetta, Nortier, et al. (2001), Henke, Laszig, Ruebe, et al. (2003), and Leyland-Jones (2003), respectively.

Due to recent concern with increased epoetin exposure (see Scope and Key Questions), a major question of interest for clinicians and their patients is whether there is a clear distinction between FDA-recommended ("labeled") and "unlabeled" use. As listed in Table 1 of the Introduction, the current product labels include recommended doses and hemoglobin or hematocrit levels at which to reduce dose or temporarily stop administration, although no starting Hb level is specified. We identified 3 studies (Cascinu, Fedeli, Del Ferro, et al., 1994; Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995), constituting 5.6% of all patients evaluated for survival in included trials, that most closely met current labeled criteria for use and compared these to all other trials in a subgroup analysis. These studies used labeled (Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995) or slightly lower doses (Cascinu, Fedeli, Del Ferro, et al., 1994) and stopped administration when Hb reached 13 g/dL. Dose reduction strategies were slightly different from labeled recommendations. Of these, one trial (Cascinu, Fedeli, Del Ferro, et al., 1994) reported no deaths in either arm and thus does not contribute to the analysis.

For this subgroup analysis there was no evidence of heterogeneity within subgroups or overall (I^2 = 0-1%). Subgroup meta-analysis results (Figure 11) are as follows:

- Labeled: HR 0.91, 95% CI 0.47; 1.78
- Unlabeled: HR 1.12, 95% CI 1.01; 1.24
- Unclear[14]: HR 0.56, 95% CI 0.23; 1.39

The HR for death is not significantly different from 1.0 for labeled use of epoetin. For unlabeled use, the HR for death is greater than 1.0. A major limitation of this analysis is that data are scant from studies closely approximating labeled recommendations for epoetin use; as a

[14] Oberhoff, Neri, Amadori, et al., 1998; Kurz, Marth, Windbichler, et al., 1997; Throuvalas, Antonadou, Boufi, et al., 2000; the latter two studies have 0-1 events.

result, the 95 percent confidence interval is extremely wide and the labeled use subgroup cannot be statistically distinguished from unlabeled use. Two studies (Rose, Rai, Revicki, et al., 1994; P-174, 2004) used a value of Hb at which to stop epoetin administration only slightly higher (at or near 13.3 g/dL or hematocrit 40%) than the currently recommended 13 g/dL. Including these studies in the labeled use subgroup changes the HR to 1.08 (95 percent CI, 0.63; 1.84) but still accounts for only about 11 percent of the overall study weight and affords no clearer distinction between subgroups.

Because FDA considers a high Hb target a potential risk factor for greater mortality, and because trial protocols have tested various Hb values at which epoetin is discontinued ("stopping value"), we also conducted an analysis by Hb stopping value in 1 g/dL increments (Figure 12). By visualizing the data at different stopping points, this analysis asked whether the data form a continuum, or whether there is a discernable Hb cutoff value separating risk from no risk of increased mortality. The results (Table 27) show that for Hb stopping values above the labeled value of $\leq$13 g/dL, the HR point estimate tends to increase but differences among the subgroup HR point estimates are not statistically significant (p=0.11) and a trend analysis was also not significant (p=0.6709) Data are concentrated at stopping values >13 and $\leq$15 and there are no useful data at stopping points higher than 16.

Figure 11. Meta-Analysis of Data on Survival: Labeled vs. Unlabeled Criteria for Use in Trials Comparing Epoetin to Control

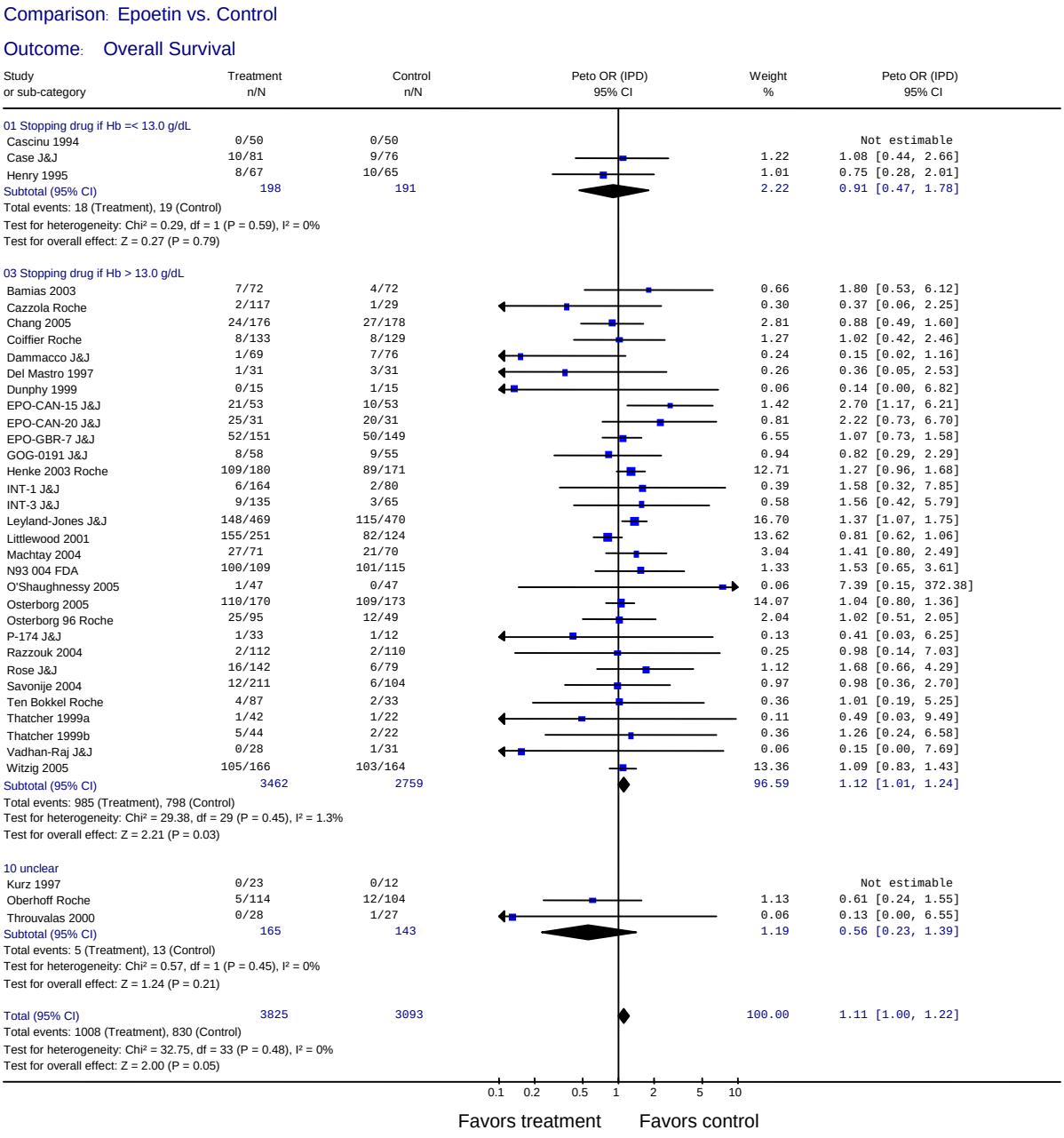


Figure 12. Meta-Analysis of Data on Survival by 1 g/dL Hb Unit Increments for Treatment Stopping Point in Trials Comparing Epoetin to Control

Comparison: Epoetin vs. Control

Outcome: Overall Survival

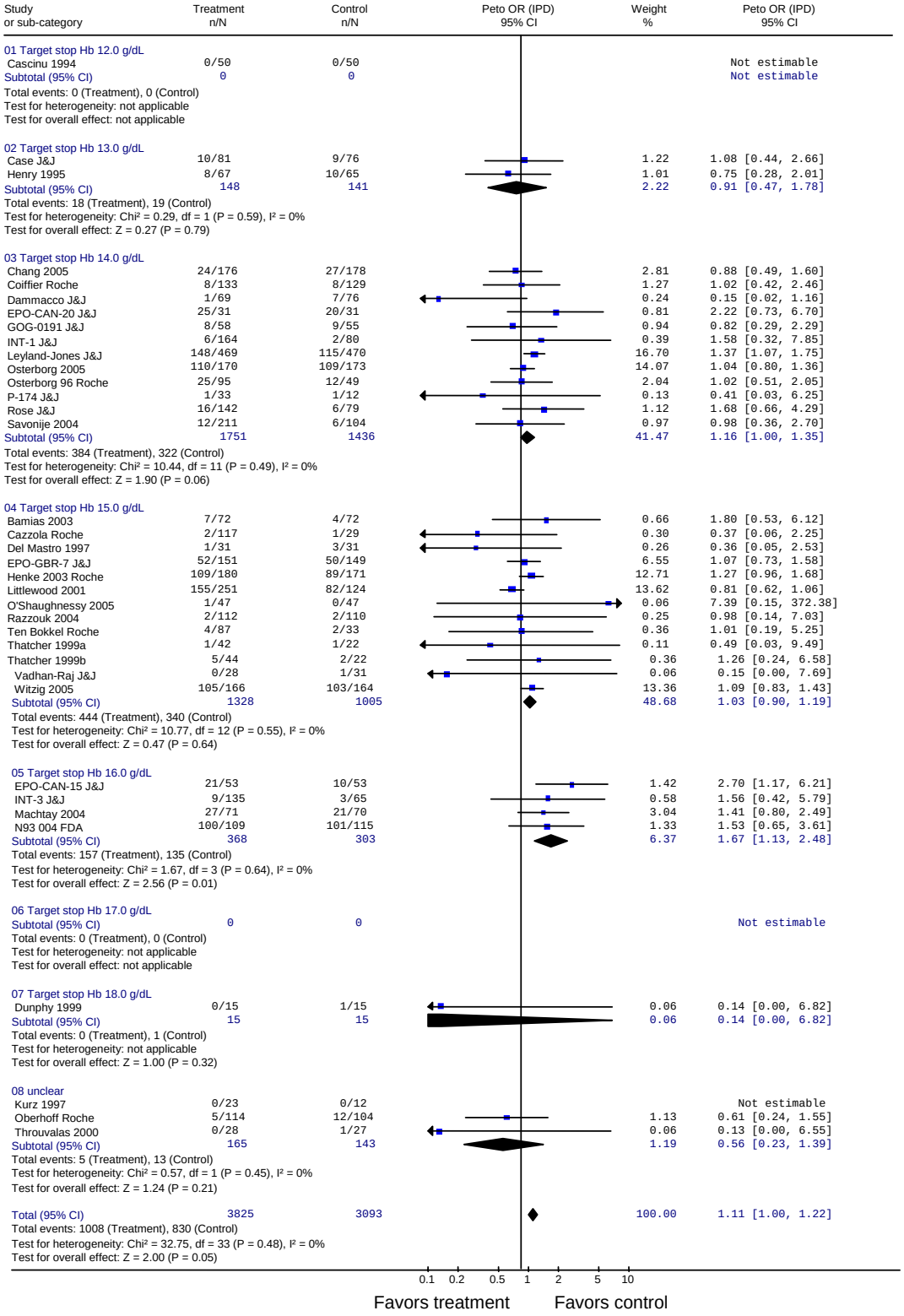


Table 27. Meta-Analysis of Hazard Ratio for Death by Hb Stopping Value in 1 g/dL Increments

Hb Stopping Value	#Treated Patients	#Control Patients	Hazard Ratio for Death	95% CI
≤ 12 g/dL	50	50	not estimable (0 events)	
>12 and ≤ 13 g/dL	148	141	0.91	0.47; 1.78
>13 and ≤ 14 g/dL	1751	1436	1.16	1.00; 1.35
>14 and ≤ 15 g/dL	1328	1005	1.03	0.90; 1.19
>15 and ≤ 16 g/dL	368	303	1.67	1.13; 2.48
>16 and ≤ 17 g/dL	0	0	(no studies)	
>17 and ≤ 18 g/dL	15	15	0 Tx events, 1 Ctl event	
(Unclear)	165	143		

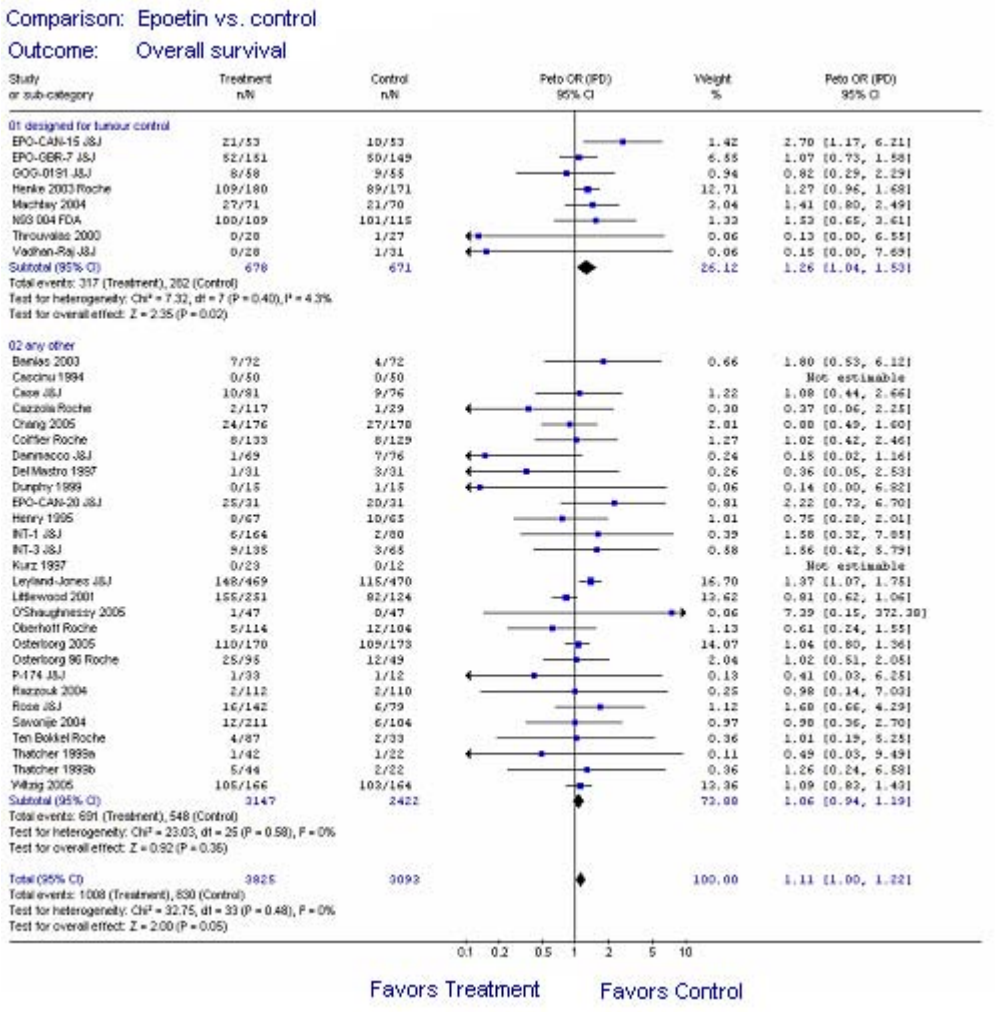
As noted, most studies included in our analyses of survival were not designed to evaluate survival as a primary outcome. The FDA Oncologic Drugs Advisory Committee identified study design factors of importance to test the effect of products on survival: enrolling patients with homogeneous primary tumor types and treatment regimens; sufficient duration of followup within the investigator-controlled course of the study; and sufficient patient numbers such that significant differences in survival, or in surrogate measures such as progression-free survival or tumor response can be detected (see Scope and Key Questions). We compared a subgroup of studies that met homogeneous tumor and treatment criteria, whether or not they were originally designed to evaluate survival outcomes, with the larger subgroup that did not (Figure 13).

Studies meeting these criteria suggest a statistically significant, detrimental effect of epoetin on survival while the pooled effect is not significant across those studies that do not meet criteria. However, the results for the two subgroups overlap considerably and cannot be clearly differentiated. Note that the Leyland-Jones (2003) study, while designed by the investigators to evaluate survival as the primary outcome, did not ensure homogeneous treatment regimens for malignancy, and thus the study does not meet criteria for homogeneous tumor type and treatment regimen.

Darbepoetin versus Control. Four trials (N=973; 583 randomized to darbepoetin, 390 randomized to control) reported survival (Hedenus, Hansen, Taylor, et al., 2002; Hedenus, Adriansson, San Miguel, et al., 2003; Kotasek, Steger, Faught, et al., 2003; Vansteenkiste, Pirker, Massuti, et al., 2002). However, for one study (Hedenus, Hansen, Taylor, et al., 2002) the hazard ratio could not be estimated because there were no events in either study arm. Characteristics of reporting studies are enumerated in Table 25.

Trials that reported survival differed with respect to several variables. Baseline characteristics of study populations differed by average baseline Hb concentration and type of malignancy. Treatment protocols differed by therapies for malignancy and iron supplementation. Trials also varied with respect to publication type and duration of followup. Two studies (Hedenus, Hansen, Taylor, et al., 2002; Kotasek, Steger, Faught, et al., 2003) were designed as dose-finding studies, but reported survival only for pooled treatment arms.

Figure 13. Meta-Analysis of Epoetin Trial Data on Survival: Studies Meeting Homogeneous Tumor and Treatment Criteria vs. Those that Did Not



Results. A test for heterogeneity across trials included for survival outcomes was strongly significant (p=0.03, I^2=72%). Therefore, a random-effects meta-analysis was also performed.

Meta-analysis of data from 4 trials (Figure 14) yielded:

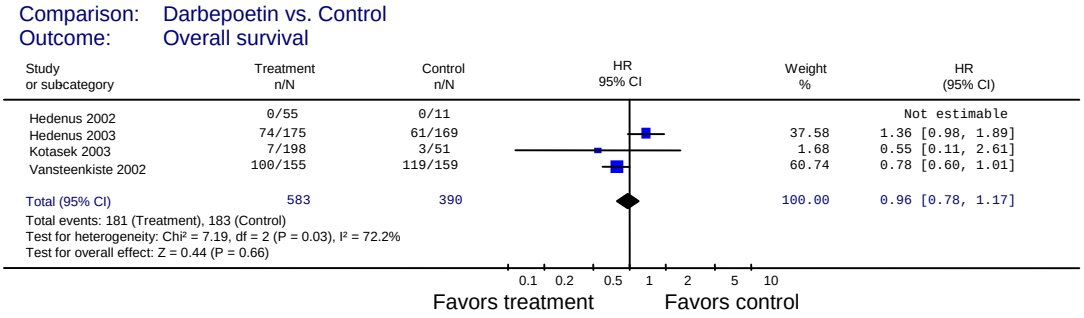
- Fixed-effects: HR 0.96 (95% CI 0.78; 1.17), p=0.66
- Random-effects: HR 0.97 (95% CI 0.59; 1.58), p=0.90
- Total event rates were 31% for epoetin treatment arms and 47% for control arms.
- Hazard ratios ranged from 0.55 (Kotasek, Steger, Faught, et al., 2003) to 1.36 (Hedenus, Adriansson, San Miguel, et al., 2003).

Our combined summary estimate of effect is nearly identical to the results of a recently published meta-analysis (Hedenus, Vansteenkiste, Kotasek et al., 2005), which included the

same four trials but likely had access to different data sources for some of the trials. They reported HR=0.95 (95% CI 0.78; 1.16).

No conclusion can be drawn from the limited evidence on the effect of darbepoetin on survival. In the two studies that contributed >98 percent of the weight to the meta-analysis, hazard ratio point estimates showed opposite effects but neither was significantly different from 1.0 (Hedenus, Adriansson, San Miguel, et al., 2003: HR 1.36, 95% CI 0.98; 1.89 and Vansteenkiste, Pirker, Massuti, et al., 2002: HR 0.78 95% CI 0.60; 1.01). The two dose-finding trials (Hedenus, Hansen, Taylor, et al., 2002; Kotasek, Steger, Faught, et al., 2003) contributed very little weight to the meta-analysis and thus did not influence the results. Too few trials were available for subgroup analyses to be meaningful.

Figure 14. Meta-Analysis of Data on Survival from 4 RCTs of Darbepoetin versus Control



Evidence Regarding the Class of Erythropoiesis-Stimulating Products

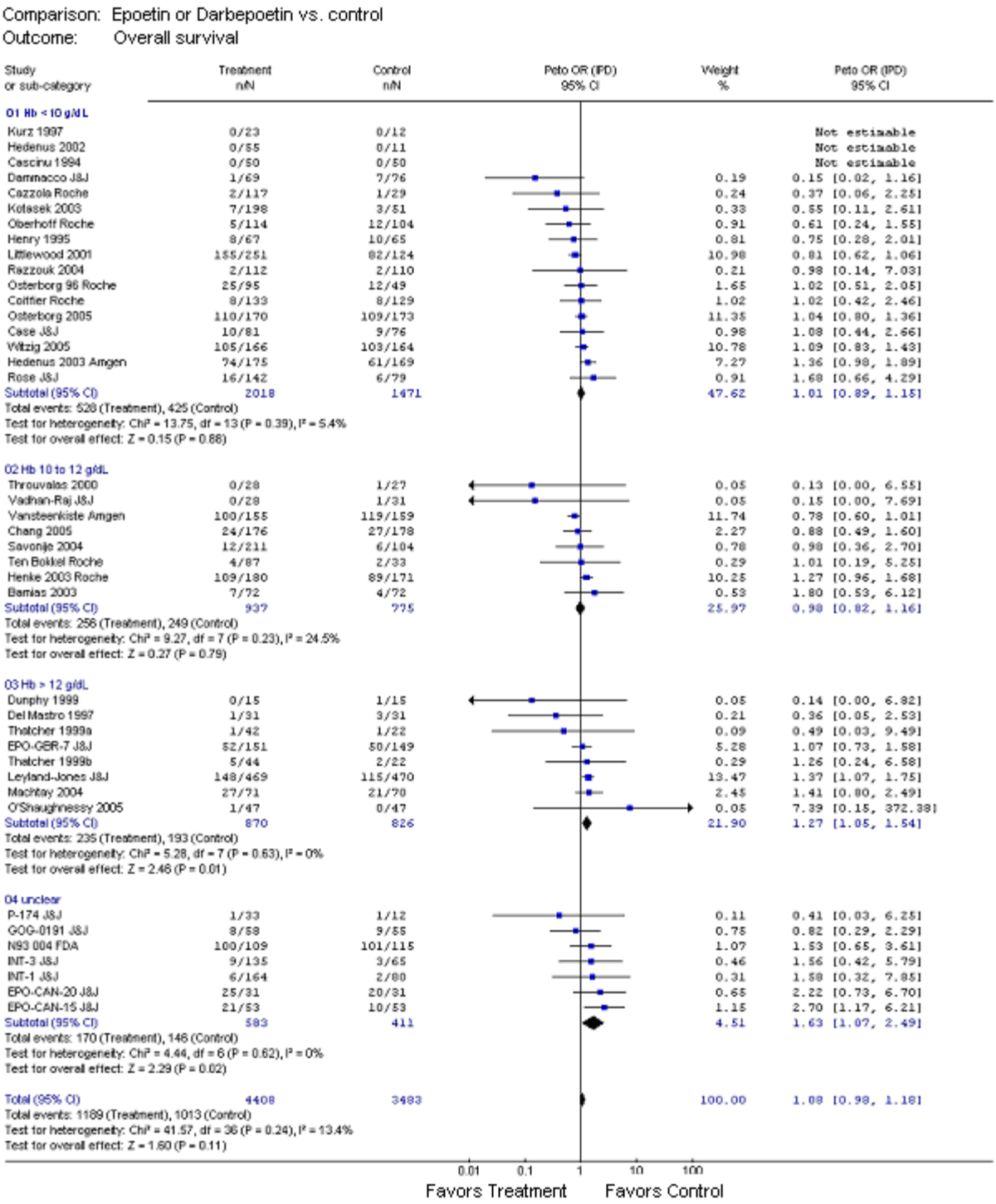
Combined Analysis of Epoetin versus Control and Darbepoetin versus Control. Erythropoiesis-stimulating products are considered to have similar pharmacodynamic properties and class effects (Food and Drug Administration Oncologic Drugs Advisory Committee Meeting Briefing Information, 2004); therefore we conducted a combined analysis of trials reporting survival outcomes for more robust results. When we combined studies of epoetin or darbepoetin versus control, the overall hazard ratio changed little in value and not at all in interpretation (Table 28 and Figure 15). While heterogeneity is high (I^2=72.2 percent) for darbepoetin vs. control because of few studies, heterogeneity is minimal (I^2=13.4 percent) for the combined analysis. Planned subgroup analyses were inconclusive due to lack of information from several studies.

Additional analyses of trials by labeled vs. unlabeled use, by target hemoglobin 1 g/dL increments, and by homogeneous tumor and treatment criteria were also conducted. In each case, combined results for epoetin and darbepoetin trials were similar to those for epoetin trials alone, as shown in Table 28. Notably, the combined subgroup of trials meeting homogeneous tumor and treatment criteria for analysis of survival outcomes showed less differentiation from the subgroup of trials not meeting those criteria.

Table 28. Overall Survival Meta-Analyses: Epoetin vs. Control; Darbepoetin vs. Control; and Epoetin vs. Control and Darbepoetin Combined

Parameter	Epoetin vs. control meta-analysis	Darbepoetin vs. control meta-analysis	Epoetin + darbepoetin combined meta-analysis
Number of studies	35	4	39
Patients analyzed	6,918	973	7,891
HR (95% CI)	1.11 (1.00; 1.22) p=0.05 I^2=0%	0.96 (0.78; 1.17) p=0.66 I^2=72.2%	1.08 (0.98; 1.18) p=0.11 I^2=13.4%
HR (95% CI) for subgroups:			
Labeled use	0.91 (0.47; 1.78)		0.91 (0.47; 1.78)
Unlabeled use	1.12 (1.01; 1.24)		1.09 (0.99; 1.19)
HR (95% CI) for subgroups:			
Max Hb target 12 g/dL	(no events)		(no events)
Max Hb target 13 g/dL	0.91 (0.47; 1.78)		0.91 (0.47; 1.78)
Max Hb target 14 g/dL	1.16 (1.00; 1.35)		1.16 (1.00; 1.35)
Max Hb target 15 g/dL	1.03 (0.90; 1.19)		1.01 (0.90; 1.13)
Max Hb target 16 g/dL	1.67 (1.13; 2.48)		1.67 (1.13; 2.48)
HR (95% CI) for subgroups:			
Homogeneous tumor + tx	1.26 (1.04; 1.53)		1.06 (0.91; 1.24)
Not homogeneous tumor + tx	1.06 (0.94; 1.19)		1.08 (0.97; 1.21)

Figure 15. Meta-Analysis of Data on Survival from 35 RCTs of Epoetin versus Control Combined with Four RCTs of Darbepoetin versus Control



KQ1 Outcome V. Tumor Response and Progression

Investigators have hypothesized opposite effects of epoetin or darbepoetin on malignancies. Some proposed that by improving tumor oxygenation, these drugs might enhance cytocidal effects of certain chemotherapy regimens and/or radiation therapy (e.g., Glaspy 2002). Alternatively, tumor cells with erythropoietin receptors (e.g., Westenfelder and Baranowski, 2000; Acs, Zhang, Rebbeck, et al., 2002) might proliferate and progress more rapidly if either drug is present. The first hypothesis suggests erythropoietic stimulants might increase tumor response rate to therapy, which could then increase survival. The second suggests they may decrease response duration or increase progression, which could then reduce survival. These are not mutually exclusive possibilities; but testing each hypothesis requires different outcomes that must be analyzed separately.

This report defines tumor response as the proportion of patients with a complete response (CR) to nonsurgical treatment of their malignancy (see Introduction/Scope; Appendix C Table C33). We focus on CR since for many malignancies, achieving CR is a prerequisite for long term survival without additional treatment. Several studies also reported overall response (OR), which is the sum of CR plus partial response (PR) rates. Studies were excluded unless prospectively designed to assess tumor response in a homogeneous population (i.e., one malignancy) given a protocol-specified cancer treatment regimen.

Outcomes related to response duration (e.g., time to progression, progression-free survival) were abstracted if available from studies that met the same selection criteria (one malignancy; protocol-specified regimen). They are summarized in Results (see below) and included in Appendix C tables, but cannot be pooled with tumor response for meta-analysis.

Darbepoetin versus Epoetin. Trials that directly compared darbepoetin versus epoetin did not report tumor response rate or duration-related outcomes.

Epoetin versus Control. Five trials (EPO GBR-07, 2004; Machtay, Pajak, Suntharalingam, et al., 2004; N93 004, 2004[15]; Throuvalas, Antonadou, Boufi, et al., 2000; Vadhan-Raj, Skibber, Crane, et al., 2004) reported tumor response rate as defined for this review (N=788 randomized, 688 evaluated; 344 from epoetin arms, 344 from control arms). Table 29 enumerates variables prespecified for subgroup analysis (Fig. 1) from these five trials. Two of these trials (EPO GBR-07, 2004; Machtay, Pajak, Suntharalingam, et al., 2004) plus three others (GOG-191, 2004; Henke, Laszig, Ruebe, et al., 2003; EPO-CAN-15, 2004) reported time to progression (TTP), progression-free survival (PFS), or disease-free survival (DFS). Table 30 lists noteworthy features of all eight studies that reported tumor response rate or a duration-related outcome.

Among trials that reported tumor response rate (Table 29), characteristics of study populations differed only by average baseline Hb concentration. Each trial reporting this outcome enrolled only adult patients with solid tumors. Three trials studied head and neck cancer, two each treated small cell lung cancer or gynecologic tumors, and the remaining trial investigated gastric and rectal tumors (Table 30). Treatment protocols differed by therapies for

[15] As this report was released, a full-text version of this trial was published (Grote, Yeilding, Castillo et al., 2005).

Table 29. Study Characteristics of RCTs Reporting Tumor Response Rates

Outcome Subgroup	Epoetin versus Control				
	# Studies	**# Total Patients**	**# Epo/# Ctl Patients**	**Relative Risk**	**95% CI (p-value)**
Tumor Response – HR	5	688	344/344	1.00	0.92; 1.10
(Heterogeneity)					0.94
Subgroups: Patient Baseline Characteristics					
Bsln Hb <10					
Bsln Hb 10-12	2	195	99/96		
Bsln Hb >12	3	493	245/248		
Bsln Hb ?					
(Group difference[1])					
Solid tumors	(all)	688	344/344		
Hematologic					
Mixed					
(Group difference[1])					
Children					
Adults	(all)	688	344/344		
(Group difference[1])					
Subgroups: Treatment Protocols					
Chemo, all plat	1	224	109/115		
Chemo, some plat					
Chemo, no plat					
Chemo, plat unknown					
Chemo+RT or RT	4	464	235/229		
Unknown					
(Group difference[1])					
Iron, fixed					
Iron, as needed	1	54	28/26		
Iron unknown	4	634	316/318		
(Group difference[1])					
Epo tx 6-9 weeks	2	195	99/96		
Epo tx 12-16 weeks	3	493	245/248		
Epo tx >20 weeks					
Epo tx ? Weeks					
(Group difference[1])					
Subgroups: Reporting and Quality					
High quality	2	274	135/139		
Low quality	3	414	209/205		
(Group difference[1])					
Data from full text					
Data from abstract					
Data unpublished	1	54	28/26		
Data from FDA	4	634	316/318		
(Group difference[1])					

[1] p value for differences among subgroup categories calculated by inverse variance method (see Methods/Data Extraction and Analysis/Statistical Data Analysis)

malignancy, iron supplementation and duration of epoetin treatment. Trials also varied with respect to publication type and overall quality rating. Epoetin dosage and dose adjustments also varied (Table 30). While five of the eight trials initiated epoetin treatment with FDA-recommended dosages, none conform to currently recommended dose adjustments or Hb targets.

Table 30. Features of Studies Reporting Tumor Response or Duration-Related Outcomes

STUDY: feature:	N93-004[1]	Throuvalas 2000	Machtay 2004	Vadhan-Raj 2004	EPO GBR-7	EPO CAN-15	GOG-0191	Henke 2003
Control N	115	26	70	22	111	53	55	171
EPO N	109	28	71	22	114	53	58	180
malignancy	SCLC (limited or extensive)	cervix or bladder	head&neck (no mets., unresected)	gastric or rectal	head&neck (stages I-IV)	SCLC (limited only)	cervix cancer	head&neck (stages III or IV)
Tx regimen	cisplatin + etoposide	Pt chemo + radioTx	chemo (?) + radioTx	5FU + radioTx	radioTx	Pt chemo + radioTx	Pt chemo + radioTx	adjuvant radioTx
Tx duration	NR	5-6 weeks	NR	NR	NR	NR	NR	6-7 weeks
outcome	CR, OR	CR	CR, PFS	CR	CR, OR, DFS	median TTP	PFS	PFS
when assessed	after last cycle	2-3 mos after Tx	12 mos median	NR	CR: wk 12 DFS: 3 yrs	NR	NR	~2 yrs
EPO dose	150 IU/kg 3X/wk	10,000 IU 5X/wk	150 IU/kg 3X/wk	40,000 IU/wk	10,000 IU 3X/wk	40,000 IU/wk	40,000 IU/wk	300 IU/kg 3X/wk
EPO duration	12 wks	5-6 wks	9-10 wks	16 wks	throughout radioTx	12-24 weeks	NR	throughout radioTx
baseline Hb (cont/EPO)	12.8/13.0 g/dL	11.1/11.5 g/dL	12.2/12.0 g/dL	13.0	13.4 g/dL	NR	NR	11.7/11.8 g/dL
Hb target, UL	16 g/dL	NR	14 g/dL (F) 16 g/dL (M)	15 g/dL	15 g/dL	16 g/dL	14 g/dL	14 g/dL (F) 15 g/dL (M)
re-start if Hb<	14 g/dL	NR	12.5 (F) 13.5 (M)	14 g/dL	12.5 g/dL	14 g/dL	13 g/dL	14 g/dL (F) 15 g/dL (M)

[1] As this report went to press, a full-text version of this trial was published (Grote, Yeilding, Castillo et al., 2005).

Results. Five of eight trials reported CR rate (the most frequently reported tumor response outcome); epoetin did not affect CR rate in any trial (Figure 16). Two studies reported OR rate, with no significant differences between epoetin and control arms (EPO-GBR-07; N93 004).

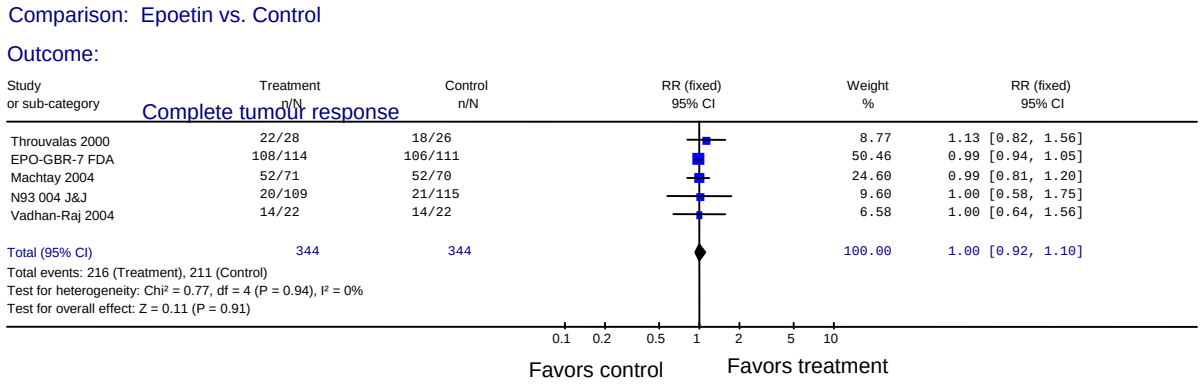
Relative risk (likelihood) to achieve CR ranged from 0.99 to 1.13 across reporting trials (EPO-GBR-7 FDA and Machtay, Pajak, Suntharalingam, et al., 2004; Throuvalas, Antonadou, Boufi, et al., 2000). Each 95 percent CI included 1.0. A test for heterogeneity across trials included for tumor response was not statistically significant ($p=0.94$, $I^2 = 0$ percent). An I^2 value of zero percent indicates no observed statistical heterogeneity, thus only a fixed-effects meta-analysis was done.

Fixed-effects meta-analysis of CR data from the five trials (Figure 16) yielded:

- relative risk (RR) = 1.00 (95% CI: 0.92, 1.10), p=0.91
- pooled CR rates (range by trial): epoetin arms, 63% (18% to 95%); control arms, 61% (18% to 96%)

Subgroup analyses were not done since the five trials were quite homogeneous for prespecified variables.

Figure 16. Meta-Analysis of Data on Relative Risk (Likelihood) of Achieving CR from Five RCTs of Epoetin versus Control



In one of five studies reporting outcomes related to response duration or tumor progression, with 351 head and neck cancer patients undergoing radiation therapy alone, locoregional PFS was significantly worse among those randomized to epoetin than among controls (RR = 1.62; 95 percent CI: 1.22, 2.14; p=0.0008; Henke, Laszig, Ruebe, et al., 2003). The other four studies reported no significant differences between arms in PFS (Machtay, Pajak, Suntharalingam, et al., 2004; GOG 0191), median TTP (EPO-CAN-15), or DFS (EPO-GBR-7). However, these trials likely lacked adequate statistical power to detect a difference, since only one randomized >100 patients per arm (EPO GBR-7). Additionally, three of the four (GOG 0191; EPO-CAN-15; EPO-GBR-7) closed before meeting accrual targets, due to excess thromboembolic events and following reports from Henke, Laszig, Ruebe, et al. (2003) and Leyland-Jones (2003) that survival decreased relative to controls in epoetin arms. Note also that FDA labeling for epoetin products comments on the Leyland-Jones trial as follows: "At four months, death attributed to disease progression also was higher (6% vs. 3%) in women receiving Epoetin alfa."

Darbepoetin versus Control. No trials comparing darbepoetin versus control reported CR or OR rates. One trial reported PFS and the proportion of patients whose tumors progressed (Vansteenkiste, Pirker, Massuti, et al., 2002; N=320 randomized; 314 evaluated; 155 from darbepoetin arm, 159 from control arm; see Appendix C Table C34).

This trial enrolled adult patients with small cell and non-small cell lung cancers, whose mean baseline Hb was just above 10 g/dL; used platinum-based chemotherapy for all patients, and administered darbepoetin at the labeled dose of 2.25 mcg/kg per week for 12 weeks, but did not conform to current recommendations for dose adjustments. Darbepoetin was discontinued if Hb rose above 15 g/dL for males or 14 g/dL for females, and was reinstated (at half the dose) if Hb fell below 13 g/dL for either sex. The trial did not report on iron use; and was rated a high-quality study, published in full text, and updated at the May 2004 ODAC meeting.

Results. PFS over 24 months' followup reportedly did not differ significantly between arms (HR = 0.81; 95 percent CI: 0.64, 1.03). Cox proportional hazards analysis reportedly showed less frequent tumor progression over 12 months median followup in the darbepoetin than in the control arm (HR = 0.70; 95 percent CI: 0.53, 0.92).

KQ1 Outcome VI. Thromboembolic Events

Thromboembolic events were not well defined in the reports of included trials; definitions in general did not appear to be prespecified.[16] Studies usually did not provide a detailed definition of thromboembolic events. Most studies did not provide information on severity of reported events. Only 10 studies reported detailed lists of thromboembolic events (Razzouk, Hockenberry, Hinds, et al 2003; Rosenzweig, Bender, Lucke, et al., 2004; Henke, Laszig, Ruebe, et al., 2003; Witzig, Silberstein, Loprinzi, et al., 2005; Ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; EPO GBR-07 2004; GOG-191 2004; N93-004 2004; EPO-CAN-15 2004; Vadhan-Raj, Skibber, Crane, et al., 2004). Given these difficulties, we required neither grade nor elaboration of different types of thromboembolic events for inclusion in the analysis. Events for this review included: thrombosis or related complications such as transient ischemic attacks, stroke, pulmonary embolism or myocardial infarction. However, given the lack of detailed reporting, it was not possible to quantify the frequency of specific thromboembolic events.

Discrepancies among data for the same study from different sources also posed a problem. Twelve of the 30 studies of epoetin vs. control evaluated for thromboembolic complications and contributing 72.7 percent of the weight to the overall analysis were reported in two or more documents (e.g. abstracts, full publications, FDA reviewer documents and reports submitted by the pharmaceutical companies for the FDA ODAC hearing in May 2004) and thromboembolic event data did not agree. The discrepancies were resolved for Henke, Laszig, Ruebe, et al., 2003 (the journal publication reported hypertension and thromboembolic events together whereas the Roche FDA ODAC document reported the events separately) and for Leyland-Jones 2003 (the journal publication reported thromboembolic events during the first four months, the FDA reviewer summary listed deaths following thromboembolic event during the first 4 months, and clinically relevant events were reported in the J&J FDA ODAC document/slides; the latter was chosen for the analysis). For the other 10 studies (EPO-CAN-20 2004; EPO-GBR-07, 2004; GOG-191, 2004; Machtay, Pajak, Suntharalingam, et al., 2004; N93-004 2004; Witzig, Silberstein, Loprinzi, et al., 2005; Vadhan-Raj, Skibber, Crane, et al., 2004; Littlewood, Bajetta, Nortier, et al., 2001; Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995) it was not possible to resolve the data discrepancies. For these studies, we employed a predefined rule: the most complete data set (largest sample size) OR data with consistent outcome definitions across trials were chosen for analysis.

Evidence for Comparative Safety

Darbepoetin versus Epoetin. Characteristics of reporting studies are enumerated in Table 31. Three RCTs directly compared thromboembolic event rates after darbepoetin or epoetin treatment (N = 1,879; 948 to darbepoetin, 931 to epoetin) (Glaspy, Berg, Tomita, et al., 2005;

[16] For example, Johnson & Johnson applied the following definition in their document prepared for the FDA ODAC hearing (Food and Drug Administration Oncologic Drugs Advisory Committee. May 4, 2004, Meeting Briefing Information) and from which several sets of study data were abstracted: "The list of general TVEs [thrombovascular {i.e., thromboembolic} event] is the Sponsor's broadest approach for identifying TVEs, and includes all superficial TVEs, all catheter related TVEs and events that could but not necessarily would, be caused by an underlying thrombovascular event and where no information was available to prove the contrary. General TVEs are also subclassified as clinically relevant, a definition that is broader than the generally accepted clinically important TVEs (e.g. DVT, PE, stroke/TIA, and MI)." We found no consistent definitions for data abstracted from the Roche FDA ODAC hearing document or hearing documents prepared by FDA reviewers.

Table 31. Characteristics and Subgroup Analyses of RCTs Reporting Thromboembolic Events

Outcome Subgroup	Darbepoetin versus Epoetin					Epoetin versus Control				
	# Studies	#Total Patients	#Darb/ #Epo Patients	Point Estimate	95% CI (p-value)	# Studies	#Total Patients	#Epo/#Ctl Patients	Point Estimate	95% CI (p-value)
Thromboembolism – RR	3	1,879	948/931	0.86	0.61; 1.21	30	6,092	3,355/2,737	1.69	1.36; 2.10
(Heterogeneity)					(0.98)					(0.67)
Subgroup Analyses: Patient Baseline Characteristics										
Bsln Hb <10						10	2,172	1,205/967	1.53	0.98; 2.39
Bsln Hb 10-12	(all)					7	1,394	782/612	1.78	1.12; 2.83
Bsln Hb >12[1]						5	1,505	771/734	1.71	1.08; 2.70
Bsln Hb unclear[1]						8	1,021	597/424	1.74	1.18; 2.56
(Group difference[2])										(0.93)
Solid tumors	2	670	337/333			20	4,108	2,200/1,908	1.70	1.33; 2.16
Hematologic						5	898	509/389	3.00	1.10; 8.12
Mixed/unknown	1	1,209	611/598			5	1,086	646/440	1.33	0.76; 2.32
(Group difference[2])										(0.34)
Subgroup Analyses: Treatment Protocols										
Children										
Adults	(all)					(all)				
(Group difference[2])										
Chemo, all plat						9	1,439	861/578	1.15	0.77; 1.71
Chemo, some plat	(all)					2	478	237/241	2.02	0.83; 4.89
Chemo, no plat						7	2,494	1,362/1,132	1.46	1.04; 2.05
Chemo, plat unknown										
Chemo+RT or RT						8	1,187	601/586	3.00	1.77; 5.10
Unknown						4	494	294/200	3.99	1.28; 12.41
(Group difference[2])										**(0.036)**
Iron, fixed						1	333	168/165	1.47	0.54; 4.05
Iron, as needed						14	2,730	1,513/1,217	1.56	1.09; 2.23
Iron unknown	(all)					15	3,029	1,674/1,355	1.80	1.35; 2.39
(Group difference[2])										(0.97)
Epo tx 6-9 weeks						4	646	329/317	1.91	0.78; 4.64
Epo tx 12-16 weeks	(all)					15	2,836	1,546/1,290	1.48	1.11; 1.98
Epo tx >20 weeks						8	1,953	1,107/846	1.85	1.26; 2.72
Epo tx ? Weeks						3	657	373/284	2.89	1.11; 7.55
(Group difference[2])										(0.43)
Subgroup Analyses: Reporting and Quality										
High quality						18	4,224	2,292/1,932	1.55	1.21;1.99
Low quality	(all)					12	1,868	1,063/805	2.18	1.38; 3.44

(Group difference[2])										(0.21)
Data from full text	1	312	157/155			9	1,388	764/624	1.73	1.01; 2.95
Data from abstract	2	1,567	791/776			4	732	422/310	3.61	1.21; 10.74
Data unpublished										
Data from FDA						17	3,972	2,169/1,803	1.59	1.25; 2.03
(Group difference[2])										(0.29)

[1]The N93-004 trial was published in full in December, 2005 (Grote, Yeilding, Castillo, et al., 2005) and included information on baseline Hb which classified it into subgroup Hb >12. A re-categorized analysis resulted in subgroup Hb>12 HR 1.36 (95% CI, 0.97; 1.89); the p-value for the group difference changed from 0.93 to 0.1381. Because this did not alter the interpretation of results, we did not alter our presentation of the overall analysis.
[2]p-value for differences among subgroup categories calculated by inverse variance method (see Methods/ Data Extraction and Analysis/ Statistical Data Analysis).

Schwartzberg, Yee, Senecal, et al., 2004; Waltzman, Croot, Williams, 2005). Patients varied only by type of malignancy across studies; treatment protocols did not differ. Trials varied with respect to type of publication.

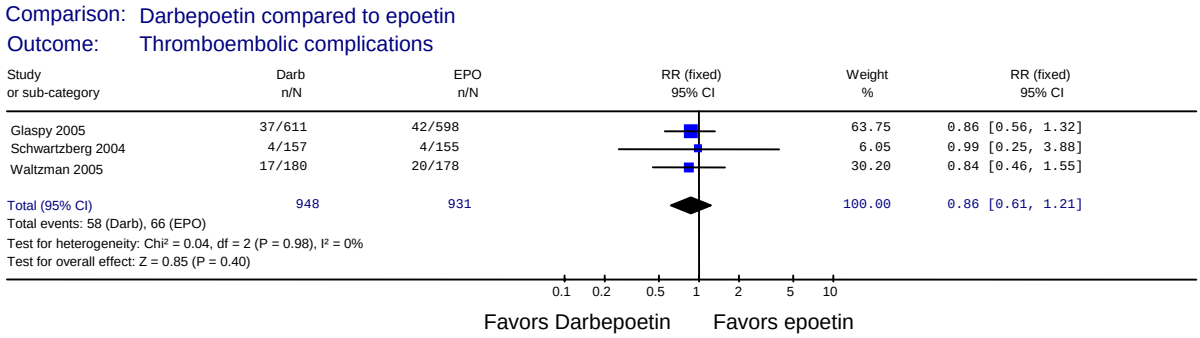
Results. No single trial reported a statistically significant difference in thromboembolic events between epoetin and darbepoetin trial arms. A test for heterogeneity across included trials for thromboembolic events was not statistically significant (p=0.98, I^2=0 percent). An I^2 value of 0 percent indicates no observed statistical heterogeneity.

Fixed-effects meta-analysis of data from these studies (Figure 17) showed:

- Relative risk (RR) = 0.86 (darbepoetin to epoetin; 95 percent CI 0.61; 1.21), p=0.40
- Pooled event rates (ranges across trials): darbepoetin, 6.1 percent (2.6 percent to 9.4 percent); epoetin, 7.1 percent (2.6 percent to 11.2 percent)
- RRs ranged from 0.84 to 0.99.

Pooled analysis did not show evidence of a statistically significant difference in rates of thromboembolic events for epoetin vs. darbepoetin. Subgroup analyses were not done since differences between trials were minimal. Given limited direct evidence from only three trials, indirect evidence (epoetin vs. control, darbepoetin vs. control) was evaluated for effect on thromboembolic events.

Figure 17. Meta-Analysis of Data on Thromboembolic Event Rates from Three RCTs of Darbepoetin versus Epoetin



Epoetin versus Control. Characteristics of reporting studies are enumerated in Table 31. Thirty RCTs (N=6,092; 3,355 to epoetin, 2,737 to control) reported thromboembolic events (Bamias, Aravantinos, Kalofonos, et al., 2003; Cascinu, Fedeli, Del Ferro, et al., 1994; Case, Bukowski, Carey, et al., 1993; Chang, Couture, Young, et al., 2005; Dammacco, Castoldi, Rodjer, et al., 2001; EPO-CAN-15, 2004; EPO-CAN-20, 2004; EPO-GBR-07, 2004; GOG-191, 2004; Henke, Laszig, Ruebe, et al., 2003; Henry, Brooks, Case, et al., 1995; Leyland-Jones, 2003; Littlewood, Bajetta, Nortier, et al., 2001; Machtay, Pajak, Suntharalingam, et al., 2004; N93 004, 2004; Osterborg, Boogaerts, Cimino, et al., 1996; Osterborg, Brandberg, Molostova, et al., 2002; Razzouk, Hockenberry, Hinds, et al., 2004; Rose, Rai, Revicki, et al., 1994; Rosenzweig, Bender, Lucke, et al., 2004; Savonije, Van Groeningen, Van Bochove, et al., 2004; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Thatcher, De Campos, Bell, et al., 1999; Throuvalas, Antonadou, Boufi, et al., 2000; Vadhan-Raj, Skibber, Crane, et al., 2004; Welch, James, Wilkinson, 1995; Witzig, Silberstein, Loprinzi, et al., 2005; EPO-INT-1, 2004; EPO-INT-3, 2004; P-174, 2004).

Trials that reported thromboembolic events differed with respect to several variables prespecified for subgroup analysis. Baseline characteristics of study populations differed by average baseline Hb concentration and type of malignancy. Treatment protocols differed by therapies for malignancy, iron supplementation, and epoetin treatment duration. Trials also varied with respect to publication type and overall quality rating.

Results. Although most trials (25 of 33 comparisons[17]; see Figure 18) reported thromboembolic events in a larger proportion of patients randomized to epoetin than of controls, only one trial reported a statistically significant increase in relative risk (EPO-CAN-15 FDA report; RR=8.00 favoring controls; 95% CI: 1.93, 33.09). A test for heterogeneity across included trials was not statistically significant (p=0.67, I^2=0%).

Fixed-effects meta-analysis of data from all 30 RCTs (Figure 18) yielded:

[17] Three RCTs compared two arms given different epoetin doses (ten Bokkel 1998; Thatcher 1999) or a fixed versus a titrated dosing regimen (Osterborg 1996) against one control arm per study. Together, these studies contributed N=394 (7.1%) to the total number of evaluated patients. For the meta-analysis, each control arm was split artificially and randomly into two groups, each entered with one experimental arm as a separate study. As this might influence weighting of the studies, the analysis was repeated with both experimental arms of each study merged and compared to that study's full control arm. The original (unmerged) result (RR = 1.69; 95% CI: 1.36, 2.10) was nearly identical to the result using merged experimental arms (RR = 1.70; 95% CI: 1.37, 2.12).

- RR = 1.69 (95% CI: 1.36, 2.10), p<0.00001
- Pooled event rates (range by trial): epoetin, 6.5% (0 to 30%); control, 4.1% (0 to 22.6%)
- RR for a thromboembolic event ranged from 0.33 to 5.5 with extreme values of 8.0 and 8.4.

RR was not estimable in three small trials because no events occurred in either arm (Cascinu 1994; P-174 J&J; Thatcher 1999a). Pooled results indicate that thromboembolic events are statistically significantly more likely to occur in patients administered epoetin than controls.

We calculated number needed to harm (NNH; Table 32) from the meta-analytic point estimate, which depends on baseline risk of thromboembolic event in untreated controls. Baseline risk is influenced by: tumor type, extent of cancer, treatment regimen, extrinsic factors (e.g., surgery, immobilization), and prior history. Data from Figure 18 showed that event rates in control arms of included RCTs ranged from zero (reported from 11 RCTs; next lowest rate was 0.67%) to 22.6% (next highest rate was 12.31%). NNH ranged from 7 to 58 for baseline risk values of 20% to 2.5%. Thus, at a baseline thromboembolic event risk of 2.5%, one additional thromboembolic event would occur in every 58 patients treated; at a baseline risk of 20%, one additional thromboembolic event would occur in every seven patients.

Table 32. Number of Patients that Must Be Treated with Epoetin to Cause One Extra Thromboembolic Event, as a Function of Baseline Event Risk

Baseline Risk[1]	NNH	lower limit 95% CI	upper limit 95% CI
2.5%	58	36	111
5%	29	18	56
10%	15	9	28
20%`	7	5	14

[1] To put baseline risk in clinical context, we used a recent review on thrombosis and cancer (Levine, Lee and Kakkar, 2005). The review tabulated data on thrombosis incidence reported from published studies (mostly case series), but did not include confidence intervals. The following table summarizes these findings by incidence range:

Footnote Table. Thrombosis incidence in various malignancies

Incidence Range (%)	Malignancies (Regimens)
<2.5%	Early stage breast cancer (without chemotherapy)
2.5% to ≤5%	Early stage breast cancer (e.g. FAC, CMF); cervix cancer (cisplatin + radiation); lung cancer (not specified);
5% to ≤10%	Early stage breast cancer (CMFVP); lymphoma (not specified); germ cell tumors (not specified)
10% to ≥20%	ovarian (not specified); malignant glioma (not specified)

Abbreviations: CMF(VP) = cyclophosphamide, methotrexate, fluorouracil, (vincristine, prednisone); FAC = fluorouracil, doxorubicin, cyclophosphamide

Univariate subgroup analyses resulted in RR point estimates that were greater than 1.0 (i.e., increased risk in the epoetin arms) for every subgroup evaluated and that in most cases were statistically significant. Cancer treatment regimen was the only statistically significant predictor of a thromboembolic event from epoetin treatment (p=0.0361, Table 31). Trials with regimens including radiation therapy (RR=3.00; 95% CI: 1.77, 5.10), and those that did not report the type

of regimens utilized (RR=3.99; 95% CI: 1.28, 12.41), had the largest increases in relative risk. However, it is uncertain whether this finding is clinically meaningful or a result of confounding by other factors such as tumor type.

Additional Analyses. As for survival outcomes, additional analyses not anticipated in the original protocol were conducted to answer specific questions or explore new hypotheses. We conducted an influence analysis to identify those studies that most strongly influenced the pooled RR for thromboembolic events. We conducted a subgroup analysis of those studies that administered epoetin according to current FDA-recommended ("labeled") criteria vs. those studies that used criteria exceeding the labeled limits of dose or target Hb value. We also compared RR for thromboembolic events among subgroups defined by 1 g/dL increments in maximum Hb target value (Table 33).

Figure 18. Meta-Analysis of Relative Risk of Thromboembolic Events from RCTs of Epoetin versus Control

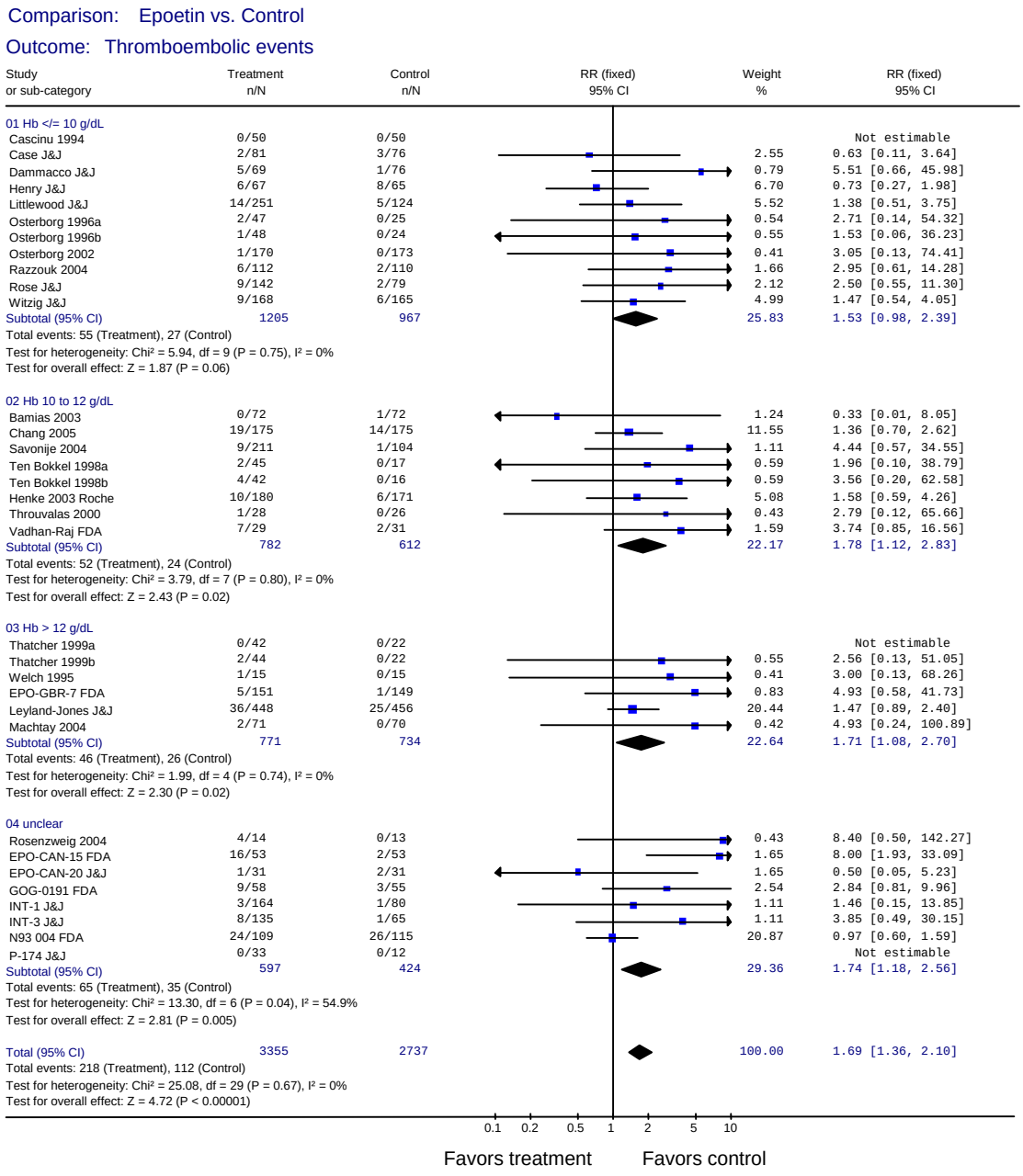


Table 33. Meta-Analysis of Risk Ratio for Thromboembolic Event by Hb Stopping Value in 1 g/dL Increments

Hb Stopping Value	#Treated Patients	#Control Patients	Risk Ratio for Thromboembolic Event	95% CI
≤ 12 g/dL	50	50	not estimable (0 events)	
>12 and ≤13 g/dL	148	141	0.70	0.29, 1.67
>13 and ≤14 g/dL	1,596	1,290	1.71	1.23, 2.40
>14 and ≤15 g/dL	1,151	914	1.92	1.22, 3.02
>15 and ≤16 g/dL	368	303	1.66	1.08, 2.54
>16 and ≤17 g/dL	0	0	(no studies)	
>17 and ≤18 g/dL	0	0	(no studies)	
(Unclear)	42	39	5.59	0.71, 43.94

The results of the influence analysis, in which each study is omitted, one at a time, and the remaining studies are pooled, are shown in Figure 19. The two studies most strongly influencing the meta-analysis are EPO-CAN-15 (2004) and N93-004 (2004). Interestingly, both studies enrolled patients with small cell lung cancer, used standard epoetin doses, and targeted a Hb value of 16 g/dL, but each study influenced the meta-analysis in the opposite direction. However, summary point estimates are not markedly changed by omission of either study, and remain statistically significant.

Three studies (Cascinu, Fedeli, Del Ferro, et al., 1994; Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995), constituting 6.4% of all patients evaluated for thromboembolic events, that most closely met current labeled criteria for use were compared to all other trials in a subgroup analysis. These studies used labeled (Case, Bukowski, Carey, et al., 1993; Henry, Brooks, Case, et al., 1995) or slightly lower epoetin doses (Cascinu, Fedeli, Del Ferro, et al., 1994) and stopped administration when Hb reached 13 g/dL, as recommended on the product label. Dose reduction strategies were slightly different from labeled recommendations.

For this subgroup analysis there was no evidence of heterogeneity within subgroups or overall (I^2 =0%). Subgroup meta-analysis results (Figure 20) are as follows:

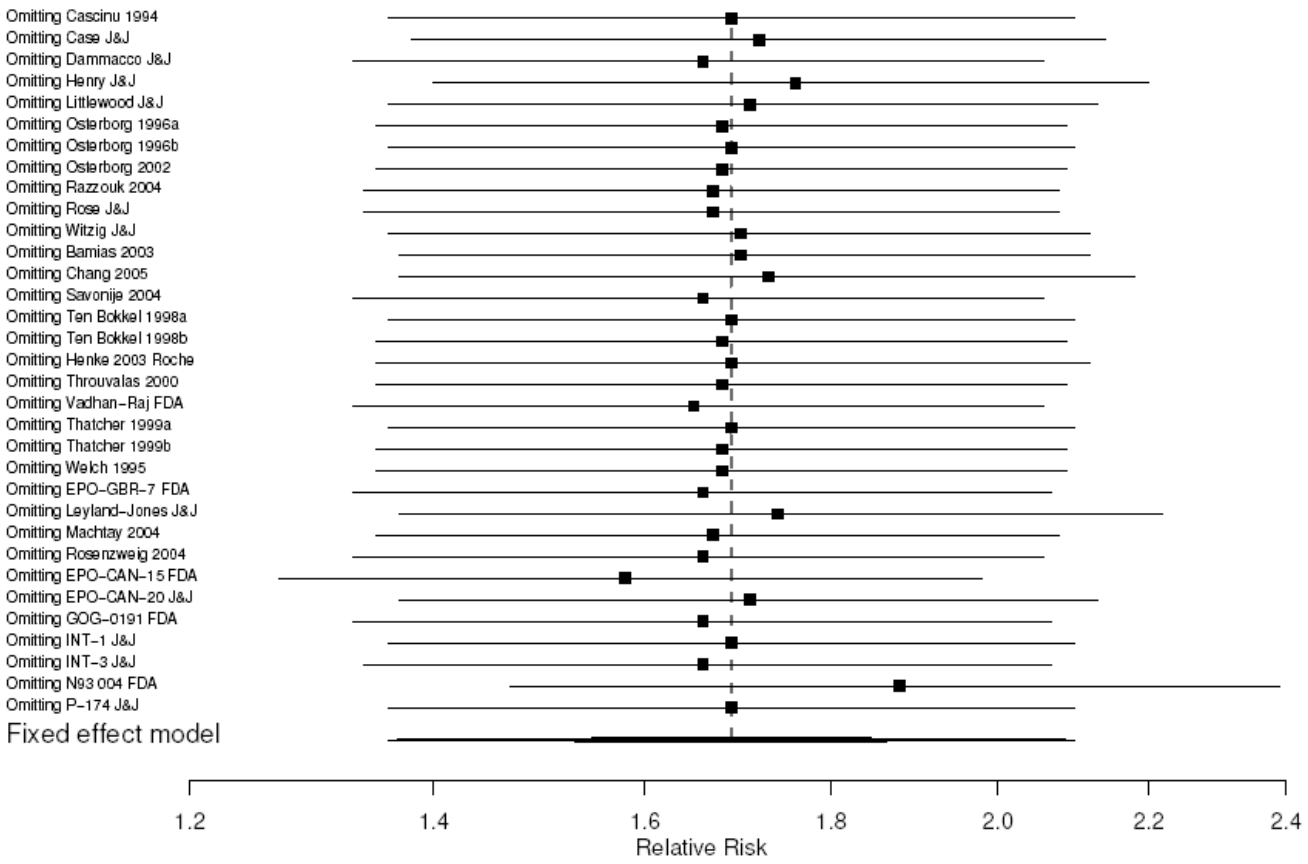
- Labeled: RR 0.70, 95% CI 0.29; 1.67
- Unlabeled: RR 1.75, 95% CI 1.40; 2.20
- Unclear[18]: RR 5.59, 95% CI 0.71; 43.9

The labeled and unlabeled groups differ significantly from each other (p=0.046), consistent with the explanation that targeting higher than recommended Hb values increases thromboembolic event risk. However, given the small number of studies and patients comprising the labeled group (3 studies, N=389) versus the unlabeled group (25 studies, N=5,622), these results could be confounded by other characteristics that affect risk, such as tumor type or treatment regimen.

Visualizing the data by 1 g/dL increments in upper limit of target Hb (Figure 21) again suggests that beyond the labeled target of 13 g/dL, thromboembolic event risk is greater, but

[18] Rosenzweig, Bender, Lucke, et al. (2004) and Throuvalas, Antonadou, Boufi et al. (2000) have 0-4 events per arm and together contribute only 1.7% of the total weight to the analysis.

Figure 19. Influence Analysis: Relative Risk for Thromboembolic Event Recalculated after Omission of One Study at a Time; Point Estimates (Squares) and 95% Confidence Intervals (Lines)



there is no clear relationship between increasing Hb and increasing risk and the trend is not statistically significant (p=0.742). Limitations to this representation are similar in that studies targeting 13.0 g/dL or less are few and results may be confounded by other factors.

Darbepoetin versus Control. Only one trial compared darbepoetin versus control and reported the proportion of participants with a thromboembolic event (Vansteenkiste, Pirker, Massuti, et al., 2002; n=320 randomized; 314 evaluated; 155 from darbepoetin arm, 159 from control arm). This trial enrolled adult patients with solid tumors whose mean baseline Hb was just above 10 g/dL; used platinum-based chemotherapy for all patients; administered darbepoetin for 12 weeks, but did not report on iron use; and was rated a high-quality study, published in full text, and updated at the May, 2004 ODAC meeting.

Results. The point estimate was not statistically significant for an increased relative risk of thromboembolism (RR = 1.44; 95% CI: 0.47, 4.43). Reported event rates were 4.5% in the darbepoetin arm and 3.1% in controls.

Evidence Regarding the Class of Erythropoiesis-Stimulating Products

Combined Analysis of Epoetin versus Control and Darbepoetin versus Control. Erythropoiesis-stimulating products are considered to have similar pharmacodynamic properties when used at recommended doses (Food and Drug Administration Oncologic Drugs Advisory Committee Meeting Briefing Information, 2004); therefore we conducted a combined analysis of trials reporting thromboembolic events for more robust results. However, because there is only one trial of darbepoetin vs. control, the result changed little (RR, 1.68; 95% CI: 1.36, 2.08) and the additional influence analysis and analysis by 1 g/dL Hb increments are not presented.

Figure 20. Meta-Analysis of Data on Thromboembolic Events: Labeled versus Unlabeled Criteria for Use in Trials Comparing Epoetin to Control

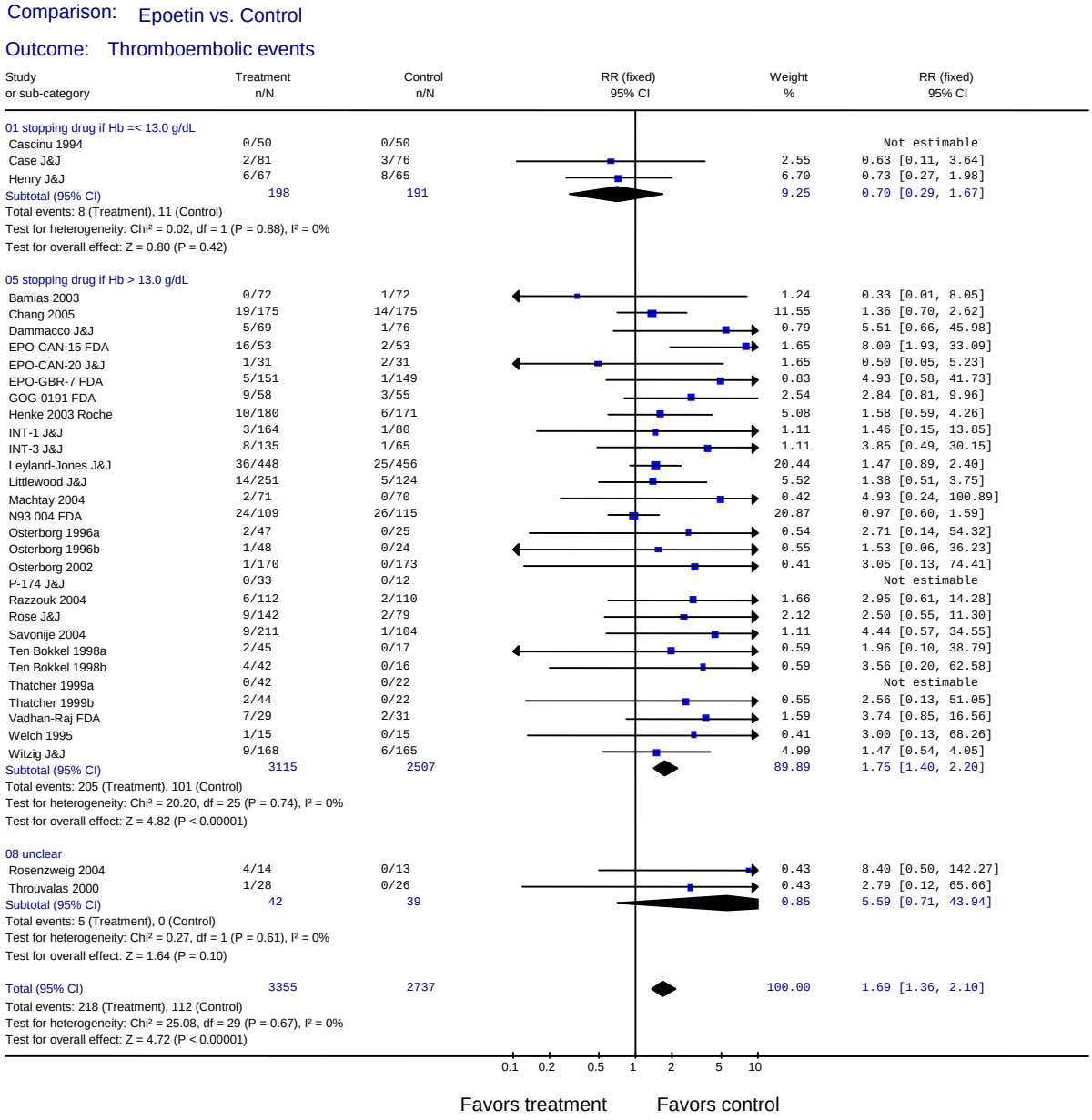
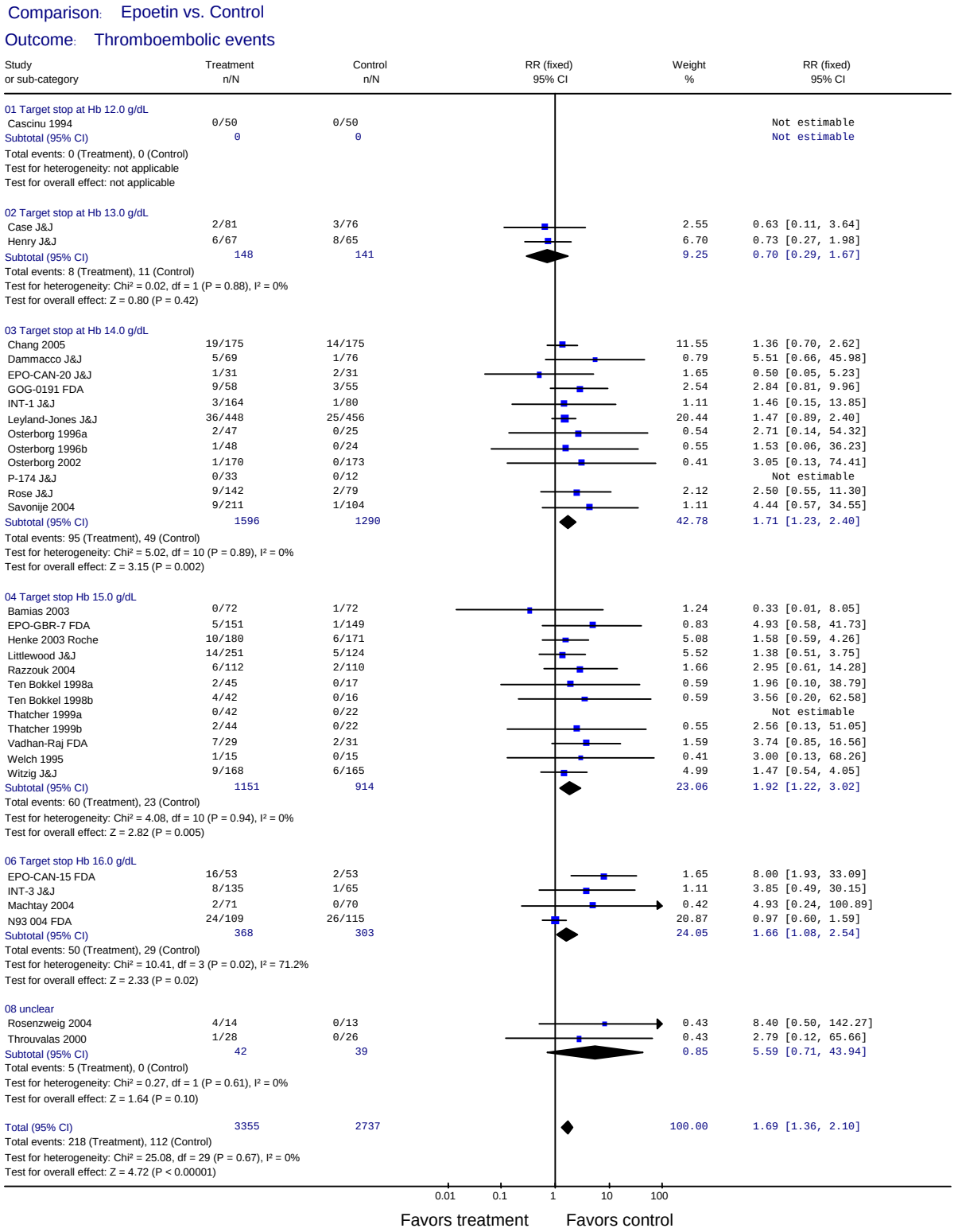


Figure 21. Meta-Analysis of Data on Survival by 1 g/dL Hb Unit Increments for Treatment Stopping Point in Trials Comparing Epoetin to Control



KQ1 Outcome VII. Other Adverse Events

Adverse events other than thromboembolism reported separately by study arm from multiple RCTs include: hypertension (16 trials), thrombocytopenia and/or hemorrhage (nine trials), rash (six trials), and seizures (three trials). Also summarized here are published data from RCTs on development of antibodies to epoetin or darbepoetin that might neutralize natural erythropoietin.

Darbepoetin versus Epoetin

One direct comparative study (Glaspy, Jadeja, Justice, et al., 2003) reported there were no seizures in either study arm. No other trials that directly compared darbepoetin versus epoetin reported rates of these adverse events separately by study arm.

Antibodies. Three trials that directly compared darbepoetin versus epoetin tested for antibodies to either product (Schwartzberg, Yee, Senecal, et al., 2004; Glaspy, Berg, Tomita, et al., 2005; Glaspy, Jadeja, Justice, et al., 2003). Another comparative RCT only tested for antibodies to darbepoetin (Glaspy, Jadeja, Justice et al., 2002). Antibodies were not detected in any patients.

Epoetin versus Control

FDA-approved Prescribing Information.

Hypertension, thrombocytopenia/hemorrhage, rash and seizures were not included in tables listing adverse experiences that occurred in >10 percent of patients from either arm of FDA-reviewed trials with cancer patients on chemotherapy. Sections on Information for Patients with cancer on chemotherapy note that "Hypertension, associated with a significant increase in hemoglobin, has been noted rarely in patients treated with…" Epogen® or Procrit®. While these sections do not estimate the frequency of hypertension, they recommend that blood pressure "…should be monitored carefully, particularly in patients with an underlying history of hypertension or cardiovascular disease." These sections also note that seizures occurred in 3.2 percent of those treated with the thrice-weekly regimen in double blind, placebo-controlled trials reviewed by FDA, and in 2.9 percent of placebo-treated controls. In similar trials using the weekly dosing regimen, seizures occurred in 1.2 percent of those given Epogen® or Procrit® and 1 percent of placebo-treated controls.

Evidence from Published Trials.

Table 34 summarizes available evidence and overall results for adverse events other than thromboembolism reported by multiple RCTs. Since heterogeneity was not statistically significant (i.e., each I^2 was well below 25 percent), data were pooled using fixed-effects meta-analysis (separately for each adverse event). Subgroup analyses were not done for any adverse event, since event rates were not reported separately for subgroups with different malignancies or other baseline characteristics.

Table 34. Other Adverse Events Reported by RCTs of Epoetin versus Control

Outcome	# studies reporting	Total N evaluated	N to epoetin	N to control	RR	95% CI	p-value overall effect	heterogeneity	
								p value	I^2
hypertension	15	1,949	1,156	793	1.22	0.98; 1.52	0.07	0.36	8.2%
thrombocytopenia &/or hemorrhage	9	1,422	830	592	1.08	0.76; 1.53	0.66	0.74	0%
rash	6	522	306	216	1.77	0.82; 3.81	0.14	0.66	0%
seizures	3	389	198	191	1.19	0.33; 4.35	0.79	0.74	0%

Hypertension. Only two of 15 reporting studies defined hypertension in their published Methods sections (ten Bokkel Huinink, de Swart, van Toorn et al., 1998; Kunikane, Watanabe, Fukuoka et al., 2001). Reviewers extracted definitions from details of results reported by two additional trials (Welch, James, Wilkinson, 1995; Thatcher, De Campos, Bell et al., 1999). Reviewers also extracted definitions from clinical study reports made available by sponsors of two other trials, each of which specified thresholds for systolic and diastolic hypertension (Rose, Rai, Revicki et al., 1994; Dammacco, Castoldi, Rodjer, et al., 2001). Trials differed with respect to hypertension thresholds, ranging from 140 to 180 mm Hg for systolic pressure, and from 95 to 105 mm Hg for diastolic pressure (Appendix C Table C39). The remaining nine trials did not report definitions or details for hypertension. Thus, severity of hypertension could not be ascertained.

Among 19 comparisons[19] (see Figure 22 and Appendix C Table C39), point estimates of relative risk (RR) for hypertension were not estimable in two (i.e., no events in either arm; Cascinu, Fedeli, Del Ferro, et al., 1994; Iconomou, Koutras, Rigopoulos, et al., 2003), <1 (i.e., favoring epoetin) in four (Kunikane 2001a and b; Rose, Rai, Revicki, et al., 1994; Henry, Brooks, Case, et al., 1995), and >1 (i.e., favoring control) in 13 (Bamias, Aravantinos, Kalofonos, et al., 2003; Case, Bukowski, Carey, et al., 1993; Dammacco, Castoldi, Rodjer, et al., 2001; Littlewood, Bajetta, Nortier, et al., 2001; Osterborg 1996 a and b; Rosenzweig, Bender, Lucke, et al., 2004; Silvestris, Romito, Fanelli, et al., 1995; ten Bokkel Huinink, de Swart, van Toorn et al., 1998a and b; Thatcher, De Campos, Bell et al., 1999a and b; Welch, James, Wilkinson, 1995).

Meta-analysis[19] of 15 reporting RCTs (see Figure 22; N=1,949; 1,156 to epoetin, 793 to control) showed:

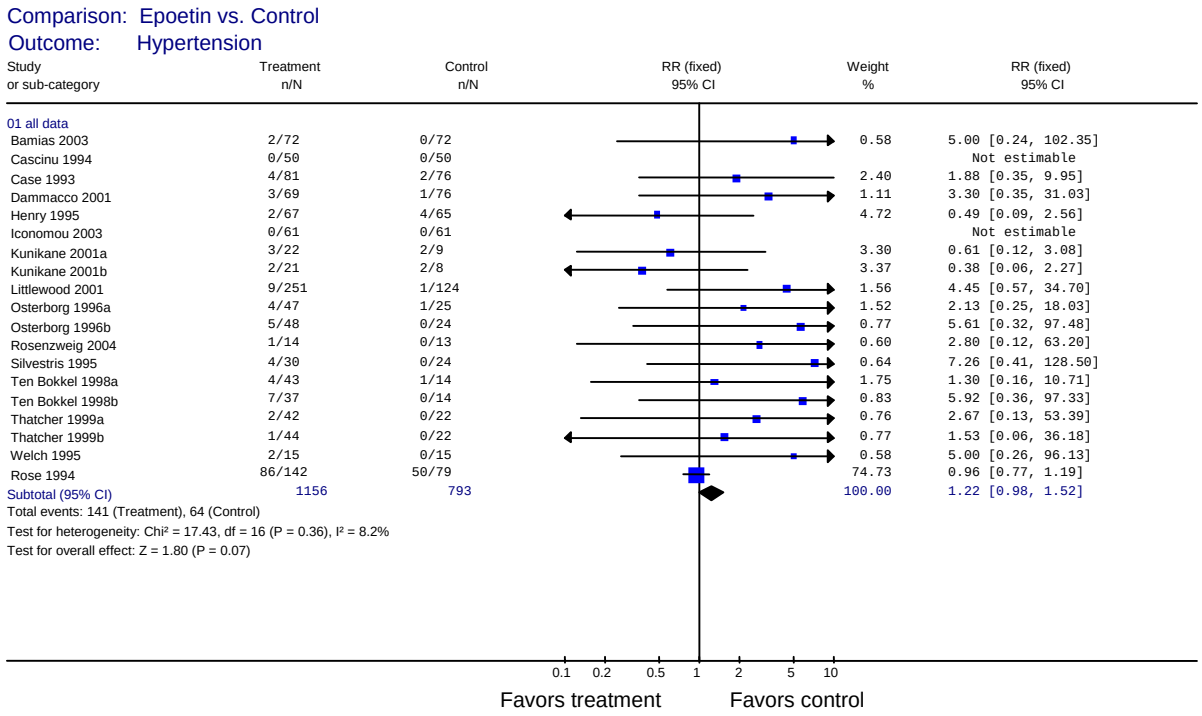
- Increased risk for hypertension in epoetin arms was not statistically significant (RR=1.22; 95 percent CI: 0.98, 1.52; p=0.07)

- Pooled event rates: epoetin, 12.2 percent; controls, 8.1 percent

[19] Four studies compared two arms given different epoetin doses (ten Bokkel 1998; Kunikane 2001; Thatcher 1999) or a fixed versus a titrated dosing regimen (Osterborg 1996) against one control arm per study. For the meta-analysis, each control arm was split artificially and randomly into two groups, each entered with one experimental arm as a separate study. As this might influence weighting of the studies, the analysis was repeated with both experimental arms of each study merged and compared to that study's full control arm. Results with each study's experimental groups merged (RR=1.24; 95% CI: 0.99, 1.54) were similar to results with the control groups split (RR=1.22; 95% CI: 0.98, 1.52).

Several aspects of the evidence available on hypertension limit interpretability and conclusions from the meta-analysis. Although 15 RCTs reported, one trial with 11.3 percent of the total patient population but 66.3 percent of events contributes 75 percent weight and thus likely dominates the analysis' results (Rose 1994). Additionally, reporting trials used a wide range of thresholds to define hypertension. Furthermore, only a minority of RCTs on epoetin versus control reported on hypertension (15 of 48 with 23.6 percent of randomized patients).

Figure 22. Meta-Analysis of 15 Epoetin-versus-Control RCTs that Reported Hypertension



Thrombocytopenia and/or Hemorrhage. Among 12 comparisons (see Figure 23), point estimates for relative risk (RR) of thrombocytopenia and/or hemorrhage were not estimable in one (i.e., no events in either arm; Cascinu, Fedeli, Del Ferro, et al., 1994), <1 (i.e., favoring epoetin) in two (Boogaerts, Coiffier, Kainz, 2003; Osterborg, Boogaerts, Cimino, et al., 1996b), indistinguishable from 1.0 in two (Del Mastro, Venturini, Lionetto, et al., 1997; Littlewood, Bajetta, Nortier, et al., 2001), and >1 (i.e., favoring control) in seven (Thatcher 1999a and b; Osterborg 1996a; Kunikane 2001 a and b; Dammacco, Castoldi, Rodjer, et al., 2001; Bamias, Aravantinos, Kalofonos, et al., 2003).

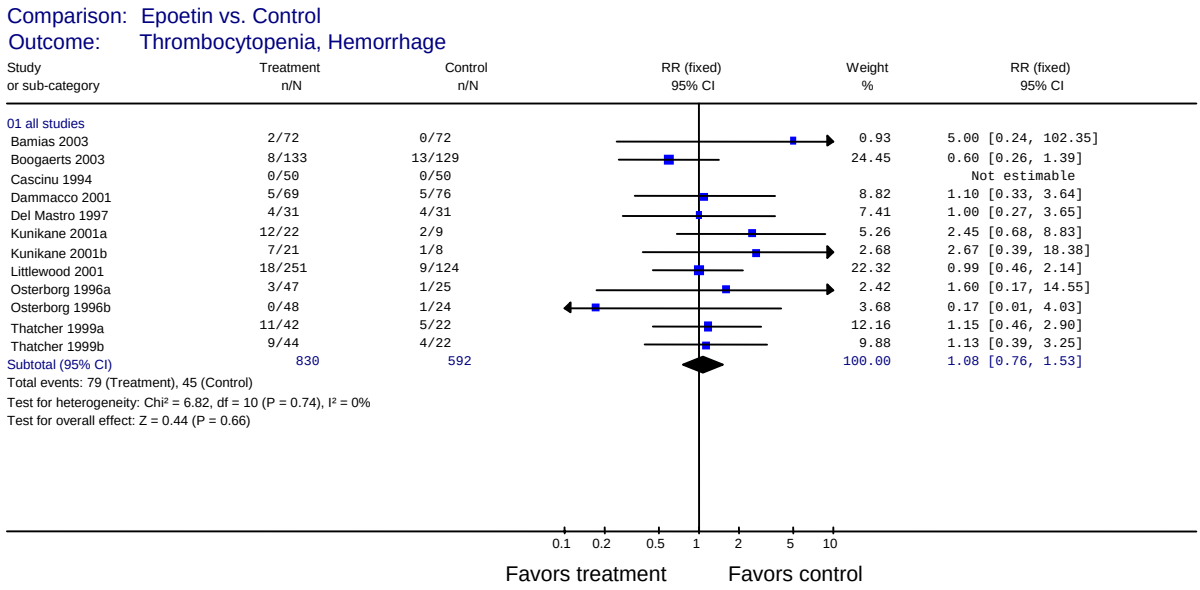
Meta-analysis[20] of nine reporting RCTs (see Figure 23; N=1,422; 830 to epoetin, 592 to control) showed:

[20] Three studies compared two arms given different epoetin doses (Kunikane 2001; Thatcher 1999) or a fixed versus a titrated dosing regimen (Osterborg 1996) against one control arm per study. For the meta-analysis, each control arm was split artificially and randomly into two groups, each entered with one experimental arm as a separate study. As this might influence weighting of the studies, the analysis was repeated with both experimental arms of each study merged and compared to that study's full control arm. Results with each study's experimental groups merged (RR=1.08; 95% CI: 0.74, 1.57) were similar to results with the control groups split (RR=1.19; 95% CI: 0.80, 1.76).

- Increased relative risk for thrombocytopenia and/or hemorrhage in epoetin arms was not statistically significant (RR=1.08; 95 percent CI: 0.76, 1.53; p=0.66)

- Pooled event rates: epoetin, 9.5 percent; controls, 7.6 percent

Figure 23. Meta-Analysis of Seven Epoetin-versus-Control RCTs that Reported Thrombocytopenia and/or Hemorrhage



Rash. Among eight comparisons (see Figure 24), point estimates for relative risk (RR) for rash were not estimable in one (i.e., no events in either arm; Kurz, Marth, Windbichler, et al., 1997), <1 (i.e., favoring epoetin) in one (Thatcher 1999b), and >1 (i.e., favoring control) in six (Thatcher 1999a; Osterborg 1996a and b; Henry, Brooks, Case, et al., 1995; Del Mastro, Venturini, Lionetto, et al., 1997).

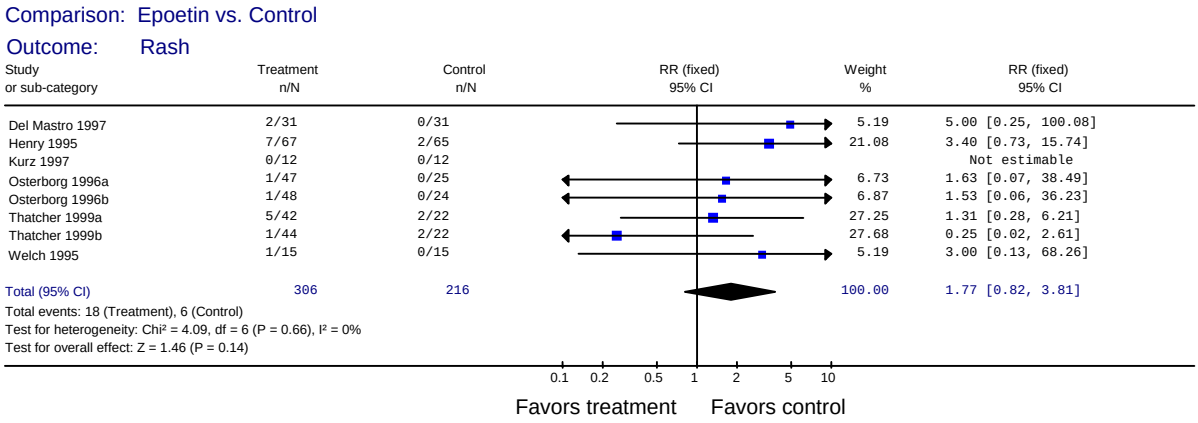
Meta-analysis[21] of six reporting RCTs (see Figure 24; N=522; 306 to epoetin, 216 to control) showed:

- Increased relative risk for rash in epoetin arms was not statistically significant (RR=1.77; 95 percent CI: 0.82, 3.81; p=0.14)

- Pooled event rates: epoetin, 5.9 percent; controls, 2.8 percent

[21] Two studies compared two arms given different epoetin doses (Thatcher 1999) or a fixed versus a titrated dosing regimen (Osterborg 1996) against one control arm per study. For the meta-analysis, each control arm was split artificially and randomly into two groups, each entered with one experimental arm as a separate study. As this might influence weighting of the studies, the analysis was repeated with both experimental arms of each study merged and compared to that study's full control arm. Results with each study's experimental groups merged (RR=1.86; 95% CI: 0.84, 4.09) were similar to results with the control groups split (RR=1.77; 95% CI: 0.82, 3.81).

Figure 24. Meta-Analysis of Six Epoetin-versus-Control RCTs that Reported Rash



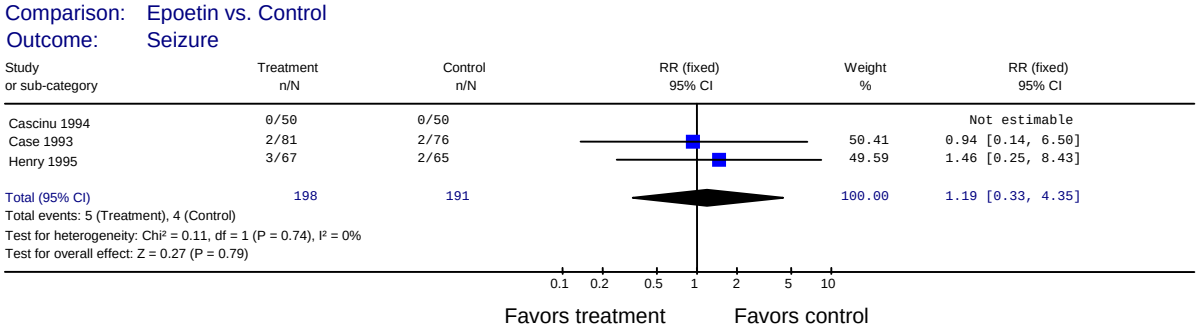
Seizures. Among three reporting trials, (see Figure 25), point estimates for relative risk (RR) of seizure were not estimable in one (i.e., no events in either arm; Cascinu, Fedeli, Del Ferro, et al., 1994), just below 1 (i.e., favoring epoetin) in a second (Case, Bukowski, Carey, et al., 1993), and >1 (i.e., favoring control) in the third (Henry, Brooks, Case, et al., 1995).

Meta-analysis of the three reporting RCTs (see Figure 25; N=389; 198 to epoetin, 191 to control) showed:

- Increased relative risk for seizure in epoetin arms was not statistically significant (RR=1.19; 95 percent CI: 0.33, 4.35; p=0.79)

- Pooled event rates: epoetin, 2.5 percent; controls, 2.1 percent

Figure 25. Meta-Analysis of Three Epoetin-versus-Control RCTs that Reported Seizures



Antibodies. Six trials of epoetin versus control tested for antibodies to erythropoietin (Chang, Couture, Young, et al., 2005; Henry, Brooks, Case, et al., 1995; Oberhoff, Neri, Amadori, et al., 1998; Thatcher, De Campos, Bell, et al., 1999; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Osterborg, Brandberg, Molostova, et al., 2002). Antibodies were not detected in any tested patient.

Darbepoetin versus Control

FDA-approved Prescribing Information.

Tables summarizing adverse events in cancer patients receiving chemotherapy enrolled in FDA-reviewed trials reported incidence of hypertension, rash, and seizures or convulsions (Table 35). The tables did not report incidence of thrombocytopenia and/or hemorrhage.

Table 35: Incidence of Selected Adverse Events in FDA-Reviewed Trials of Aranesp®

	Aranesp®	controls
N	873	221
hypertension	3.7%	3.2%
rash	7%	3%
seizures or convulsions	0.6%	0.5%

Evidence from Published Trials.

One trial (Vansteenkiste, Pirker, Massuti, et al., 2002) reported that hypertension occurred in nine of 155 patients (5.8 percent) receiving darbepoetin, and in six of 159 controls (3.8 percent) (RR 1.54, 95 percent CI 0.56; 4.22, n=314). The between-arm difference was not statistically significant (p=0.40). Investigators did not report a definition for hypertension.

No studies that compared darbepoetin versus control reported data separately by study arm on rates of thrombocytopenia and/or hemorrhage, rash, or seizures.

Antibodies. Each included trial of darbepoetin versus control tested for antibodies to that product and found none in any patients.

KQ1 Discussion and Conclusions

Erythropoietic stimulants effectively increase Hb levels and reduce transfusion risk. This review did not identify evidence demonstrating that either of the available erythropoietic stimulants (epoetin or darbepoetin) achieves hematologic response or reduces transfusion risk in a larger proportion of patients than the other. Meta-regression results for transfusion risk suggest that the magnitude of the benefit varies with type of tumor and with treatment duration.

Evidence for the effect of erythropoietic stimulants on quality of life and on survival and associated outcomes is much more difficult to evaluate and interpret, and is therefore the major focus of this discussion.

Quality of Life

One large study found that the difference in the FACT-An and FACT-fatigue QoL assessments during treatment between darbepoetin- and epoetin-treated study arms was not statistically significant, suggesting no difference in impact on QoL measures targeted to anemia symptoms. Evidence from studies of epoetin or darbepoetin vs. control suggest that patients treated with erythropoietic stimulants show improvement from baseline in QoL assessments,

particularly on symptom-specific scales. Whether patients experience perceptible improvement in QoL is less clear for the following reasons (details follow):

- Factors other than Hb are associated with cancer fatigue;
- Empirically based estimates of the minimally important difference (MID) in QoL scales are not fully developed;
- FACT-fatigue subscale trial results have been compared to MID estimates anchored to ECOG and Karnovsky performance scores. FACT-fatigue improvements may achieve clinical significance in some of the few studies that adequately report this measure;
- Our conclusions regarding quality of life benefits disagreed with a recent meta-analysis of selected epoetin trials (Jones, Schenkel, Just, et al., 2004), which concluded that epoetin significantly improves QoL in patients with cancer. However, results of this other study could be biased by an analysis heavily weighted by inclusion of uncontrolled studies, and by the considerable amount of QoL data missing in some studies.

Other factors that may influence the effect of erythropoietic stimulants on QoL

Fallowfield, Gagnon, Zagri, et al. (2002) conducted a multivariate analysis of the Littlewood, Bajetta, Nortier, et al. (2001) QoL data that confirmed the statistically significant results of the univariate analysis, but showed that significant improvements were limited to patients without disease progression. Wisloff, Gulbrandsen, Hjorth et al. (2005) examined the impact of Hb concentration on EORTC QLQ-C-30 scores for 745 multiple myeloma patients while adjusting for disease characteristics including response/progression. The statistical significance of the effect of Hb change on the 3-item fatigue component of QoL was reduced by a factor of 10 when adjusted for response to therapy. Thus, only a subset of patients may be able to realize a QoL benefit with epoetin or darbepoetin treatment. In another study (Nieboer, Buijs, Rodenhuis, et al., 2005), of patients treated with chemotherapy for breast cancer, fatigue, an important component of the FACT-An assessment of QoL, was strongly correlated with mental health and with muscle and joint pain, but not with hemoglobin status, suggesting that multiple causes of fatigue need to be taken into account. Other studies similarly indicate that Hb values alone do not fully account for perceived fatigue (Holzner, Kemmler, Greil et al., 2002; Okuyama, Akechi, Kugaya et al., 2000).

Clinical significance of statistically significant changes in QoL

Whether statistically significant improvements detected in QoL assessments are clinically significant and meaningful to the patient is inadequately answered by the data presented here. The FACT scales and subscales most often used are, as designed, symptom specific. Treatment-associated improvement on these scales refers to fatigue and other aspects of anemia-related QoL. Seven of ten studies using global QoL scales (including FACT-G but not FACT-An, which contains a substantial proportion of symptom-specific questions) found nonsignificant changes with treatment, suggesting that the less-sensitive global scales may not reflect the changes seen in the anemia symptom-specific FACT scales. Alternatively the improvements reported may not

be large enough to be detected as a change in overall quality of life. However, this question is not answered sufficiently by the data as only 4 studies used both symptom-specific and global scales and result patterns were different for each.

To determine the clinical significance of improvements on the FACT-An and its subscales, a clear, empirically-based estimation of the minimum clinically important difference (MID) is needed for each scale. Anchor-based and distribution-based methods can be employed to estimate the MID. Anchor-based methods evaluate the relationship between change in the QoL scale of interest (target) and an independent measure (anchor). Required qualities of the anchor are, first, that it is an accepted clinical measure for which the clinical significance of change in the measure is well understood. The anchor should also measure QoL in some way. Second, there should be an association between the anchor and the target (Yost and Eton, 2005; Guyatt, Osoba, Wu et al., 2002); associations of 0.5 or greater are strongly recommended (Guyatt, Norman, Juniper et al., 2002). This information should be included in reports of MID studies. Distribution-based methods rely on QoL score statistical distributions, and may use standard deviation (SD) or standard error of measurement (SEM) as the criterion for clinical significance. Because anchor-based approaches are difficult to validate, and distribution-based methods are statistical, rather than clinical, in nature, current recommendations are to estimate MID with more than one anchor; distribution-based methods may supplement but should not substitute for anchor-based methods (Guyatt, Osoba, Wu, et al., 2002; Osoba, Rodrigues, Myles et al., 1998).

Both anchor- and distribution-based methods have been used to estimate MID for FACT-An and subscales in cancer patients treated with epoetin (Cella, Eton, Lai, et al., 2002; Patrick, Gagnon, Zagari, et al., 2003). Using change in Hb as an anchor; Patrick, Gagnon, Zagari, et al. (2003) reported correlations of 0.26 (FACT-G) and 0.29 (FACT-fatigue subscale) between QoL scale and a Hb increase of 1 g/dL.[22] No correlation information was provided by Cella, Eton, Lai, et al. (2002), using the same anchor, nor was additional information on interpretation of the Hb change anchor provided in either study. Given correlations between anchor and target that are not strong, and no documented validation of the anchor's interpretability, it is unclear what the identified minimal change in the anchor of 1 g/dL means to the perceived QoL of the patient. Furthermore, whether or not increased Hb is interpretable as a measure of QoL is part of the question at hand: does the use of epoetin or darbepoetin, which increase Hb levels, improve QoL? Thus, change in Hb is not an informative anchor.

Anchoring changes in FACT scales to performance scores, however, is more persuasive. Cella, Eton, Lai, et al. (2002) also used ECOG and Karnovsky performance scores as anchors in their study.[23] The authors did not report information on the correlation of either performance scores with target QoL scales, or on the interpretability of change in the performance scores. However, as these scores reflect physical function, changes are likely to be more closely linked to the physical aspects of QoL in epoetin and darbepoetin-treated patients. This is supported by data from an unrelated study of chemotherapy in patients with lung cancer, where baseline ECOG performance score was strongly correlated with the EORTC QLQ C-30 scales at -0.52 (physical function), -0.63 (global health status), and 0.52 (fatigue) (Bircan, Berktas, Bayiz et al., 2003). Similar published information could not be found for FACT scales in epoetin-treated patients.

[22] These correlations are similar to those reported by studies included in this review (e.g. 0.35 for change in FACT-fatigue subscale and Hb, Iconomou, Koutras, Rigopoulos 2003; 0.26 (FACT-G) and 0.29 (FACT-fatigue subscale for change in QoL scale and change in Hb, Littlewood, Bajetta, Nortier 2001).

[23] Due to few patients in categories of considerable disability, Cella, Eton, Lai, et al. (2002) collapsed such categories into a single category for both ECOG and Karnovsky performance scales.

Interpreting the results of this review

When results of the most commonly reported QoL scale, the symptom-specific FACT-fatigue subscale, are compared to MID estimates from anchoring to performance scores, the results are not strong. The estimated MID range is 3.5-8.8 (Cella, Eton, Lai, et al., 2002). As shown in Table 36, four of six absolute mean change differences in scores between epoetin and control arms fall within the lower half of the estimated MID range of 3.5-8.8, while the other 2 are below that range. When these results are translated into effect size, most effect sizes would be considered small.[24] Thus, this analysis suggests that in the small sample of studies that reported results for the FACT-fatigue subscale, some improvements in QoL may be clinically significant (depending on the "true" MID value) but the magnitude of the effect is likely to be small.[25]

Results were similar using a distribution-based method (Cella, Eton, Lai, et al., 2002). For the FACT-fatigue subscale, the average MID based on SEM was 2.6 while the average MID based on 0.5 SD was 5.8. Thus, if 2.6 was used as the MID, results from 4 of 6 studies in Table 36 would be clinically significant, whereas if 5.8 was the MID, none of the studies would be clinically significant.

Thus, for purposes of this review, the true MID is not known with certainty, only a few studies reporting QoL results can be evaluated in this way, and the clinical significance of their results remains unclear. Additional limitations on interpretation are the unknown effects of potential bias due to substantial missing data in included studies and other concerns regarding study validity, including lack of blinding and of information on QoL instrument administration.

Table 36. FACT-Fatigue Subscale Mean Change Differences Between Epoetin and Control Arms in 6 Included Studies and Corresponding Effect Sizes

Study	FACT-fatigue subscale difference in change from baseline, Epoetin - Control	Effect size	p-value for comparison of change
Boogaerts 2003	5.2	0.45	<0.05
Littlewood 2001	5.5	(cannot be calculated)	0.004
Osterborg 2002	2.2	0.20	>0.05
Iconomou 2003	3.6	0.32	0.022
Witzig 2005	2.4	0.11	0.18
Chang 2005	4.6	0.41	<0.001

Other analyses of the effects of erythropoietic stimulants on QoL. While our analysis relies on a non-quantitative vote-counting method due to the lack of sufficient published information for quantitative analysis, Jones, Schenkel, Just, et al. (2004) conducted a quantitative meta-analysis of change from baseline score on a variety of QoL measures reported in published and unpublished studies. For example, the authors report a mean change of 4.6 for the FACT-fatigue subscale after adjustment for potential confounders, which would be within the MID

[24] Cohen (1988) arbitrarily defined effect sizes of 0.2 as "small," 0.5 as "moderate," and 0.8 as "large."

[25] As this report went to press, an analysis of the clinical significance of QoL data from Hedenus, Adriansson, San Miguel et al. (2003) was published (Littlewood, Kallich, San Miguel et al. (2006). In this analysis, treatment and control arms were pooled; patients who improved by at least 3 points on their FACT-fatigue subscale score were significantly more likely to show improvement in other FACT scales (except social well-being), in Brief Symptom Inventory Depression and Anxiety subscales, and in numeric rating scales of Energy, Activity, and Overall Health.

range of 3.5-8.8 estimated by Cella, Eton, Lai, et al. (2002). However, difficulties with this study include an analysis heavily weighted by inclusion of uncontrolled studies, which are subject to bias. Although the authors report that statistical significance was retained when the analysis was repeated without large cohort ("community") studies, the resulting score change was not reported. Because the authors include cohort studies, they also analyze treatment and control arms of randomized controlled trials as separate cohorts, losing the advantage of within-study comparison to control. The authors report that statistical significance is retained for some measures, when the analysis is "controlled" for placebo effect, but again do not report the resulting score change. Because factors other than epoetin intervention may affect outcomes, randomized controlled trials are necessary for accurate, within-study comparison to placebo. Finally, there is no mention of the considerable amount of QoL data missing in some studies and the resulting potential for bias.

Survival, Thromboembolic Events, and Tumor Response

Because these outcomes are interrelated, they are discussed together. Limited evidence from trials directly comparing epoetin to darbepoetin found no significant differences in survival or thromboembolic events; tumor response was not reported. The majority of the evidence for these outcomes is derived from trials of epoetin or darbepoetin versus control. Major topics discussed include:

- Results of other recent evidence summaries;
- Results of large trials designed for survival outcomes and FDA analysis for the Oncologic Drugs Advisory Committee;
- Potential confounding variables and how they may affect interpretations of results;
- Limitations of the data.

Recent evidence summaries

Prior to this review, major summaries on survival outcomes of erythropoietin product administration include a review conducted by the Cochrane Haematological Malignancies Group (http://www.cochrane.org/reviews/en/ab003407.html; Bohlius, Langensiepen, Schwarzer et al., 2005), a systematic review from the National Institute for Health and Clinical Excellence (Wilson, Yao, Rafferty, et al., 2005), and a review of safety conducted by the FDA Oncologic Drugs Advisory Committee (ODAC) on May 4, 2004 (Food and Drug Administration Oncologic Drugs Advisory Committee Meeting Briefing Information, 2004). The Cochrane review included studies published through December 2001, none of which were designed to evaluate survival as the primary outcome. Rather, survival was a secondary outcome, often collected retrospectively after the close of the study and after patient treatment was no longer controlled by the study protocol. In many cases results were not included in the trial's published report but were available only from investigators responding to the authors' request for supplementary data. The pooled, unadjusted hazard ratio for death was 0.84 (95% CI, 0.69-1.02), favoring epoetin

treatment. The result, however, was not statistically significant. Given the limitations of the evidence, the results were considered inconclusive.

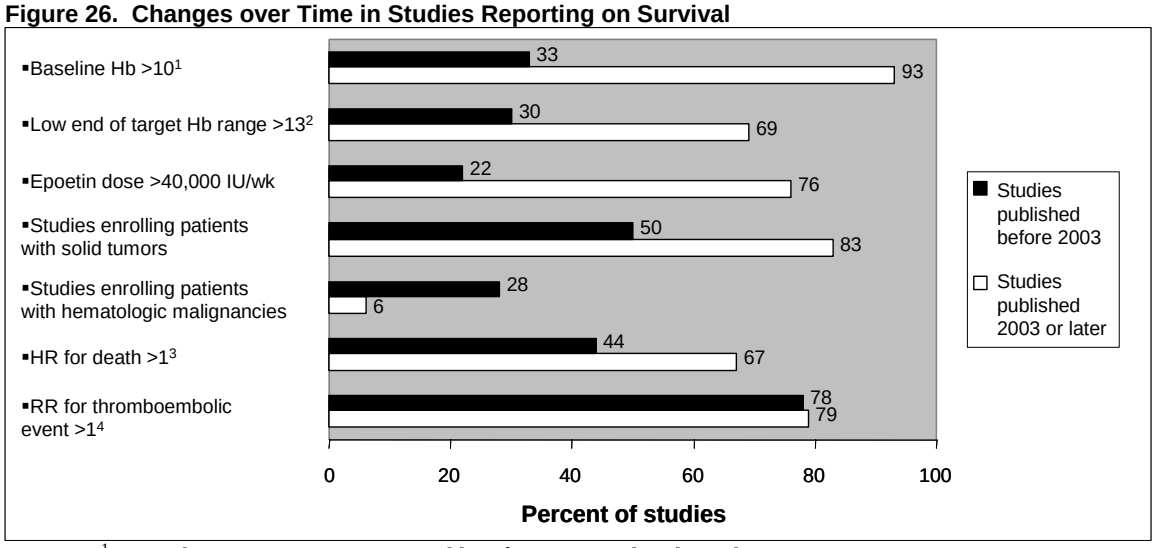
The NICE report updated the Cochrane review with 9 new studies reporting survival outcomes and published through September 2004. The pooled HR for death was 1.03 (95% CI 0.92-1.16), also not statistically significant and suggesting no effect on survival. The report's authors commented that "The marked change in the results is due to the fairly extreme results favouring no treatment/placebo in the newer studies."

Changes in study characteristics over time are illustrated in Figure 26, which shows the percentages of studies published before 2003 and after 2003 with the listed characteristics. Studies published in 2003 or later enrolled patients with higher baseline Hb, used higher epoetin doses and/or targeted higher final Hb levels compared to studies published before 2003. Later studies also tended to enroll patients with solid tumors rather than hematologic tumors, likely affecting chemotherapy regimen. Later studies were more likely to have a HR for death greater than 1. Of 18 studies published in 2003 or later and included in this review, 12 reported HR for death >1 and 11 reported RR for thromboembolic event >1 (Appendix C Table C28).

Trials designed for survival outcomes precipitating FDA analysis

Two recent and larger trials designed for overall or progression-free survival (Leyland-Jones 2003; Henke, Laszig, Ruebe, et al., 2003) had increased mortality in the epoetin study arms; two other trials designed to measure survival outcomes (EPO-CAN-15; GOG-191) and 1 trial designed for local tumor response (Vadhan-Raj, Skibber, Crane, et al., 2004) had significant increases in thromboembolic events in the treatment arms and were consequently closed prematurely. The adverse events reported in these trials prompted the FDA to examine the safety of higher doses of erythropoietic stimulants or higher Hb target levels, in an ODAC meeting on May 4, 2004. Both the Leyland-Jones (2003) and Henke, Laszig, Ruebe, et al. (2003) studies were intended to assess survival and tumor response outcomes. Results showed shorter overall survival; shorter progression-free survival; and increased incidence of thrombotic/cardiovascular events in the patients receiving epoetin (Table 37). Particularly troubling was the increased mortality due to thrombotic vascular and cardiovascular adverse events in the epoetin-treated arm of the Leyland-Jones study at 4 months' followup.

Figure 26. Changes over Time in Studies Reporting on Survival



[1]For studies reporting any interpretable information on baseline Hb.
[2]For studies reporting information on Hb target range.
[3]For studies with an estimable HR for death.
[4]For studies reporting on survival that also reported on thromboembolic events.

Table 37. Summary of Adverse Events, Tumor Response, and Survival Outcomes reported in the FDA Briefing Document, Oncologic Drugs Advisory Committee meeting, May 4, 2004

Trial [Epo Product, dose]	Trial Description	Epo Target	Thrombosis/ Cardiovascular (CV) Outcomes	Epoetin Arm Results	Placebo Arm Results	Disease Progression/ Tumor Response	Epoetin Arm Results	Placebo Arm Results	Survival Outcome	Epoetin Arm Results	Placebo Arm Results
Leyland-Jones 2003 Breast Cancer Erythropoietin Trial (BEST) [EPREX, 40,000 IU qw]	RCT of epoetin in 939 women with metastatic breast cancer; **designed to assess overall survival**	initiation of epoetin when Hb ≤13 g/dL, to target Hb 12-14 g/dL	fatal thrombotic or CV events in the first 4 mo	2.3%	0.4%	disease progression	6%	3%	12-mo OS rates	70%	76%
			fatal thrombotic or CV events after 4 mo	0.6%	1.5%	early mortality (by 4 mo)	8.7%	3.4%	Hazard ratio, 12 mo followup	HR = 1.37 95% CI, 1.07-1.74 p=0.012	
Henke et al. (2003) [NeoRecormon, 300 IU/kg tiw]	351 head and neck cancer patients receiving radiotherapy; **designed to assess locoregional progression-free survival**	≤14 g/dL for women and ≤15 g/dL for men	hypertension, hemorrhage, venous thrombosis, pulmonary embolism or CV event	11%	5%	locoregional progression-free survival	RR = 1.62[1] 95% CI, 1.22-2.14 p=0.0008		Median OS	605 days	928 days
										p=0.09 (logrank)	
									Relative Risk of death	RR = 1.4[1] 95% CI, 1.05-1.84 p=0.02 (Cox)	
			died of "cardiac disorders"	5%	3%	locoregional progression	RR = 1.69 95% CI, 1.16-2.47 p=0.007				
N93-004[2] [Procrit, 150 IU/kg tiw]	post-marketing, non-inferiority RCT of epoetin in **224 patients**[3] with small cell lung cancer undergoing first line therapy; **powered at n=400 to assess tumor response**	epoetin dose was not reduced until Hb ≥16 g/dL	incidence of thrombotic vascular events[4]	22%	23%	CR+PR response rate after 3 chemo cycles	72%	67% p=NS	Median OS, 3 yr followup	10.5 mos.	10.4 mos.
									Overall mortality rate, 3 yr followup	92%	88%
									Hazard ratio, 3 yr followup	HR[5] = 1.53 95% CI, 0.65-3.61	

[1] Adjusted for stage and randomization stratum

[2] The N93-004 epoetin trial was published in full in December, 2005 (Grote, Yeilding, Castillo, et al., 2005)

[3] 66% of pts in epoetin arm (n=109) had extensive stage SCLC cf. 59% of pts in placebo arm (n=115); else no differences in baseline characteristics; trial terminated early for poor accrual

[4] Incidences of specific subtypes of thrombotic vascular events similar except for chest pain (7% epoetin; 14% placebo) and extracardiac vascular disorders (10% epoetin, 4% placebo)

[5] Not available from FDA Briefing Document; abstracted from Industry-supplied summaries for ODAC meeting.

Table 37. Summary of Adverse Events, Tumor Response, and Survival Outcomes reported in the FDA Briefing Document, Oncologic Drugs Advisory Committee meeting, May 4, 2004 (continued)

Trial [Epo Product, dose]	Trial Description	Epo Target	Thrombosis/ Cardiovascular (CV) Outcomes	Epoetin Arm Results	Placebo Arm Results	Disease Progression/ Tumor Response	Epoetin Arm Results	Placebo Arm Results	Survival Outcome	Epoetin Arm Results	Placebo Arm Results
980297 (Vansteenkiste 2002) [Aranesp, 2.25 mcg/kg qw]	320 anemic patients with lung cancer being treated with platinum chemotherapy; **powered for transfusion outcomes**	Epo dose was not adjusted until Hb ≥14 g/dL for women and ≥15 g/dL for men.	thrombotic events	5%	3%	disease progression over median 12 mo	HR = 0.71[6] 95% CI, 0.54-0.94		Median time to death	43 wks	35 wks
						locoregional PFS, over median 12 mo	HR = 0.74[6] 95% CI, 0.57-0.97		Hazard ratio, 11 mo median followup	HR = 0.80[6] 95% CI, 0.58-1.11	
Studies halted prematurely by Johnson & Johnson											
EPO-CAN-15 [Procrit, 40,000 IU qw]	106 patients with SCLC receiving chemoradiation therapy	Hb 14-16 g/dL	thrombotic vascular events	34%	6%				Hazard ratio, ?followup	HR[5] = 2.70 95% CI, 1.17- 6.21	
GOG-191 [Procrit, 40,000 IU qw]	113 patients with cervical cancer receiving chemo-radiation	Hb 13-14 g/dL	thrombotic vascular events	16%	5%				Hazard ratio, ?followup	HR[5] = 0.82 95% CI, 0.29-2.29	
PR00-03-006 (Vadhan-Raj 2004) [Procrit, 40,000 IU qw]	60 patients with gastric or rectal cancer undergoing preoperative chemoradiation	Hb 14-15 g/dL	thrombotic vascular events	24%	6%				Hazard ratio, ?followup	HR[5] = 0.15 95% CI, 0.00-7.69	

[6] Adjusted for tumor type and region

Complicating the analysis was a concomitant decrease in progression-free survival in the treatment arms of the Leyland-Jones and Henke, Laszig, Ruebe, et al. (2003) trials. The FDA analysis of these studies could not determine whether epoetin potentiates tumor progression. The other studies analyzed were not powered for survival outcomes, but thrombosis or vascular events were more frequent in the treatment arms of most.

Some have questioned the generalizability of the Henke, Laszig, Ruebe, et al. (2003) results based on the number of protocol violations (60 radiotherapy violations and 20 medication violations among N=180 assigned to epoetin; 54 radiotherapy violations and 8 medication violations among 171 assigned to control; nature and direction of violations unspecified). However, the relative risk for locoregional progression-free survival remained significantly in favor of control if analysis was restricted to patients given correct radiotherapy (RR=1.42; 95 percent CI: 1.01, 2.01). For all three outcomes shown in Table 37, results favored control although statistical significance was lost in per-protocol analyses. Thus, the protocol violations do not clearly explain the unfavorable results. Additionally, published comments on both the Henke, Laszig, Ruebe, et al. (2003) and Leyland-Jones (2003) trials noted some imbalances in baseline characteristics, suggesting that the epoetin arms in both trials had slightly greater proportions of patients with poor prognostic factors. However, these imbalances were detected by a retrospective chart review, something that was not done for studies reporting more favorable survival outcomes with administration of erythropoiesis-stimulating products. Therefore this reporting of imbalances is selective and may bias the comparison of Henke, Laszig, Ruebe, et al. (2003) and Leyland-Jones (2003) to other studies.

Variables that contribute to survival outcome

Survival depends upon several interrelated factors such as cancer type and stage, treatment, and presence of other co-morbidities. Potential effects of erythropoietic stimulants on tumor progression may be positive, negative, or neutral depending on type of cancer, density of erythropoietin receptors, and cancer treatment regimen. The individual risk of thromboembolic events also varies with tumor type and extent, and additionally with type of anticancer therapy, previous history of thrombosis, and presence of other risk factors such as surgery or immobilization (Levine, Lee, Kakkar 2005). Risk appears to be higher with certain types of chemotherapy (e.g. cisplatin) and with drug combinations (e.g., chemotherapy plus tamoxifen) (Weiss 2001). There is evidence that the presence of metastatic disease and number of comorbidities influences risk (Alcalay, Wun, Khatri 2006). Other significant risk factors may include prechemotherapy platelet count, and use of white cell growth factors (Khorana, Francis, Culakova 2005). It is against this background variability that we attempt to define the influence of erythropoietic stimulants on survival, tumor progression, and thromboembolic risk.

The limited evidence available does not support the hypothesis that erythropoietic stimulants increase rates of solid tumor response to therapy. However, other observations raise the possibility that erythropoietic stimulants may accelerate progression of solid tumors expressing erythropoietin receptors. For example, Dr. Michael Henke and colleagues tested tumor samples from a subset of trial patients for erythropoietin receptors and found that epoetin administration to patients with receptor-expressing tumors correlated with shorter progression-free intervals (personal communication; manuscript submitted).

Our analysis of outcomes from 30 studies of epoetin treatment versus control found that erythropoietic stimulation increases relative risk for a thromboembolic event in anemic oncology

patients undergoing cancer therapy. Whether survival and thromboembolic event outcomes in these studies were adversely affected by use of epoetin doses and/or Hb target levels higher than recommended in the product label is unclear. Prior studies on patients with chronic renal failure (CRF) and concurrent cardiovascular disease given erythropoietic stimulants dosed to achieve and maintain target Hb above current recommendations reported an increased risk of cardiovascular and thromboembolic adverse events and death (Besarab, Bolton, Browne, et al., 1998). FDA-conducted exploratory analyses of data from the licensing studies of darbepoetin (which included a comparison group treated with epoetin) suggested that increasing thrombosis/ischemic events in patients treated with epoetin or darbepoetin were associated with increasing rate of rise in Hb, but not with absolute Hb concentration (Food and Drug Administration Oncologic Drugs Advisory Committee Meeting Briefing Information, 2004). However, an FDA-requested analysis by Amgen of the Aranesp Integrated Summary of Safety (ISS) database (873, 115, and 221 cancer patients who received darbepoetin, epoetin, and placebo, respectively) found no evidence of an association between maximum achieved Hb level, or the rapidity of increase in Hb, and risk of cardiovascular or thrombotic adverse events. The analysis did indicate that the highest rate of death events in patients receiving darbepoetin or epoetin was in the category of patients with the highest rate of Hb increase. But most patients died of tumor progression, rather than thromboembolic events. While the data from cancer patients are less clear than those in ESRD patients, the accumulation of adverse events in trials using higher doses and/or achieving higher maximum Hb levels resulted in the recommendation to target maximum Hb levels during treatment no higher than 12 g/dL, and to adhere to recommended doses and dose adjustments to avoid a rapid Hb increase. The current product labels reflect these recommendations.

Additional analysis of end-stage renal disease Medicare patient data suggests that the highest mortality rates may be associated with total exposure to erythropoietic stimulants. Zhang, Thamer, Stefanik et al. (2004) report high inter-patient variation in epoetin dose requirements to attain defined hematocrit levels, and for the same achieved hematocrit, there is a wide variation in survival. For every hematocrit cohort studied, patients administered higher doses of epoetin had significantly lower hematocrit values and greater mortality rates. The association between hematocrit and survival may be confounded by patients' ability to respond; patients who are better able to respond may achieve better outcomes regardless of intervention. Two possible explanations may account for the data: 1) resistance to erythropoietic stimulants could be a marker for undefined comorbidities explaining high mortality rates among trial participants who did not achieve the target hematocrit; or 2) there are side effects of erythropoietic stimulants independent of their effect on hematocrit that may be more pronounced in nonresponders who are administered more product. An accurate investigation of the effects of erythropoietic stimulant exposure on survival in cancer patients receiving therapy would require a patient-level meta-analysis to account for dose adjustments during the course of treatment, information that at present is not publicly available.

Limitations of the data

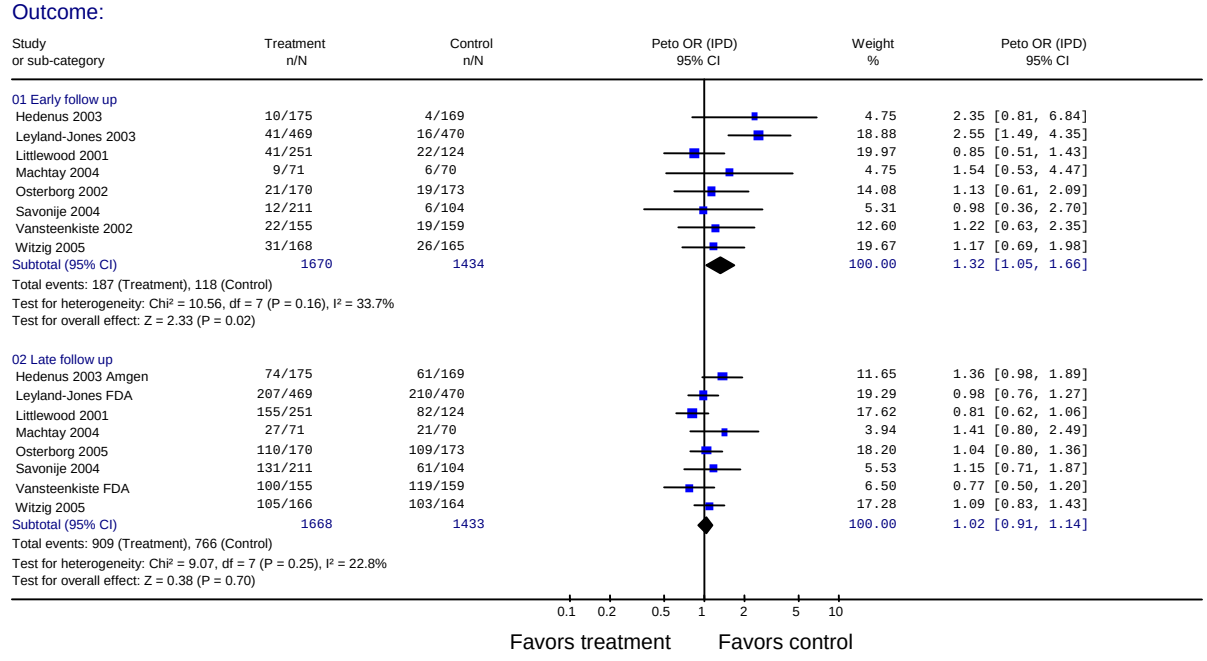
Pooled results from the evidence included in this review do not show improved survival with administration of erythropoietic stimulants. The data from included trials have several limitations. Few studies were designed to evaluate survival outcomes, or met limited criteria of homogeneous tumor types and treatment regimens, to avoid confounding from these important variables. For many trials, particularly the older ones, survival outcomes may have been

collected beyond the stipulated followup period of the randomized controlled trial, when patient management was not controlled by the trial protocol.

Two studies originally intended to evaluate survival outcomes (Henke, Laszig, Ruebe, et al., 2003; Leyland-Jones 2003) found poorer survival with epoetin administration and were two of the most influential studies in the pooled analysis. The studies were notable in that Henke, Laszig, Ruebe, et al. (2003) used a higher than recommended dose, and both targeted a maximum Hb well above current recommendations. However, these studies were not unique in these attributes; the study reported by Littlewood, Bajetta, Nortier, et al. (2001), also a strongly influential study in the pooled analysis, targeted a higher than recommended Hb level yet survival outcomes favored epoetin treatment. The Littlewood study used a recommended epoetin dose, as did the Leyland-Jones study, but survival was not a primary outcome and data were collected after study completion.

Various subgroup analyses of important study attributes (labeled vs. unlabeled use; maximum target Hb by 1 g/dL increments; and homogeneous tumor type and treatment regimen vs. not homogeneous) did not distinguish studies that showed an adverse effect on survival from those that did not. Because tumor progression over longer followup times may dilute the effects of erythropoietic stimulant treatment, examination of survival outcomes at shorter followup times (e.g., during study period) vs. later followup times (1-3 years) might be more informative. Figure 27 shows the results of such an exploratory analysis; the results suggest greater adverse effects of erythropoietic stimulant treatment on survival at earlier time points. However, since the available data are extremely limited and not representative of all included studies, no conclusions can be drawn.

Figure 27. Exploratory Meta-Analysis of Data on Survival at Early vs. Late* Timepoints from Trials with Available Data Comparing Epoetin to Control



* Early followup: during study period (Hedenus 2003, Osterborg 2002, Savonije 2004, Vansteenkiste 2002) or during study plus 30 days (Littlewood 2001, Machtay 2004, Witzig 2005) or in the first 4 months (Leyland Jones 2003); late followup: 1 to 3 years after start of study.

Key Question 2: How do alternative dosing strategies affect the comparative efficacy and safety of epoetin and darbepoetin?

Overview of Evidence and Findings for KQ2

Dosing of erythropoietic stimulants can be individualized based on weight or identical for all regardless of weight (fixed dosing). The same dose can be given in fewer or more frequent injections over time. The amount per unit time can be constant throughout treatment; start high then decrease (front-loading); or adjusted to hematologic response (titrated). They can be given subcutaneously or intravenously. Nineteen trials addressing seven different comparisons (only two done separately for epoetin and darbepoetin) met selection criteria for Key Question 2 (see Table 38). Table 39 summarizes major findings for each comparison.

Table 38. Evidence for Direct Comparison of Alternative Doses, Frequencies, Regimens or Routes

comparison	erythropoietic stimulant	# trials available	total N randomized (all study arms)	N evaluated in non-control arms	outcomes reported					
					Hb response	transfusions	QOL	thromboembolli	other AEs	costs
smaller (S) versus intermediate (I) versus larger (L) weight-based doses	darbepoetin	3	485	S: 76 I: 211 L: 94	✓	✓				
	epoetin	3	324	S: 103 I: 0 L: 104		✓	1 of 3	1 of 3	2 of 3	
smaller (S) versus intermediate (I) versus larger (L) fixed doses	darbepoetin	0								
	epoetin	5	676	S: 280 I: 89 L: 278	✓	✓	4 of 5	2 of 5	2 of 5	
weight-based (W) versus fixed dose (F) regimens	Darbepoetin[1]	1	242	W: 120 F: 122		✓				
	epoetin	1	546	W: 264 F: 268	✓	✓		✓	✓	
same total/unit time given in more (M) versus less (L) frequent dosing	epoetin	2	602	M: 302 L: 300	✓	✓	1 of 2	1 of 2		
front-loaded (F) versus constant (C) weight-based dosing	darbepoetin	2	854	F: 420 C: 399	1 of 2	1 of 2	✓			
titrated (T) versus constant (C) fixed dosing	epoetin	1	144	T: 48 C: 47	✓	✓		✓	✓	
intravenous (I) versus subcutaneous (S) administration	darbepoetin	1	120	I: 59 S: 59		✓				

✓= reported by each relevant trial

[1] As this report was released, a new trial was published reporting similar outcomes with weekly (2.25 mcg/kg) versus every third week (500 mcg) darbepoetin (Canon, Vansteenkiste, Bodoky et al., 2006).

Table 39. Major Findings of Trials Comparing Doses, Regimens, Schedules, or Routes

Drug	# Trials (arms/ trial	Total N (N per arm)	Comparisons	Hb Response	Transfusion Risk
Different Weight-Based Doses					
Darbepoetin	2 (5 & 4)	226 (11-33)	3.0, 5.0, 7.0, or 9.0 mcg/kg/week versus 40,000 IU/week of epoetin	**Similar** at ≥2.25 mcg/kg/ week	**Similar** at ≥2.25 mcg/kg/ week
			1.0, 2.25 or 4.5 mcg/kg/week versus placebo		
Darbepoetin	1 (7)	259 (17-46)	4.5, 6.75, 9.0, 12.0, 13.5, or 15.0 mcg/kg every third week versus placebo	**Greater** with 12-15 than 4.5 mcg/kg	**Similar** at all doses
Epoetin	3 (3)	324 (16-45)	100 versus 200 IU/kg 3x/week versus placebo; or 150 versus 300 IU/kg 3x/week versus untreated (two trials)	not reported separately by dose	**Similar** in each dose pair, 2 of 3 trials
Different Fixed Doses					
Epoetin	5 (2-5)	676 (26-90)	1K, 2K, 5K or 10K IU/day versus untreated; 2K versus 10K IU thrice weekly; 1K versus 5K IU thrice weekly (2 trials); or 9K versus 18K versus 36K IU once weekly	**Greater** at highest dose(s), each trial	**Similar** at all doses compared
Weight-Based versus Fixed Doses					
Darbepoetin	1 (2)	242	4.5 mcg/kg weekly (N=120)	**Similar**	**Similar**
			325 mcg weekly (N=122)		
Epoetin	1 (2)	546	150 IU/kg thrice weekly (N=264)	**Similar**	**Similar**
			10,000 IU thrice weekly (N=268)		
More versus Less-Frequent Dosing					
Epoetin	1 (2)	237	10,000 IU thrice weekly (N=119)	**Similar**	**Similar**
			30,000 IU once weekly (N=118)		
Epoetin	1 (2)	365	40,000 IU weekly (N=183)	**Greater**	**Similar**
			120,000 IU every third week (N=182)	—	
Front-Loaded Regimens					
Darbepoetin	1	723	4.5 mcg/kg 1X, weeks 1-4, 7, 10, 13, 16 (N=356)	**Similar**	**Similar**
			2.25 mcg/kg weekly (N=367)		
Darbepoetin	1 (4)	127 (31-32)	Various front-loaded regimens	**Similar**	Not Reported
Titrated versus Constant-Dose Regimens					
Epoetin	1 (3)	144 (47-19)	Treatment titrated by Hb changes versus 10,000 IU/day versus untreated	**Similar** in treated arms	**Similar** in treated arms
Intravenous versus Subcutaneous Administration					
Darbepoetin	1 (2)	120 (60)	4.5 mcg/kg weekly, intravenous	—	**Similar**
			4.5 mcg/kg weekly, subcutaneous	**Greater**	

Trials for Key Question 2 differed with respect to several variables prespecified for subgroup analysis (see Figure 3). Since each comparison has few relevant trials, study and population parameters are summarized below, with results for each comparison. Transfusion rate was the outcome most consistently reported; no studies reported costs or other economic measures.

Detailed Analysis

KQ2 Comparison I. Different Weight-Based Doses

A. Darbepoetin

Characteristics of Available Studies. Three studies randomized 485 patients to one of multiple darbepoetin doses adjusted by body weight or to epoetin (Glaspy and Tchekmedyian, 2002B) or placebo-treated controls (Hedenus, Hansen, Taylor, et al., 2002; Kotasek, Steger, Faught, et al., 2003). Table 40 summarizes doses compared, evaluable sample sizes, and differences in study and population characteristics between these trials. Each trial studied adult patients with mean baseline Hb ≤10 g/dL given darbepoetin for 12 weeks, and was published as a full paper.

Table 40. Designs and Populations of Studies Comparing Weight-Based Doses of Darbepoetin

Study	lower doses[1]	N eval.[1]	higher doses[1]	N eval.[1]	malignancy type	cancer therapy	iron use
Glaspy 2002 Part B	3.0 mcg/kg/2wk 5.0	A: 33 B: 31	7.0 mcg/kg/wk 9.0	C: 32 D: 32	solid tumors	unspecified chemotherapy	not reported
Hedenus 2002	1.0 mcg/kg/wk 2.25	A: 11 B: 22	4.5 mcg/kg/wk	C: 22	hematologic	unspecified chemotherapy	as needed
Kotasek 2003	4.5 mcg/kg/3wk 6.75 mcg/kg/3wk 9	A: 32 B: 17 C: 46	12 mcg/kg/3wk 13.5 mcg/kg/3wk 15	D: 28 E: 35 F: 40	solid tumors	chemotherapy, some platinum	not reported

[1] Letters denoting study arms in Table 40 correspond to letters denoting study arms compared in Figures 28 and 29.

Results. Figure 28 shows likelihood (relative risks) for a Hb response, comparing each pair of darbepoetin doses in the same study. Each arm of Hedenus, Hansen, Taylor, et al. (2002) and Kotasek, Steger, Faught, et al. (2003) is compared with placebo controls (from the corresponding trial) in Figure 5 of Key Question 1. Results from Glaspy and Tchekmedyian (2002) comparing each darbepoetin arm to epoetin controls are summarized in Table 13. Meta-analysis was not done since doses varied substantially within and across trials. The only statistically significant findings favoring higher doses were in the Kotasek, Steger, Faught, et al. (2003) study, in which Hb responses were more frequent with every third week doses ≥12 mcg/kg than at the lowest dose (4.5 mcg/kg).

Figure 29 shows relative risks of transfusion for the same dose comparisons. Since each 95% CI included 1.0, none of the dose-pair comparisons showed a statistically significant difference

in transfusion risk. Relative risks of transfusion at each darbepoetin dose also were not statistically significant when compared to epoetin controls for Glaspy and Tchekmedyian (2002) (Figure 6 of Key Question 1), and when compared to placebo controls for Hedenus, Hansen, Taylor, et al. (2002) and Kotasek, Steger, Faught, et al. (2003) (Figure 8).

The three trials for this comparison did not report other outcomes.

Figure 28. Darbepoetin Dose and Likelihood (Relative Risk) of Hematologic Response

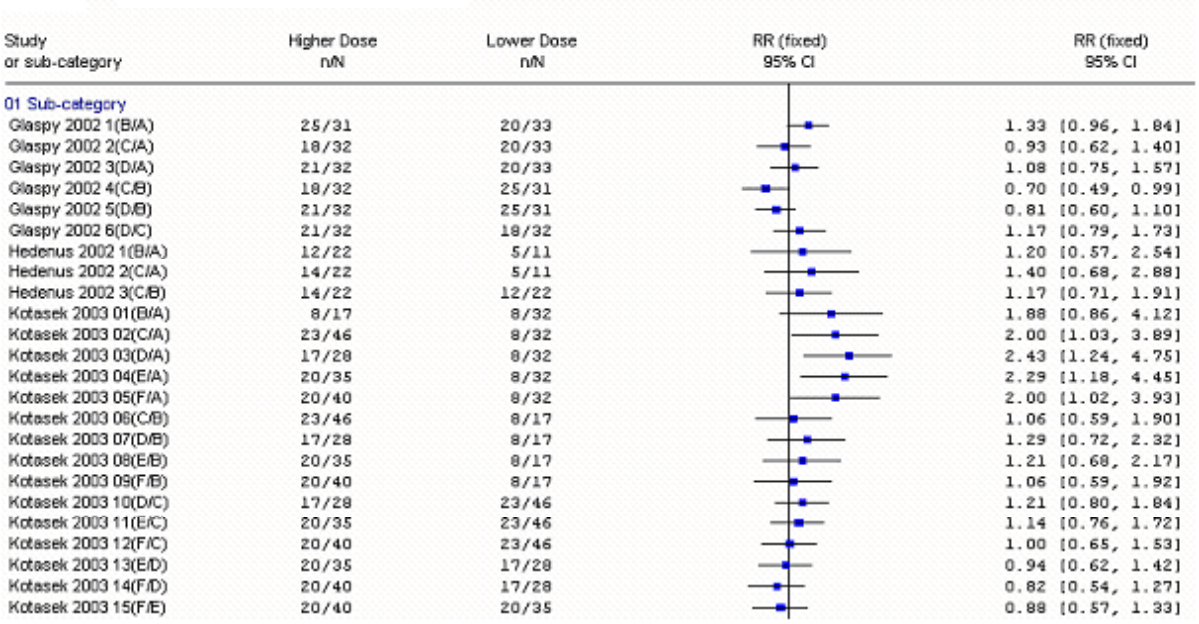
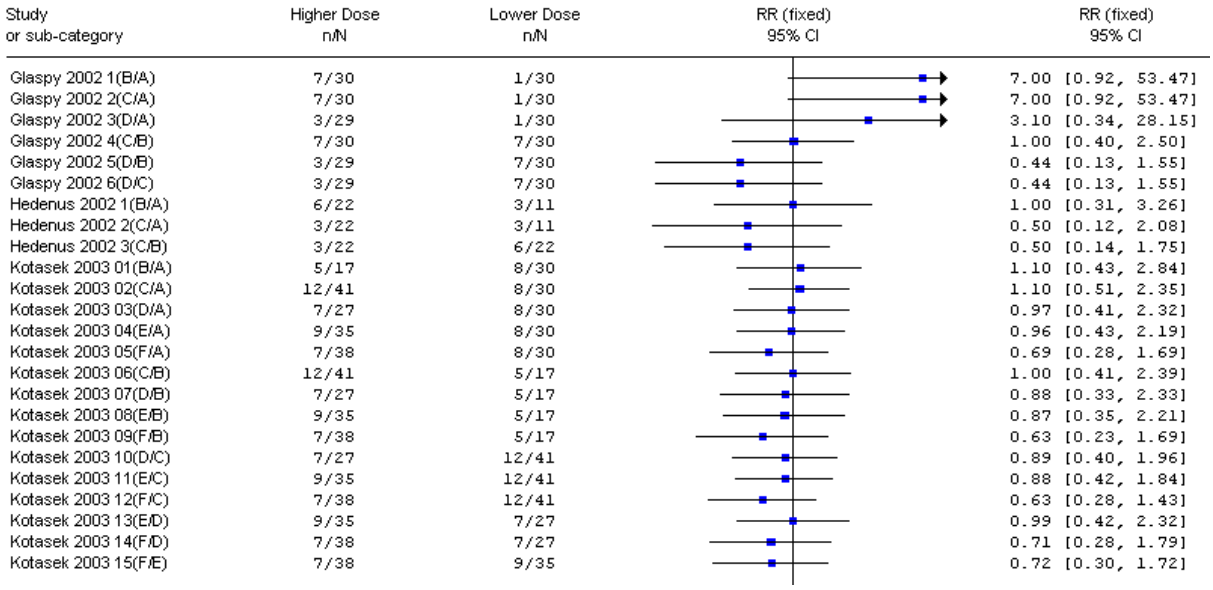


Figure 29. Darbepoetin Dose and Relative Risk of Transfusion



Summary. Two trials (combined N=226) suggest that weekly darbepoetin doses greater than recommended by FDA (2.25 mcg/kg) do not increase Hb responses or decrease transfusion rate. When patients are treated every third week, one trial (n=259) suggests Hb responses are more likely at 12-15 mcg/kg than at 4.5 mcg/kg, although transfusion risks did not differ.

B. Epoetin

Characteristics of Available Studies. Three studies randomized 324 patients to one of two epoetin doses adjusted by body weight or to untreated controls (Kunikane, Watanabe, Fukuoka et al., 2001; ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Thatcher, De Campos, Bell, et al., 1999). Table 41 summarizes doses compared, evaluable sample sizes, and differences in study and population characteristics between these trials. Each trial treated adult patients with solid tumors using platinum-based chemotherapy for most or all.

Table 41. Design and Populations of Studies Comparing Weight-Based Doses of Epoetin

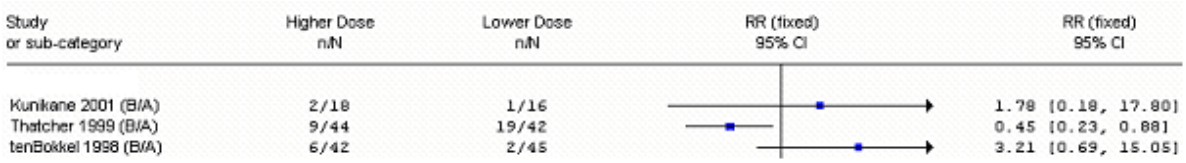
Study	low dose[1]	N eval.[1]	high dose[1]	N eval.[1]	baseline Hb category	EPO Tx duration	iron use
Kunikane 2001	100 IU/kg thrice/wk	16	200 IU/kg thrice/wk	18	≥12 g/dL	8 wks	not reported
ten Bokkel 1998	150 IU/kg thrice/wk	45	300 IU/kg thrice/wk	42	>10-<12 g/dL	>24 wks	as needed
Thatcher 1999	150 IU/kg thrice/wk	44	300 IU/kg thrice/wk	42	≥12 g/dL	26 wks	as needed

[1] Each low dose arm corresponds to arm A, and each high dose arm to arm B, in comparisons of Figure 30

Results. None of the trials reported Hb response rates separately for the different epoetin dose arms. Figure 30 compares relative risk of transfusion of the high- versus low-dose arms, which significantly favored the high-dose arm in only one trial (Thatcher, De Campos, Bell, et al., 1999The decrease in relative risk of transfusion was not statistically significant (RR=0.70; 95% CI: 0.40, 1.25; p=0.23) when data from trials that compared identical doses (ten Bokkel Huinink, De Swart, Van Toorn, et al., 1998; Thatcher, De Campos, Bell, et al., 1999) were pooled for meta-analysis (not shown). Each arm is compared to controls in Figure 7 of Key Question 1.

Only the Thatcher, De Campos, Bell, et al. (1999) study evaluated quality of life outcomes, but the only measures utilized were LASA scale items and differences between dose arms were not statistically significant. Only the ten Bokkel Huinink, De Swart, Van Toorn, et al. (1998) study reported thromboembolic complications, which occurred in 9.5% of the high-dose arm and 4.4% of the low-dose arm (RR=2.14; 95% CI: 0.41, 11.10). Kunikane, Watanabe, Fukuoka et al., (2001) reported hypertension was more frequent in the low-dose arm (13.6% versus 9%) while ten Bokkel et al. (1998) reported effects in the opposite direction (2% in the low-dose arm versus 7% in the high-dose arm). However, neither difference was statistically significant.

Figure 30. Weight-Based Epoetin Dose and Relative Risk of Transfusion



Summary. Three trials (combined N=324) suggest that higher initial weight-based doses of epoetin are not more effective than a starting dose of 150 IU/kg three times weekly (the FDA-recommended weight-based initial dose).

KQ2 Comparison II. Different Fixed Doses

A. Darbepoetin

No studies compared different fixed doses of darbepoetin.

B. Epoetin

Characteristics of Available Studies. Five studies randomized 676 patients to one of multiple fixed epoetin doses (i.e., not based on body weight). Of these, only the Cazzola, Messinger, Battistel, et al. (1995) trial included (untreated) controls. Table 42 summarizes doses compared, evaluable sample sizes, and differences in study and population characteristics between these trials. Each trial studied adult patients and all but Sakai, Ohashi, Hirashima, et al. (2004) (reported in two meeting abstracts) were published as full papers. Patients enrolled in the Johansson, Wersall, Brandberg, et al. (2001) study were characterized as elderly.

Table 42. Design and Population Differences of Studies Comparing Fixed Epoetin Doses

Study	lower doses[1]	N eval.[1]	higher doses[1]	N eval.[1]	baseline Hb category	malignancy type	cancer therapy	EPO Tx duration	iron use
Cazzola 1995	1,000 IU/day 2,000	A: 31 B: 29	5,000 IU/day 10,000	C: 31 D: 26	≤10 g/dL	hematologic	unspecified chemotherapy	8 wks	as needed
Glimelius 1998	2,000 IU 3X/wk	41	10,000 IU 3X/wk	43	>10-<12 g/dL	solid tumors	unspecified chemotherapy	18 wks	as needed
Johansson 2001	1,000 IU 3X/wk	90	5,000 IU 3X/wk	90	≤10 g/dL	solid	not reported	12 wks	fixed
Olsson 2002	1,000 IU 3X/wk	90	5,000 IU 3X/wk	90	≤10 g/dL	solid	not reported	24 wks	fixed
Sakai 2004	9,000 IU 1X/wk 18,000	A: 28 B: 29	36,000 IU 1X/wk	C: 29	≤10 g/dL	mixed	unspecified chemotherapy	12 wks	fixed

[1] Letters denoting study arms in Table 42 correspond to letters denoting study arms compared in Figures 31 and 32.

Results. Figure 31 shows likelihood (relative risks) for a Hb response, comparing each pair of epoetin doses in the same multi-arm study. Meta-analysis was not done since doses varied substantially within and across trials. Cazzola, Messinger, Battistel, et al. (1995) studied daily dosing, and reported that raising doses from 1,000 to 5,000 IU daily increases the likelihood of Hb response. However, response likelihood did not change when daily dose increased to 10,000 IU. The two highest doses (arms C and D) are compared with controls in Figure 4 of Key Question 1.

Results of Johansson, Wersall, Brandberg, et al. (2001) and Olsson, Svensson, Sundstrom, et al. (2002) agreed with Cazzola, Messinger, Battistel, et al. (1995) that patients given 5,000 IU thrice weekly were more likely to achieve Hb responses than those given 1,000 IU thrice weekly. Glimelius, Linne, Hoffman, et al. (1998) reported responses are more likely after 10,000 IU thrice weekly than after 2,000 IU thrice weekly. Finally, Sakai, Ohashi, Hirashima, et al. (2004) reported inconsistent dose-response behavior for likelihood of Hb response after single weekly doses of 9,000, 18,000, and 36,000 IU

Figure 31. Fixed Epoetin Dose and Relative Risk of Hematologic Response

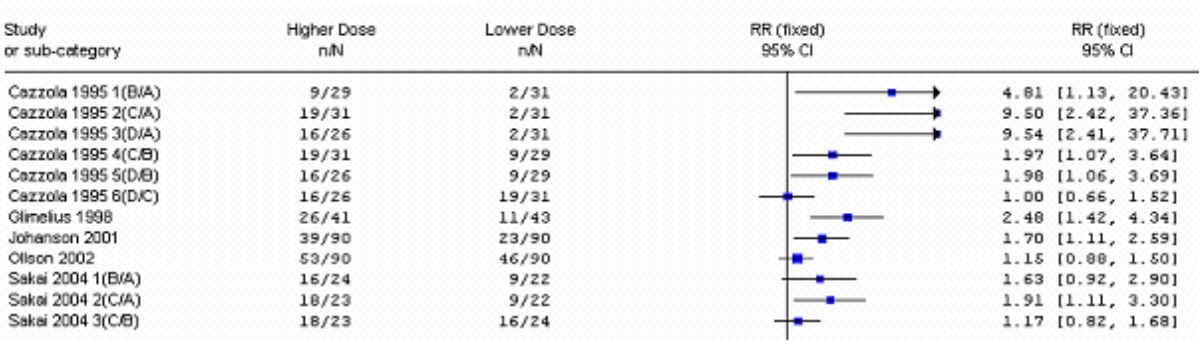
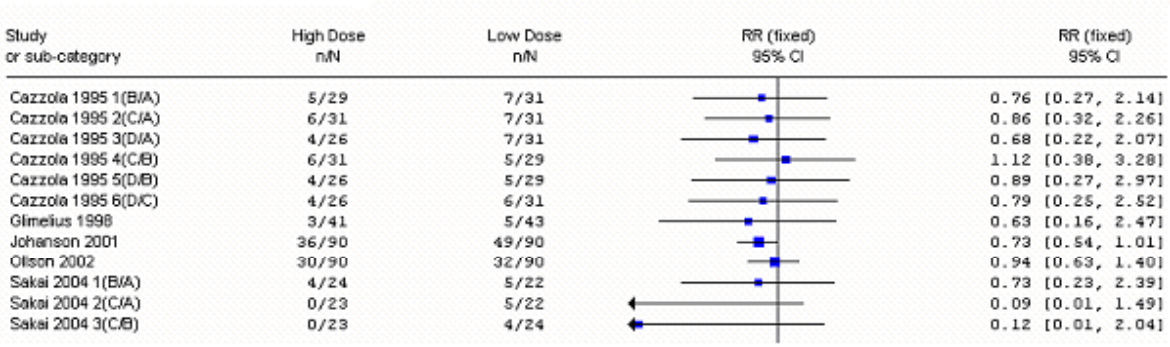


Figure 32 shows data from these trials on relative risks of transfusion at different fixed doses of epoetin. Arms C and D of Cazzola are compared with controls in Figure 7 of Key Question 1. Despite significantly greater likelihood (relative risk) to achieve Hb response, each paired comparison of doses for relative risk of transfusion was not statistically significant. This may be due to small sample sizes and inadequate statistical power in each comparison. Meta-analysis was possible only for the Johansson, Wersall, Brandberg, et al. (2001) and Olsson, Svensson, Sundstrom, et al. (2002) studies, which compared the same two thrice-weekly doses (not shown). Relative risk (5,000 versus 1,000 IU) was not statistically significant (RR=0.81; 95% CI: 0.64, 1.05; p=0.11).

Figure 32. Fixed Epoetin Dose and Relative Risk of Transfusion



Four of these trials assessed quality of life using the EORTC QLQ-C30 (Glimelius, Linne, Hoffman, et al., 1998; Johansson, Wersall, Brandberg, et al. 2001; Olsson, Svensson, Sundstrom, et al., 2002) or FACT-Fatigue (Sakai, Ohashi, Hirashima, et al., 2004) measures. None reported statistically significant increases in QOL scores at the higher epoetin doses.

Two trials reported more frequent thromboembolic events with larger fixed epoetin doses. Glimelius, Linne, Hoffman, et al. (1998) reported events in 14.6% of n=41 treated with 10,000 IU thrice weekly and in 7% of n=43 given 2,000 IU thrice weekly. Johansson, Wersall, Brandberg, et al. (2001) reported events in 12.2% of n=90 given 5,000 IU thrice weekly, and in 4.4% of n=90 given 1,000 IU thrice weekly. However, the relative risk for an event was not statistically significant in either trial (not shown). Both trials also measured effects of epoetin on hypertension at different doses, but reported hypertension was not observed in any patients.

Summary. Increasing fixed daily doses of epoetin from 1,000 to 5,000 IU, or thrice-weekly doses from 1,000-2,000 IU to 5,000-10,000 IU, increased the likelihood of hematologic response. However, the larger doses apparently did not significantly reduce relative risk of transfusion compared with the smaller doses. No trials compared these doses with the weekly fixed dose recommended by FDA (40,000 IU). Two trials reported more frequent thromboembolic complications at the higher doses, but the between-arm differences were not statistically significant.

KQ2 Comparison III. Weight-Based versus Fixed-Dose Regimens

A. Darbepoetin

Characteristics of available study. One trial[26] published as a full paper compared a weight-based versus a fixed-dose regimen of weekly darbepoetin (Hesketh, Arena, Patel, et al., 2004). This study randomized 242 adult patients with mean baseline Hb just above 10 g/dL (10.2±1), undergoing chemotherapy for one of various solid tumors or hematologic malignancies, to 4.5 mcg/kg (n=120) or 325 mcg (n=122) once weekly for 16 weeks. Patients were supplemented with iron at the treating physician's discretion.

Results. Hesketh, Arena, Patel, et al. (2004) defined Hb response to include those who achieved Hb concentrations ≥12 g/dL or a 2 g/dL rise, and reported similar response rates: 84% in the weight-based arm versus 86% in the fixed-dose arm by Kaplan-Meier analysis of time-to-response curves. Transfusion rates were also very similar in the two arms: 18.9% of 122 patients in the fixed-dose arm versus 15.8% of 120 in the weight-based arm (RR=1.19; 95% CI: 0.68, 2.07). Other outcomes were unavailable.

Summary. One RCT (n=242) suggests that outcomes are similar after weight-based (4.5 mcg/kg) or fixed-dose (325 mcg) regimens of once-weekly darbepoetin.

B. Epoetin

Characteristics of available study. One trial published as a full paper compared a weight-based versus a fixed-dose regimen for thrice-weekly treatment with epoetin (Granetto, Ricci, Martoni, et al., 2003). This study randomized 546 adult patients with mean baseline Hb <10g/dL, and with solid tumors undergoing platinum-based chemotherapy, to 150 IU/kg (n=264) or 10,000 IU (n=268) thrice weekly for 12 weeks. Patients were supplemented with iron as needed (transferrin saturation <20%).

Results. The likelihood for hematologic response was similar in both arms: 53% of 230 evaluable in the weight-based arm versus 50.5% of 218 evaluable in the fixed-dose arm (RR=0.95; 95% CI: 0.80, 1.14). Transfusion rates also were similar in the two arms: 12.6% of 238 in the weight-based arm versus 16.4% of 225 in the fixed-dose arm (RR=1.30; 95% CI: 0.84, 2.04). Subgroup analyses comparing the regimens in smaller-sized (45-63 kg) and larger-sized (70-100 kg) patients also found no significant differences in their effects on Hb responses or relative transfusion risks. Granetto, Ricci, Martoni, et al. (2003) also reported no significant differences between regimens in the rates of thromboembolic events or hypertension.

Summary. One RCT (n=546) suggests that outcomes are similar after weight-based (150 IU/kg) or fixed-dose (10,000 IU) regimens of thrice weekly epoetin treatment.

[26] As this report was released, a new trial was published reporting similar outcomes with weekly weight-based (2.25 mcg/kg) versus every third week fixed-dose (500 mcg) darbepoetin (Canon, Vansteenkiste, Bodoky et al., 2006).

KQ2 Comparison IV. More- versus Less-Frequent Dosing

Characteristics of available studies. Two trials[27] investigated fixed-dose epoetin regimens that gave the same total dose as a single bolus or as several fractions over time.

- Cazzola, Beguin, Kloczko, et al. (2003) randomized patients to 30,000 IU/week given as either three (n=122) or one (n=119) injections per week. Treatment duration was 16 weeks.

- Steensma, Molina, Sloan, et al. (2005)[28] randomized patients to 120,000 IU/3 weeks, injected either in three weekly fractions (n=183) or as a single bolus (n=182). Treatment duration was 21 weeks.

Results. Cazzola, Beguin, Kloczko, et al. (2003) reported similar proportions of patients in each arm achieved Hb responses: 85 of 118 (72%) in the arm given 30,000 IU once weekly and 89 of 119 (75%) in the arm given 10,000 IU thrice weekly (RR=0.96; 95% CI: 0.83, 1.12). Steensma, Molina, Sloan, et al. (2005) reported more frequent Hb responses in the arm given 40,000 IU once weekly (128 of 183, 70%) than in the arm given 120,000 once every three weeks (109 of 182, 60%), a statistically significant result (RR=0.86; 95% CI: 0.74, 1.00; p=0.04).

Differences between arms in transfusion rates were not statistically significant in either study, but neither trial was designed to test a non-inferiority hypothesis. Cazzola, Beguin, Kloczko, et al. (2003) reported transfusions in 10 of 115 patients (8.7%) given 30,000 IU once weekly and 16 of 114 (14%) patients given 10,000 IU thrice weekly (RR=0.62; 95% CI: 0.29, 1.31). In Steensma, Molina, Sloan, et al. (2005), 35 of 183 patients (19%) given 40,000 once weekly and 29 of 182 patients (16%) given 120,000 once every three weeks were transfused (RR=0.83; 95% CI: 0.53, 1.30).

Steensma, Molina, Sloan, et al. (2005) measured QoL and reported differences between groups at baseline that favored the every three weeks regimen, although QoL scores at end of study were equivalent in each arm. Thromboembolic complication rates were similar in both arms of Cazzola, Beguin, Kloczko, et al. (2003): 18 of 118 (15%) in the arm given 30,000 IU per week and 21 of 119 (17.7%) in the arm given 10,000 IU thrice weekly (RR=0.86; 95% CI: 0.49, 1.54).

Summary. Two RCTs suggest outcomes are similar with either more- or less-frequent dosing to achieve the same total amount of epoetin per week (N=237) or per three weeks (N=365). While one trial reported significantly more Hb responses with weekly than with every-three-weeks dosing, a second trial found no significant difference between thrice-weekly and once-weekly dosing. Neither trial reported a statistically significant difference in transfusion rates.

[27] As this report was released, a new trial was published reporting similar outcomes with weekly weight-based (2.25 mcg/kg) versus every third week fixed (500 mcg) darbepoetin (Canon, Vansteenkiste, Bodoky et al., 2006).
[28] As this report was released, a full-text version of this trial was published (Steensma, Molina, Sloan et al. 2006).

KQ2 Comparison V. Front-Loaded versus Reduced or Constant Dosing

Characteristics of available studies. Two studies investigated front-loaded regimens of darbepoetin.

- Glaspy, Jadeja, Justice, et al. (2003) randomized patients to 4.5 mcg/kg/week for four weeks, followed by eight more weeks at either 2.25 mcg/kg/week (n=32; arm B), 3 mcg/kg/2 weeks (n=32; arm C) or to 4.5 mcg/kg/week until Hb >12 g/dl, then 1.5 mcg/kg/week until week 12 (n=32; arm A; dose reduced). A control arm (N=31) received 40,000 IU epoetin once weekly.

- Kotasek, Canon, San Miguel, et al. (2004) randomized patients to four weeks of 4.5 mcg/kg/week, then every third week, weeks 5-16 (n=356), or to 2.25 mcg/kg/week (n=367) for 16 weeks.

Results. Glaspy, Jadeja, Justice, et al. (2003) reported no significant differences between arms in Hb response rates, but did not report transfusion rates. Kotasek, Canon, San Miguel, et al. (2004) reported no incremental benefit on Hb response rates from front-loading, and no significant difference between arms in transfusion rates. Both trials measured QoL using FACT scales but did not report clinically meaningful differences in QoL scores between different treatment schedules.

Summary. One trial (n=127) suggests outcomes of different front-loaded darbepoetin regimens are similar to each other; and another trial (n=723) shows outcomes of a front-loaded regimen are similar to outcomes of a constant dose regimen.

KQ2 Comparison VI. Titrated versus Constant Dosing

Characteristics of available study. Osterborg, Boogaerts, Cimino, et al. (1996) compared an initial dose of 2,000 IU/day for eight weeks, increasing to 5,000 IU/day for seven weeks if Hb <11 g/dl, and increasing again to 10,000 IU/day for seven weeks if Hb <11 g/dL at week 12 (n=48), versus a constant dose of 10,000 IU/day until Hb reached 11 g/dL (n=47). Treatment duration was up to 24 weeks.

Results. Differences in hematologic response rates between arms were not statistically significant: 23 of 38 evaluable (60.5%) in the titrated arm compared with 21 of 44 evaluable (47.7%) in the constant-dose arm (RR=1.27; 95% CI: 0.85, 1.90). Differences in transfusion rates also were not statistically significant: 31 of 48 (64.6%) in the titrated arm compared with 27 of 47 (57.5%) in the constant-dose arm (RR=1.12; 95% CI: 0.81, 1.55). Osterborg, Boogaerts, Cimino, et al. (1996) reported two pulmonary emboli and five cases of hypertension in the titrated group, and one pulmonary embolus and four cases of hypertension in the constant-dose group. These differences also were not statistically significant.

Summary. One trial (n=144) suggests outcomes are similar with either titrated or constant-dose regimens of epoetin.

KQ2 Comparison VII. Intravenous versus Subcutaneous Dosing

Characteristics of available study. Justice, Kessler, Jadeja, et al. (2005) compared 4.5 mcg/kg per week of darbepoetin administered subcutaneously versus intravenously (N=60 in each arm). After the first six weeks of treatment, dosing frequency decreased from weekly to every three weeks for the remaining 18 weeks.

Results. Justice et al. (2005) defined Hb response as achieving either a Hb concentration of 12 g/dL or a 2 g/dL increase. Responses were reported in 40 of 59 evaluable (67.8%) in the intravenous arm and 47 of 59 evaluable (79.7%) in the subcutaneous arm. Transfusion rates were similar in the two arms: 21 of 59 (35.6%) in the intravenous arm and 19 of 59 (32.2%) in the subcutaneous arm. Other outcomes were not reported quantitatively; however, the published study report states that adverse event rates were similar in both arms.

Summary. One trial (n=120) suggests outcomes are similar with either the intravenous or subcutaneous routes of darbepoetin administration.

KQ2 Discussion and Conclusions

Of 19 studies included for this Key Question, only one darbepoetin trial, a pilot dose-finding study, included epoetin controls (Glaspy and Tchekmedyian, 2002B). It is uncertain whether the absence of statistically significant between-arm differences reflects small sample size per arm (N=31-33) or true similarity of outcomes. Thus, available evidence is insufficient to determine whether comparative efficacy or safety is altered by changes in dose, schedule, regimen, or route.

Aside from pilot dose-finding studies, the objective of trials that compare different doses, regimens, schedules or routes of administration is to determine whether the alternatives differ in efficacy or safety. For erythropoietic stimulants, comparing hematologic response rates tests whether the alternatives differ in their ability to elicit the predicted physiologic response. Comparing transfusion rates and changes in quality of life tests whether they differ in clinical benefits. Comparing adverse event rates tests whether they differ in safety. Absent differences in clinical benefit or safety, choice between alternatives may be driven by costs, convenience or a balance between these factors.

Except for low-dose arms in some early dose-finding studies, the evidence reviewed here showed no between-arm differences in transfusion rate for any comparisons of different doses, schedules, regimens, or routes of administration. None of the eight studies that reported changes in quality of life measures found a significant difference between the alternatives compared. Differences in thromboembolic event rates, and in rates of one or two other adverse events, also were not statistically significant. However, a minority of trials (six of 19) reported on adverse events. No trials reported costs, and none reported data on amounts of erythropoietic stimulant consumed per patient.[29] Thus, it remains uncertain whether some of the dosing strategies compared in the studies reviewed here might be superior to an alternative with respect to safety

[29] A new trial, published as this report was being released, that compared weekly weight-based (2.25 mcg/kg) versus every third week fixed (500 mcg) darbepoetin, reported data on planned and delivered weekly average doses, weight-adjusted average weekly doses, and mean cumulative doses in each arm (Canon, Vansteenkiste, Bodoky et al., 2006).

or costs. It seems evident, though, that dosing strategies are optimally convenient when they minimize office visits (e.g., every third week in patients undergoing chemotherapy cycles of three weeks each).

The conclusions of this section are as follows:

- With either weight-based or fixed dosing of erythropoietic agents, incremental benefit from doses exceeding those recommended in FDA labeling (or commonly used by clinicians in the U.S.) appears limited. While some trials report modest increases in Hb response rates at higher doses, none report significantly larger reductions in transfusion risks or significantly larger improvements in quality of life.

- Comparisons of weight-based versus fixed-dose regimens; more- versus less-frequent injection schedules; and front-loaded or titrated versus constant-dose regimens showed similar transfusion rates in both arms, including the trials reporting statistically significant but modest differences in Hb response rates.

- Transfusion rates are similar with the subcutaneous or intravenous routes for darbepoetin, although Hb responses may be more frequent after subcutaneous administration.

- Reporting on adverse events is incomplete. Data on thromboembolic events are available for five comparisons, and on other adverse events for 4 comparisons. A minority (six of 19) of trials report on thromboembolic events and a similar minority (six of 19) on other adverse events. Although some trials reported higher rates of thromboembolic events at the highest doses tested, rate differences between arms were not statistically significant.

Key Question 3: How do alternative thresholds for initiating treatment, or alternative criteria for discontinuing therapy or duration of therapy, affect the efficacy and safety of erythropoietic stimulants?

Overview of Evidence and Findings for KQ3

Three unblinded trials, presented at meetings but not yet published, compared treatment with an erythropoietic stimulant at mean hemoglobin concentrations of ~11 g/dL (2 trials) and ~13 g/dL (1 trial) versus treatment only if hemoglobin fell below thresholds of 9 g/dL (1 trial) or 10 g/dL (2 trials) (Table 43). While all received erythropoietic stimulant in one arm of each trial, delayed arm patients were untreated if hemoglobin stayed above threshold. In all trials, delayed therapy was accompanied by higher transfusion rates without statistically significant between-arm differences. One trial reported statistically significant between-arm differences that favored immediate therapy in change from baseline to end of study of FACT quality of life measures; however significantly more thromboembolic events occurred with immediate therapy.

Table 43. Summary of Findings on Thresholds for Initiating Treatment

study	N randomized		malignancy	drug and treatment duration	Hb when EPO/DARB initiated		% patients given EPO/DARB		% transfused		between-arm differences in Δ (FACT measures) from baseline	thrombo-embolic events (%)	
	I[1]	D[1]			I[1]	D[1]	I[1]	D[1]	I[1]	D[1]		I[1]	D[1]
Straus 2003	135	134	hematologic	epoetin 16 weeks	11.1 ± 0.7[3]	<9	100	19.4	17.8	26.1	FACT-An*, FACT-fatigue*	11	3*
Rearden 2004	102	102	mixed solid or hematologic	darbepoetin 12 weeks[2]	11.1 ± 0.7[4]	<10	100	62.7	17.2	26.5	NS	2	1
Crawford 2003	109	107	non-small cell lung cancer	epoetin 16 weeks	13.1 ± 1.0[3]	<10	100	44	12.3	21.0	NS	NR[5]	NR[5]

[1] I = erythropoietic stimulant therapy begun immediately after randomization; D = erythropoietic stimulant therapy delayed until Hb falls to threshold; [2] transfusion data include 22 weeks followup as patients received chemotherapy throughout; [3] mean ± standard deviation; [4] mean ± standard error; [5] "…no differences between groups in frequency or pattern of adverse events…"; * statistically significant difference; NS=no significant difference; NR=not reported

Comparative data are unavailable to determine how the safety and benefits of darbepoetin or epoetin are affected by criteria for discontinuing therapy or duration of therapy (see Table 1, Introduction and Scope, for current recommendations in the FDA-approved package inserts).

Detailed Analysis

A. Alternative Hb Thresholds for Initiating Treatment

This section evaluates outcomes of different thresholds for initiating treatment with an erythropoietic stimulant from RCTs comparing each threshold with treatment initiated immediately after randomization.

Characteristics of Available Studies. Three studies randomized patients to immediate treatment with epoetin or darbepoetin, versus treatment delayed until Hb fell below a threshold. Key aspects of study design and populations are summarized in Table 44.

Table 44. Characteristics of Studies on Thresholds for Initiating Treatment

study	N		malignancy	erythropoietic stimulant	initial dose	treatment duration	baseline Hb		Hb threshold
	I[1]	D[1]					I[1]	D[1]	
Straus 2003	135	134	hematologic	epoetin	40,000 IU weekly	16 weeks	11.1 ± 0.7[2]	11.2 ± 0.6[2]	9 g/dL
Rearden 2004	102	102	mixed solid or hematologic	darbepoetin	300 mcg every third week	12 weeks	11.1 ± 0.7[3]	11.2 ± 0.7[3]	10 g/dL
Crawford 2003	109	107	non-small cell lung cancer	epoetin	40,000 IU weekly	16 weeks	13.1 ± 1.0[2]	13.0 ± 1.2[2]	10 g/dL

[1] I = erythropoietic stimulant therapy begun immediately after randomization; D = erythropoietic stimulant therapy delayed until Hb falls to threshold; [2] mean ± standard deviation; [3] mean ± standard error

- Straus, Testa, Riggs, et al. (2003) randomized adults with hematologic (lymphoid) malignancies to epoetin at mean baseline Hb=11.1 g/dL (n=135), or to epoetin delayed until Hb fell below 9 g/dL (n=134). Epoetin dosage was 40,000 IU weekly for 16 weeks.

- Rearden, Charu, Saidman, et al. (2004) randomized adults to darbepoetin at mean baseline Hb=11.1 g/dL (n=102) or to darbepoetin delayed until Hb fell below 10 g/dL (n=102). Patients were undergoing treatment for hematologic (lymphoid) malignancies or solid tumors (including breast, lung, gastrointestinal, genitourinary, gynecologic, or other cancers). Darbepoetin dosage was 300 mcg every third week for 12 weeks.

- Crawford, Robert, Perry, et al. (2003) randomized adults with non-small cell lung cancer to epoetin treatment at mean baseline Hb=13.1 g/dL (n=109) or to epoetin delayed until Hb fell below 10 g/dL (n=107). Epoetin dosage was 40,000 IU weekly for 16 weeks.

The Crawford, Robert, Perry, et al. (2003) trial enrolled patients with Hb concentrations from 11 to 15 g/dL. In contrast, the other trials enrolled patients with Hb concentrations from 10 (Straus, Testa, Riggs, et al., 2003) or 10.5 (Rearden, Charu, Saidman, et al., 2004) to 12 g/dL. No trial was blinded, placebo-controlled, or specified a transfusion trigger. Crawford, Robert, Perry, et al. (2003) supplemented patients with iron as needed, while the other trials did not report on iron use. Patients in each trial received concurrent chemotherapy, with some in the Rearden, Charu, Saidman, et al. (2004) study and most in the Crawford, Robert, Perry, et al. (2003) study given platinum-based regimens. Since these trials were presented at meetings and published only as abstracts, we could not determine whether randomization methods and allocation concealment were adequate. Each was judged a low-quality trial since they were unblinded and inadequately reported (Appendix Table C55).

Another trial, comparing epoetin initiated at Hb within the normal range versus delayed until Hb fell below 10 g/dL, was excluded since only interim results were reported (Richart, Petruska, Klebert, et al., 2002).

Results. Table 45 compares data from the three available trials on Hb at start of treatment, proportion given an erythropoietic stimulant, and transfusion and thromboembolic event rates.

Table 45. Transfusion and Thromboembolic Event Rates from Trials on Thresholds for Initiating Therapy

study	N evaluated (for transfusion)		drug and treatment duration	Hb when EPO/DARB initiated		% patients given EPO/DARB		% transfused		relative risk of transfusion	95% confidence interval	thrombo-embolic events (%)	
	I[1]	D[1]		I[1]	D[1]	I[1]	D[1]	I[1]	D[1]			I[1]	D[1]
Straus 2003	135	134	epoetin 16 weeks	11.1 ± 0.7[3]	<9	100	19.4	17.8	26.1	0.68	0.43, 1.08	11	3*
Rearden 2004	99	102	darbepoetin 12 weeks[2]	11.1 ± 0.7[4]	<10	100	62.7	17.2	26.5	0.65	0.38, 1.11	2	1
Crawford 2003	106	105	epoetin 16 weeks	13.1 ± 1.0[3]	<10	100	44	12.3	21.0	0.59	0.31, 1.10	NR[5]	NR[5]

[1] I = erythropoietic stimulant therapy begun immediately after randomization; D = erythropoietic stimulant therapy delayed until Hb falls to threshold; [2] transfusion data include 22 weeks followup as patients received chemotherapy throughout; [3] mean ± standard deviation; [4] mean ± standard error; [5] "…no differences between groups in frequency or pattern of adverse events…"; * statistically significant difference; NR=not reported

We did not pool outcome data for meta-analysis because differences in Hb eligibility criteria and in Hb thresholds for treatment of patients in the delayed arms yielded a greater than two-fold range in the proportion untreated with an erythropoietic stimulant across these arms. Each trial compared a different pair of alternatives for initiating treatment.

All patients in the immediate arms were treated with an erythropoietic stimulant, but those in the delayed arms were not if Hb concentration remained above threshold throughout the study (Appendix Table C53). With mean baseline Hb at 11.2 g/dL and 9 g/dL as threshold, 80.6% of patients remained above and were not treated (Straus, Testa, Riggs, et al., 2003). With mean baseline Hb at 11.2 g/dL and 10 g/dL as threshold, fewer (37.3%) remained above and untreated, perhaps because followup and study duration was longer (22 versus 16 weeks) (Rearden, Charu, Saidman, et al., 2004). With mean baseline Hb at 13.0 g/dL and 10 g/dL as threshold, 56% remained above and untreated (Crawford, Robert, Perry, et al., 2003).

Hb responses. Only Rearden, Charu, Saidman, et al., (2004) reported Hb responses as defined for this review (Hb increase $\geq$2 g/dL; see Methods), but some randomized patients were not evaluated. They reported Hb responses in 19 of 94 patients (20.2%) in the arm treated at mean Hb of 11.1 g/dL and 16 of 86 (18.6%) in the arm delayed to a threshold of 10 g/dL (RR=1.09; 95%CI: 0.60, 1.97), a non-significant difference. Straus, Testa, Riggs, et al. (2003) included achieving Hb $\geq$12 in their definition of response, and reported responses in 95 of 135 (70.4%) of those in the arm treated at mean Hb of 11.1 g/dL and 34 of 134 (25.4%) of those in the arm delayed to a threshold of 9 g/dL ($p<0.001$). Crawford, Robert, Perry, et al. (2003) did not report Hb response rates, but Hb concentrations remained above 10 g/dL without transfusion for 82% of those in the arm treated at mean Hb of 13.1 g/dL versus 56% of those in the arm delayed to a threshold of 10 g/dL ($p=0.0001$).

Transfusion rate. Each trial reported fewer transfusions in the arm treated at randomization, although differences were not statistically significant in any trial. Rearden, Charu, Saidman, et al. (2004) reported transfusions in 17 of 99 (17.2%) patients treated with darbepoetin at mean Hb of 11.1 g/dL, and in 27 of 102 (26.5%) treated once Hb fell below 10 g/dL (RR=0.65; 95% CI: 0.38, 1.11). Straus, Testa, Riggs, et al. (2003) transfused 24 of 135 (17.8%) patients treated with epoetin at mean Hb of 11.1, and 35 of 134 (26.1%) patients treated once Hb fell below 9 g/dL (RR=0.68; 95% CI: 0.43, 1.08). Crawford, Robert, Perry, et al. (2003) reported transfusions in 13 of 106 (12.3%) patients treated with epoetin at mean Hb of 13.1, and in 22 of 105 (21.0%) treated after Hb fell below 10 g/dL (RR=0.59; 95% CI: 0.31, 1.10). Transfusion rates were 8% to 9% lower when treatment was initiated upon randomization. The trials were not pooled for meta-analysis for reasons discussed above.

Quality of Life. Each trial compared score changes on FACT scales and/or subscales following therapy with an erythropoietic stimulant at randomization or after Hb fell to a threshold. While the Crawford, Robert, Perry, et al. (2003) study did not show FACT scores, investigators reported "…scores were generally slightly higher throughout the study…" for the group treated at a mean Hb of 13.1 g/dL, versus the group delayed to Hb below 10 g/dL. However, they also noted that between-arm differences in scores and mean changes from baseline were not statistically significant.

The Straus, Testa, Riggs, et al. (2003) epoetin trial reported small positive changes in the physical and functional components of FACT-G, in the FACT-fatigue subscale, and in the

FACT-An total score for patients treated at mean Hb of 11.1 g/dL (Appendix Table C60). In contrast, this study reported small negative changes for each measure in the arm delayed to Hb below 9 g/dL, with statistically significant between-arm differences for each change measure reported. However, absolute changes in either direction were very small compared to baseline score. Whether such small changes are clinically significant is unclear. In this study, baseline data were missing on 13% and 16% of patients randomized to immediate or delayed epoetin respectively. Although there were no losses to followup after the baseline assessment, it is not known how the lost patients affected the baseline comparison of study arm evaluable patient characteristics.

In contrast, Rearden, Charu, Saidman, et al. (2004) reported no statistically significant differences between arms at weeks 13 or 22 in FACT-fatigue change scores, comparing patients given darbepoetin at mean Hb of 11.1 g/dL versus those treated after Hb fell below 10 g/dL (Appendix Table C61). However, data were missing from 13% and 27% of immediate patients (weeks 13 and 22, respectively), and 29% and 49% of delayed patients. It is not known how many patients were missing from the initial FACT-fatigue assessment, and thus how many of these were lost to followup after the initial assessment. If losses to followup were substantial and not random (e.g., if patients with poorer quality of life were more likely to drop out), the results could be significantly biased.

Each trial was published in abstract form only; details of study design, missing data on FACT evaluation, losses to followup, and blinding to Hb concentration at the time of FACT instrument administration to patients were unavailable. Given somewhat inconsistent results between the trials and the lack of detailed information, conclusions are not possible with respect to changes in quality of life.

Thromboembolic events. Crawford, Robert, Perry, et al. (2003) did not report on thromboembolic events, but noted "…no differences between groups in frequency or pattern of adverse events." Straus, Testa, Riggs, et al. (2003) reported 15 undefined thromboembolic events (11% of N=135) in the arm given epoetin at mean Hb of 11.1 g/dL, and only four events (3% of N=134) in the arm with treatment delayed until Hb fell below 9 g/dL (RR=3.72; 95% CI: 1.27, 10.92). Rearden, Charu, Saidman, et al. (2004) reported one case of atrial fibrillation and two cases of deep venous thrombosis, with two of these events in the arm given darbepoetin at mean Hb of 11.1 g/dL (2% of N=99) and one in the arm with treatment delayed until Hb fell below 10 g/dL (1% of N=102).

B. Alternative criteria for discontinuing therapy or duration of therapy

No randomized controlled trials were identified that fulfill the inclusion criteria of this review; thus, no results can be presented.

Duration of therapy: subgroup analyses for Question 1

Meta-analyses on studies of epoetin versus control included for Question 1 were conducted for treatment duration subgroups: 6-9 weeks; 12-16 weeks; and >20 weeks. Results for Hb response rates (Table 46) suggest a greater likelihood of response when treatment is limited to 6-9 weeks, but this result is based on a single study evaluating only 86 patients. The likelihood of response at 12-16 vs. >20 weeks is similar. No conclusions can be reached from these data.

Results for transfusion rates (Table 47) more strongly suggest a difference among treatment duration subgroups, with lower risk of transfusion at the shortest treatment duration. In the meta-regression analysis of transfusion outcomes, the final model included the covariate "duration of treatment."

However, these data cannot be used to reach firm conclusions about differences in treatment duration because the observation periods in these studies were generally the same as treatment duration, thus results from different observation time points are being compared. To answer this question, studies would require different durations of treatment and ideally evaluate outcomes at the same time points during and/or after treatment.

Table 46. Question 1: Treatment Duration Subgroup Analysis for Hb Response

Outcome: Hb response Subgroup: treatment duration	**Epo v Ctl: # Studies**	**# Total Patients**	**#Epo/#Ctl Patients**	**Point Estimate (RR for response)**	**95% CI (p-value)**
Epo tx 6-9 weeks	1	86	57/29	8.91	2.30; 34.50
Epo tx 12-16 weeks	11	2,560	1,376/1,184	3.31	2..91; 3.77
Epo tx >20 weeks	4	647	411/236	3.65	2.62; 5.05
Epo tx ? Weeks					
(Group difference[1])					(0.1509)

Table 47. Question 1: Treatment Duration Subgroup Analysis for Transfusion

Outcome: Transfusion Subgroup: treatment duration	**Epo v Ctl: # Studies**	**#Total Patients**	**#Epo/#Ctl Patients**	**Meta-analysis Point Estimate (RR for transfusion)**	**95% CI (p-value)**
Epo tx 6-9 weeks	5	320	182/138	0.43	0.28; 0.65
Epo tx 12-16 weeks	18	3,189	1,689/1,500	0.64	0.59; 0.69
Epo tx >20 weeks	10	1,329	802/527	0.67	0.60; 0.75
Epo tx ? Weeks	1	372	186/186	0.4	0.23; 0.67
(Heterogeneity)					**(0.0062)**

KQ3 Discussion and Conclusions

Three trials compared treatment with erythropoietic stimulants upon randomization (at varying mean baseline Hb concentrations), versus therapy delayed until Hb fell below threshold, in patient populations undergoing chemotherapy for hematologic malignancies or solid tumors. There were no trials identified outlining criteria for discontinuing therapy or duration of therapy.

In the three trials evaluating thresholds to initiate therapy, mean baseline Hb was ~11 g/dL in two trials, and was ~13 g/dL in the third. Markedly fewer patients received erythropoietic stimulant when treatment was delayed to a threshold than when it began at randomization. All patients in immediate arms of each trial were treated with erythropoietic stimulant. When immediate treatment was initiated at 11.1 g/dL and threshold for delay was 9 g/dL, 19% of delayed patients were treated with erythropoietic stimulant (Straus, Testa, Riggs, et al., 2003); when immediate treatment was initiated at mean Hb of 13.1 g/dL and threshold for delay was 10

g/dL, 44% of delayed patients were treated; and when immediate therapy was at 11.1 g/dL and threshold for delay was 10 g/dL, 63% were treated (Rearden, Charu, Saidman, et al., 2004).

Fewer patients were transfused if treated with erythropoietic stimulant upon randomization than when treatment was delayed until Hb concentration declined below threshold, although between-arm differences were not statistically significant in any trial. The lack of blinding may have biased these results, and the absence of information on a transfusion trigger suggests another potential source of bias. Absolute rates of transfusion were 12%-18% of those randomized to immediate treatment and 21%-26% in delayed arms. Absolute between-arm differences in transfusion rates were 8%-9% in each trial. Thus, in the available trials, treating an additional 37% to 81% of patients with erythropoietic stimulant spared between 8% and 9% additional patients from transfusions.

Thromboembolic events were more frequent with treatment at randomization in two trials, with the difference statistically significant in one of these (delayed to Hb=9; Straus, Testa, Riggs, et al., 2003), while the third trial did not report on thromboembolic events. Note that eligibility for the trial reporting significantly more thromboembolic events in the immediate treatment arm (Straus, Testa, Riggs, et al., 2003) required baseline Hb $\leq$12 g/dL (consistent with current labeling) and mean baseline Hb was ~11 g/dL. In two trials (Rearden, Charu, Saidman, et al., 2004; Crawford, Robert, Perry, et al. 2003), between-arm differences in changes of FACT scores with time were not statistically significant; differences in change scores were statistically significant favoring the arm treated at randomization in a third trial (Straus, Testa, Riggs, et al., 2003).

Indirect comparison of the three available trials did not establish an optimal threshold for initiating therapy with an erythropoietic stimulant. Additionally, since the trials compared three different sets of paired alternatives, it also remained uncertain whether the balance of outcomes favors treatment early in a course of chemotherapy or only after Hb concentration falls to a threshold. We also sought to address these questions using another indirect approach, by comparing outcomes of trials on epoetin or darbepoetin versus control and grouping trials by mean baseline hemoglobin concentration (see Key Question 1). Univariate analysis suggested a larger difference between treated and control arms for trials with mean baseline Hb >10 and <12 g/dL than for trials with mean baseline Hb $\leq$10 g/dL (Figure 7; Table 15). However, multivariate regression analysis suggested this univariate result was confounded by other factors, and that mean baseline Hb was not an independent predictor for the effect of treatment on transfusion rates (Table 17). Additionally, univariate analyses on survival (Figures 9 and 15; Table 25) and thromboembolic events (Figure 18; Table 31) did not suggest significantly different effects of treatment between trials with mean baseline Hb >10 and <12 g/dL, compared with trials with mean baseline Hb $\leq$10 g/dL.

Further trials might determine whether the balance of outcomes from treatment delayed until hemoglobin falls to a threshold is more or less favorable than the balance of outcomes from immediate treatment, and if more favorable, what the optimal threshold might be. However, it is unlikely that baseline hemoglobin is the only factor affecting risks of either transfusion or adverse events from erythropoietic stimulant therapy. For example, each risk varies with tumor type and extent and type of anticancer therapy, while risk of thromboembolic events also depends on previous history of thrombosis, and presence of other factors such as surgery or immobilization (Levine, Lee, Kakkar 2005). Patient-level meta-analyses on trials included in this report, plus systematic reviews of literature unrelated to erythropoietic stimulant intervention

may provide more complete understanding of risks for transfusion and thromboembolic events in cancer patients (see Future Research).

Key Question 4. Are any patient characteristics at baseline or early hematologic changes useful to select patients or predict responses to treatment with erythropoietic stimulants?

Overview of Findings for KQ4

This review included twenty-six cohort studies or randomized clinical trials (total N treated with epoetin or darbepoetin = 10,836) evaluating potential predictive factors measured at baseline (e.g., serum erythropoietin level or observed/predicted ratio (O/P ratio); serum ferritin) or early after starting treatment (e.g., Hb increase, serum ferritin, reticulocyte increase). In general, most of the studies/analyses included fewer than 120 patients; no study defined a refutable hypothesis; and no study was designed to test predictive factors as the primary objective nor used predictive factors prospectively to select treatment. Study quality and reporting was poor to moderate at best with regard to prediction outcomes.

Available evidence does not identify any one factor as clinically useful to select patients or guide treatment decisions; individual factors had mostly weak or no ability to discriminate between responders and non-responders to epoetin or darbepoetin treatment.

Algorithms combining multiple factors, potentially more useful to predict Hb response, are each presently supported only by one exploratory study. Larger studies do not report sufficient predictive ability for any algorithm to establish clinical utility for selecting treatment.

Overview of Studies

Note that methods and materials for Key Question 4 are somewhat different than those for Key Questions 1-3 (see Methods). In particular:

- Randomized and nonrandomized controlled clinical trials; prospective cohort studies; as well as analyses based on data derived from such studies were all allowable. A key inclusion criterion was that the study be designed to prospectively test predictive factors for hematologic response (as defined for Key Question 1) as primary or secondary outcome measures. In addition, the study must have included and reported a direct comparison of patient characteristics at baseline or early hematologic changes in the first four weeks after therapy began for patients responding and not responding to therapy.

- Possible classification systems for predictive factor studies analogous to the different phases of clinical trials evaluating interventions (phase I-IV) have been developed but agreement on a standard system is lacking. Therefore, a 3-level classification system was

developed for this review (see Methods). According to this system, studies are classified as level I-III in Appendix C Table C64. See also Table 48 for definitions.

In addition to study classification, for the purposes of this review, a list of 19 quality assessment criteria was developed based on several proposed quality assessment tools for predictive factor studies (see Methods). All studies were assessed for each criterion as met, not/partially met, or not applicable.

Of 145 potentially relevant studies reviewed in full text, 26 were included in this review for Key Question 4. All but one study (Hedenus, Hansen, Taylor, et al., 2002) treated patients with epoetin. Because some studies did not report the number of patients initially enrolled, the total number enrolled cannot be accurately calculated. Additionally, because the number of patients in the predictive factor analyses was often not available or unclear, numbers of study patients quoted in this key question refer to total who received treatment. McKenzie, Lefebvre, Rosberg, et al. (2004) reported a study in an abstract that used patient data from three prospective cohort studies (per personal communication with Dr. McKenzie): Demetri, Kris, Wade, et al. (1998); Gabrilove, Cleeland, Livingston et al. (2001); and Shasha, George, and Harrison (2003). Patient totals presented here count patients from Demetri, Kris, Wade, et al. 1998, also included in this analysis, only once.

Table 48. Characteristics of Included Studies for KQ4

	Number of studies	Number of patients treated[1]
Study Design		
RCTs	14	2,194
Prospective cohort studies	10	4,802
Phase I/II	1	21
Review of studies	1	5,934[2]
Publication Source		
Full Text	22	6,684[2]
Full Text (Letter)	1	143
Abstract	3	6045[3]
Study classification		
I - exploratory study i.e., no clear statement if possible predictive factors had been defined before the study and/or analysis started, no refutable hypotheses	25	10,662[4]
II - prospective evaluation of potential predictive factors i.e., a restricted set of factors had been defined before the study started, refutable hypotheses	1	174
III - fulfills standards defined by Simon & Altman (1994) or an RCT employing a predictive model in one arm and standard epoetin therapy in the other arm	0	0
% of applicable quality criteria met		
average of all studies	32	
range across studies	16-50	

[1]Patients evaluated for primary outcomes; number evaluated for prediction not always reported.
[2]2,289 patients reported in Demetri, Kris, Wade, et al. 1998 (cohort study) also included in McKenzie, Lefebvre, Rosberg, et al., 2004 review.
[3]Includes 5,934 patients in McKenzie, Lefebvre, Rosberg, et al., 2004 review, 2,289 of which were originally reported in the Demetri 1998 publication.
[4]Patients reported in Demetri, Kris, Wade, et al. 1998 and in McKenzie, Lefebvre, Rosberg, et al., 2004 counted only once.

The total number of treated patients in these studies is approximately 10,836 (many studies did not specify the number evaluable for predictive factors). This estimate also does not include patients that did not receive epoetin or darbepoetin treatment (i.e. controls; patients enrolled but not treated). The number of original study patients ranges from 10 (Garton, Gertz, Witzig, et al., 1995) to 2,289 (Demetri, Kris, Wade, et al., 1998); data from 5,934 patients in 3 different trials were analyzed in the McKenzie, Lefebvre, Rosberg, et al. (2004) review.

Hematologic response was the only outcome assessed in trials reporting on predictive factors. However, the definition of response varied widely between studies. Nine studies (Miller, Platanias, Mills, et al., 1992; Cascinu, Fedeli, Del Ferro, et al., 1994; Garton, Gertz, Witzig, et al., 1995; Glaspy, Bukowski, Steinberg, et al., 1997; Musto, Falcone, D'Arena, et al., 1997; Fjornes, Wiedemann, Sack, et al., 1998; Glimelius, Linne, Hoffman, et al., 1998; Gonzalez, Ordonez, Jua, et al., 1999; Chang, Couture, Young, et al., 2005) did not define Hb increase as >2 g/dL as required in the protocol for this review, but in view of the limited available evidence, these studies were included.

To evaluate study quality systematically, studies were categorized according to study design (primarily prospective cohort study vs. RCT), source, study classification (using the classification system described in Methods and in Table 48), and evaluated for 19 desirable study quality characteristics. Table 48 summarizes the overall results. Most studies were exploratory; only one study (Witzig, Silberstein, Loprinzi, et al., 2004) was classified as phase II, even though it did not report a refutable hypothesis, since it was a RCT with the secondary objective of evaluating a previously published prediction algorithm. Most studies met less than 50% of desirable quality criteria (however, we did not employ a summary score); in general, study quality is poor to moderate, at best.

A large number of potential predictive factors were explored in the included studies. The predictive factors and the number of studies and patients evaluated for each are summarized in Table 49. Few factors were evaluated in more than 5 publications. In addition, in some cases, (e.g., Ludwig, Fritz, Leitgeb, et al., 1994; Gonzalez-Baron, Ordonez, Franquesa, et al., 2002; Littlewood, Zagari, Pallister, et al., 2003) many factors were evaluated within a single study, making it likely that some would be statistically significant by chance alone. Because studies used various statistical methods to evaluate possible predictive factors (univariate, descriptive, multivariate, etc.) comparability of results was limited.

Main Findings

For predicting Hb response, negative predictive value (NPV) is an important parameter; predictive factors should result in a very high NPV for the factor to be clinically useful to identify patients who will not receive treatment because they are so unlikely to respond. Positive predictive value (PPV) should also be high so that the majority of patients selected by the predictive factor for continued treatment with erythropoiesis-stimulating agents would be expected to respond. Where sensitivity and specificity were reported or could be calculated, PPV and NPV were calculated based on the assumption of an overall response rate of 60% (see Hematologic Outcomes in this review).

Table 49. Predictive Factors Measured, Numbers of Studies Reporting, and Numbers of Patients Treated

Predictive factor	Number of studies	Number of patients treated [1]
Measured at baseline		
Serum erythropoietin level	22	6,547
Serum erythropoietin observed/predicted ratio (O/P ratio) as described in Beguin et al. 1992	7	1,125[2]
Serum ferritin	10	1,457
Serum iron	4	267
Serum transferrin	4	324
Serum transferrin saturation	4	872
Soluble transferrin receptor (sTFR)	3	149
Leukocyte count	3	662
Neutrophil count	3	548
Platelet count	6	697
Reticulocyte count	6	822
Serum creatinine	4	457
Creatinine clearance	1	22
Various other factors[3]	4	228
Measured early after initiation of treatment (2, 3, or 4 weeks)		
Hemoglobin/Hematocrit (absolute)	1	117
Hemoglobin/Hematocrit increase	10	9,379
Serum erythropoietin (absolute)	2	197
Serum erythropoietin increase	2	197
Serum ferritin (absolute)	4	932
Serum ferritin increase	2	197
Reticulocyte count (absolute)	1	117
Reticulocyte count increase	4	901
Various other factors[4]	5	1,031
Algorithms	7	22-2030/algorithm

[1]Actual number evaluated for each predictor could not be determined for all trials; patients evaluated in the McKenzie 2004 combined analysis of trials were not included if the original trial populations were already counted.

[2] Some studies used their own controls to establish predicted epo levels (e.g. Musto 1997, Glimelius 1998).

[3]C-reactive protein; interleukin-1 and -6; tumor necrosis factor alpha and beta; neopterin; alpha-1 antitrypsin; interferon gamma; stem cell factor; number of circulating erythropoietic blast-forming units (BFE-E); percent hypochromic erythrocytes; undefined "hemogram," "chemistry," "renal failure."

[4]Increase in: serum neopterin; serum C-reactive protein; serum sTFR; serum transferrin; transferrin saturation; serum iron; alpha-1 antitrypsin; interleukin-1 and -6; tumor necrosis factor; interferon-gamma; stem-cell factor; leukocytes; platelets; reticulocyte hemoglobin. Absolute levels of: serum iron; transferrin; transferrin saturation.

Predictive Factors Measured at Baseline

Measures of endogenous erythropoietin. Of twenty-two studies measuring baseline endogenous erythropoietin levels, thirteen comparing levels of serum erythropoietin in responders compared to non-responders reported no significant difference (Miller, Platanias, Mills, et al., 1992; Case, Bukowski, Carey, et al., 1993; Cascinu, Fedeli, Del Ferro, et al., 1994; Garton, Gertz, Witzig, et al., 1995; Glaspy, Bukowski, Steinberg, et al., 1997; Kasper, Terhaar, Fossa, et al., 1997; Demetri, Kris, Wade, et al. 1998; Glimelius, Linne, Hoffman, et al., 1998; Oberhoff, Neri, Amadori, et al., 1998; Gonzalez-Baron, Ordonez, Franquesa, et al., 2002; Hedenus, Hansen, Taylor, et al., 2002; Katodritou, Speletas, Kapetanos, et al., 2004; Witzig, Silberstein, Loprinzi, et al., 2004)., In contrast, 3 studies making the same comparison reported significant correlations (Ludwig, Fritz, Leitgeb, et al., 1994; Fjornes, Wiedemann, Sack, et al., 1998; Cazzola, Beguin, Kloczko, et al., 2003).

Seven studies tested the use of specific cutoff values of serum erythropoietin to discriminate responders from non-responders; results are shown in Table 50.

Seven studies tested serum erythropoietin observed/predicted ratio (O/P ratio); results are shown in Table 51.

Results for both endogenous erythropoietin and serum O/P levels as predictors of Hb response are in some cases statistically significant. Overall, however, test sensitivities and specificities, where reported, do not result in high enough predictive power to be clinically useful. Many studies enrolled a small number of patients and were likely underpowered for prediction analysis; however, study size or design did not appear to be related to results.

Measures of iron metabolism. Study results testing measures of iron metabolism as predictors of Hb response are shown in Table 52. Studies were mostly small cohorts and were likely underpowered for predictive factor analysis. Only one larger evaluation of RCT results (Littlewood, Zagari, Pallister, et al., 2003) identified a significant predictor of response in ferritin $\leq$400 ng/mL but the resulting predictive power was low.

Cell Counts. Eleven studies evaluated various cell counts as factors possibly predicting Hb response; results are shown in Table 53. Two relatively small RCTs using 100,000/uL platelets as a cutoff found significantly more responders above the cutoff than below the cutoff. However, the differences between groups are relatively small and not likely to be of clinical use. No other studies identified significant cell count predictors of Hb response.

Table 50. Prediction of Hemoglobin Response to Erythropoietic Stimulants Based on Use of a Baseline Serum Erythropoietin Cutoff Value

Study (RCT unless otherwise indicated)	**#Patients treated**	**Serum erythropoietin cutoff tested**	**Comparison of % patients responding below cutoff vs. above**	**Predictive value**
Cazzola 1995	117	<50 IU/L	<cutoff significantly more likely to respond	
Boogaerts 2003	133	<50 IU/L	<cutoff significantly more likely to respond	
Glimelius 1998	99	<50 IU/L	no significant difference	
Osterborg 1996	77	<50 IU/L vs. $\geq$ 400 IU/L	76% vs. 9% response	
Littlewood 2003	561	<100 IU/L	p=0.004	sensitivity = 75%, specificity = 43%, PPV= 66%, NPV = 53%
Ludwig 1994 (cohort)	80	<100 IU/L	no significant difference	
Henry 1995	143	<100 IU/L		sensitivity = 62% specificity = 53% PPV = 66% NPV = 48%

Table 51. Prediction of Hemoglobin Response to Erythropoietic Stimulants Based on Use of a Baseline Serum Erythropoietin Observed to Predicted Ratio (O/P) Cutoff Value

Study (RCT unless otherwise indicated)	**#Patients treated**	**Serum erythropoietin O/P ratio cutoff tested**	**Comparison of % patients responding below cutoff vs. above**
Musto 1997 (cohort)	37	0.8	p=0.001
Glimelius 1998	99	0.8	no significant difference
Cazzola 1995	117	0.8	Patients with O/P ratio <cutoff more likely to respond
Boogaerts 2003	133	0.9	predictive only for patients with solid tumors: RR=1.9; 95% CI, 1.0-3.7; p<0.001
Littlewood 2003	561	0.9	no significant difference
Osterborg 1996	77	N/A	O/P ratio = only significant predictor in multivariate analysis: HR 0.84, p<0.01
Oberhoff 1998	101	N/A	No significant correlation between O/P ratio and response

Table 52. Prediction of Hemoglobin Response to Erythropoietic Stimulants Based on Use of Various Measures of Iron Metabolism

Study (RCT unless otherwise indicated)	#Patients treated	Predictor of Hb response				
		Baseline ferritin	**Baseline iron**	**Baseline serum transferrin**	**Transferrin saturation**	**Soluble transferrin receptor**
Miller 1992 (cohort)	21	NS				
Ludwig 1994 (cohort)	80	NS	NS	NS		NS
Henry 1995	143	Using ferritin cutoff values: 400 ng/mL / 500 ng/mL Sensitivity 60% / 68% Specificity 58% / 56% PPV 68% / 70% NPV 50% / 54%				
Cazzola 1995					NS using a cutoff of 40%	
Osterborg 1996					NS	
Kasper 1997 (cohort)	48	NS	NS	NS		
Musto 1997						O/P ratio cutoff of <0.8: Sensitivity = 92% Specificity = 13% PPV = 61% NPV = 52%
Fjornes 1998 (cohort)	22	NS	NS			
Gonzalez 1999 (cohort)	79	NS		NS		
Gonzalez-Baron 2002 (cohort)	117	NS	NS	NS	NS	
Littlewood 2003	561	Using cutoff value of <400 ng/mL: Significant relationship with greater Hb response in a multivariate regression model (p=0.0002); Sensitivity = 61% Specificity = 50% PPV = 65% NPV = 46%			NS using cutoff values of ≤40% or >20%	
Katodritou 2004 (cohort)	32	NS				NS
Chang 2005	354	NS				

Abbreviations: NS, no significant correlation.

Table 53. Evaluation of Cell Counts as Predictors of Hb Response

Cell type	#Studies, Study type	#Patients treated	Cell count significantly corresponds with responder status?	Significant discrimination of responder status using cutoff value?
Leukocytes[1]	2-cohort 1-RCT	21, 80 561	cohort studies: no	RCT: ≤2000/uL vs. >2000/uL, p=0.2
Neutrophils[2]	3-RCT	117, 77, 354	no	
Platelets[3]	3-cohort 3-RCT	21, 80, 48 117, 77, 354	cohort studies: no RCT (N=354): no	RCTs ≤100,000/uL vs. >100,000/uL: 1) 13% vs. 38% responders, p=0.04 2) 39% vs. 72% responders, p<0.01
Reticulocytes[4]	4-cohort 2-RCT	80, 22, 117, 32 10, 561	cohort studies: no RCT (N=10): no	RCT (N=561): >2.5% vs. ≤2.5%, p=0.6

[1]Miller 1992; Ludwig 1994; Littlewood 2003
[2]Cazzola 1995; Osterborg 1996; Chang 2004
[3]Miller 1992; Cazzola 1995; Ludwig 1994; Osterborg 1996; Kasper 1997; Chang 2004
[4]Ludwig 1994; Garton 1995; Fjornes 1998; Gonzalez-Baron 2002; Littlewood 2003; Katodritou 2004

Measures of renal function. Only one study (Fjornes, Wiedemann, Sack, et al., 1998; cohort study, n=22) reported on creatinine clearance, finding a significant difference in values between responders and non-responders. This study also reported a significant difference in serum creatinine between response groups. However, three RCTs (Cazzola, Messinger, Battistel, et al, 1995, n=117; Osterborg, Boogaerts, Cimino, et al., 1996; n=77; Cazzola, Beguin, Kloczko, et al., 2003; n=241) were unable to confirm serum creatinine as a significant predictor.

Other baseline measures. Of several other factors investigated in four studies (see Table 49, footnote 2), only three showed significant correlation with responder status in one cohort study of 37 patients (Musto, Falcone, D'Arena, et al., 1997): increased number of circulating erythropoietic burst-forming units (BFU-E; p<0.01); decreased levels of interleukin-1 (p<0.001) and tumor necrosis factor (p<0.001). Selected cutoff values for each parameter resulted in 2 groups with the following percentages of responders: 17% vs. 67%; 14% vs. 63%; and 11% vs. 61%, respectively. No additional data confirms the significance of these potential predictors.

Predictive Factors Measured Early After Initiation of Treatment

Hemoglobin/Hematocrit. Ten studies measured the early increase in Hb (and/or equivalent Hct) and determined the correlation with eventual full hematologic response; data were reported in various formats without sufficient information to transform them into a common format. Results are shown in Table 54. Several studies reported some degree of discrimination between eventual hematologic responders and non-responders using specified increases in Hb over 2-4 weeks after start of treatment. However, where sufficient information was available on performance characteristics, at best the results are PPVs of 80-89% and NPVs of 65-71%, which are not likely clinically useful to determine which patients should continue to receive erythropoiesis-stimulating agents and which should not.

Table 54. Prediction of Hemoglobin Response to Erythropoietic Stimulants Based on Use of Early Changes in Serum Components

		Predictor of Hb response					
Study (RCT unless otherwise indicated)	**#Patients treated**	**Early increase in hemoglobin, cutoff = 0.5 g/dL**	**Early increase in hemoglobin, cutoff = 1 g/dL**	**Hemoglobin, other**	**Serum erythropoietin**	**Serum Ferritin**	**Reticulocyte count**
Ludwig 1994 (cohort)	80	Hb increase ≥ vs. <cutoff at 2 wks: R^2=0.39; p<0.001			Increase at 2 wks: r=0.55, p<0.01 Absolute level at 2 wks: r=0.39, p<0.01	<400 ng/mL after 2-4 wks: r=0.32, p<0.01 increase after 2 wks: r=-0.32, p<0.01 absolute level after 2 wks: r=0.37, p<0.02	increase at 2 wks: r=0.28, p<0.05
Henry 1995	143	Hb increase ≥cutoff at 2 wks: 64% responders					increase >40,000/mcL at 2 wks: 59% responders
Glaspy 1997 (cohort)	2,030		Hb increase ≥ vs. <cutoff at 4 wks: 75% vs. 30% responders; Sensitivity = 75%, Specificity = 72%, PPV = 80%, NPV = 65%				
Demetri 1998 (cohort)	2,289		Hb increase ≥cutoff at 4 wks: 81% responders				
Glimelius 1998	99	Hb increase ≥ vs. <cutoff at 2-3 wks: 79% vs. 45% responders; PPV=76%, NPV=58%					

Table 54. Prediction of Hemoglobin Response to Erythropoietic Stimulants Based on Use of Early Changes in Serum Components (continued)

		Predictor of Hb response					
Study (RCT unless otherwise indicated)	**#Patients treated**	**Early increase in hemoglobin, cutoff = 0.5 g/dL**	**Early increase in hemoglobin, cutoff = 1 g/dL**	**Hemoglobin, other**	**Serum erythropoietin**	**Serum Ferritin**	**Reticulocyte count**
Gonzalez-Baron 2002 (cohort)	117	Hb increase ≥ vs. <cutoff at 4 wks: PPV=89%, NPV=71%		no discriminatory ability for absolute Hb or Hct levels measured early after the start of treatment	no discriminatory power for either absolute concentration of or increase in erythropoietin at 2 or 4 wks	no significant discrimination between responders and non-responders for absolute level or increase at 2 or 4 wks	no significant discrimination between responders and non-responders for absolute or count increase at 2 or 4 wks
Littlewood 2003	561	Hb increase ≥ vs. <cutoff at 2 wks: 77% vs. 62% responders, p=0.001	Hb increase ≥ vs. <cutoff at 4 wks: 88% vs. 52% responders, p<0.001; Sensitivity = 59%, Specificity = 82%			≤400 vs. >400 ng/mL at 2 wks: 75% vs. 57% responders, p=0.04	increase > vs. < 0.8% at 2 or 4 wks: 72–73% vs. 61-63%, p=0.016-0.21
Cazzola 2003	117			Hb increase ≥0.1 g/dL at 3 wks, HR=1.05, p<0.05			
Witzig 2004	174		Hb increase ≥ vs. <cutoff at 4 wks: 77% vs. 62% responders; Sensitivity 60%, Specificity 59%			<400 vs. ≥400 ng/mL at 2 wks: 77% vs. 39%, Sensitivity 76% Specificity 63%	
McKenzie 2004 (review)	5,934[1]		Hb increase ≥ vs. <cutoff at 4 wks: 79-84% vs. 44-49% responders, p<0.0001				

[1]Includes 2,289 from Demetri 1998.

Serum erythropoietin. Results of 2 small cohort studies (Table 54) evaluating absolute concentration or increase in serum erythropoietin at 2-4 weeks are either negative or, where positive, correlate only weakly with Hb response.

Serum ferritin. Various measures of early changes in serum ferritin have been tested as predictors of Hb response (Table 54). Two small cohort studies report weak or no correlation between absolute concentration of serum ferritin at 2-4 weeks after start of treatment with Hb response. Two RCTs found that absolute concentration of ferritin <400 ng/mL after 2 weeks predicted a significantly better response, but corresponding predictive values are unlikely to be clinically useful.

Reticulocyte counts. Based on the available evidence (Table 54), the ability of absolute or increased reticulocyte counts to discriminate between responders and non-responders is poor in four studies, and unlikely to be clinically useful for determining which patients should be administered erythropoiesis-stimulating agents.

Other factors. Several other factors, measured early after start of treatment, have been investigated in 5 studies (Ludwig, Fritz, Leitgeb, et al., 1994; Gonzalez-Baron, Ordonez, Franquesa, et al., 2002; Littlewood, Zagari, Pallister, et al., 2003; Cazzola, Beguin, Kloczko, et al., 2003; Katodritou, Speletas, Kapetanos, et al., 2004) for ability to effectively discriminate between Hb responders and non-responders (see Table 49, footnote 3). A small cohort study (Katodritou, Speletas, Kapetanos, et al., 2004; N=32) reported 100% sensitivity and 80% specificity for increase in reticulocyte Hb at 2 weeks after start of treatment; at 60% prevalence of response, PPV would be 88% and NPV 100%. These results might be considered clinically useful to select patients for treatment if confirmed among larger study populations and tested prospectively.

Cazzola, Beguin, Kloczko, et al. (2003; RCT, n=241) investigated various cutoff values for soluble transferrin receptor measured after 2-3 weeks with significant discrimination between eventual responders and non-responders, but modest hazard ratios of 1.6-1.7 and lower confidence limits close to 1. Other potential predictive factors showed either non-significant discriminatory capacity or differences between predictive groups were not sufficiently different to be clinically useful.

Of several other baseline factors investigated, only one (increase in reticulocyte hemoglobin at 2 weeks) showed potentially clinically useful predictive power; no additional data supports these results.

Predictive algorithms. Seven studies reported results for different algorithms attempting to predict which patients will have a hematologic response to erythropoiesis-stimulating agents; results are shown in Table 55 (Ludwig, Fritz, Leitgeb, et al., 1994; Cazzola, Messinger, Battistel, et al., 1995; Henry, Brooks, Case, et al., 1995; Glaspy, Bukowski, Steinberg, et al., 1997; Fjornes, Wiedemann, Sack, et al., 1998; Littlewood, Zagari, Pallister, et al., 2003; Witzig, Silberstein, Loprinzi, et al., 2004). NPVs for response in these studies ranged from 42–90%. PPVs for response ranged from 70-100%. The highest predictive values all came from small cohort studies whereas larger cohort studies and RCTs tended to result in lower predictive values, suggesting that some studies are likely underpowered for testing algorithms.

Table 55. Results of Algorithms Combining Various Parameters to Predict Hb Response

Study	Algorithm predicting response	Algorithm predicting non-response	%Responders meeting response/ non-response criteria	Sensitivity/ Specificity	PPV/NPV (assuming 60% prevalence of Hb response)
Ludwig 1994 cohort N=80	Baseline erythropoietin level < 100 IU/l and/or Hb increase after 2 weeks ≥ 0.5 g/dl	Baseline erythropoietin level ≥ 100 IU/l and Hb increase after 2 weeks < 0.5 g/dl	80% / 6%	76% / 95%	96% / 72%
	Baseline erythropoietin level < 100 IU/l and Hb increase > 0.5 g/dl after 4 weeks	Baseline erythropoietin level ≥ 100 IU/l and/or Hb increase ≤ 0.5 g/dl after 4 weeks	100% / 38%	39% / 100%	100% / 52%
Cazzola 1995 RCT N=117	Step 1: baseline erythropoietin level ≤ 50 IU/L or erythropoietin O/P ratio ≤ 0.9 Step 2: after 2 weeks increase of Hb ≥ 0.3 g/dl	Step 1: baseline erythropoietin level > 50 IU/L or erythropoietin O/P ratio > 0.9 Step 2: after 2 weeks increase of Hb < 0.3 g/dl	Step 1: 75% / 12% Step 2: 88% / 0%	Step 1: 97% / 41% Step 2: 100% / 60%	71% / 90%
Henry 1995 RCT N=143	Hb increase ≥ 0.5 g/dl and reticulocytes increase ≥ 40000/µl after 2 weeks	Hb increase < 0.5 g/dl and reticulocytes increase < 40000/µl after 2 weeks	For response: 67% / 53% For non-response: 52% / 39%	For response: 19% / 88% For non-response: 53% / 57%	For response: 70% / 42% For non-response: 45% / 64%
	Hb increase ≥ 1 g/dl and reticulocytes increase ≥ 40000/µl after 4 weeks	Hb increase < 1 g/dl and reticulocytes increase < 40000/µl after 4 weeks	For response: 84% / 46% For non-response: 64% / 33%	For response: 38% / 91% For non-response: 52% / 77%	For response: 86% / 49% For non-response: 60% / 71%
Glaspy 1997 cohort N=2030	Hb increase after 4 weeks ≥ 1 g/dl and no transfusion requirement during first 4 weeks	Hb increase < 1 g/dl and transfusion requirement during first 4 weeks	For response: 81% / 34% For non-response: 78% / 43%	For response: 62% / 84% For non-response: 17% / 96%	For response: 85% / 60% For non-response: 74% / 63%

Table 55. Results of Algorithms Combining Various Parameters to Predict Hb Response (continued)

Study	Algorithm predicting response	Algorithm predicting non-response	%Responders meeting response/ non-response criteria	Sensitivity/ Specificity	PPV/NPV (assuming 60% prevalence of Hb response)
Fjornes 1998 cohort N=22	Baseline erythropoietin level < 75 IU/l and serum creatinine > upper limit of normal and creatinine clearance < 60 ml/min	Baseline erythropoietin level ≥ 75 IU/l and serum creatinine ≤ upper limit of normal and creatinine clearance ≥ 60 ml/min	100% / 14%	80% / 100%	100% / 77%
Littlewood 2003 RCT N=561	12 algorithms tested/reported incorporating two or three factors per algorithm				(All algorithm results essentially no better than single factors)
	Example (modification of Ludwig 1994): Baseline erythropoietin <u>≤</u>100 mU/mL and Hb increase at week 4 >1.0 g/dL	Baseline erythropoietin >100 mU/mL and Hb increase at week 4 <u>≤</u>1.0 g/dL	88% / 44%	74% / 66%	88% / 44%
Witzig 2004 RCT N=174 (modification of Ludwig 1994 algorithm)	Erythropoietin level < 100 IU/l <u>and/or</u> Hb increase after 4 weeks ≥ 0.5 g/dl	Erythropoietin level ≥ 100 IU/l <u>and</u> Hb increase after 4 weeks < 0.5 g/dl;	72% / 50%	19% / 92%	78% / 43%
	Erythropoietin level < 100 IU/l <u>and</u> Hb increase ≥ 0.5 g/dl after 4 weeks	Erythropoietin level ≥ 100 IU/l <u>and/or</u> Hb increase < 0.5 g/dl after 4 weeks	84% / 55%	60% / 75%	78% / 56%

All studies were exploratory and only one algorithm, originally tested in a small cohort study (Ludwig, Fritz, Leitgeb, et al., 1994) was re-tested in a larger RCT (Ludwig, Fritz, Leitgeb, et al., 1994) with resulting lower predictive power. Littlewood, Zagari, Pallister, et al. (2003) also tested a version of this algorithm, changing the cutoff value for Hb increase after 4 weeks from 0.5 g/dL to 1.0 g/dL, with similarly reduced predictive power. Thus, most algorithms are supported by only one, exploratory, and often small study and results do not indicate sufficient predictive power to be of clinical use in selecting treatment.

Based on the available evidence for individual predictive factors summarized above, it is not possible to identify any single factor as a clinically relevant predictive factor for Hb response. As noted, none of the algorithms tested appear to have sufficient predictive power to warrant further testing. Rather, a comprehensive multivariate analysis of pooled data for individual predictors may be needed to evaluate possible predictive factors for a complex algorithm that meets published quality standards (Simon and Altman, 1994; Concato, Feinstein, and Holford, 1993). Factors to be evaluated might include: baseline Hb, baseline erythropoietin, reticulocytes,

platelets, and Hb increase after 2–4 weeks. In addition, other patient characteristics such as age and tumor type may need to be included.

Several different algorithms for predicting Hb response or non-response have been tested in exploratory studies, but none have been rigorously studied. Based on the available evidence, none have sufficient predictive power to be clinically useful in making treatment decisions.

KQ 4 Discussion and Conclusions

Many individual potential predictive factors, measured at baseline (e.g., serum erythropoietin level; serum erythropoietin observed/predicted ratio (O/P ratio); serum ferritin) or early after the start of treatment (e.g., Hb increase, serum ferritin, reticulocyte increase), were evaluated in 26 studies with mostly weak or no statistical clinical significance and overall poor predictive power for Hb response. Few factors were evaluated by more than 5 studies, and for all studies there are quality limitations to the evidence. Most studies were exploratory and did not identify predictive factors or hypotheses in advance; many were small and likely underpowered. In addition, some studies evaluated a large number of factors within a single study, making it likely that some would be statistically significant by chance alone. Thus, based on the available evidence, it is not possible to identify any single factor as a clinically relevant predictive factor for Hb response that could be used to make treatment decisions.

Predictive algorithms combining multiple factors are potentially more useful for predicting Hb response. Presently, however, most algorithms are supported by only one, exploratory, and often underpowered study. Results from larger studies do not indicate sufficient positive or negative predictive power for any particular algorithm to be of clinical use in selecting treatment and thus do not warrant additional studies. Rather, a comprehensive multivariate analysis of pooled data for individual predictors may be needed to evaluate possible predictive factors for a complex algorithm that meets published quality standards. Factors to be evaluated might include: baseline Hb, baseline erythropoietin, reticulocytes, platelets, and Hb increase after 2-4 weeks. In addition, other patient characteristics such as age and tumor type may need to be included.

Chapter 4. Future Research

The present review incorporates not only published literature, but also abstracts and presentation materials from major specialty meetings through spring of 2005. Research on the use of erythropoietic stimulants to manage cancer therapy-related anemia is ongoing.[30] Following are recommendations for future research.

1. Reporting of adverse events should be complete and consistent in all trials.

The first AHRQ evidence report (Seidenfeld, Aronson, Piper, et al., 2001) found no statistically significant differences in reported adverse events for epoetin compared to controls. Of 22 trials (N=1,927) that reported on efficacy outcomes, nine (N=722) reported on hypertension and 6 (N= 580) on deep vein thrombosis or thromboembolism. However, it is now clear that erythropoietic stimulants do increase the risk of thromboembolic events.

While reporting of adverse events has improved, it is far from complete and consistent. Adverse events should be clearly classified with respect to severity and occurrence, or absence of events explicitly stated in all reports. In the present review, 30 of 48 trials of epoetin versus control reported on thromboembolic events, as did one of four trials of darbepoetin versus control, and three of seven trials comparing epoetin and darbepoetin. Reporting is markedly less consistent for other adverse events. For example, approximately 25 percent of all trials of epoetin reported on thrombocytopenia or hemorrhagic events; no trials of darbepoetin compared to control or epoetin reported on this outcome.

Trials that compare alternative dosing strategies do not adequately address the possibility that risk of adverse events may be greater with some dosing strategies than others. Of nine dosing strategy comparisons addressed in this review (19 trials), reports of thromboembolic events were available for only five strategies (six trials). For other adverse events, data were available for only four comparisons (six trials).

2. Unpublished studies should be made available as full-text publications.

Many of the trials investigating the effects of erythropoietic stimulants on tumor response and survival have not been published as full-text reports. Much of the evidence that suggests detrimental effect on tumor response and survival was available for the present review only from briefing information presented to the Food and Drug Administration Oncologic Drugs Advisory Committee, May 4, 2004. In the absence of the FDA briefing, this important evidence would not have been available to clinicians, researchers and the public.

[30] Appendix G provides a summary of trials listed on clinicaltrials.gov investigating treatment of cancer patients with darbepoetin or epoetin, including some testing effects on survival and disease progression endpoints.

3. The following steps should be taken to improve the quality of evidence available from trials reporting quality of life (QoL) outcomes:

- Methods for evaluating clinically significant change in quality of life measures should be refined and used consistently in all reports to support interpretation of findings. To determine the clinical significance of improvements on the FACT-An and its subscales, a clear, empirically-based estimation of the minimum clinically important difference (MID) is needed for each scale. Because anchor-based approaches are difficult to validate and distribution-based methods are statistical, rather than clinical in nature, methodologists currently recommend that MID should be estimated with more than one well established anchor. Distribution-based methods may supplement but should not substitute for anchor-based methods.

- There should be a consensus among researchers as to the core QoL measures. Even for the FACT instrument, the variety of modules used in the present literature makes it difficult to quantitatively compare results. Use of general QoL measures would assist in interpreting anemia-specific measures.

- Investigators should evaluate change in QoL by comparison of change between study and control arms from RCTs. RCTs should be double-blinded and study protocols should minimize bias in administering QoL measures to patients. Results should be reported as the proportion of patients in each study arm achieving the MID.

- Authors should clearly state, by study arm, the numbers of study participants to which QoL results apply. QoL analyses should clearly identify losses to followup and reasons. Prospectively planned statistical analysis should adequately minimize the impact of losses to followup and explore the impact of alternative assumptions about missing data mechanisms as part of their analysis strategies.

- Investigators should give absolute numbers as well as percentages, with measures of variance, when reporting QoL results.

4. Collect and report economic outcomes, particularly when comparing doses, frequencies of treatment, and alternative dosing strategies.

Economic outcomes were not reported in any of the trials included in this review. Economic data could support the development of strategies to maximize value and reduce cost of using erythropoietic stimulants in the management of cancer-related anemia. The present review found no evidence to show that one dosing strategy was superior to another. If outcomes of alternative regimens are equivalent, lower cost may be the deciding factor in selecting one over another.

5. Additional research on single predictors of response is unlikely to be fruitful. Algorithms combining multiple factors might be more useful, but none tested thus far has been shown to have clinical utility.

6. Systematically review existing evidence on baseline and ongoing risks for transfusion and for adverse events, to individualize clinical decisions.

Clinicians need better information to estimate and balance potential benefits (reduced transfusion risk) versus potential harms (increased risk of serious thromboembolic events; other adverse events) based on individual patient characteristics (e.g., type of malignancy, prior treatment history, current regimen, age, sex, comorbidities, etc.). More complete understanding of risks for transfusion and thromboembolic events in cancer patients could be obtained from a systematic review of literature unrelated to erythropoietic stimulant intervention. A patient-level meta-analysis of completed trials on erythropoietic stimulants could delineate risks in better-described, more homogeneous patient categories; questions regarding risk differences in specific patient populations (e.g., younger versus older adults; women versus men) could also be addressed. A patient-level meta-analysis could also determine whether risk of adverse events increases with increasing exposure to erythropoietic stimulants, particularly in non-responding patients who are given higher doses over time. Synthesis of all this information would support decision analysis to aid clinical decisions.

Currently available evidence is insufficient to compare the balance of risk versus benefit of treatment in children versus adults; more trials are needed in pediatric populations.

Cited References and Included Studies

Abels RI. Use of recombinant human erythropoietin in the treatment of anemia in patients who have cancer. Semin Oncol 1992; 19(3 Suppl 8):29-35.

Acs G, Zhang PJ, Rebbeck TR, et al. Immunohistochemical expression of erythropoietin and erythropoietin receptor in breast carcinoma. Cancer 2002; 95:969-981.

Akaike H. Fitting autoregressive models for prediction. Ann Inst Statist Math 1969; 21(2):243-7.

Alcalay A, Wun T, Khatri V, et al. Venous thromboembolism in patients with colorectal cancer: incidence and effect on survival. J Clin Oncol 2006; 24(7):1112-8.

Alexopoulos CG, Kotsori AA. A randomized comparison of rHuEPO with darbepoetin for cancer related anemia. Ann Oncol 2004; 15(Suppl 3). Abstract.

Altman DG. Systematic reviews of evaluations of prognostic variables. In: Egger M, Smith GD, Altman DG (eds): Systematic Reviews in Health Care - Meta-Analysis in Context. London, BMJ Publishing Group, 2001, pp 228-247.

Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000; 19:453-473.

Amgen, Inc. Aranesp® (darbepoetin alfa) prescribing information. Thousand Oaks, CA. Issue date: March 24, 2006. Available online at http://www.aranesp.com/pdf/aranesp_PI.pdf. Last accessed April 2006.

Amgen, Inc. Epogen® (epoetin alfa) prescribing information. Thousand Oaks, CA. Issue date: October 26, 2005. Available online at http://www.epogen.com/pdf/epogen_pi.pdf. Last accessed March 2006.

Aravantinos G, Linardou H, Makridaki D, et al. Recombinant human erythropoietin for platinum-based chemotherapy-induced anaemia: A single-centre randomised study. J BUON 2003; 8(2):127-32.

Aronson N, Seidenfeld J, Piper MA, et al. Use of erythropoietin in oncology. Evidence report/technology assessment No. 30. Prepared by the Blue Cross and Blue Shield Association Evidence-based Practice Center under Contract No. 290-97-0015. AHRQ Publication No. 01-E009. Rockville (MD): Agency for Healthcare Research and Quality; June 2001.

Bamias A, Aravantinos G, Kalofonos C, et al. Prevention of anemia in patients with solid tumors receiving platinum-based chemotherapy by recombinant human erythropoietin (rHuEpo): A prospective, open label, randomized trial by the Hellenic Cooperative Oncology Group. Oncology 2003; 64(2):102-10.

Besarab A, Bolton WK, Browne JK, et al. The effects of normal as compared with low hematocrit values in patients with cardiac disease who are receiving hemodialysis and epoetin. N Engl J Med 1998; 339(9):584-90.

Bircan A, Berktas B, Bayiz H, et al. Effects of chemotherapy on quality of life for patients with lung cancer. Turkish Respiratory Journal 2003; 4:61-66.

Blajchman MA, Hebert PC. Red blood cell transfusion strategies. Transfus Clin Biol 2001; 8(3):207-10.

Bohlius J, Langensiepen S, Schwarzer G, et al. Erythropoietin for patients with malignant disease. Cochrane Database Syst Rev 2004;(3):CD003407.

Bohlius J, Langensiepen S, Schwarzer G, et al. Recombinant human erythropoietin and overall survival in cancer patients: results of a comprehensive meta-analysis. J Natl Cancer Inst 2005; 97(7):489-98.

Bokemeyer C, Aapro MS, Courdi A, et al. EORTC guidelines for the use of erythropoietic proteins in anaemic patients with cancer. Eur J Cancer 2004; 40(15):2201-16.

Bokemeyer C, Oechsle K, Hartmann JT, et al. Treatment-induced anaemia and its potential clinical impact in patients receiving sequential high dose chemotherapy for metastatic testicular cancer. Br J Cancer 2002; 87:1066-1071.

Boogaerts M, Coiffier B, Kainz C. Epoetin Beta QOL Working Group. Impact of epoetin beta on quality of life in patients with malignant disease. Br J Cancer 2003; 88(7):988-95.

Boracchi P, Biganzoli E. Markers of prognosis and response to treatment: ready for clinical use in oncology? A biostatistician's viewpoint. Int J Biol Markers 2003; 18(1):65-9.

Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003; 138:40-44.

Busch MP, Kleinman SH, Nemo GJ. Current and emerging infectious risks of blood transfusions. JAMA 2003; 289(8):959-62.

Canon JL, Vansteenkiste J, Bodoky G, et al. Randomized, double-blind, active-controlled trial of every-3-week darbepoetin alfa for the treatment of chemotherapy-induced anemia. J Natl Cancer Inst 2006;98(4):273-84.

Carabantes FJ, Benavides M, Trujillo R, et al. Epoetin alfa in the prevention of anemia in cancer patients undergoing platinum-based chemotherapy (CT). A prospective randomized study. Proc Am Soc Clin Oncol 1999;18:596a. Abstract 2303.

Caro JJ, Salas M, Ward A, et al. Anemia as an independent prognostic factor for survival in patients with cancer: a systemic, quantitative review. Cancer 2002; 91(12):2214-21.

Cascinu S, Fedeli A, Del Ferro E, et al. Recombinant human erythropoietin treatment in cisplatin-associated anemia: a randomized, double-blind trial with placebo. J Clin Oncol 1994; 12:1058-62.

Case DC, Bukowski RM, Carey RW, et al. Recombinant human erythropoietin therapy for anemic cancer patients on combination chemotherapy. J Natl Cancer Inst 1993; 85(10):801-6.

Cazzola M, Beguin Y, Kloczko J, et al. Once-weekly epoetin beta is highly effective in treating anaemic patients with lymphoproliferative malignancy and defective endogenous erythropoietin production. Br J Haematol 2003; 122(3):386-93.

Cazzola M, Messinger D, Battistel V, et al. Recombinant human erythropoietin in the anemia associated with multiple myeloma or non-Hodgkin´s lymphoma: dose finding and identification of predictors of response. Blood 1995; 86(12):4446-53.

Cella D, Eton DT, Lai JS, et al. Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. J Pain Symptom Manage 2002; 24(6):547-61.

Chang J, Couture F, Young S, et al. Weekly epoetin alfa maintains hemoglobin, improves quality of life, and reduces transfusion in breast cancer patients receiving chemotherapy. J Clin Oncol 2005; 23(12):2597-605.

Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. Ann Intern Med 1993; 118:201-210.

Crawford J, Robert F, Perry M, et al., for the Anemia Prevention in NSCLC Study Group: Epoetin alfa 40,000 IU once weekly maintains hemoglobin in advanced non-small-cell lung cancer patients receiving first-line chemotherapy. Proc Am Soc Clin Oncol 2003; 22:628, abstract 2527.

Abstract and full poster available online at http://www.asco.org/portal/site/ASCO/menuitem.34d60f5624ba07fd506fe310ee37a01d/?vgnextoid=76f8201eb61a7010VgnVCM100000ed730ad1RCRD&vmview=abst_detail_view&confID=23&abstractID=101646. Last accessed March 2006.

Dammacco F, Castoldi G, Rodjer S. Efficacy of epoetin alfa in the treatment of anaemia of multiple myeloma. Br J Haematol 2001; 113(1):172-9.

Deicher R, Horl WH. Differentiating factors between erythropoiesis-stimulating agents: a guide to selection for anaemia of chronic kidney disease. Drugs 2004; 64(5):499-509.

Del Mastro L, Venturini M, Lionetto R, et al. Randomized phase III trial evaluating the role of erythropoietin in the prevention of chemotherapy-induced anemia. J Clin Oncol 1997; 15(7):2715-21.

Demetri GD, Kris M, Wade J, et al. Quality-of-life benefit in chemotherapy patients treated with epoetin alfa is independent of disease response or tumor type: results from a prospective community oncology study. Procrit Study Group. J Clin Oncol 1998; 16(10):3412-25.

Donaldson GW, Moinpour CM. Learning to live with missing quality-of-life data in advanced-stage disease trials. J Clin Oncol 2005; 23(30):7380-7384.

Dunphy FR, Harrison BR, Dunleavy TL, et al. Erythropoietin reduces anemia and transfusions. Cancer 1999; 86:1362-7.

Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315:629-34.

Engels EA, Schmid CH, Terrin N, et al. Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses. Stat Med 2000; 19:1707-28

Engert A. Recombinant human erythropoietin as an alternative to blood transfusion in cancer-related anaemia. Dis Manag Clin Outcomes 2000; 8(5):259-72.

EPO-CAN-15. Protocol EPO-CAN-15: "A randomized, double-blind placebo controlled study to evaluate the impact of maintaining haemoglobin levels using EPREX (Epoetin alfa) in limited disease small cell lung cancer (LD SCLC) patients receiving combined chemotherapy and radiation therapy." (The LEGACY Trial). Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/4037s2.htm). Online sources last accessed July 26, 2005.

EPO-CAN-20. "A randomized trial of epoetin alfa in patients with advanced non-small cell carcinoma of the lung." Reported in Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

EPO-GBR-07. Protocol EPO-GBR-7: "Open-label randomized, comparative group evaluation of the effect of epoetin alfa on local disease-free survival and quality of life in head and neck cancer patients receiving radical radiotherapy." Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

Estrin JT, Schocket L, Kregenow R, et al. A retrospective review of blood transfusions in cancer patients with anemia. Oncologist 1999; 4(4):318-24.

Fallofield L, Gagnon D, Zagari M, et al. Multivariate regression analyses of data from a randomised, double-blind, placebo-controlled study confirm quality of life benefit of epoetin alfa in patients receiving non-platinum chemotherapy. Br J Cancer 2002; 87:1341-53.

Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC). Recommandations pour la pratique clinique: Standards, Options et Recommandations 2003 pour l'utilisation de l'érythropoïétine recombinante (époetine alfa et bêta, darbepoetine alfa, EPO) dans la prise en charge de l'anémie en cancérologie, mise à jour. Federation Nationale des centres de lutte contre le cancer. Paris (France): FNCLCC; November 2003. Available online at http://www.fnclcc.fr/sor/SORSpecialistes/Traitements/ Medicaments/Util-EPO/TT_MED_Util-EPO_int/html/pdf/TT_MED_Util_EPO_int.pdf. Last accessed March 2006.

Fjornes T, Wiedemann GJ, Sack K, et al. Serum erythropoietin and creatinine concentrations as predictive factors for response to recombinant human erythropoietin treatment in anaemic tumour patients on chemotherapy. Oncol Rep 1998; 5(1):81-6.

Food and Drug Administration Oncologic Drugs Advisory Committee. May 4, 2004, Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

Gabrilove JL, Cleeland CS, Livingston RB, et al. Clinical evaluation of once-weekly dosing of epoetin alfa in chemotherapy patients: improvements in hemoglobin and quality of life are similar to three-times-weekly dosing. J Clin Oncol 2001; 19(11):2875-82.

Garton JP, Gertz MA, Witzig TE, et al. Epoetin alfa for the treatment of the anemia of multiple myeloma. Arch Intern Med 1995; 155:2069-74.

Glaspy JA. The potential for anemia treatment to improve survival in cancer patients. Oncology (Williston Park) 2002; 16(9 Suppl 10):35-40.

Glaspy J, Berg R, Tomita D, et al. Final results of a phase 3, randomized, open-label study of darbepoetin alfa 200 mcg every 2 weeks (Q2W) versus epoetin alfa 40,000 U weekly (QW) in patients with chemotherapy-induced anemia (CIA). Proc Am Soc Clin Oncol 2005; 23:760s. Abstract 8125.

Glaspy J, Bukowski R, Steinberg D, et al. Impact of therapy with epoetin alfa on clinical outcomes in patients with nonmyeloid malignancies during cancer chemotherapy in community oncology practice. Procrit Study Group. J Clin Oncol 1997; 15(3):1218-34.

Glaspy JA, Jadeja JS, Justice G, et al. A randomized, active-control, pilot trial of front-loaded dosing regimens of darbeopetin-alfa for the treatment of patients with anemia during chemotherapy for malignant disease. Cancer 2003; 97:1312-20.

Glaspy JA, Jadeja JS, Justice G, et al. Darbepoetin alfa given every 1 or 2 weeks alleviates anaemia associated with cancer chemotherapy. Br J Cancer 2002; 87(3):268-76.

Glaspy JA, Tchekmedyian NS. Darbepoetin alfa administered every 2 weeks alleviates anemia in cancer patients receiving chemotherapy. Oncology (Huntingt) 2002B; 16(10 Suppl 11):23-9.

Glimelius B, Linne T, Hoffman K, et al. Epoetin beta in the treatment of anemia in patients with advanced gastrointestinal cancer. J Clin Oncol 1998; 16(2):434-40.

GOG-191. Study GOG-191 (PR01-04-005): "A phase III trial to evaluate the efficacy of maintaining Hgb levels above 120 g/l with erythropoietin versus above 100 g/l without erythropoietin in anemic patients receiving concurrent radiation and cisplatin for cervical cancer." Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

Gonzalez M, Ordonez A, Jua VO, et al. Predicting factors to epoetin alfa treatment. Proc ASCO 1999. Abstract 2308.

Gonzalez-Baron M, Ordonez A, Franquesa R, et al. Response predicting factors to recombinant human

erythropoietin in cancer patients undergoing platinum-based chemotherapy. Cancer 2002; 95(11):2408-13.

Goodnough LT. Risks of blood transfusion. Anesthesiol Clin North Am 2005; 23(2):241-52.

Granetto C, Ricci S, Martoni A, et al. Comparing the efficacy and safety of fixed versus weight-based dosing of epoetin alpha in anemic cancer patients receiving platinum-based chemotherapy. Oncol Rep 2003; 10(5):1289-96.

Groopman JE, Itri LM. Chemotherapy-induced anemia in adults: incidence and treatment. JNCI 1999; 91(19):1616-34.

Grote T, Yeilding AL, Castillo R, et al. Efficacy and safety analysis of epoetin alfa in patients with small-cell lung cancer: a randomized, double-blind, placebo-controlled trial. J Clin Oncol 2005; 23(36):9377-86.

Gurevitch J, Hedges LV. Statistical issues in ecological meta-analyses. Ecology, 1999; 80:1142–1149.

Guyatt GH, Norman GR, Juniper EF, et al. A critical look at transition ratings. J Clin Epidemiol 2002; 55:900-8

Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. Mayo Clin Proc, 2002; 77:371-83.

Halstenson CE, Macres M, Katz SA, et al. Comparative pharmacokinetics and pharmacodynamics of epoetin alfa and epoetin beta. Clin Pharmacol Ther 1991; 50(6):702-12.

Hedenus M, Adriansson M, San Miguel J, et al. Efficacy and safety of darbepoetin alfa in anaemic patients with lymphoproliferative malignancies: a randomized, double-blind, placebo-controlled study. Br J Haematol 2003; 122(3):394-403.

Hedenus M, Hansen S, Taylor K, et al. Randomized, dose-finding study of darbepoetin alfa in anaemic patients with lymphoproliferative malignancies. Br J Haematol 2002; 119:79-86.

Hedenus M, Vansteenkiste J, Kotasek D, et al. Darbepoetin alfa for the treatment of chemotherapy-induced anemia: disease progression and survival analysis from four randomized, double-blind, placebo-controlled trials. J Clin Oncol 2005; 23(28):6941-8.

Hedges LV, Olkin I. Vote counting methods in research synthesis. Psychol Bull 1980; 88:359-369.

Henke M, Guttenberger R, Barke A, et al. Erythropoietin for patients undergoing radiotherapy: a pilot study. Radiother Oncol 1999; 50(2):185-90.

Henke M, Laszig R, Ruebe C, et al. Erythropoietin to treat head and neck cancer patients with anaemia undergoing radiotherapy: randomised, double-blind, placebo-controlled trial. Lancet 2003; 362:1255-60.

Henry D, Abels R, Larholt K. Prediction of response to recombinant human erythropoietin (r-HuEPO/epoetin-alpha) therapy in cancer patients. Blood 1995a; 85:1676-1678.

Henry DH, Brooks BJ Jr, Case DC Jr, et al. Recombinant human erythropoietin therapy for anemic cancer patients receiving cisplatin chemotherapy. Cancer J Sci Am 1995; 1(4):252-60.

Henze G, Michon J, Morland B, et al. Phase III randomized study: efficacy of epoetin alfa in reducing blood transfusions in newly diagnosed pediatric cancer patients receiving chemotherapy. Proc Am Soc Clin Oncol 2002; 21. Abstract 1547.

Hesketh PJ, Arena F, Patel D, et al. A randomized controlled trial of darbepoetin alfa administered as a fixed or weight-based dose using a front-loading schedule in patients with anemia who have nonmyeloid malignancies. Cancer 2004; 100(4):859-68.

Higgins J, Thompson S, Deeks J, et al. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy, 2002; 7(1):51-61

Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ, 2003; 327:557-60

Hollander N, Schumacher M. On the problem of using "optimal" cutpoints in the assessment of quantitative prognostic factors. Onkologie 2001; 24:194-199.

Holzner B, Kemmler G, Greil R, et al. The impact of hemoglobin levels on fatigue and quality of life in cancer patients. Ann Oncol 2002; 13: 965–973.

Huddart RA, Welch RS, Chan S, et al. A prospective randomised comparative-group evaluation of epoetin alfa for the treatment of anaemia in UK cancer patients receiving platinum-based chemotherapy. Ann Oncol 2002; 13 (Suppl 5):177.

Iconomou G, Koutras A, Rigopoulos A, et al. Effect of recombinant human erythropoietin on quality of life in cancer patients receiving chemotherapy: results of a randomized, controlled trial. J Pain Symptom Manage 2003; 25(6):512-8.

Ihaka R, Gentleman R. A language for data analysis and graphics. J Comput Graph Stat 1996; 5(3):299-314.

Infante-Rivard C, Villeneuve JP, Esnaola S. A framework for evaluating and conducting prognostic studies: an application to cirrhosis of the liver. J Clin Epidemiol 1989; 42(8):791-805.

INT-1. Study of epoetin supportive care in ovarian cancer. Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing

Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

INT-3. Study of epoetin supportive care in multiple myeloma. Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

Janinis J, Dafni U, Aravantinos G, et al. Quality of life (QoL) outcome of epoietin-alfa (EPO-A) in anemic cancer patients undergoing platinum or non-platinum-based chemotherapy: a randomized study conducted by the Hellenic Cooperative Oncology Group. Proc Am Soc Clin Oncol 22, 789. 2003. Abstract.

Johansson JE, Wersall P, Brandberg Y, et al. Efficacy of epoetin beta on hemoglobin, quality of life, and transfusion needs in patients with anemia due to hormone-refractory prostate cancer--a randomized study. Scand J Urol Nephrol 2001; 35(4):288-94.

Jones M, Schenkel B, Just J, et al. Epoetin alfa improves quality of life in patients with cancer. Results of a meta-analysis. Cancer 2004; 101:1720-32

Joy MS. Darbepoetin alfa: a novel erythropoiesis-stimulating protein. Ann Pharmacother 2002; 36:1183-92.

Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999; 130:515-524.

Justice G, Kessler JF, Jadeja J, et al. A randomized, multicenter study of subcutaneous and intravenous darbepoetin alfa for the treatment of chemotherapy-induced anemia. Ann Oncol 2005; 16(7):1192-1198.

Kasper C, Terhaar A, Fossa A, et al. Recombinant human erythropoietin in the treatment of cancer-related anaemia. Eur J Haematol 1997; 58(4):251-6.

Katodritou E, Speletas M, Kapetanos D, et al. Evaluation of traditional and novel predictive markers of anemia response to recombinant human erythropoietin (rh-EPO) in patients with multiple myeloma and lymphoma: emerging role of functional iron deficiency. Blood 2004; 104(11). Abstract 3690.

Khorana AA, Francis CW, Culakova E, et al. Risk factors for chemotherapy-associated venous thromboembolism in a prospective observational study. Cancer 2005; 104(12):2822-9.

Knight K, Wade S, Balducci L. Prevalence and outcomes of anemia in cancer: a systematic review of the literature. Am J Med 2004; 116(Suppl. 7A):11S-26S.

Koeller JM. Clinical guidelines for the treatment of cancer-related anemia. Pharmacotherapy 1998; 18(1):156-69.

Kotasek D, Canon JL, San Miguel J, et al. Correction/maintenance dosing (front loading) of darbepoetin alfa: final results from a randomized phase 3 active controlled trial. Proceedings of the American Society of Hematology. Blood 2004; 104:455a, abstract number 1636.

Kotasek D, Steger G, Faught W, et al. Darbepoetin alfa administered every 3 weeks alleviates anaemia in patients with solid tumours receiving chemotherapy; results of a double-blind, placebo-controlled, randomised study. Eur J Cancer 2003; 39(14):2026-34.

Kunikane H, Watanabe K, Fukuoka M, et al. Double-blind randomized control trial of the effect of recombinant human erythropoietin on chemotherapy-induced anemia in patients with non-small cell lung cancer. Int J Clin Oncol 2001; 6:296-301.

Kurz C, Marth C, Windbichler G, et al. Erythropoietin treatment under polychemotherapy in patients with gynecologic malignancies: a prospective, randomized, double-blind placebo-controlled multicenter study. Gynecol Oncol 1997; 65(3):461-6.

L'Abbe KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. Ann Intern Med 1987; 107(2):224-33.

Levine MN, Lee AY, Kakkar AK. Thrombosis and cancer. American Society of Clinical Oncology Educational Book. MC Perry, ed. Alexandria, VA; 2005: American Society of Clinical Oncology. pp 748-57.

Leyland-Jones B. Breast cancer trial with erythropoietin terminated unexpectedly. Lancet Oncol 2003 Aug; 4(8):459-60.

Leyland-Jones B, Semiglazov V, Pawlicki M, et al. Maintaining normal hemoglobin levels with epoetin alfa in mainly nonanemic patients with metastatic breast cancer receiving first-line chemotherapy: A survival study. J Clin Oncol 2005; 23:5960-5972.

Littlewood TJ, Bajetta E, Nortier JW, et al. Effects of epoetin alfa on hematologic parameters and quality of life in cancer patients receiving nonplatinum chemotherapy: results of a randomized, double-blind, placebo-controlled trial. J Clin Oncol 2001; 19(11):2865-74.

Littlewood TJ, Kallich JD, San Miguel J, et al. Efficacy of darbepoetin alfa in alleviating fatigue and the effect of fatigue on quality of life in anemic patients with lymphoproliferative malignancies. J Pain Symptom Manage 2006; 31:317-25.

Littlewood TJ, Zagari M, Pallister C, et al. Baseline and early treatment factors are not clinically useful for predicting individual response to erythropoietin in anemic cancer patients. Oncologist 2003; 8(1):99-107.

Ludwig H, Fritz E, Leitgeb C, et al. Prediction of response to erythropoietin treatment in chronic anemia of cancer. Blood 1994; 84:1056-63.

Ludwig H, Sundal E, Pecherstorfer M, et al. Recombinant human erythropoietin for the correction of cancer associated anemia with and without concomitant cytotoxic chemotherapy. Cancer 1995; 76(11):2319-29.

Luksenburg H, Weir A, Wagner R. FDA Briefing Document May 4, 2004 Oncologic Drugs Advisory Committee: Safety Concerns Associated with Aranesp (darbepoetin alfa) Amgen, Inc. and Procrit (epoetin alfa) Ortho Biotech, L.P., for the Treatment of Anemia Associated with Cancer Chemotherapy. Available online at http://www.fda.gov/ohrms/dockets/ac/04/briefing/4037B2_04_FDA-Aranesp-Procrit.htm. Last accessed January 2006.

Machtay M, Pajak T, Suntharalingam M, et al. Definitive radiotherapy erythropoietin for squamous cell carcinoma of the head and neck: preliminary report of RTOG 99-03. Proceedings from the 46th annual meeting of the American Society for Therapeutic Radiology and Oncology. Atlanta GA. 2004. Abstract 5.

McEvoy GK, editor. Epoetin alfa. In: AHFS Drug Information 2005. Bethesda: American Society of Health-system Pharmacists; 2005. p. 1486-98.

McGuire WL. Breast cancer prognostic factors: evaluation guidelines. J Natl Cancer Inst 1991;83(3):154-5.

McKenzie RS, Lefebvre P, Rosberg J, et al. Comparison of anemia outcomes associated with early response to epoetin alfa (EPO) with QW and TIW dosing. J Clin Oncol 2004; 22 (14S). Abstract 6141.

Miller CB, Platanias LC, Mills SR, et al. Phase I-II trial of erythropoietin in the treatment of cisplatin-associated anemia. J Natl Cancer Inst 1992; 84(2):98-103.

Musto P, Falcone A, D'Arena G, et al. Clinical results of recombinant erythropoietin in transfusion-dependent patients with refractory multiple myeloma: role of cytokines and monitoring of erythropoiesis. Eur J Haematol 1997; 58(5):314-9.

N93 004. Final Study Report, Protocol N93-004 ("The Effect of r-HuEPO in Patients with Small Cell Cancer (SCLC): A Randomized, Double blind, Placebo-controlled Trial), Johnson & Johnson, October 25, 2002. Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/4037s2.htm). Online sources last accessed July 26, 2005.

National Cancer Institute (NCI) Cancer Therapy Evaluation Program (CTEP). Common toxicity criteria (Version 2.0). Published April 30, 1999. Available online at: http://ctep.cancer.gov/forms/CTCv20_4-30-992.pdf. Last accessed January 2006.

National Institute for Health and Clinical Excellence. Appraisal Consultation Document: Erythropoietin for anaemia induced by cancer treatment. June 2005. Available online at http://www.nice.org.uk/page.aspx?o=263492. Last accessed March 2006.

NCCN (National Comprehensive Cancer Network). Practice Guidelines in Oncology: Cancer- and Treatment-Related Anemia. Available online at www.nccn.org/professionals/physician_gls/PDF/anemia.pdf. Last accessed January 25, 2006.

Nieboer P, Buijs C, Rodenhuis S, et al. Fatigue and relating factors in high-risk breast cancer patients treated with adjuvant standard or high-dose chemotherapy: a longitudinal study. J Clin Oncol 2005; 23(33):8296-8304.

Oberhoff C, Neri B, Amadori D, et al. Recombinant human erythropoietin in the treatment of chemotherapy- induced anemia and prevention of transfusion requirement associated with solid tumors: a randomized, controlled study. Ann Oncol 1998; 9(3):255-60.

Okuyama T, Akechi T, Kugaya A et al. Factors correlated with fatigue in disease-free breast cancer patients: application of the Cancer Fatigue Scale. Support Care Cancer 2000; 8: 215–222.

Olsson AM, Svensson JH, Sundstrom J, et al. Erythropoietin treatment in metastatic breast cancer--effects on Hb, quality of life and need for transfusion. Acta Oncologica 2002; 41(6):517-24.

Ortho-Biotech Products. Procrit® (epoetin alfa) prescribing information. Raritan, NJ. Issue date: November 2005. Available online at http://www.procrit.com/common/prescribing_information/PROCRIT/PDF/ProcritBooklet.pdf. Last accessed March 2006.

O'Shaughnessy JA. Effects of epoetin alfa on cognitive function, mood, asthenia, and quality of life in women with breast cancer undergoing adjuvant chemotherapy. Clin Breast Cancer 2002; 3 Suppl 3:S116-20.

O'Shaughnessy JA, Vukelja SJ, Holmes FA, et al. Feasibility of quantifying the effects of epoetin alfa therapy on cognitive function in women with breast cancer undergoing adjuvant or neoadjuvant chemotherapy. Clin Breast Cancer 2005; 5(6):439-46.

Osoba D, Rodriguew G, Hyles J, et al. Interpreting the significance of changes in health-related quality of life scores. J Clin Oncol 1998; 16:139-144.

Osterborg A, Boogaerts MA, Cimino R, et al. Recombinant human erythropoietin in transfusion-dependent anemic patients with multiple myeloma and non-Hodgkin's lymphoma - A randomized multicenter study. Blood 1996; 87(7):2675-82.

Osterborg A, Brandberg Y, Hedenus M. Impact of epoetin-beta on survival of patients with lymphoproliferative malignancies: long-term follow up of a large randomized study. Br J Haematol 2005; 129(2):206-9.

Osterborg A, Brandberg Y, Molostova V, et al. Randomized, double-blind, placebo-controlled trial of recombinant human erythropoietin, epoetin beta, in hematologic malignancies. J Clin Oncol 2002; 20(10):2486-94.

P-174. Study of epoetin supportive care in CLL. Reported in Food and Drug Administration Oncologic Drugs Advisory Committee May 4, 2004 Meeting Briefing Information (available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/ 4037b2.htm) and Slides (available at http://www.fda.gov/ohrms/dockets/ac/04/slides/ 4037s2.htm). Online sources last accessed July 26, 2005.

Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998; 17:2815-34.

Patrick DL, Gagnon DD, Zagari MJ, et al. Epoetin Alfa Study Group. Assessing the clinical significance of health-related quality of life (HrQOL) improvements in anaemic cancer patients receiving epoetin alfa. Eur J Cancer 2003; 39(3):335-45.

Pealer LN, Marfin AA, Petersen LR, et al. Transmission of West Nile virus through blood transfusion in the United States in 2002. N Engl J Med 2003; 349(13):1236-45.

Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford: Oxford University Press; 2003.

Poole C, Greenland S. Random-effects meta-analyses are not always conservative. Am J Epidemiol 1999; 150(5):469-75.

Quirt I, Bramwell V, Charette M, et al., and the Systemic Treatment Disease Site Group. The role of erythropoietin in the management of cancer patients with non-hematologic malignancies receiving chemotherapy. Practice Guideline Report #12-1. #12-1. 2005. Available online at http://www.cancercare.on.ca/pdf/pebc12-1f.pdf. Last accessed March 2006.

Quirt I, Micucci S, Moran LA, et al. The role of recombinant human erythropoietin (EPO) in reducing red blood cell transfusions and maintaining quality of life (QOL) in patients with lymphoma and solid tumors receiving cytotoxic chemotherapy. Results of a randomized, double-blind, placebo-controlled clinical trial. Blood 1996; 88(10 Suppl 1):347a.

Razzouk BI, Hockenberry M, Hinds PS, et al. A double-blind, placebo-controlled study of once-weekly epoetin alfa in children with cancer undergoing myelosuppressive chemotherapy. J Clin Oncol 2004; 22(14S). Abstract 8527.

Rearden TP, Charu V, Saidman B, et al. Results of a randomized study of every three-week dosing (Q3W) of darbepoetin alfa for chemotherapy-induced anemia (CIA). J Clin Oncol 2004; 22(14 Suppl). Abstract 8064.

Richart JM, Petruska PJ, Klebert KS, et al. A phase III trial of epoetin alfa: early vs late administration in patients treated with chemotherapy. Blood 2002; 100:175b. Proceedings of the 44th Annual Meeting of the American Society of Hematology; December 6-10, 2002; Philadelphia, PA. Abstract 4199.

Rizzo JD, Lichtin AE, Woolf SH, et al. Use of epoetin in patients with cancer: evidence-based clinical practice guidelines of the American Society of Clinical Oncology and the American Society of Hematology. J Clin Oncol 2002; 20(19):4083-107.

Rose E, Rai K, Revicki D, et al. Clinical and health status assessments in anemic chronic lymphocytic leukemia (CLL) patients treated with epoetin alfa (EPO). Blood 1994; 84(10 Suppl 1):526a.

Rosenzweig MQ, Bender CM, Lucke JP, et al. The decision to prematurely terminate a trial of R-HuEPO due to thrombotic events. J Pain Symptom Manage 2004; 27(2):185-90.

Sakai H, Ohashi Y, Hirashima K, et al. Once weekly epoetin beta to increase hemoglobin and improve quality of life in anemic cancer patients receiving chemotherapy: A randomized, double-blind, dose-finding study. J Clin Oncol, 2004 ASCO Annual Meeting Proceedings (Post-Meeting Edition). 22(14S, July 15 Suppl). Abstract 8169.

Samol J, Littlewood TJ. The efficacy of rHuEPO in cancer-related anaemia. Br J Haematol 2003 Apr;121(1):3-11.

Savonije J, van Groeningen C, van Bochove A, et al. Early intervention with epoetin-alfa during platinum-based chemotherapy. J Clin Oncol 2004; 22(14S). Abstract 8111.

Savonije JH, van Groeningen CJ, van Bochove A, et al. Effects of early intervention with epoetin alfa on transfusion requirement, hemoglobin level and survival during platinum-based chemotherapy: Results of a multicenter randomised controlled trial. Eur J Cancer 2005; 41(11):1560-9.

Schumacher M, Hollander N, Schwarzer G, et al. Prognostic factor studies. In: Crowley J, editor. Handbook of statistics in clinical oncology. New York: Marcel Dekker; 2001. p. 321-78.

Schwartzberg LS, Yee LK, Senecal FM, et al. A randomized comparison of every-2-week darbepoetin alfa and weekly epoetin alfa for the treatment of chemotherapy-induced anemia in patients with breast, lung, or gynecologic cancer. Oncologist 2004; 9(6):696-707.

Seidenfeld J, Piper M, Flamm C, et al. Epoetin treatment of anemia associated with cancer therapy: a systematic review and meta-analysis of controlled clinical trials. J Natl Cancer Inst 2001; 93(16):1204-14.

Shasha D, George MJ, Harrison LB. Once-weekly dosing of epoetin-alpha increases hemoglobin and improves quality of life in anemic cancer patients receiving radiation therapy either concomitantly or sequentially with chemotherapy. Cancer 2003; 98(5):1072-9.

Silvestris F, Romito A, Fanelli P, et al. Long-term therapy with recombinant human erythropoietin (rHu-EPO) in progressing multiple myeloma. Ann Hematol 1995; 70:313-8.

Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. Br J Cancer 1994; 69:979-985.

Stasi R, Amadori S, Littlewood TJ, et al. Management of cancer-related anemia with erythropoietic agents: doubts, certainties and concerns. Oncologist 2005; 10:539-54.

Steensma DP, Loprinzi CL. Erythropoietin use in cancer patients: a matter of life and death? J Clin Oncol 2005; 23(25):5865-8.

Steensma DP, Molina R, Sloan JA, et al. A phase III randomized trial of two different dosing schedules of erythropoietin (epo) in patients with cancer-associated anemia: North Central Cancer Treatment Group (NCCTG) Study N02C2. J Clin Oncol 2005. Proceedings of the 41th Annual Meeting of the American Society of Oncology;13-17 May 2005; Orlando FL; Abstract 8031.

Steensma DP, Molina R, Sloan JA, et al. Phase III study of two different dosing schedules of erythropoietin in anemic patients with cancer. J Clin Oncol 2006; 24(7):1079-89.

Storring PL, Tiplady RJ, Gaines Das RE, et al. Epoetin alfa and beta differ in their erythropoietin isoform compositions and biological properties. Br J Haematol 1998; 100:79-89.

Straus DJ, Testa M, Riggs SA, et al. Early treatment with epoetin alfa improves anemia, quality of life, and productivity in patients with hematologic malignancies and mild anemia during chemotherapy. Blood 2003; 102:497a. Abstract 1811.

Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. Can J Psychiatry. 2002; 47(1):68-75.

ten Bokkel Huinink WW, De Swart CA, Van Toorn DW, et al. Controlled multicentre study of the influence of subcutaneous recombinant human erythropoietin on anaemia and transfusion dependency in patients with ovarian carcinoma treated with platinum-based chemotherapy. Med Oncol 1998; 15(3):174-82.

Thatcher N, De Campos ES, Bell DR, et al. Epoetin alpha prevents anaemia and reduces transfusion requirements in patients undergoing primarily platinum-based chemotherapy for small cell lung cancer. Br J Cancer 1999; 80(3-4):396-402.

Thomas DR. Anemia and quality of life: unrecognized and undertreated. J Gerontol A Biol Sci Med Sci 2004; 59(3):238-41.

Thomas H, McAdam KF, Thomas RJ, et al. Early intervention with epoetin alfa for treatment of anaemia and improvement of quality of life in cancer patients undergoing myelotoxic chemotherapy. Ann Oncol 2002; 13(suppl 5):177. Abstract.

Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002; 21:1559-73.

Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999; 18:2693-708.

Throuvalas NA, Antonadou D, Boufi M, et al. Erythropoietin decreases transfusion requirements during radiochemotherapy. Proc ASCO 2000; 19:394a. Abstract 1558.

Vadhan-Raj S, Mirtsching B, Charu V, et al. Assessment of hematologic effects and fatigue in cancer patients with chemotherapy-induced anemia given darbepoetin alfa every two weeks. J Support Oncol 2003; 1(2):131-8.

Vadhan-Raj S, Skibber JM, Crane C, et al. Randomized, double-blind, placebo-controlled trial of epoetin alfa (Procrit) in patients with rectal and gastric cancer undergoing chemo-radiotherapy (CT/RT) followed by surgery: early termination of the trial due to increased incidence of thrombo-embolic events (TEE). Blood 2004; 104(pt 1):797a. Abstract 2915.

Vansteenkiste J, Pirker R, Massuti B, et al. Double-blind, placebo-controlled, randomized phase III trial of darbepoetin alfa in lung cancer patients receiving chemotherapy. J Natl Cancer Inst 2002; 94:1211-20.

Waltzman R, Croot C, Justice GR, et al. Randomized comparison of epoetin alfa (40,000 U weekly) and darbepoetin alfa (200 microg every 2 weeks) in anemic

patients with cancer receiving chemotherapy. Oncologist 2005; 10(8):642-50.

Waltzman RJ, Croot C, Williams D. Final hematologic results: epoetin alfa (EPO) 40,000 U QW vs. darbepoetin alfa (DARB) 200 mcg Q2W in anemic cancer patients (pts) receiving chemotherapy (CT). Proc Am Soc Clin Oncol 2005; 23:736s. Abstract 8030.

Watine J, Bouarioua N. (2002). Anemia as an independent prognostic factor for survival in patients with cancer. Cancer. 94(10):2793-6.

Weiss RB. Section 8, Miscellaneous Toxicities, Chapter 55, Adverse Effects of Treatment. In: Cancer: Principles & Practice of Oncology, DeVita VT Jr, Hellman S, Rosenberg SA, eds. 6th edition, 2001. Lippincott Williams & Wilkins, Philadelphia. p 2964.

Welch RS, James RD, Wilkinson PM. Recombinant human erythropoietin and platinum-based chemotherapy in advanced ovarian cancer. Cancer J Sci Am 1995 Nov;1(4):261.

Westenfelder C, Baranowski RL. Erythropoietin stimulates proliferation of human renal cell carcinoma cells. Kidney Int 2000; 58:647-657.

Wilson J, Yao G, Raftery J, et al. A systematic review and economic evaluation of epoetin alfa, epoetin beta and darbepoetin alfa in anaemia associated with cancer, especially that attributable to cancer treatment. National Institute for Health and Clinical Excellence, UK, March 2005. Available at: http://www.nice.org.uk/pdf/Anaemia_Assessment_Report.pdf. Last accessed March 2006.

Wisloff F, Gulbrandsen N, Hjorth M, et al. Quality of life may be affected more by disease parameters and response to therapy than by haemoglobin changes. Eur J Haematol 2005; 75:293-8.

Witzig TE, Silberstein PT, Loprinzi CL, et al. Phase III, randomized, double-blind study of epoetin alfa versus placebo in anemic patients with cancer undergoing chemotherapy. J Clin Oncol 2004; 23(12):2606-17.

Wun T, Law L, Harvey D, et al. Increased Incidence of symptomatic venous thrombosis in patients with cervical carcinoma treated with concurrent chemotherapy, radiation, and erythropoietin. Cancer 2003; 98:1514–20.

Wurnig C, Windhager R, Schwameis E, et al. Prevention of chemotherapy-induced anemia by the use of erythropoietin in patients with primary malignant bone tumors (a double-blind, randomized, phase III study). Transfusion 1996; 36(2):155-9.

Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005; 28(2):172-91.

Zhang Y, Thamer M, Stefanik K, et al. Epoetin requirements predict mortality in hemodialysis patients. Am J Kidney Dis 2004; 44:866-876.

List of Acronyms/Abbreviations

AE	adverse events
AHRQ	Agency for Healthcare Research and Quality
AJCC	American Joint Committee on Cancer
ALL	acute lymphocytic leukemia
ASCO	American Society of Clinical Oncology
ASH	American Society of Hematology
AUC	area under the curve
BCBSA	Blue Cross and Blue Shield Association
ca	cancer
CCOPG	Cancer Care Ontario Practice Guidelines
CDC	Centers for Disease Control and Prevention
CDER	Center for Drug Evaluation and Research
CERA	continuous erythropoiesis-receptor activator
chemo	chemotherapy
CLAS	cancer linear analog scale
CLL	chronic lymphocytic leukemia
CR	complete response
CT	chemotherapy
darb	darbepoetin
DBP	diastolic blood pressure
dL	deciliter
EORTC	European Organization for Research and Treatment of Cancer
EPC	Evidence-based Practice Center
epo	epoetin
est	estimated
FACT	Functional Assessment of Cancer Therapy, including G-General; F-Fatigue; An-Anemia
FDA	Food and Drug Administration
FNCLCC	Federation Nationale des Centres de Lutte Contre le Cancer
g	grams
G-CSF	granulocyte colony-stimulating factor
GI	gastrointestinal
GU	genitourinary
Gy	Gray
gyne	gynecologic
H&N	head and neck
Hb	hemoglobin
Hct	hematocrit
HD	Hodgkin's disease
hematol	hematologic
HG	mercury
HIV	human immunodeficiency virus
HR	hazard ratio
ID	identification
IPD	individual patient data
ITT	intention-to-treat
IU	international units
IV	intravenous
J&J	Johnson and Johnson
kg	kilogram
K-M	Kaplan-Meier
KQ	key question
LASA	linear analog self-assessment
MA	meta-analysis
malign	malignancy

MDACC	M.D. Anderson Cancer Center
MM	multiple myeloma
n, N	number
NCCN	National Comprehensive Cancer Network
NESP	novel erythropoiesis-stimulating protein
NHL	non-Hodgkin's lymphoma
NICE	National Institute for Health and Clinical Excellence
NNH	number needed to harm
NNT	number needed to treat
NR	not reported
NS	not significant
NSCLC	non-small cell lung cancer
ODAC	Oncologic Drugs Advisory Committee
plat	platinum
PR	partial response
pub	publication
q2w	every two weeks
QLQ	Quality of life Questionnaire
QoL	quality of life
qw	every week
radio	radiotherapy
random	randomized
RBC	red blood cell
RBCT	red blood cell transfusion
RCT	randomized, controlled trial
RR	relative risk
SBP	systolic blood pressure
SC	subcutaneous
sc	subcutaneous
SCLC	small cell lung cancer
SD	standard deviation
TEC	Technology Evaluation Center
tiw	three times weekly
tx	treatment
U	units
U.K.	United Kingdom
U.S.	United States
VAS	visual analog scales
WHO	World Health Organization

U.S. Department of Health and Human Services
Mike Leavitt, *Secretary*

Office of Public Health and Science
Richard H. Carmona, M.D., M.P.H., F.A.C.S., *Surgeon General of the United States*

Agency for Healthcare Research and Quality
Carolyn M. Clancy, M.D., *Director*

U.S. Department of
Health and Human Services

Public Health Service
Agency for Healthcare Research and
Quality

AHRQ Publication No. 06-EHC008-EF
May 2006