Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment: Comparative Effectiveness Update*

Draft review available for public comment from June 17, 2011 to July 15, 2011

Research Review Citation: Grant MD, Piper M, Bohlius J, Tonia T, Robert N, Vats V, Bonnell C, Ziegler KM, Aronson N. Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment: Comparative Effectiveness Update. Comparative Effectiveness Review No. 113. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC077-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1480
Published Online: April 25, 2013

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Executive Summary	None	
	Introduction	None	
	Methods	None	
	Results	We have several suggestions for the decision analysis portion of the report: 1) edit the description to reach a broader audience, which makes it more usable for individuals who may not be familiar with decision analysis techniques, 2) strengthen the rationale to better support assumptions of the model, and 3) provide more explicit descriptions for the derivation/sources of estimations. This is especially important regarding explanations for the 5% annual mortality rate following the on-study period (noted on page 96 of the report) and application of that assumed rate across the broad range of indications for chemotherapy (adjuvant through refractory/metastatic disease) and ESA use. With better descriptions of assumptions and limitations in this section, readers can make more informed choices about how to use this report and its important conclusions. The report does not adequately discuss a crucial question of keen interest to most practitioners in this area: the mechanism of harm attributed to erythropoiesis-stimulating agents (ESAs). We understand there is a paucity of evidence in the medical literature, but we believe a brief acknowledgement of this would strengthen the report. It would also be appropriate to state more explicitly that despite the large number of trials of these agents across diverse populations over more than a decade, the potential mechanism of action for the harms attributed to ESA use remains poorly understood. Current evidence does not inform the mechanism of action or clearly identify which patients will experience benefits or harms from use of ESAs. This is a serious limitation related to the literature about ESAs - one that makes it particularly difficult to provide meaningful guidance concerning their safe or appropriate use.	We have extensively revised the decision analysis for clarity and accessibility including detail of model assumptions. Sources for parameters are referenced with justification including the annual mortality rates. As noted in the report, annual mortality rates following therapy vary considerably. From those data we have focused on 2 scenarios—treatment with curative intent (annual mortality following treatment of 5% or near at the lowest bound) and not treated with curative intent (annual mortality of 50% or at the median of the distribution). While biologic mechanisms are beyond the scope of the report, we hope that directions suggested for future research may partly assist elucidating them—e.g., dosing strategies and subgroups for whom risks of ESA treatment are low.
	Discussion	None	
	Conclusion	None	
	Figures	None	
	References	None	
	Appendix	None	

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Executive Summary	None	
	Introduction	There are two errors of scientific fact on page 1 of the introduction. First, while it is true that it takes 2 weeks to make a red cell, it is not the case that after 2 weeks of ESA therapy, the number of red cells produced is sufficient to increase the hematocrit; that takes four weeks. What is interpreted as a hematocrit increase after two weeks of therapy, is actually plasma volume contraction, another component of ESA physiology. Second, while erythropoietin and thrombopoietin share a 20 % homology, ESAs do not stimulate thromboppoiesis and thrombopoietin does not stimulate erythopoiesis because of receptor specificity. This well-documented scientifically. The refernce cited (2) to support the contention in the text is not scientifically valid. Any platelet increment seen with ESA use is clincially trivial and due to induced tissue iron deficiency.	These have been changed to be consistent with the prescribing information stating 2 to 6 weeks. This has been deleted.
	Methods	The methodology employed was appropriate, comprehensive and well-documented.	No response needed
	Results	The results are clearly displayed in the appropriate detail. With respect to message, however, the extent to which the data up to 2006 biased the studies thereafter is unclear in this analysis. That is to say it is well recognized that studies such as those of Henke and Leyland-Jones violated the guidelines (in the name of research) of how to use ESAs safely. What is unclear is the extent to which such studies could have influenced the data with respect to shortened survival during ESA exposure. This is the important; thrombosis risk was always well recognized; shortened survival during therapy is a different issue. In this regard, as alluded in this report, ESA dose may be the important issue, not the hematocrit achieved.	The reviewer properly points out uncertainties that are acknowledged. ESA dosing may be the important issue. These data, however, do not allow examining it—both because of the lack of dose information and individual patient data. As noted in the Future Research section, this question should and could be addressed with observational data.
	Discussion	The implications of the major findings are clearly stated but the underlying assumptions are open to question. For example, it is assumed for the analysis on page 96 that a 2 gm increase in hemoglobin will occur after 4 weeks of ESA therapy; this is not highly likely. Additionally, it is stated on page 43 that an ESA is superior to a transfusion strategy for avoiding transfusion, which is stating the obvious but what is not discussed is that no head-to-head trial of transfusion to a chosen hemoglobin versus ESA therapy to same has ever been performed which would eliminate important forms of bias.	The decision analysis attempts to assume a best-case scenario with respect to gain in QALYs (e.g., a 2 gm increase in Hb at 4 weeks) and so any bias favors ESAs. The reviewer's noting that a trial comparing ESA and transfusion to a target is spot on and would eliminate some biases. That said, our interpretation is that biases, even if present, are unlikely to explain the effect magnitude.

Commentator & Affiliation	Section	Comment	Response
	Conclusion	Finally, I would not agree that future research section was clear with respect to potential new research. Certainly, from many considerations it is unlikley that prospective RCTs will be conducted in this area but how best to proceed otherwise was not well delineated.	The Future Research section does describe what observational data should be collected and how it should be analyzed to answer the important questions raised—" A large registry with accurate and precise information on ESA dose (amount, frequency, duration, escalation), Hb (baseline, and all recorded values preferably at times specified by protocol), stage of malignancy, treatment regimen and response, and outcomes (including but not limited to thromboembolism, myocardial infarction, death including underlying and contributory causes) would provide the best opportunity to examine these questions. The Dosing and Outcomes Study of Erythropoiesis-Stimulating Therapies (DOSE) is one example. While deriving conclusions from appropriate analytical methods—inverse probability weighting, G-methods, and marginal structural models—requires some assumptions for inference, they are approaches most able to address unanswered questions.'. This outlines the broad design and analysis of a registry that could likely address the important unanswered questions.
	Figures	None	
	References	None	
	Appendix	None	
	General	This is a comprehensive and well-conducted study and, therefore, by definition, will be of most interest to scholars working in this area. Clinicians are much more likely to read or consult the ASH/ASCO ESA guidelines because they represent a condensed version translated into clinically meaningful guidelines for everyday practice. Overall, this is a carefully performed and clearly presented report but the conclusions are not different than previous reports and, therefore, are not more informative with respect to policy or practice decision making.	We concur, as the review is not a guideline. No response needed.
Peer Reviewer #3	Executive Summary	None	

Commentator & Affiliation	Section	Comment	Response
	Introduction	Page 8, line 13: The Conclusion section states that “harms appear greater than benefits when ESAs are used to manage anemia in patients undergoing chemotherapy or radiation for malignancy”. The authors may consider the following regarding this statement: 1)- state what the specific harms are suggested by the pooled analysis data as it is (mortality risk etc) rather than a general statement that harms are greater than benefits and, 2)- add a “Limitations” paragraph discussing the potential ability to generalize the findings of this type of pooled analysis when dealing with a specific patient or specific patient populations since there is evidence that certain groups of patients may not necessarily be harmed by an adverse effect on survival (see examples of RCTs and citations provided below), 3)- ESAs are not used “to manage anemia in patients undergoing …. radiation for malignancy” . The clinical trials in head-neck cancer patients were designed to improve outcomes by increasing tumor oxygenation even in non-anemic patients. ESAs have never been indicated for use in patients treated with radiation therapy.	We appreciate the comment and have followed the recommendation for (1) 1) The conclusion has been replaced by “Since the 2006 review, evidence remains consistent that ESAs reduce the need for transfusions and increase the risk of thromboembolism. FACT-Fatigue scores are better with ESA use but the magnitude is less than the minimal clinically important difference. An increase in mortality accompanies the use of ESAs.”A) 2) The length of the structured abstract limits the ability to include all issues. Whether evidence supports concluding that some patient groups may not be harmed is uncertain—e.g., VTE occurrence. For example, as pertains to mortality, because the relative risk is small, patients at low underlying risk have corresponding small increases in absolute risk. Still, we point out that those trials reporting no adverse mortality effect severely underpowered to detect the magnitude of relative risk estimated here (i.e., would require samples of 10,000 or more). 3) The reviewer correctly notes issues regarding radiotherapy and ESA use. However, that group was part of the patient population defined for the review. Prompted by the comment, we have addressed the radiotherapy subgroup in analyses (excluding those trials and in meta-regressions). Their consideration or exclusion would not alter any conclusions. Results excluding those trials are noted now in the tables and stated in the discussion. We also refer Figure 1 at the end of this document and comments accompanying the Figure on later (Page 32).

Commentator & Affiliation	Section	Comment	Response
	Introduction continued	The limitations of many of the individual clinical trials of ESAs are well known and documented in the literature and in FDA briefings, hence the indication for systematic reviews and pooled analyses. It is important, however, to also emphasize that the general conclusion stating that "harms appear greater than benefits" is in the context of this systematic review of the literature and pooled analysis of the evidence from very diverse trials, many of which investigated ESA use outside of the typical indication for which the agents are used in routine clinical practice. The best example of this is that ESA use has not been indicated for managing anemia (prevention or treatment of anemia or theoretically raising tumor oxygenation to increase radiation sensitivity) in patients undergoing radiation therapy, even though the signals demonstrating an adverse ESA effect on survival initially emerged in the Henke trial (and subsequently confirmed by other head-neck cancer trials and a Cochrane head-neck review) involving non-anemic patients with head-neck cancer, a malignancy that is primarily treated by radiation therapy. There were 5 radiation therapy and 5 chemo-radiation therapy trials included in the overall survival analysis (page 71, Table 28). While any and all mortality and adverse progression-free survival signals related to ESA use are very significant even if associated with off-label uses, the clinically more relevant and important risk is in the cohort of patients with chemotherapy-induced anemia. It is important to also note that ESAs may not necessarily exhibit uniform effects in patients with all types and stages of solid tumors treated with different chemotherapy regimens in the palliative treatment setting. The ability to generalize the stated conclusion (page 8, line 13) becomes less clear when a physician is considering the risks-benefits of ESA therapy in the context of a specific patient, with a specific type and stage of malignancy, treated with a specific chemotherapy regimen. This issue is also relevant to the decision analysis and model discussed on page 117 (see additional comments below).	The consequences of including radiotherapy trials is address in detail on page 32 and in Figure 1. Language acknowledging this has been edited in the text. For example, see discussion paragraph beginning "Much of the evidence included here was obtained under treatment protocols that used higher baseline and target Hb levels than applied in current practice…." Note that this conclusion has been changed. We concur that applying results to an individual patient lies in the domain of clinical practice.

Commentator & Affiliation	Section	Comment	Response
	Introduction continued	The authors may wish to consider including a discussion of the limitations and applicability of this conclusion emerging from this review as to its ability to inform the decision when dealing with an individual patient. This is particularly important since there are published clinical data in the literature demonstrating that ESA use is not always associated with increased mortality even when the anticipated outcome of cancer treatment is cure. For instance, in a large RCT involving patients with Hodgkin's disease treated with an intensive chemotherapy regimen with curative intent, ESA therapy was not associated with increased mortality (Engert et al J. Clin. Oncol. 28: 2239–2245, 2010). There are other examples in the palliative chemotherapy setting in patients with specific types of cancer that carry a poor prognosis. For instance, in two RCTs involving small cell lung cancer patients undergoing chemotherapy, ESA therapy targeted to Hb levels >12 g/dL may reduce transfusion requirements without a significant negative impact on mortality (Grote et al J. Clin. Oncol. 23: 9377–9386, 2005 and Pirker et al-J. Clin. Oncol. 26:2342–2349 2008).	We did consider the recommendation, but judged it somewhat beyond our scope. Patients with low underlying mortality risk will experience corresponding small increases in absolute risk of mortality—patients treated with curative intent are the least likely to die due to ESAs owing to the low underlying absolute mortality risk
	Methods	None	
	Results	Page 19, line 37: "The evidence is lacking to determine whether immediate treatment versus delayed treatment produces better outcomes ". Does better outcomes refer specifically to transfusion-sparing effect ? If so, this could be stated here. Page 19, line 50: "Under circumstances representative of patients included in these trials, a decision analysis shows ESA use is always accompanied by a net loss of lifeyears due to increased mortality during the active treatment period." There are limitations to the ability to generalize based on analyses of pooled data of ESA outcome from diverse studies (chemotherapy, radiation therapy, different tumor types, different stages, single trials involving mixed non-myeloid tumors, anemic and non-anemic patients, palliative versus curative treatment intent). ESA use may not necessarily always be accompanied by a net loss of life when looking at specific cancer types and treatment regimens as discussed above (Grote et al J. Clin. Oncol. 23: 9377– 9386, 2005, Pirker et al-J. Clin. Oncol. 26:2342–2349 2008, Engert et al J. Clin. Oncol. 28: 2239–2245, 2010). Given these examples of exceptions, the statement that ESAs are always accompanied by a net loss of life becomes inaccurate.	This refers to all outcomes as in KQ2 so has not been changed. The reviewer's comments are well made. We and have edited this paragraph and deleted "always." The values are those expected for a cohort .
	Results continued	Page 20, line 7: "Whether there are subgroups at higher and lower risk of adverse events and mortality is unclear." As mentioned above, there are subgroups of cancer patients already identified with no increase in mortality. For instance, as mentioned above, two RCTs involving small cell lung cancer patients undergoing chemotherapy demonstrated that ESA therapy targeted to Hb levels >12 g/dL may reduce transfusion requirements without a negative impact on mortality in this cohort of patients with overall poor prognosis (Grote et al J. Clin. Oncol. 23: 9377–9386, 2005 and Pirker et al- J. Clin. Oncol. 26:2342–2349	While relative risk of on-study mortality varied according to underlying absolute risk, we were otherwise not able to identify subgroups based on those characteristics evaluated. Again we are cautious drawing conclusions based on individual trial results given issues of power

Commentator & Affiliation	Section	Comment	Response
		2008). In the curative treatment setting, in a large trial involving patients with Hodgkin's disease treated with an intensive chemotherapy regimen, ESA therapy was not associated with increased mortality (Engert et al J. Clin. Oncol. 28: 2239–2245, 2010). Page 23, line 27: This discussion focuses on hematologic malignancies. Most of the actual clinical use of ESAs in practice are in patients with solid tumors. Information on anemia prevalence in non-myeloid malignancies could be included here. Page 25, line 25: " Too few trial results were available to perform a subgroup analysis conforming to label recommendations." This is a key issue that could be emphasized more in this manuscript because it is relevant to clinical practice and it is the reason for ongoing safety concerns leading to use restriction. The actual impact of ESAs on tumor progression and/or survival remains poorly characterized for many specific types/stages of tumors when ESAs are given at the minimum required doses to reduce/avoid red cell transfusions (rather than targeting an arbitrary hemoglobin level) in cancer patients receiving myelosuppressive chemotherapy in the palliative setting.	This has been revised to include non-myeloid malignancies. The referenced statement referred to the previous CER. The questions of applicability raised are relevant, and are addressed directly in the discussion "Much of the evidence included here was obtained under treatment protocols that used higher baseline and target Hb levels than those used in current practice. While it is possible that adverse event rates might be somewhat different with lower baseline and target Hb levels, we found little difference in effect when baseline Hb was less than, or exceeded 10 g/dL. This result is similar to a prior individual patient data meta-analysis.4 Additionally, five trials included in KQ 1 enrolled patients undergoing radiotherapy. Although not an FDA-approved indication for ESA use those results were included in the synthesis here because the population of interest was patients undergoing treatment for cancer. Moreover, we did not find those trials trial results influential in the synthesis. While some uncertainty remains, given that the adverse consequences are life threatening, the current evidence does not suggest that by following new guidelines, adverse event rates and relative effects will be substantially different."
	Results continued	Page 25, line 57: has the 2008 FDA / ODAC briefing been included as a reference ? http://www.fda.gov/ohrms/dockets/ac/08/briefing/2008-4345b2-01-FDA.pdf Page 52, line 50: "Reported target Hb levels ranged from 11.5 g/dL to 14 g/dL (mean 12.6 g/dL), but was in only one trial lower than 12 g/dL and in two trials higher than 13 g/dL." How many trials reported the achieved hemoglobin ? What was the relationship between achieved hemoglobin and outcomes ? Did the data	In this historical perspective, the 2006 CER had not yet been completed. ODAC documents were included in the review. The question is relevant and considered. However, unfortunately these data do not allow examining it both because of reporting and being study-level. Evaluating the question requires

Commentator & Affiliation	Section	Comment	Response
		allow identification of cohorts of patients who responded to ESA versus non-responders and was there a difference in outcomes between ESA responders and non-responders as has been suggested in ESA trials in the chronic kidney disease setting ? Page 56, table 14: Was the data from the PREPARE neoadjuvant breast cancer trial actually included ? In Table 14 Untch et al is included but on page 241 it is also among the excluded studies and then again in Table E1 on page 245 and page 252 (line 38) it seems to be included. Also, ODAC 2008 is not included in the Table on page 252 in the Appendix. http://www.fda.gov/ohrms/dockets/ac/08/briefing/2008-4345b2-01-FDA.pdf Page 118: line 28: The on-study mortality rate included here is 15% based on the HR from the pooled analyses of all patients as illustrated on pages 76-77 in Table 26 and Figure 11. The model assumes a 12-week course of ESA during chemotherapy, but the analysis where the 15% risk is derived from includes the head-neck cancer / radiation therapy trials by Henke, Machtay and Hoskin. What was the on-study mortality data for chemotherapy only-treated patients? In the Bohlius 2009 IPD meta-analysis, in the subgroup of patients receiving chemotherapy only, the observed increase in mortality risk was lower than in the analyses involving the entire cohort and did not reach statistical significance (HR:1.10, 95% CI 0.98–1.24, p=0.12).	individual patient data. The related issue examined here was use of dose escalation. The study excluded was Untch 2005. Data from PREPARE (Untch 2011) were included. We have used the PREPARE publication as reference. The noted exclusion was the duplicate abstract presentation for PREPARE. The subgroups (chemotherapy versus radiotherapy) are now addressed in the report more clearly and in other detail (as relates to underlying mortality risk). When evaluated in a meta-regression (not accounting for underlying absolute mortality risk) we did not find evidence for effect modification. Moreover, when taking into account underlying absolute mortality risk (Figure 1 of this document) the chemotherapy and radiotherapy trials appear to have similar effects.
	Results continued	Page 120, line 40: “While some uncertainty remains, given that the adverse consequences are life threatening, the current evidence does not suggest that by following new guidelines, adverse event rates and relative effects will be substantially different”. This reviewer respectfully disagrees with this statement. There is uncertainty but it cannot be presumed that ESAs will always be associated with adverse effects (mortality and/or tumor progression) in patients with all types of cancer treated with chemotherapy in the palliative setting. There are already published examples of RCTs conducted in patients with specific tumor types (indicated in comments above), demonstrating absence of increased mortality associated with ESA therapy. The new guidelines state that ESA use is an option only in the palliative chemotherapy setting in a manner that minimizes ESA dose and exposure just enough to avoid or reduce red cell transfusions rather than targeting an arbitrary hemoglobin level and initiate therapy only after a risk-benefit discussion with the patient following the REMS/APPRISE procedure. The guidelines also recommend discontinuation and not dose increase within a few weeks of therapy in non-responders. As stated above, in the Bohlius 2009 IPD meta-analysis, in the subgroup of patients receiving chemotherapy only, the observed increase in mortality risk was lower and did not reach statistical significance (HR:1.10, 95% CI 0.98–1.24, p=0.12). The guidelines (and the FDA) clearly indicate that ESAs should not be used in the curative chemotherapy	The critique is fair. We have deleted this statement. (Response 22)

Commentator & Affiliation	Section	Comment	Response
		setting to avoid any potential for harm, although in the future there may be exceptions to this given the Hodgkin's disease data from Germany (Engert et al). Page 121, line 22: "Furthermore, patients with the worst prognosis experience the greatest loss." What is the evidence for this statement? There is evidence from clinical trials to the contrary. Two RCTs mentioned above involving small cell lung cancer patients undergoing chemotherapy demonstrated that ESA therapy targeted to Hb levels >12 g/dL may reduce transfusion requirements without a negative impact on mortality in this cohort of patients with overall very poor prognosis. Page 121, line 35: "The confidence and credible intervals for the estimated relative increase in mortality span a range values—the true relative increase in risk could be higher or lower than 1.15 estimated here." Previously published data indicates that the risk is likely to be lower if only the chemotherapy-induced anemia patients were to be included in the analysis (as in Bohlius 2009 mentioned above) and when the radiation-therapy treated patients are excluded.	A decision model based on the summary evidence revealed that the expected quality of life gains with ESA treatment must be traded for fewer life-years—3.6 per 1,000 patients treated with curative intent and 9.2 per 1,000 patients not treated with curative intent. We take a cautious approach interpreting negative results from one or two trials. Trials reporting no adverse mortality effect were severely underpowered to detect the magnitude of relative risk estimated here. Excluding the radiotherapy trials had little impact on the estimated effect. We also refer to Figure 1 below that shows radiotherapy trials quite consistent with the chemotherapy ones when underlying risk is taken into account.
	Results continued	Page 121, line 43: "It is therefore important to address whether there are patient subgroups with low risk of harm and how dosing practices influence harms. Unfortunately, these questions present complexities not addressed even in the most carefully designed trials." As mentioned above, patients with small cell lung cancer and Hodgkin's disease treated with specific chemotherapy regimens may be at low risk or even no risk for harm based on clinical trial data. Page 128 (and the entire references section)- reference 95 and reference 177 (Hernandez et al, 2009) are the same. References 43, 106, 142 (Overgaard et al) are the same. References 191 and 196 (Glaspy et al) are the same. Are there any other duplicate / triplicate references?	We take a more cautious approach to interpreting trials failing to identify a risk of mortality. Individual trials reporting no adverse mortality effect were severely underpowered to detect the magnitude of relative risk estimated here. All duplicate references have been corrected. Appendix F details trials and references when there were subsequent reports.
	Discussion	None	
	Conclusion	None	
	Figures	None	
	References	None	
	Appendix	None	
	General	This well-written review updates the CER 2006 analyzing the comparative benefits and harms of erythropoiesis-stimulating agent (ESA) strategies and non-ESA strategies to manage anemia in patients undergoing chemotherapy or radiation for malignancy. Specific comments are listed below. There are many duplicate/triplicate references in the References section beginning on page 124. Some of the specific examples of the duplicate / triplicate references are indicated in the comments below but it is difficult for a reviewer to address all of these.	Duplicate references have been corrected.

Commentator & Affiliation	Section	Comment	Response
		Page16, line 13: There is no standard definition of “extremely” low rate of infection although transfusion-related infection transmission risk is generally low. Consideration may be given to deleting “extremely”. Page 23, line 52: “Anemia may also have be associated with outcomes or have direct effects on the tumor itself” this statement needs to be reworded. Page 25, line 38: what does “five trials” in brackets refer to when the sentence states 6 trials ? Page 32, line 56 & page 118, line 22: typing errors noted. There are a number of other typing errors throughout the manuscript. Proofreading is recommended.	We agree that there is no standard definition and has been changed to “very”; actual rates quantified in the report background. Corrected Corrected Corrected
Peer Reviewer #4	Executive Summary	None	
	Introduction	Very detailed and helpful (better than most similar reports)	
	Methods	No concerns	
	Results	No concerns	
	Discussion	In general, conclusions are very appropriate and connected to data. The exception would be the decision analysis section. This section can be made more clearer (particularly the description of this section in the abstract or executive summary which is quite vague). It maybe that investigator have much more expertise in evidence synthesis (conducting SR and MA) than decision analysis. I recommended more clarity.	The decision analysis has been extensively revised, expanded, for clarity and presented in a conventional manner.
	Conclusion	None	
	Figures	None	
	References	None	
	Appendix	None	
	General	The report is methodologically sound with clear inclusion criteria, research questions and plan of analysis No concerns. Note that there are a bunch of typos (e.g., line 5 page 26, “with” is misspelled as “will”)	Corrected.
Peer Reviewer #5	Executive Summary	Page v, line 27 and ES-1, line 50:inserted: most comprehensive Page ES-1, line 56: It seems a bit misleading to cite the numbers randomized here, since not all those randomized to control arms were transfused, and some of those randomized to ESA were also transfused. What may really matter is the total number of patients actually transfused (and perhaps even the total number of units transfused) in these studies regardless of which arm they were in, and that there were no adverse effects reported attributable to all those transfusions.	The sentence identified the data source, was not intended to judge comprehensiveness. The has been deleted

Commentator & Affiliation	Section	Comment	Response
	Introduction	Overall, the Introduction is well written, provides a thorough and accurate historical review, and also offers a convincing rationale for why an updated review was needed and for the decisions made by the Effective Health care program and the EPC to revise some key questions and delete others. Page1, Line 37: inserted: in response to hypoxia Page 1, Line 38: The second sentence seems redundant, since production in kidney is already mentioned in first sentence. Page 1, Line 55: Suggest moving the parenthetical to follow "thrombopoiesis" since these two processes work sequentially to produce platelets, and since thrombocytes are platelets. Page 2, Line 57: Uncertain how or why the text/content of this page disappeared. Two minor changes to suggest on this page (below). The first is to delete the word "have" on the first line of the last paragraph (in other words, change "...may also have be associated..." to "...may also be associated..." The second is to insert the word "negative" in front of "factors" on the last line of the page. Page 2, Line 50: inserted text Page 2, Line 53: typing errors noted Page 2, Line 55: inserted text Page 2, Line 56: typing errors noted Page 5, Line 3: This may be an incorrect citation. The initial IPD publication is the 2009 Lancet paper by Bohlius et al that is presently cited as number 58 on p. 105 in the list of references. Page 5, Line5: typing errors noted	None Inserted "in response to tissue hypoxia" Deleted sentence Done Adding the word negative might be redundant so we have maintained existing usage. I nserted Amended as suggested Inserted Amended as suggested Corrected citations Revised sentences for clarity
	Methods	Page 13, Line 22: It might be more technically correct to say "...published subsequent to the end date of the literature review for the 2006 report,..." Page 22, Line 12: Since previously stated that observational studies were included only if they were comparative, a row should probably be added to the "Types of studies" section of Table 4 that clarifies the criteria used to distinguish comparative from non-comparative observational studies.	Revised. We have indicated that included observational studies were required to be comparative, which is also reflected by text and analysis methods. We

Commentator & Affiliation	Section	Comment	Response
		Page 25, Line 10: See above note re criteria used to distinguish comparative versus non-comparative observational studies	believe it clearer to maintain Table 4 as it is. Treatment assignment being necessary in observational studies requires their being observational. For this reason we have not added further criteria.
		Page 25, Line 39: typing errors noted	Corrected
		Page 26, Line 9: What about conflicts of interest? (for HRQoL too)	Our study quality assessment (Higgins and Green, 2008), which was specified in our protocol, did not include funding as a variable.
		Page, 28, Line 28: inserted text	Inserted
		Page 29, line 36: The Methods used for this review (searches, study selection, data extraction, analyses) are well-described, rigorous, and appropriate to the body of evidence and key questions.	No response needed.
	Results	Page 31, Line 26: typing errors noted	Corrected.
		Page 31, Line 47: Should either be "Trials'..." (ie, possessive) or "Trial..."	Changed to 'Trial.'
		Page 31, Line 52: It's unclear how this mean was calculated. Was each individual trial weighted equally, or were trials weighted in some way by sample size? How useful or meaningful is the mean value, if only 3 trials reported targets other than 12 or 13? Also, unable to find entries in either Parts I or II of the Study Characteristics tables in Appendix C that clearly listed the target Hb level for each trial. Finally, it might be useful to identify by citation number the sole trial with a target below 12 and the two trials with a target above 13.	The mean is unweighted. The mean and range are intended to describe the distribution. We defined target Hb in the Data Analysis section, which provides guidance to identify target Hb in Appendix Table C1. Citations we feel are not required as the intent is to provide a general description of trials and distribution of target Hb--not to draw attention to specific trials at this point in the text.
		Page 38, Line 9: Again, uncertain how to interpret this mean across trials.	Similarly, no means are weighted.
		Page 38, Line 53: Reference 87 is identical to reference 84. This paragraph leaves readers with the incorrect impression that there are a total of five separate trials, when there are only 4. It might also be useful to attach citation numbers to the entries in Table 17. e.g., Had to search through the reference list (inconvenient when reading on line) to find that 2 of the 4 trials listed in Table 17 corresponded to references 83 or 84 and the other two corresponded to 85 and 86.	Corrected reference and citation duplication; with citations correctly numbered, table and footnotes correspond with text; citations added to all tables.

Commentator & Affiliation	Section	Comment	Response
		Page 39, Line 44: Readers might find this a bit confusing. Did the sensitivity analysis add the two trials included in the previous review but excluded in this review, or did it add the two trials included unchanged but not pooled because they used different definitions for response. Again, using the citation numbers might help clarify this for readers.	Edited to clarify that two excluded trials from the previous review were added for the sensitivity analysis.
	Results continued	Page 40, Line 15: Here again, it would help readers to state more explicitly whether these five trials are among (or in addition to) the 16 trials mentioned in the first paragraph on this page, as well as to provide the citation numbers for those 16 trials in the first paragraph. Page 40, Line 22: inserted text Page 40, Line 53: These paragraphs and Appendix Tables detailing specific changes from the 2006 review for each outcome are likely to be very helpful to readers. Page 41, Line 11: Would suggest verifying that values are identical to "current and excluded." - Note also, per forest plot the CI be 2.7, 4.4 (rounding). Page 41, Line 46: This style or approach for citations makes it crystal clear to readers which specific trials the paragraph will summarize. It would help to apply the same citation style in the section summarizing Hb response. Page 43, Line 14: Recommend adding the citation numbers for these 28 trials, if possible Page 43, Line 19: Might be useful to mention whether any of these trials blinded physicians making decisions on transfusion to the treatment arm to which patients were assigned. Page 43, Line 22: Might be useful to add a sentence here on the comparison of RR and CI for transfusion between high-quality and low-quality trials, and mention the analyses shown in Table 25 (p 46). Page 44, line 37: inserted text	Inserted “Five additional trials . . . “ and citations added to all tables. Inserted No response required Corrected Adding citations to all tables should, we believe, also prove helpful. Citations have been added to all tables, including Table 14 referred to in this sentence. Added “nine unblinded trials . . .” Heterogeneity is due to a variety of characteristics; to single out one source potentially infers too much importance; Table 25 summarizes all identified sources of heterogeneity. Inserted

Commentator & Affiliation	Section	Comment	Response
	Results continued	Page 45, Line 8: This figure is a particularly useful way to visually convey effects of treatment across all trials reporting each outcome. It makes it easy for readers to see why authors conclude that ESA treatment consistently increases Hb response (but does not achieve response in all patients in any trials) and consistently decreases transfusion rates (but doesn't protect all patients from transfusion in any trial).	No response required
		Page 46, Line 5: Given the heterogeneity between trials (i.e., both patient and study characteristics), would a multivariate meta-regression have been more appropriate? (This comment applies to all outcomes.)	While a multivariate meta-regression could be performed the purpose of heterogeneity analysis was to explore reasons. A multivariate analysis has potential to be difficult to interpret.
		Page 47, Line 49: Was on-study mortality defined as such across all studies? Are there any limitations with available data?	Revised the paragraph to clarify. Only 2 included studies in the draft were not part of Bohlius et al on-study mortality analysis, see Table 14.
		Page 50, Line 13: Recommend adding the citation numbers for the 37 trials here.	Added citations to all tables.
		Page 50, Line 36: Recommend adding the citation numbers for these 7 trials	Added citations to all tables.
		Page 53, Line 43: Might be worth noting here that trials directly comparing epoetin versus darbepoetin were excluded from the IPD meta-analysis since they lacked no-ESA controls, and that's why on-study mortality data were unavailable.	Added comment.
		Page 53, Line 45: Recommend adding the citation numbers for these 31 trials	Added citations to all tables.
		Page 53, Line 54: Many readers would benefit from having the difference between CI and CrI explained somewhere in the Methods chapter.	Added explanation in Methods.
	Results continued	Page 53, bottom: Since it was not examined in the 2006 review, it might be useful to compare the results obtained here with those reported by Bohlius et al. in 2009.	Added comment on comparison to end of Results paragraph in Meta-analysis of Survival Outcomes section.
		Page 54, Line ~10: With such wide confidence limits and a hazard ratio considerably closer to 1.0 than for epoetin, this may overstate the analytic result. It seems more reasonable (because it's a bit less definitive) to say it does not appear to be inconsistent with an increased risk of mortality than to say it is consistent with an increased risk.	Saying “consistent with” given the consistent findings we believe is appropriate.
		Page 56, Line 5: It might be useful to provide an explanation or rationale for choice to explore these specific trial characteristics as sources of heterogeneity (here, and for other outcomes). For example, baseline Hb was explored but not	Subgroups were chosen based factors likely to clinically impact outcomes and as potential sources of bias (as listed in Methods Data analyses). Blinding was singled out not only because a potentially critical element of study

Commentator & Affiliation	Section	Comment	Response
		target Hb or maximal achieved Hb. Also, since blinding is one of the factors that determine study quality, why were both these characteristics evaluated? Page 56, Line 42: Shouldn't the influence of baseline (control arm) risk on the absolute increase in risk attributable to ESA treatment also be evaluated and described? Is it appropriate to focus exclusively on relative risk here? It seems at least possible that when the baseline risk is small, that absolute increase in risk may also be smaller even if relative risk exceeds that seen with a larger baseline risk. Page 57, Line 51: inserted text “more” suggested Page 58, Line 10: Given the wide confidence intervals, perhaps better to rate precision as moderate or medium? Page 58, line 53: inserted text “more”	quality. These are still, being study level, exploratory and should be interpreted cautiously. The issue of dependence of relative effect on control group absolute risk is one common to many meta analyses and so addressed here. The issue of relative and absolute effects is relevant and we have added numbers needed to harm that we believe addresses these points. We believe this is the informative result, so elected not to change text. Regarding application of GRADE, results are sufficiently precise according to current AHRQ Guidance. “.. when the total sample size across the body of evidence is reasonably large (e.g., 4000 patients), EPCs can consider the estimate to be precise because even with a low number of total events, prognostic factors are likely to be evenly distributed.” For a modest increase in risk our interpretation of the CI width as precise is in keeping with the above. This has been rephrased.
	Results continued	Page 59, line 37: Here, and throughout the review, the authors do a thorough job of informing readers about which studies they excluded and why they were excluded. Page 61,line3: Suggest adding a few more details and/or a footnote to this table that show limitations: e.g.,- sample/disease characteristics of included studies - sources of data (e.g., didn't Glaspy et al use a mix of published, updated, and IPD) Page 64, line 18: Should include citation numbers for the 2 trials that favored control. Page 64, line 25: What are the potential implications for including all of these	No response needed. While a reasonable suggestion, we have elected to focus on the AMSTAR domains. Added citations Inserted additional language to clarify.

Commentator & Affiliation	Section	Comment	Response
		varied outcomes? Page 64, line 24, 29: inserted text Page 65, line 52: Recommend adding the citation numbers for these 29 trials Page 66, line 21: Recommend adding the citation numbers for these 5 trials Page 66, line 48: inserted text “absolute” Page 67, line 14: Recommend adding "absolute" to title of Fig 13 and to label for the horizontal axis Page 72, line 10: A brief summary paragraph at the end of this section would be useful to readers. Currently, it may leave the reader uncertain whether there is or is not adequate information to establish a "benchmark" for clinically meaningful change in HRQoL scores that used FACT-An. Page 74, line 4: Table 48 does not inform readers which 17 trials reported FACT-Fatigue scores and of the 17, which 7 reported no statistically significant improvement in scores for the ESA arms. Adding citation numbers to the text (at least for the 14 included in the meta-analysis) might be helpful to most readers.	 Inserted Added citations Added citations We have not added as risk differences are by definition absolute. Inserted summary text Changed Table 48 to Table 50, which indicates the studies reporting statistically significant results; citations added to all tables as suggested.
	Results continued	Page 74, line 12: inserted text Page 75, line 25: What about other implications, such as over-/under-estimating the effect? Page 75, line 41: Given all the limitations mentioned in the preceding paragraph, rating the available evidence as a "precise" estimate for HRQoL effects of ESA treatment seems inappropriate. Page 79, lines 10, 11, 16, 19, 20: typing errors noted Page 86, lines 17, 37, 43, 52: Recommend adding citation numbers for these three trials, and for the trials mentioned in the two sections below.	Inserted The potential for bias is addressed in GRADE evidence table. Precision is based on the confidence interval of the estimate; the estimate can be precise, even if potentially biased. The discussion points out the limitations of the estimate and states that the overall strength of evidence is low. GRADE summary tables were revised to include strength of evidence. Revised sentences for clarity. Citations added to all tables.

Commentator & Affiliation	Section	Comment	Response
		Page 89, line 50: inserted text Page 90, lines 56, 57: typing errors noted Page 92, lines 34, 42: If results don't achieve statistical significance, perhaps ratings of Consistent and Precise somewhat overstate the strength of evidence? Even if point estimates are all in the same direction, wide confidence intervals should caution against over-emphasizing the directionality of their midpoints.	Inserted Revised sentences for clarity The text in the 'Evidence GRADE for Central Outcomes' states that the "overall strength of evidence [is] low." GRADE summary tables were revised to include strength of evidence. We have conformed to AHRQ guidance on GRADE.
	Results continued	Page 95: The decision analysis section lacks adequate rationales for the assumptions of the model. Furthermore, the sources of the estimations need to be better described and made more explicit. This is especially important for the 5% annual mortality rate following the on-study period (noted on page 96 of the report). With better descriptions of the assumptions in this section, and the basis thereof, readers will be able to make more informed judgments about the validity of this analysis. Page 95, line 34: Note, this also assumes that the mortality rate over the period is the same for those treated with chemotherapy for adjuvant intent (even curative) and those treated for recurrent/refractory/metastatic disease. Or perhaps this represents an average of the rates, in which case this should be better explained. Also the assumption that the mortality rate is constant over the subsequent year should be justified/defended. Page 95, line 37: Should probably cite the 3 individual studies, rather than relying on citation of the Wilson et al systematic review (79) from the paragraph above. Page 96, line 22, 26: typing errors noted	References and rationale have been added. The reviewer correctly points out limitations, reflecting that the expected values obtained are approximate. Unless mortality rates are differentially not constant following ESA or no ESA, using an average will not bias the result. The review does not suggest that is the case. The individual studies were provided by the manufacturers to Wilson et al and are not otherwise citable. Revised sentences for clarity
	Results continued	Page 96, line 30: This assumption is unclear. Underlying risk of mortality in control arms combines the risk of death from progressive malignancy and risk of death from adverse effects of cancer treatment. For those undergoing adjuvant therapy after successful local treatment, these risks would decrease substantially once adjuvant therapy ends. Note that the RCTs used to derive estimates for mortality and other outcomes were done before FDA limited the labeled indication to patients NOT undergoing treatment for cure, and thus in many cases included patients receiving adjuvant therapy. For those undergoing treatment for advanced, recurrent, or metastatic disease, chemotherapy often continues well	The section has been extensively revised to address the issues raised. We elected to define the base case as one most representative. It was not our intent to cover all circumstances, but to examine the tradeoffs. The issue of labeling and relevance of results to current practice is addressed in the discussion.

Commentator & Affiliation	Section	Comment	Response
		past the 12-16 weeks such patients were enrolled on an ESA trial, and the control group's risks for mortality probably did not change much when the study ended. The underlying assumption is questionable of a single base case that lumps patients undergoing defined-term adjuvant therapy with those undergoing extended treatment for more advanced disease. Page 96, line 51: This is the on-study mortality rate which when applied to the "controls" may be the underlying rate. Suggest consistent terminology to be clear Page 96, line 53: This concept will not be well understood as written for the general reader, and could be explained in a few sentences.	 Also revised. Also revised.
	Results continued	Page 96, line 54: If this is being interpreted correctly, the relative impact of ESAs on on-study mortality is assumed to be level across all indications for chemotherapy. Therefore, in situations where the baseline mortality rates is worse, the absolute effect of ESAs on mortality will also be greater, which leads to "worse" tradeoffs in those with more disease/underlying mortality. Given that it was the later studies that demonstrated risk of death in groups of patients with metastatic/late disease, it may not be appropriate to assume the same relative impact (risk) of ESA use across all indications for chemotherapy (adjuvant vs. metastatic). At least this should be elaborated upon more in the assumptions and limitations sections. Note, it is interesting based upon this model as currently written that those with the worse prognosis have the greatest harms, which is in some ways consistent with the renal indication (greater dosing for non-responders leads to more risk). But, and you may not wish to comment upon this directly, it is also interesting in that the FDA introduced the limitation for those with diseases intended to be cured, a group with arguably the lowest baseline mortality rate. This was likely from a frame of reference to protect such a group from ANY harm. Page 97, line 22: Could not find information on how the model and assumptions used for Table 70 differed from those used for Table 69 and from those used for the tables in Appendix H. Page 97, line 36: inserted text Page 97, line 41, 44-46: confusing; on p. 95 authors state utilities are based on three studies	We have addressed these issues in the revisions. The decision model requires assumptions. Whether to employ a relative risk that is constant or one that varies with underlying absolute risk complicates those assumptions. Because readers may find the notion of varying underlying risk difficult, we chose to not to use that for the main result. However, the sensitivity analyses provide relevant results with varying relative risk. The reviewer is correct regarding tradeoffs according to underlying risk. The table has been replaced. Text has been revised. Corrected in revised text.
	Discussion	Page 99, line 11: Since pharmacokinetics may be the major difference between darbepoetin and epoetin, authors should be hesitant to say they are pharmacologically similar. Given the extended clearance time (and thus long duration of action) for darbepoetin, it's uncertain that one should even consider epoetin and darbepoetin similar with respect to pharmcodynamics. Would be	Substituted mechanistic.

Commentator & Affiliation	Section	Comment	Response
		more appropriate to state "biochemical and mechanistic similarities." Page 99, line 11, 13: Since the body of evidence on HRQoL comes from studies that were mostly rated as poor quality, can one really say that it provides sufficient certainty? Page 99, line 27: Suggest saying this as "...slightly better scores..." or "...modestly better scores..." since the magnitude is less than MCID. If it's acceptable to say that the decline for those randomized to transfusion is small, it should also be acceptable to use a modifier for the size of ESAs effect. Page 99, line 37: Since there was no summary or discussion of evidence on the mechanism(s) by which ESAs might increase mortality, it's unclear what basis there is to say it is consistent with a "...biologically plausible causal effect..." Page 99, line 40: The reasoning here is unclear. Which two relative risk estimates are being compared? It's also unclear why the inclusion of trials on patients "regardless of cancer treatment" imply a longer duration of ESA exposure. How does this argument strengthens support for a causal effect? Page 99, line 53: In reality, current evidence doesn't speak to the question.	Modified to "provide sufficient certainty to address." We have adopted the use “better.” The word “biologically” has been deleted. The intent was to discuss a plausible causal effect on mortality. Changed 'higher' to' increased'; consistent effect across different study samples, No response needed.
	Discussion continued	Page 100, line 13,17,18: typing errors noted Page 100, line 19: The conclusions in the decision model need to be better supported by citations to sources on which the assumptions of the analysis are based. Page 100, line 22: Difficulty reaching this same conclusion based on the model. More explanation about this conclusion (and model - as noted previously) is needed here. Page 100, line 23: May want to explain this concept more. Page 100, line 28: The report should state more explicitly that, despite the large number of trials of ESAs across populations, the mechanism of action attributed to ESAs remains poorly understood. The current evidence does not inform the mechanism of action and thereby suggest which patients may be at lesser or greater risk of harms from use of ESAs, and/or which patients may be more likely to experience a benefit(s) from these agents. This is a serious limitation of this	Sentences have been revised. Done. Now addressed in revision of decision model and calculation of NNH. We believe that within the sections on quality of life and decision analysis that this is addressed. It is not truly necessary to understand the mechanism of action to be able to draw conclusions from the data regarding benefits and harms. There is a plausible explanation of harms based on thromboembolic complications. Inserted

Commentator & Affiliation	Section	Comment	Response
		literature, and it makes it particularly difficult to provide guidance concerning the safe use of these agents. Page 100, line 56, 57: inserted text 'by' Page 101, line 18: inserted text 'stage of malignancy, treatment regimen and response," Page 101, line 19: This statement may be a bit too strong. It's entirely possible that registry studies could generate new hypotheses worth testing in an RCT that enrolls a more narrowly defined patient subset than used for trials to date.	We have elected to not speculate on which subgroups might be a higher or lower risk so have retained the current text. We do not tend to agree, but that decision is ultimately the purview of others based on further evidence as it may develop.
	Conclusion	None	
	Figures	None	
	References	None	
	Appendix	None	
	General	None	
Peer Reviewer #8 AMGEN	All Sections	See below	
		Comments from Amgen were sent in the form of a long letter from which we abstracted all individual comment items. Necessarily, not all of the text from the letter was reproduced in order to avoid duplication. However, no topics were avoided.	
		The core AHRQ analysis, while rigorous, includes studies that do not reflect current use or restrictions on usage associated with ESAs. Most notably, studies of patients with radiotherapy **(not** indicated in current ESA labeling) have been included in the AHRQ analysis, while results from large, well-controlled, pharmacovigilance studies in cancer patients being treated with chemotherapy are absent. When a sensitivity analysis is performed to account for the respective inclusion and exclusion of the results from these studies, Amgen finds that the estimate of on-study mortality is neutral (i.e., no on-study mortality risk associated with ESA use). Amgen has included the results of this important analysis on pages 4-6 of our more detailed response.	The reviewer is correct that FDA labeling does not approve ESA use in the setting of radiotherapy alone, while this CER focuses on "patients undergoing cancer treatment." In our analyses, while the relative risk of on-study mortality was higher in the radiotherapy trials there was no evidence of effect modification (Table 36). While Table 36 reports the result from a meta-regression with radiotherapy as a covariate, excluding the radiotherapy studies diminished the relative risk even less (1.16 [1.03, 1.30]). Accordingly, excluding those trials would not alter any conclusion. These results support the premise that the mortality risk of ESAs is not modified. Also, please note that we have pointed this issue out in both Executive Summary and Discussion. "Much of the evidence included here was obtained under treatment protocols that used higher

Commentator & Affiliation	Section	Comment	Response
			baseline and target Hb levels than those used in current practice. While it is possible that adverse event rates might be somewhat different with lower baseline and target Hb levels, we found little difference in effect when baseline Hb was less than, or exceeded 10 g/dL. This result is similar to an individual patient data meta-analysis.[4] Additionally, three trials included in Key Question 1 enrolled patients predominantly undergoing radiotherapy. Although not an FDA-approved indication for ESA use, those results were included because the population of interest was patients undergoing treatment for cancer. Moreover, we did not find those trial results influential in these analyses."
		The broad AHRQ conclusion also stands in contrast to conclusions drawn by the Food and Drug Administration (FDA) and other global regulatory bodies that have continued approval of ESA therapy for use in appropriate cancer patients with anemia due to concomitant myelosuppressive chemotherapy. The FDA-approved labeling takes into consideration the same clinical data analyzed in the Draft Report, and reflects the acknowledged concern about risk of increased tumor progression and/or shortened overall survival (OS).	We believe these results support current black box warnings on ESA use. The report offers conclusions based on careful consideration of the evidence identified and synthesized within the report. It neither offers, nor recommends any decisions in the regulatory domain.
		The Draft Report's conclusion statement also stands in stark contrast to clinical guidelines from professional societies such as the National Comprehensive Cancer Network (NCCN) and the American Society of Clinical Oncology (ASCO) that reviewed essentially the same evidence base. For example, the most current NCCN guidelines (version 1.2012) state that "[r]ecent pharmacovigilance trials have reported no adverse effects on survival in cancer patients with chemotherapy-induced-anemia receiving ESAs." Amgen requests AHRQ review and carefully consider the alternative analyses provided before any final report is issued.	The quoted text from NCCN guidelines refers to three studies (Engert 2010, Moebus 2010, Untch 2011), all contained in the CER. The NCCN comment does not specify type of survival evaluated. We focus on on-study mortality, because during longer term post-treatment follow-up, subsequent nonrandom interventions can affect overall survival (time-dependent confounding), potentially causing a bias to the null. Engert 2010 did not report on-study mortality. Our on-study mortality analysis included Moebus 2010 and Untch 2011 (total N=1372), however these studies contributed no events from either treatment or control arms indicating unique patient populations at such low underlying risk of mortality that at the enrolled sample sizes no survival effect could be estimated. We included 15 additional studies in our analysis of on-study mortality (see Figure 11) for a total N

Commentator & Affiliation	Section	Comment	Response
			of 11,266. Subgroup analysis by platinum-based chemotherapy or by chemo- or radiotherapy did not significantly change the results of the analysis (Table 36). Moreover, studies included patient populations at a variety of baseline mortality risks. We found that relative risk of mortality is higher in trials enrolling patients at lower (but finite) risk of mortality during the active treatment period. Please note that a single study, or even three, unless exceedingly large does not provide significant support for a conclusion of no increase in risk of mortality. For example, assume that the relative risk of on-study mortality estimated in the CER is correct (ie, 1.17). A single study of 1000 patients per arm (2000 total) with a control group on-study mortality rate of 7.5% would have a power of 16% to detect a RR of 1.17. If the control group on-study mortality rate was 2.5% the power would be 7%. To achieve a power of 80% to detect a RR of 1.17 with a 2.5% on-study control arm mortality (low risk patients) and 1:1 randomization would require almost 50,000 patients.
		The inclusion of patients with radiation-induced anemia, in the absence of chemotherapy, for the meta-analysis of on-study mortality and health-related quality of life (HRQoL) is not consistent with the current FDA-approved indication. Recent results from large, randomized, well-conducted, single-tumor, pharmacovigilance studies should be considered for inclusion in a meta-analysis of ESA data.	See following comments
		Amgen believes that if radiotherapy-only studies are excluded from the analysis and on- study mortality results from important large pharmacovigilance studies are included in the analysis, the results will support the continued access for appropriate patients to this important anemia therapy. Amgen notes that the analysis provided in Figure 11 (Page 55) includes 3 studies from anemic patients with head and neck cancer receiving radiotherapy alone. The use of ESAs in these 3 trials was not to treat anemia, but rather to determine if ESA therapy would augment the anti-tumor effect of radiation therapy and in none of these 3 trials did the mean hemoglobin in the placebo arm fall to less than 12 g/dL. Amgen believes that these studies should be removed from the analysis of overall survival and HRQoL (if applicable) since treatment of anemia	The matter of appropriateness of including radiotherapy only trials in the main result is addressed generally above. We have included radiotherapy as a modifier in the meta-regression as noted. In our analysis excluding those 3 trials does not alter the effect meaningfully or change any conclusions. In addition, redrawing Figure 12 in the report according the treatment modality shows also that the 3 radiotherapy only studies fall quite consistently with the other study effects (see Figure 1 below). Should BEST be excluded or an analysis

Commentator & Affiliation	Section	Comment	Response
		due to radiotherapy is not an FDA-approved indication for ESA therapy. Amgen has done a reanalysis of the on-study mortality data that excludes radiotherapy studies and otherwise includes data as shown in Figure 11 (page 55) of the Draft Report. The effect of removing these 3 studies was minimal in the analysis of on-study mortality with an overall point estimate of 1.15 (95% CI: 1.02, 1.30). In this reanalysis, the Breast Cancer Erythropoietin Survival Trial (BEST) study had a weight of 20.3%. When the BEST study was excluded as part of a standard influence analysis (excluding 1 study at a time then re-estimating the point estimate and 95% confidence intervals) in Amgen's reanalysis, the point estimate for on-study mortality was 1.10 (95% CI: 0.98, 1.24). The impact of excluding individual studies, other than the BEST study, only minimally impacted the overall estimate of risk.	presented excluding those results? The commentators are correct regarding the weight accompanying the trial result—it is therefore influential on the simple pooled result. However, that observation is not reason to exclude the trial as outlined here. BEST was large and judged a high quality trial, notwithstanding some criticisms. For the main result we have presented the pooled result not accounting for the association with underlying risk (control group mortality). However, the relative effect is modified by underlying risk and is accordingly more accurately represented by accounting for it (e.g., see Table 37). When considered in this manner—and in Figure 2 following (symbols proportional to random effects weights in the simple pooled analysis; weighted regression meant to mirror the result obtained in the correct Bayesian analysis)—it is clear that BEST is consistent with the other trial results. When underlying risk is included in the analyses for a 7.5% 16-week on-study mortality rate, the pooled RR was 1.25 (95% CI: 1.10 to 1.43) (Table 37); excluding BEST 1.25 (95% CI: 1.07 to 1.46). While the simple pooled result does not account for underlying risk, it reflects an appropriate average effect familiar to CER readers. In conclusion, excluding BEST is not justified either based on these analyses or on clinical grounds.
		In addition, Amgen did the same type of influence analyses of the on-study mortality data in Figure 11. In the AHRQ analysis shown in Figure 11, the study (BEST) reported by Leyland-Jones et al has a weight of 19.3%. The impact of excluding individual studies, other than the BEST study, only minimally impacted the overall estimate of risk. However, when the BEST study was excluded, the point estimate for on-study mortality decreased from the reported value of 1.16 (95% CI: 1.03, 1.31) to 1.11 (95% CI: 0.97, 1.26). The BEST study was the largest randomized ESA study included in the meta-analysis, and also included the longest on-study period (12-months) while most other studies included a 3- to 6-month on-study period, resulting in the BEST study having a very large weight in any ESA meta- analysis.	Most of the points raised are addressed in the previous reply. While it is true that BEST was the longest trial, as shown in the figures here it is not an outlier. Its weight is due to its size and underlying risk (control arm mortality rate). As far as BEST having the longest follow-up it is important to note that the mortality difference was evident by month 4 and over 12 months the authors did not report any suggestion of non-proportional hazards. Given a constant relative hazard, the only impact of the longer follow-up is then a greater number of events without change in effect size (HR). The previous analyses have already shown that when accounting for

Commentator & Affiliation	Section	Comment	Response
			underlying risk, there is no suggestion of bias in the estimated effect. While one could truncate the result at 4 months and perform the pooled analyses, the result would differ only minimally in the CI, and not in the point estimate.
		Data from all pharmacovigilance studies in the chemotherapy-induced anemia (CIA) setting reported since 2008, such as those reported by Pirker et al,10 Engert et al (HD15EPO study),4 Delarue et al (GELA study, interim analysis),11,12 and Möbus et al,13 have been reassuring and have reported neutral mortality with ESA therapy. However, with respect to the AHRQ analysis of on-study mortality, the results reported by Möbus et al and Untch et al were excluded from the analysis since no on-study deaths occurred in either study. While this is appropriate when analyzing data as relative measures in accordance with the AHRQ Methods Guide, these data are clinically important to consider for safety analyses.	It is problematic to conclude that results from single studies are reassuring as they are severely underpowered to detect an effect. Finally, the other way to consider these 0 event studies is that because the underlying risk is low, the probability of demonstrating an effect is small, and even then one would expect only 1 or maybe 2 deaths. For example, see Fujisaka 2011.
		The majority of ESA studies had primary endpoints either related to hemoglobin change, transfusion reduction, or HRQoL change. Neither on-study mortality nor overall survival (OS) were endpoints for the majority of the ESA clinical trials. On-study mortality was typically determined based on adverse event data (e.g., an adverse event with a fatal outcome) or the reason for either investigational product termination or study termination. Therefore, this was not a specified endpoint and should be considered an adverse event endpoint.	Almost all on-study mortality were obtained through theindividual patient data meta-analysis by Bohlius et al. and we consider sufficiently valid for purposes of analysis.
		As mentioned above, the BEST study (n = 939) was very influential in the AHRQ meta-analysis of on-study morality and had a weight of 19.3% in the analysis. The BEST study enrolled patients with similar tumor type and stage, but allowed any chemotherapy regimen and dosing. The 2 breast cancer studies by Möbus et al (n = 643) and Untch et al (n = 729), respectively, both enrolled a patient population with similar tumor type and stage; however, unlike the BEST study,9 both of these studies randomized patients to receive specific schedules of the same chemotherapy agents. Furthermore, these 2 breast cancer studies are an integral part of the pharmacovigilance program because they directly address issues of ESA-related survival risk. Therefore, Amgen believes that the inclusion of the Möbus and Untch in the meta-analysis is important.	Based on these analyses, trials enrolling patients with low underlying risk would be anticipated to have few, or even no, events.
		In addition to the exclusion of the 2 large pharmacovigilance studies, there are limitations to the methodological approach used in the AHRQ meta-analysis of on-study mortality. Amgen proposes a meta-analysis using risk difference for evaluating on-study mortality. This is a recognized alternative method, which is considered to be robust and valid regardless of event rate or sample size, to compare observed risks that occur in clinical trials.16 The risk difference is important in the analysis of less frequently occurring events since using either relative risk (RR) or odds ratios (ORs) could inflate the perceived magnitude of the observed absolute difference between groups.17 As an example, consider a	The proposed pooling of risk differences is intuitively appealing. As noted it is not recommended in current AHRQ guidance which also states that it "may be considered when control rates are reasoanbly similar." This is clearly not the case here and even more problematic given the dependence of effect on underlying risk. While numerically feasible such an analysis is generally avoided.

Commentator & Affiliation	Section	Comment	Response
		treatment that causes a 1% increase in an event. The RR changes depending on the background rate whereas the risk difference would be constant at 1%. In this example, if the background rate is 0% (1% versus 0%), RR = ∞; if it is 1% (2% versus 1%), RR = 2; or if it is 10% (11% versus 10%) then the RR = 1.1. Use of risk-differences in a meta-analysis also allows the inclusion of studies where no events occur in either treatment arm, thereby affording incorporation of all available relevant data.18 Inclusion of these studies can meaningfully inform the overall assessment of risk and has been previously used by the FDA in meta-analyses.20,21 Amgen understands the position, per the AHRQ Methods Guide, that this method has been reported to demonstrate poor performance for rare events; however, the incidence of on- study mortality was as high as 21% in some of the included studies—an incidence that while infrequent, should not be characterized as a rare event.	(See also Deeks JJ, Altman DG. Effect Measures for Meta-analysis of Trials with Binary Outcomes. Systematic Reviews in Health Care; Meta-analysis in Context. BMJ Publishing Group, 2001)
		Amgen has also reviewed the meta-analysis of HRQoL results provided in Figure 14 (page 74) of the Draft Report and has reanalyzed these data excluding the Hoskin7 radiotherapy-only study (Figure 2, page 7). Amgen has also made corrections to the data provided for the studies based on a review of these publications/study data. Notably, the estimated difference in change in FACT-Fatigue score is 3.04 (2.00, 4.08) for the chemotherapy studies in this revised meta-analysis versus 2.72 (1.69, 3.74) reported in the AHRQ analysis. This value is greater than the medical clinically important difference (MCID) value reported by Cella et al. (2002)22 and cited in the AHRQ document (page 75). To more accurately reflect the underlying data, Amgen suggests that the authors reanalyze the HRQoL data excluding the radiotherapy study (i.e., Hoskin et al.) and to consider the corrections made to the results of the other studies.	We have applied to corrections noted and revised analyses. Hedenus 2003: the ESA and control means appear to have been taken from ODAC 2007 slides and data on file. We could not find their numbers and retained our numbers, abstracted independently by two reviewers from Figure 3 in the publication; Chang 2005: 1.55 was reported in the text of the article, 1.85 in the abstract of the same article. Initially we used 1.55 but changed to 1.85 as 1.55 appears to be a typo; Tsuboi 2009: the authors reported results with and without imputed values. Since the majority of studies did not impute data we prefer the results without imputation for consistency across studies; Witzig 2005: we agree and changed the SD from 14.48 to 11.48. We revised the meta-analysis (without omitting studies, per prior comments) and obtained an estimated difference in change in FACT-Fatigue score of 2.74 (95% CI 1.69, 3.78; I2 45%). Because QoL study results can be strongly biased by the placebo effect if study participants are not blinded to treatment, we also analyzed the subset of trials reporting FACT-Fatigue that blinded patients to ESA vs. control treatment. A meta-analysis of these 8 trials resulted in an estimate of 1.92 (95% CI 0.97, 2.86; I2 0%).delarue

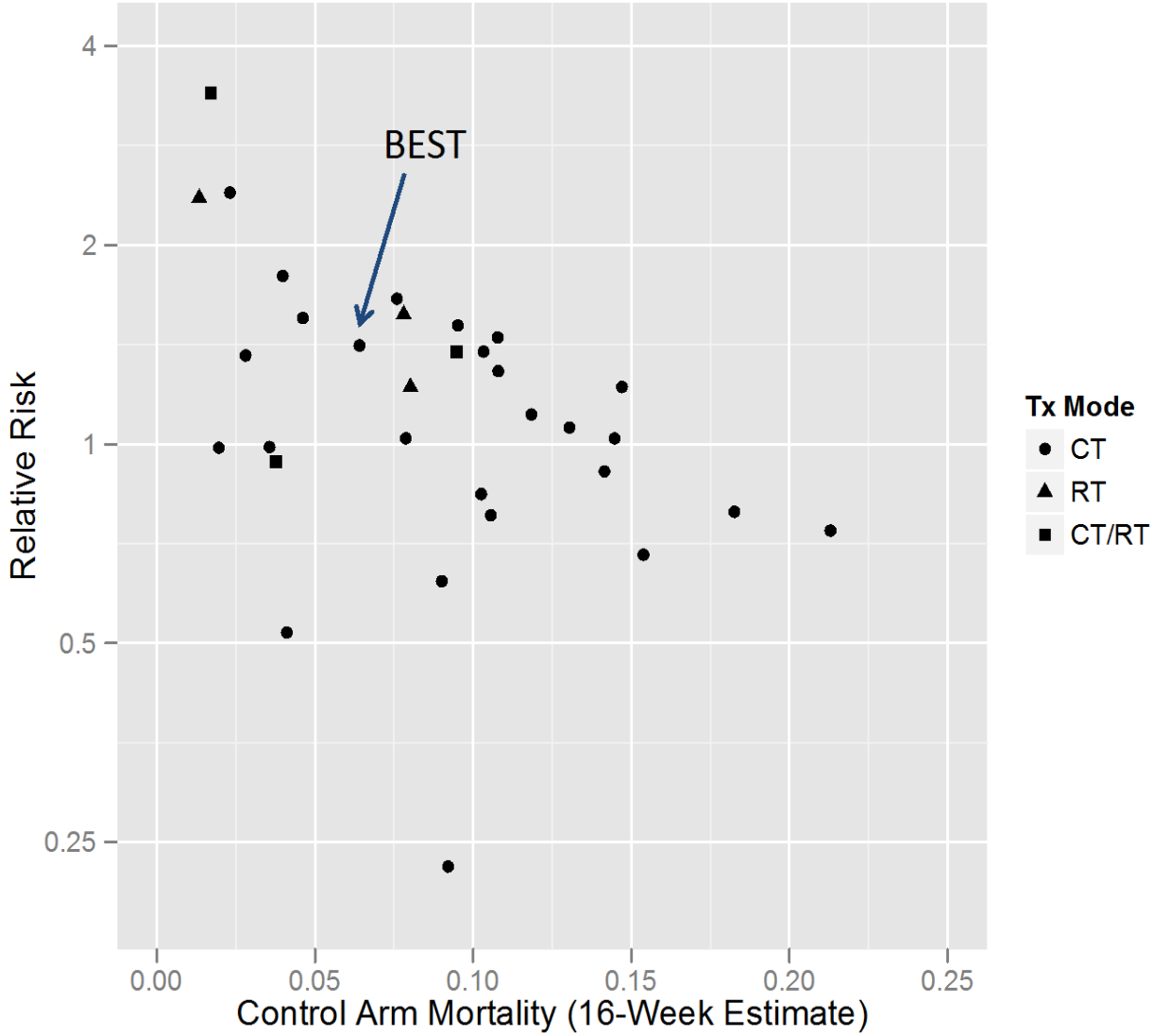
Commentator & Affiliation	Section	Comment	Response
		Page vi under "Structured Abstract," the authors indicate that 'Multiple sources (n=13) were searched for potential grey literature. A primary source for current survival evidence was a recently published individual patient data meta-analysis.' The studies in this section were initiated or completed prior to the 2007/2008 ODAC meetings where the relevant data which inform the current approved indication, limitations of use, warnings and precautions, and dosing recommendations were discussed. This selection of literature does not accurately reflect the patient population for which EPOGEN® and Aranesp® are currently indicated. In addition to patient population heterogeneity amongst the selected studies, many of these studies also incorporated hemoglobin targets or ranges that are above what is currently recommended in the FDA-approved labeling. Therefore, Amgen believes that ongoing pharmacovigilance studies will provide additional survival data with ESA use. A listing of ongoing pharmacovigilance studies provided	We have noted now more clearly in the text differences in use in the included trials and current indications. We appreciate the list of ongoing trials.
		On page ES-4, under "Executive Summary," the authors state that 'Under circumstances representative of patients included in these trials, a decision analysis shows ESA use is always accompanied by a net loss of life-years due to increased mortality during the active treatment period.' As described earlier, radiotherapy-only studies were included with chemotherapy studies in the AHRQ meta-analysis. Because the decision model results are primarily driven by estimates for on-study mortality and OS, these estimates should reflect the population for which ESAs are indicated. To evaluate model uncertainty, the decision model should incorporate additional estimates of on-study mortality, OS, and confidence intervals associated with point estimates. Thus, drawing firm conclusions based on modeled scenarios is not appropriate. Amgen suggests an update of the decision model to correspond to appropriate cancer patients with CIA in alignment with the meta-analysis.	We appreciate the comments and have mad a number of changes in response. First, we have taken out detailed reference to the decision analysis from the abstract and summary, and have used it in an ancillary role. Second, we have deleted "always" from the sentence quoted. Issues surrounding radiotherapy trials have been commented on previously and noted in Figure 1.
		Amgen replicated the AHRQ Markov decision model using different scenarios (TreeAge Pro 2011) and varying on-study mortality and OS assumptions (Table 2, page 10). In the first scenario, the life years lost for 1,000 patients was 6.7 years using on-study mortality hazard ratio (HR) from the Bohlius meta-analysis of patients with CIA, compared with 12.1 years reported in the AHRQ Draft Report. Applying the Bohlius 4-month on-study mortality HR in 1,000 patients with CIA in the second scenario resulted in a difference in life years lost of 2.0 years for ESA as compared with the control arm. In the Amgen-performed meta-analysis, the estimated risk difference across studies is 0.0 (95% CI: 0.0, 0.01) and using on-study HR from the meta-analysis resulted in no difference in life years lost for ESA patients as compared with the control arm. These sensitivity analyses did not incorporate disutilities that would be associated with differential non-Response and lower hemoglobin levels associated with the control arms.	We have also replicated the model in TreeAge Pro with results identical to those obtained using Excel. The report provides all the assumptions and parameter estimates for replication. Our judgment is that using a risk difference is not appropriate under the circumstances here as previously noted, and should not be used for a decision analysis.

Commentator & Affiliation	Section	Comment	Response
		Amgen suggests that the authors test the robustness of the decision analytic model through the use of probabilistic sensitivity analyses that evaluate assumptions based on additional mortality results.	These have been included (Tables 70 and 71)
		On page 95, under "Decision Analysis" section, the authors indicate that a Markov decision model was constructed using base-case parameters presented in Table 68 (page 96) of the Draft Report. The authors state on page 97 that 'The limitation of the current Markov model used for this analysis is that it is a basic representation of circumstances that are otherwise complex – for example, dose, escalation strategies, cancer therapies.' Amgen agrees that the model has not been fully explored as it does not accurately represent the full range of clinical scenarios and is based on critical assumptions that encompass substantial uncertainty. For example, the authors assumed that patients with a hemoglobin concentration of 9 g/dL would stay at the same level if they did not respond to either treatment and that these patients' hemoglobins would eventually improve to 11 g/dL after week 16. This unlikely scenario ignores the disutility experienced by a greater number of patients in the control arm (84% versus 46%) who do not respond and remain at a hemoglobin level of 9 g/dL or lower for an extended period of time.	The hemoglobin response rates were informed by the CER. Detail of decreasing and increasing levels by arms were not available. The response difference 52% vs. 14% is quite appropriate, and in sensitivity analyses over plausible ranges (Tables 70 an 71) can be seen to impact the results modestly. We believe the relevant points are not differences of single digits here, but that on average, 1) expected life-years are lost, 2) expected QALYs are gained, but also 3) that any expected QALYs gained are unlikely to be clinically significant for individual patients.
		III. Amgen suggests that the AHRQ Report include a balanced discussion of alternative analyses and conclusions, and critically evaluate reasons for the differences in conclusions.	The discussion was informed by CER results. We acknowledge uncertainties. At the same time, we have edited the final paragraph of the discussion (preceding Future Research) for clarity and to address our interpretation of the comment regarding a "balanced discussion." The same changes were applied to the Executive Summary.
		On page vi, the authors concluded that 'Overall, harms appear greater than benefits when ESAs are used to manage anemia in patients undergoing chemotherapy or radiation for malignancy.' The authors further state on page ES-5 under "Executive Summary" that 'Existing evidence establishes with sufficient certainty that ESA used to manage anemia in patients with cancer is accompanied by increased mortality risk' and that 'the increased risk of mortality raises questions as to whether equipoise exists to justify enrolling patients in clinical trials.' Amgen believes that the benefits of ESA therapy (when used in accordance with the current FDA-approved labeling in appropriate cancer patients with anemia due to concomitant myelosuppressive chemotherapy) outweigh the risks. Our position is based on the following:	We have substituted statements of specific benefits and harms for the reader.
		Amgen agrees with the authors of the AHRQ Report that there is consistent evidence that ESAs effectively reduce the need for RBC transfusions.	No response needed.
		ESA therapy is also reported to be associated with improvements in health-related quality of life (HRQoL).27-33	As discussed in the report.

Commentator & Affiliation	Section	Comment	Response
		The current Aranesp® label (approved by FDA on 24 June 2011) continues to support the risk:benefit of Aranesp®.	Not discussed in report. As noted, the existing evidence was not collected under conditions of use similar to current FDA-approved labeling.
		Amgen acknowledges that ESAs are associated with thromboembolic adverse reactions in patients with anemia due to concomitant myelosuppressive chemotherapy. This is well described in the approved labeling.	No response needed.
		Amgen agrees with the authors' assessment of the effect of ESAs on overall survival (OS). Prospective, well-designed studies that are homogenous with respect to tumor type and chemotherapeutic agents demonstrate a neutral effect of ESA therapy on OS and progression-free survival.	No response needed.
		Rather than a prespecified endpoint for the majority of the ESA clinical trials, on-study mortality was typically determined based on adverse event data or the reason for either investigational product determination or study termination. If the AHRQ analysis was modified to exclude radiotherapy-only studies and to include on-study mortality results from important large pharmacovigilance studies, Amgen believes the results will support the continued use of ESA therapy in appropriate patients.	Addressed in detail previously.
		Most of the data in the AHRQ analysis have been analyzed by several other groups, and have been made available to the FDA. Amgen has already updated the labeling for ESAs to account for studies assessed by AHRQ as discussed at the 2008	Relevant systematic reviews and meta-analyses were reviewed, critiqued, and judged consistent with the AHRQ results.
		Oncologic Drugs Advisory Committee (ODAC) meeting. Guidelines from professional societies, such as NCCN and ASCO, also support administration of ESAs to appropriate patients according to the current prescribing information, which describes that efficacy was demonstrated by a reduction in the proportion of patients receiving red blood cell transfusions. Since the 2008 ODAC meeting, there have been no new safety signals that would warrant the conclusion proposed in the current Draft Report.	The development or critique of guidelines is outside the report's scope. We have strived to be factual, and in fact neither said nor taken a position that ESAs should not be used. To the contrary we have presented the data synthesized in a manner that, if desired, can be used for decision-making. Importantly, we have made recommendations for future research that might help to disentangle some of the unknowns.
		A REMS has also been established to ensure that both providers and patients are aware of the risks associated with ESAs and assess the risk:benefit of these agents on an individualized basis.	The REMS program is briefly mentioned in Table 1. A discussion of REMS as it applies to benefits and harms is beyond the scope of this report.
		Therefore, Amgen suggests that the authors consider removing the radiotherapy trial data from the meta-analysis of on-study mortality. Amgen also suggests that the authors include an acknowledgement that the evidence base discussed in this AHRQ report is already incorporated into the current FDA-approved labeling for ESAs and therefore supports the conclusion that the benefits of ESA therapy outweigh the risks for appropriate cancer patients with anemia due to concomitant myelosuppressive chemotherapy.	The issue of radiotherapy trials has been discussed. The suggestion for such an acknowledgement is beyond the scope of the report.
		The clinical evidence base for use of ESAs in appropriate cancer patients with anemia due to concomitant myelosuppressive chemotherapy is considerable—	No response needed.

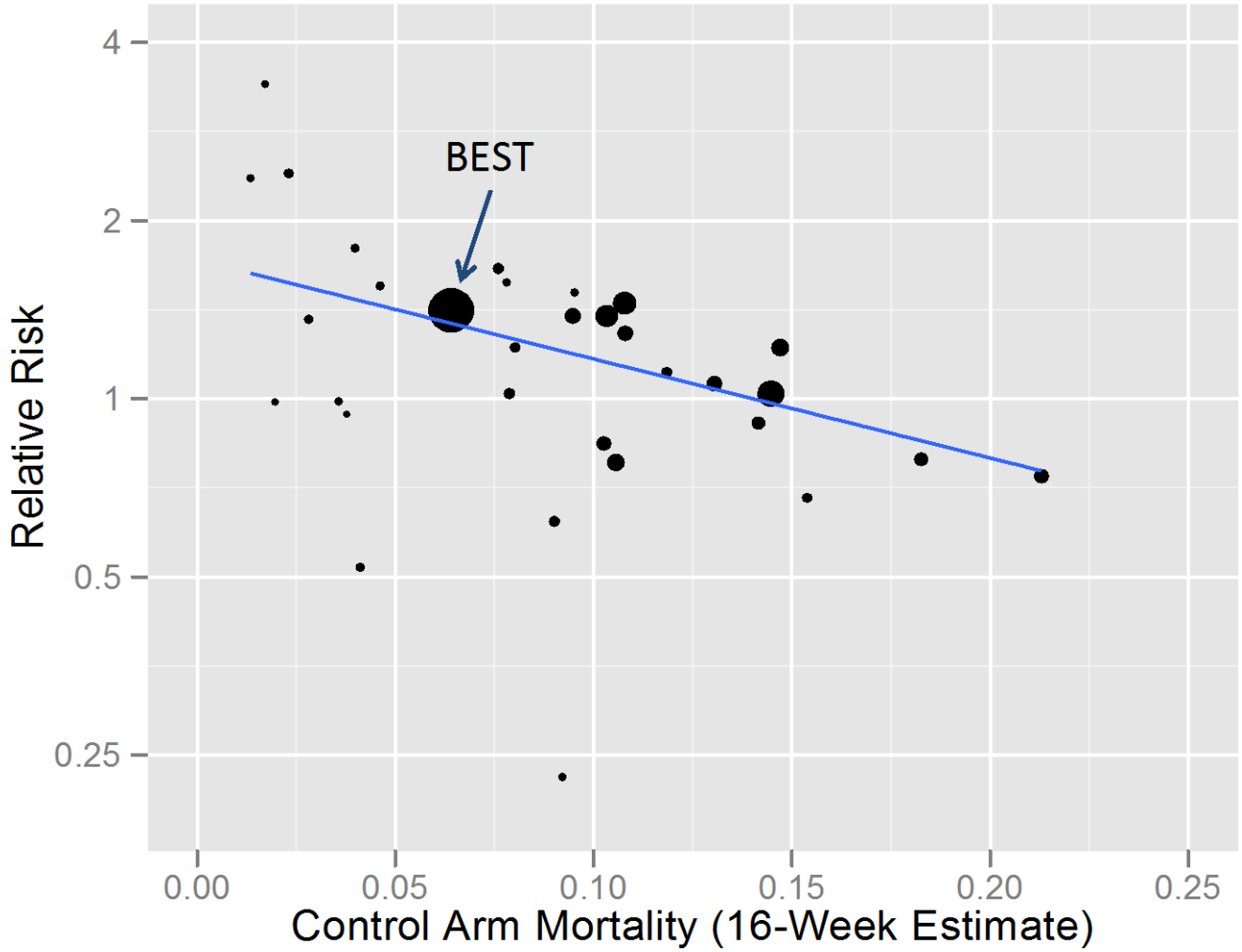
Commentator & Affiliation	Section	Comment	Response
		accumulated over 20 years in a very complicated disease state. Because of the length of time over which trials have been performed (with resultant changes in both study designs and heterogeneity), evaluating the quality of this evidence and drawing meaningful conclusions about the risk:benefit profile for ESA therapy is very complex. Amgen believes that the core AHRQ analysis, while rigorous, includes many studies that do not reflect the current use and restrictions on usage associated with ESAs. Most notably, studies of patients with radiotherapy (not indicated in current ESA labeling) have been included in the analysis, while results from large, well-controlled, pharmacovigilance studies in cancer patients being treated with chemotherapy are absent. When a sensitivity analysis is performed to account for the respective inclusion and exclusion of the results from these studies, Amgen finds that the estimate of on-study mortality is neutral (no on-study mortality risk associated with ESA use). Similarly, when sensitivity analyses are conducted with on-study mortality within the AHRQ decision analytic model structure, the net loss of life years also becomes neutral.	
		Amgen therefore firmly believes that the conclusion drawn by the Draft Report is not supported by the aggregate body of evidence. The conclusion is also inconsistent with conclusions drawn by the FDA, other global regulatory bodies, and professional medical societies. Given all of these issues and limitations of the Draft Report, Amgen suggests that the conclusion reflect these complexities and weigh the views of other agencies and groups that have analyzed essentially the same set of studies. Amgen further suggests that the authors refrain from drawing the conclusion that “harms appear greater than benefits when ESAs are used to manage anemia in patients undergoing chemotherapy or radiation for malignancy” when the evidence suggests that the ESA benefits outweigh the risk of harms in the treatment of appropriate patients with anemia due to concomitant myelosuppressive chemotherapy.	Please note previous response (..substituted statements of specific benefits and harms for the reader). We believe this addresses the first issues raised. An evaluation of processes and statements published in regulatory documents or by professional medical societies is outside the scope of this review.

Figure 1. Control arm mortality versus relative risk according to treatment modality (CT chemotherapy; RT radiotherapy; CT/RT both)



Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1480
Published Online: April 25, 2013

Figure 2. Control arm mortality versus relative risk.



Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1480
Published Online: April 25, 2013