What Caused Excess Strokes in Patients Randomized to Darbepoetin in the Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT)?

No Smoking Gun

Wolfgang C. Winkelmayer, MD, ScD

Erythropoiesis-stimulating agents (ESAs) have been the treatment of choice for patients with renal anemia for >2 decades. Epoetin alfa was first approved in 1989 on the basis of its ability to increase hematocrit swiftly and to reduce the burden of transfusion in patients with chronic kidney disease (CKD) receiving dialysis. In addition to transfusion avoidance, recombinant human erythropoietin and its longer-acting cousins were also prescribed on the basis of the common belief that reducing anemia would improve relevant hard clinical end points, including cardiovascular morbidity and mortality. However, an earlier trial of patients receiving chronic hemodialysis who also had significant cardiovascular disease was halted owing to futility and safety considerations, and seemed to signal that a more aggressive anemia treatment target (normalization) would lead to an increased risk of mortality or nonfatal myocardial infarction (hazard ratio, 1.3; 95% confidence interval, 0.9–1.9; the Table). The findings of this Normal Hematocrit Trial were readily pushed aside owing to what was perceived as an extremely selected and nonrepresentative study population and a nonsignificant primary result, and did not appear to have a sustained effect on overall prescriber behavior.

It was not until relatively recently that large randomized trials in populations with non–dialysis-dependent CKD were conducted to test for that putative cardiovascular benefit from more aggressive anemia correction. In 2006, 2 randomized trials comparing higher and lower hemoglobin treatment targets in patients with CKD were simultaneously published (for pertinent details, see the Table). The Cardiovascular Risk Reduction by Early Anemia Treatment with Epoetin Beta (CREATE) trial found no difference between treatment groups in cardiovascular risk using a comprehensive composite end point (normalization) would lead to an increased risk of mortality or nonfatal myocardial infarction (hazard ratio, 1.3; 95% confidence interval, 0.9–1.9; the Table). The findings of this Normal Hematocrit Trial were readily pushed aside owing to what was perceived as an extremely selected and nonrepresentative study population and a nonsignificant primary result, and did not appear to have a sustained effect on overall prescriber behavior.

The Correction of Hemoglobin and Outcomes in Renal Insufficiency (CHOIR) trial had a similar design and used epoetin alfa to target 2 different hemoglobin ranges; however, patients randomized to the higher hemoglobin target range experienced a 34% increased rate of the primary composite cardiovascular end point (P=0.03). On publication of these 2 trials, the US Food and Drug Administration mandated a boxed warning on the labels of epoetin alfa and darbepoetin alfa but postponed more decisive action until the results from the then-ongoing Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT) became available. TREAT was an unusual study in that patients were randomized to either treatment with darbepoetin to a hemoglobin target of 13 g/dL or placebo injections with darbepoetin rescue below a hemoglobin concentration of 9 g/dL. The primary results of TREAT, published in 2009, found no difference between the treatment groups in their rates of the primary composite cardiovascular end point (hazard ratio, 1.05; P=0.41). However, patients randomized to the darbepoetin group experienced a doubling of the rate of stroke (hazard ratio, 1.92; 95% confidence interval, 1.38–2.68), an adjudicated and prespecified end point, which was highly significant. This new finding was swiftly added to the boxed warning in the labels of both epoetin alfa and darbepoetin alfa, and speculation ensued on what mechanism might explain this unexpected excess risk of stroke in darbepoetin-treated patients.

This issue of Circulation contains a formal evaluation of potential predictors and correlates of stroke by the TREAT investigators. Skali and collaborators investigated a number of potential predictors and correlates of stroke with special emphasis on any factors that may have altered the association between darbepoetin and stroke (modifiers; eg, history of stroke) and potential downstream consequences of treatment with darbepoetin alfa that may have increased the risk of stroke (mediators): blood pressure, hemoglobin concentration and its initial rate of rise, platelet count, and darbepoetin dose. The authors approached their task using 2 analytic strategies. First, they use the entire TREAT cohort and assessed a large number of baseline characteristics for their associations with stroke, including several well-established stroke risk factors. Not surprisingly, the doubling of stroke risk from randomization to the darbepoetin arm was confirmed in the adjusted logistic regression model (odds ratio, 2.08; 95% confidence interval, 1.47–2.94). Furthermore, baseline history of cerebrovascular disease was a strong predictor of experiencing a stroke during follow-up in TREAT; patients with a history of stroke or transient ischemic attack had twice the risk of stroke compared with patients without such a history at baseline (odds ratio, 2.00; 95% confidence interval,
they obtained a study sample that included stroke and patients who remained free from stroke during follow-up, stroke event on propensity score and follow-up time to 10 propensity score modeling. Matching each subject with a small number of outcomes, 101 in the darbepoetin arm and 52 stroke in each of these 2 groups. Naturally limited by the focused on the evaluation of postrandomization factors for patients in the darbepoetin and placebo arms of TREAT and transient ischemic attack.

1.36–2.94). Remarkably, this was independent and on top of a 60% increased risk of stroke associated with history of cardiovascular disease, which was included as a separate predictor in the model (odds ratio, 1.60; 95% confidence interval, 1.03–2.49). Although one may get the impression that the excess risk in the darbepoetin arm of TREAT was more pronounced in patients with a history of stroke or that the excess risk in the darbepoetin arm of TREAT was independent and on top of cardiovascular disease, which was included as a separate predictor in the model (odds ratio, 1.60; 95% confidence interval, 1.03–2.49). Although one may get the impression that the excess risk in the darbepoetin arm of TREAT was equally present from randomization to darbepoetin was equally present from randomization to darbepoetin and strong in patients with or without a history of stroke or transient ischemic attack.

The second analysis, Skali and colleagues separated patients in the darbepoetin and placebo arms of TREAT and focused on the evaluation of postrandomization factors for stroke in each of these 2 groups. Naturally limited by the small number of outcomes, 101 in the darbepoetin arm and 52 in the placebo arm, the authors used propensity score matching on baseline variables with stroke being the outcome of propensity score modeling. Matching each subject with a stroke event on propensity score and follow-up time to 10 patients who remained free from stroke during follow-up, they obtained a study sample that included stroke and nonstroke subjects whose average characteristics at baseline were presumably balanced (although the authors neither showed nor commented on whether successful balancing of baseline characteristics between cases and controls was achieved). They then investigated whether certain characteristics measured in the most recent 90 days before the stroke event in cases and corresponding 90-day periods in controls differed between patients experiencing a stroke and those who did not. The postrandomization variables examined were systolic and diastolic blood pressures, hemoglobin concentrations, platelet count, and darbepoetin dose (darbepoetin arm) and presence of any darbepoetin rescue (placebo arm). Each of these variables was available for only a subset of cases and controls, further reducing the already limited power. However, none of the investigated factors was statistically different between cases and controls. These findings were robust across a large number of sensitivity analyses shown in the Appendix, with significantly lower hemoglobin concentrations in patients with a stroke than in corresponding controls (odds ratio, 1.60; 95% confidence interval, 1.03–2.49). Although one may get the impression that the excess risk in the darbepoetin arm of TREAT was equally present from randomization to darbepoetin and strong in patients with or without a history of stroke or transient ischemic attack.

### Table. Key Features of 4 Landmark Studies of Erythropoiesis-Stimulating Agents in Patients With Chronic Kidney Disease

<table>
<thead>
<tr>
<th>ESA used</th>
<th>NHT</th>
<th>CREATE</th>
<th>CHOIR</th>
<th>TREAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFR/ClCr, mL·min⁻¹·1.73 m²⁻¹</td>
<td>9–11</td>
<td>15–35</td>
<td>15–50</td>
<td>20–60</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>9–11</td>
<td>11–12.5</td>
<td>&lt;11</td>
<td>≤11</td>
</tr>
<tr>
<td>Other (selection)</td>
<td>Absence of advanced cardiovascular disease, ESA naive</td>
<td>Presence of diabetes mellitus, no cardiovascular events and no ESA received in previous 12 wk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary cardiovascular end point</td>
<td>Death or nonfatal MI</td>
<td>Earliest of sudden death, nonfatal MI, acute heart failure, stroke, TIA, complication of peripheral vascular disease (amputation or necrosis), angina, or cardiac arrhythmia with hospitalization for ≥24 h</td>
<td>Earliest of death, nonfatal MI, hospitalization for congestive heart failure, stroke</td>
<td>Earliest of death, nonfatal MI, congestive heart failure, stroke, hospitalization for myocardial ischemia</td>
</tr>
<tr>
<td>Follow-up, median, mo</td>
<td>14</td>
<td>36 (Mean)</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td>Hemoglobin target, g/dL</td>
<td>9–11</td>
<td>13–15</td>
<td>10.5–11.5</td>
<td>13–15</td>
</tr>
<tr>
<td>Enrolled, n</td>
<td>615</td>
<td>618</td>
<td>302</td>
<td>717</td>
</tr>
<tr>
<td>Primary end point, n</td>
<td>164</td>
<td>202</td>
<td>47</td>
<td>97</td>
</tr>
<tr>
<td>Stroke (TIA), n</td>
<td>Referent</td>
<td>0.78 (0.20)</td>
<td>Referent</td>
<td>1.34 (0.03)</td>
</tr>
<tr>
<td>Hazard ratio (P)</td>
<td>Referent</td>
<td>0.93–1.14</td>
<td>Referent</td>
<td>1.05 (0.41)</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>0.9–1.9</td>
<td>0.53–1.14</td>
<td>1.03–1.74</td>
<td>0.94–1.17</td>
</tr>
<tr>
<td>Stroke (TIA), n</td>
<td>9</td>
<td>14</td>
<td>5 (1)</td>
<td>6 (5)</td>
</tr>
<tr>
<td>Hazard ratio (P)</td>
<td>NR</td>
<td>Referent</td>
<td>1.01 (0.98)</td>
<td>Referent</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>0.45–2.25</td>
<td>1.38–2.68</td>
<td>Referent</td>
<td>(0.98)</td>
</tr>
</tbody>
</table>

NHT indicates Normal Hematocrit Trial; CREATE, Cardiovascular Risk Reduction by Early Anemia Treatment With Epoetin Beta; CHOIR, Correction of Hemoglobin and Outcomes in Renal Insufficiency; TREAT, Trial to Reduce Cardiovascular Events With Aranesp Therapy; eGFR, estimated glomerular filtration rate; ClCr, creatinine clearance; ESA, erythropoiesis-stimulating agent; MI, myocardial infarction; NR, not reported; and TIA, transient ischemic attack.
plementation and iron parameters (ferritin and transferring saturation), which were also not found to differ between cases and controls (although these findings were not presented in great detail). Thus, the authors appropriately concluded that even after close scrutiny of the rich data available in TREAT, no plausible explanation for the excess stroke risk in patients randomized to TREAT could be detected.

This study has certain weaknesses, which include the relatively small sample size for such an epidemiological exercise, few outcomes observed, and use of a trial cohort that may not be fully representative of average patients seen in typical care settings. In contrast, there are important strengths of this analysis, including prospective collection of pertinent information and adjudicated end points using a relatively stringent positivity criterion for stroke of >24 hours of neurological deficit. It would be useful, however, to see a similar analysis in a much larger cohort of ESA-treated patients that would enable more precise estimation of the studied associations and other putative associations that were not examined in this study.

As clinicians and scientists, we still have to wonder what other factors may explain the excess risk of stroke in the darbepoetin arm of TREAT. Initially, we may want to look at the 3 larger studies that had previously randomized patients with stage 3 or 4 CKD to different anemia treatment targets and thus to protocols that differed inherently in their dose of ESA used in each treatment arm (the Table). Stroke was not specifically defined as an end point in the Normal Hematocrit Trial of 1233 patients on chronic dialysis, but it was reported as a cause of death (unclear whether cause of death was adjudicated). Among 195 patients who were randomized to the normal-hematocrit group and died, cerebrovascular accident was listed as the cause of death in 14 (7%), whereas 9 of 160 deaths in the low-hematocrit group (6%) were deemed to have resulted from such a cause (overall 2% of all patients randomized). Nonfatal stroke was not reported. Stroke was a component of the prespecified primary end point of the CREATE trial and occurred in 6 (2%) and 5 (2%) patients in the higher and lower hemoglobin target arms, respectively. Stroke was also a component of the primary end point in the CHOIR trial, in which 12 strokes (2%) occurred in both study arms. Examining these 4 trials together begs the following questions: Why was there such a high number of strokes in TREAT compared with each of the other trials? And why was there such a strong safety signal for stroke in TREAT but not in any of the other trials?

Most important, TREAT was a much larger study than the other trials and enrolled more patients (n=4038) than the other 3 trials combined (total n=3268). Furthermore, follow-up was longer in TREAT (median, 29 months) compared with CHOIR (median, 16 months), so more subjects were at risk and followed up longer in TREAT, which renders stroke event rates per unit of time quite similar between these 2 studies. The Normal Hematocrit Trial had 14 months of follow-up but did not prospectively ascertain stroke. CREATE reported a mean follow-up duration of 3 years, longer than TREAT, but its composite primary end point was composed of 8 individual events, which makes it more likely that patients experience a nonstroke event and would be censored and that a subsequent stroke would not be counted. Given the overall low event rate in CREATE, this is probably a minor problem, and it appears that patients in CREATE were generally at lower cardiovascular risk than patients in CHOIR or TREAT. Thus, it is clear that only TREAT was in a position to be adequately powered to identify a (still relatively strong) signal for excess stroke risk.

Another consideration is the specific ESA used in each of these 3 trials: CREATE used epoetin beta (not marketed in the United States); CHOIR used epoetin alfa; and TREAT used darbepoetin alfa. Although it is generally assumed that these chemically distinct ESAs have similar safety profiles (at least when standardized for erythropoiesis-stimulating activity), only a few head-to-head comparisons have been conducted. These trials focused on comparing hemoglobin concentrations as the outcome and therefore were small and had relatively short follow-ups, and some were never published in a peer-reviewed journal (eg, a randomized trial comparing darbepoetin alfa with epoetin alfa in 407 blacks; NCT00111995). Thus, one may conclude that the comparative safety among ESAs is not sufficiently established and that the excess stroke risk in the darbepoetin alfa arm of TREAT could, although a remote possibility, result from an unreported association specific to this drug that may not be present in other ESAs.

Finally, one may want to consider whether the stroke finding in TREAT could have been a chance finding. After all, several significance tests were conducted on secondary and tertiary analyses in that data set, which certainly increases the risk of a false-positive finding or type 1 error. Although the possibility of a chance finding exists, it is rather remote. The hazard ratio of 1.92 and lower bound of the 95% confidence interval at 1.38 permit back-calculation of an approximate corresponding value of P=0.00012, so clearly this finding cannot be plausibly argued away as bad luck.

For the clinician, the findings from TREAT and the other trials clearly mandate a more conservative approach to the treatment of anemia in patients with CKD and careful consideration of each patient’s specific circumstances. The Food and Drug Administration publicly consulted a Cardiovascular and Renal Drugs Advisory Committee on October 18, 2010; at that meeting, the analyses published today, among other evidence and opinion, were first presented. The agency took formal action on June 24, 2011, by rather drastically changing the label for the ESAs that are available for the treatment of anemia in patients with CKD in the United States. The revised label no longer contains a hemoglobin target range for ESA treatment (previously 10–12 g/dL) and now recommends considering initiation of treatment if hemoglobin is <10 g/dL, with focus on transfusion avoidance, corresponding to the originally demonstrated benefit for approval in 1989. The extent of the label change is controversial, too conservative for some and too liberal for others, reflecting that the evidence base for any appropriate or even optimal use of ESAs remains most incomplete. For the specific question about what may be responsible for the excess stroke risk in TREAT, however, all we can say is that there is no smoking gun for now.
Disclosures
Dr Winkelmayer reports having served as a scientific advisor to Amgen Inc.

References

KEY WORDS: Editorials / anemia / kidney failure, chronic / erythropoietin / stroke
What Caused Excess Strokes in Patients Randomized to Darbepoetin in the Trial to Reduce Cardiovascular Events With Aranesp Therapy (TREAT)?: No Smoking Gun
Wolfgang C. Winkelmayer

Circulation. 2011;124:2805-2808
doi: 10.1161/CIRCULATIONAHA.111.071985

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/124/25/2805

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/