Reviewer's report:

Title: Association of dialysis facility-level hemoglobin measurement and erythropoiesis-stimulating agent dose adjustment frequencies with dialysis facility-level hemoglobin variation: a retrospective analysis

Version: 1 Date: 17 March 2011

Reviewer: Eric Weinhandl

Reviewer's report:

Abstract

1. [MINOR] In RESULTS, re: “Modeling results suggested…”; the verb tense unexpectedly shifts to the present case.

2. [MAJOR] The following criticism is arguable, but I would nonetheless appreciate thoughtful consideration of it: I would contend that the presentation of modeling results would be clearer if the contrast in frequencies for Hb measurement and ESA dose adjustment were the same. Currently, the Hb measurement contrast is once per week vs. once per month (ie, about 1:4 on the time scale), while the ESA dose adjustment contrast is once per 2 weeks vs. once per 3 months (ie, about 1:6 on the time scale). This has the effect of suggesting to the reader that ESA dose adjustment frequency has a relatively similar effect upon Hb variation as Hb measurement frequency has, when in fact the former frequency appears to confer an effect of slightly lower magnitude on Hb variation.

3. [MAJOR] In CONCLUSIONS, re: “patients being consistently maintained in the desired Hb target range.”; I am struggling with the appropriateness of this language, because this study does not appear to consider proportions of patient-time in the Hb target range of 10-12 g/dL. Instead, the study is focused on characterizing Hb within 1 or 2 g/dL of mean Hb concentration (whatever it is, as the authors do not report it).

Introduction

1. [DISCRETIONARY] Re: “Although these RCTs evaluated Hb targets higher than the 10 to 12 g/dL range (the upper Hb target in these studies ranged from 13–15 g/dL), concerns have been raised by deleterious effects observed at higher Hb levels [11,14].” I do not understand the purpose of this sentence, in light of the sentence that it follows. It seems redundant.

Methods

1. [DISCRETIONARY] Re: “The primary limitation of the dataset... and the ability to map patients to individual dialysis facilities.” This text is suitable for the Discussion, but is a distraction in the Methods.
2. [DISCRETIONARY] Re: “Only hemodialysis (HD) patients with >30 days of data during this 36-month interval were selected.” The authors should clarify whether the cohort included only thrice-weekly in-center hemodialysis patients, or whether other variants of HD treatment (daily, home, nocturnal, etc.) were included.

3. [DISCRETIONARY] Re: “Although ESA administration was not considered an inclusion/exclusion criterion, all ESA administrations in the final study cohort were in the form of intravenous epoetin alfa.” Are the authors suggesting that absolutely no subcutaneous epoetin or intravenous darbepoetin use was observed? I would not expect much use of either among in-center patients in LDOs, but no use at all defies belief.

4. [DISCRETIONARY] I have no objection to the definition of a persistent dose change. However, I do want the authors to confirm that dose changes were identified on a patient-level before the dose history was divided into mutually exclusive bins defined by calendar months, so that calendar month boundaries were not capable of interfering with the identification of a dose change.

5. [MAJOR] We need more detail about the model parameterization. The predictor of interest is obviously Hb measurement frequency (or ESA dose adjustment frequency). Let’s confine the ensuing comments to one facility at a time. To estimate cross-sectional and longitudinal effects, I would guess that the authors included one fixed-effect term with the predictor equal to mean Hb measurement frequency across all months and another fixed-effect term with the predictor equal to the difference between Hb measurement frequency in a single month and mean Hb measurement frequency across all months. Of course, other variations on the core concept of both adding and subtracting a facility-level summary of Hb measurement frequency (even including Hb measurement frequency in July 2006) are possible. In any case, I’m guessing that the specified model has the following form, based upon what the authors have written about identical sets of fixed-effect and random-effects parameters:

\[
Hb \text{ variation}_{ij} = B_0 + B_1 \times \text{Mean Frequency}_i + B_2 \times (\text{Frequency}_{ij} - \text{Mean Frequency}_i) + b_{0i} + b_{1i} \times \text{Mean Frequency}_i + b_{2i} \times (\text{Frequency}_{ij} - \text{Mean Frequency}_i) + e_{ij},
\]

where i indexes the facility and j indexes the month; B_0, B_1, and B_2 are unknown constants; b_{0i}, b_{1i}, and b_{2i} are random effects that are assumed to be normally distributed; and e_{ij} is a random error. Am I right? It’s impossible to know.

6. [MAJOR] The above comment raises two other issues that the authors have not addressed: (1) how were degrees of freedom calculated for the purpose of confidence interval construction? (2) what was the assumed covariance structure of the random effects?

7. [MAJOR] I would also like to see the results of an unweighted analysis. Use of
weights is appropriate insofar as it accounts for heterogeneity in the variance of the response. However, it is not automatically clear that large facilities and small facilities exhibit the same relationship between Hb measurement frequency and Hb variation. Use of weights may theoretically introduce bias.

8. [DISCRETIONARY] There are multiple possible causes of endogeneity here, especially insofar as the authors choose to interpret the cross-sectional parameter. I would suggest that the final paragraph of the Results is mostly unfounded speculation (though I agree that estimation of Hb variation is likely not to create endogeneity).

Results

1. [MAJOR] Percentiles of distributions nearer the median than the 1st and 99th percentiles (e.g., the quartiles) may be more relevant, insofar as we’re concerned about practice patterns. This comment pertains to several instances throughout the Results, of course.

2. [MAJOR] The Pearson correlation (roughly 0.31) between Hb measurement frequency and ESA dose adjustment frequency is not trivial, but it’s not so strong that it automatically precludes simultaneous adjustment for both frequencies, as the authors suggest ought to be done. In the Methods, I had not questioned this decision. Upon reading Results, I now do. The authors need to make a better case for why simultaneous adjustment for both frequencies is infeasible, if the manuscript is maintained in its current form.

3. [DISCRETIONARY] Re: the description of Figure 2, the authors spend considerable time opining on explanations for the data patterns. Such text is welcome in the Discussion, but largely misplaced in the Results.

4. [MAJOR] Table 1 is nicely structured, but largely non-interpretable. I have no idea how the predictors are actually parameterized. I am a statistician, so put my complaint in its proper perspective. A physician with less quantitative training would be unlikely to come as close as I can to the proper interpretation here.

5. [DISCRETIONARY] I question the utility of presenting predicted effects in the context of Hb within 2 g/dL of the mean. Such an interval is tantamount to a target range with width of 4 g/dL. No trial has contemplated a target of such width.

6. [MAJOR] Again, the contrast of 1st and 99th percentiles is awfully wide. Also, as I wrote with respect to the Abstract, it would be ideal to harmonize the contrasts for Hb measurement frequency and ESA dose adjustment frequency.

7. [MAJOR] Finally, with respect to Figure 3B, I am not convinced that the displayed domain of the predicted line is appropriate. My doubt is predicated upon my suspicion about model parameterization. If, in fact, the longitudinal predictor is a difference between one month and mean in all months, then the displayed domain ought to include both negative and positive values (i.e., differences). This figure, more than any other, has me pondering whether my
conception of your parameterization is anywhere close to being correct. But if my guess is incorrect, then how did the authors parameterize the model? Was there a time effect in the model? (The text does not suggest that.) I am more confused than I ought to be at this point in the analysis.

8. [DISCRETIONARY] I would like to see some investigation of model fit. How well did these models actually describe the data at hand?

Discussion

1. [MAJOR] In general, the Discussion is short on references. I contend that this reflects a greater emphasis on speculation as opposed to synthesis of existing literature. In particular, I want to see references to and discussion of literature concerning the effects of protocol specification on anemia management.

2. [MAJOR] Re: “The cross-sectional and longitudinal parameter estimates were not markedly different from one another, suggesting that even a purely cross-sectional analysis for assessing these associations may be considered as being somewhat moderated from confounding influences.” Perhaps, but this is merely a hypothesis. With that said, as the authors probably know, it’s very reasonable to collapse the specified model to a purely cross-sectional one if the estimated parameters for the cross-sectional and longitudinal effects are not significantly different. The authors report no such test. Perhaps they should.

3. [DISCRETIONARY] I wonder the potential effect of cross-sectional adjustment for cumulative hospitalization days among all patients within each facility-month. It seems that the authors have the data to do this. We know that hospitalizations are associated with fluctuations in hemoglobin concentration. This may go a long way toward the predictable retort that disease severity was not fully adjusted here.

[Level of interest]: An article whose findings are important to those with closely related research interests

[Quality of written English]: Acceptable

[Statistical review]: Yes, and I have assessed the statistics in my report.

[Declaration of competing interests]:

I am employed by a non-profit organization that has received unrestricted research grants from Amgen Inc.