

*Editorial Comment*

## Understanding recent haemoglobin trials in CKD: methods and lesson learned from CREATE and CHOIR

Adeera Levin

Nephrology, Education and Research, Vancouver BC V6Z 1Y6, Canada

Two randomized control trials (RCTs), CREATE and CHOIR, recently published in the *New England Journal of Medicine*, have received intense attention and discussion by the nephrology community [1,2]. Anaemia and its treatment have been the focus of multiple randomized control studies in the recent past, most small and relatively short, so why this amount of discussion now [3–7]?

These two new studies by Drueke *et al.* [1] and Singh *et al.* [2] represent the largest studies in non-dialysis patients published to date; they both tested the hypothesis that higher haemoglobin would lead to improved patient outcomes, and both were negative.

The purpose of this editorial is to critically review the two studies, with special attention to the methodology and analysis of both, and by so doing to understand what we can and cannot conclude from them. Importantly, the readership should be informed as to the specific issues related to analysis of randomized control trials, issues of dropout, bias and the impact of censoring, so that interpretation of these trials can be appropriately applied to clinical care. Of note, both are open label, randomized control trials, in non-dialysed CKD patients with primary endpoints that were composite CVD event endpoints. They are similar in many respects, but differ in some: mostly in the authors' conclusions and applicability of the results.

The CREATE study was conducted in Europe, Asia and Mexico, enrolled 600+ patients, and tested the specific hypothesis that treatment of Hgb values to prevent anaemia (defined as <11 g/dl) would lead to improved CVD outcomes. This is a well-conducted RCT, using Erythropoietin (EPO) beta, in patients with stage 4 CKD. Those assigned to the high and low Hgb groups were not statistically significantly different on any parameters at baseline. The separation between Hgb values in the two groups was achieved, 13.5 vs 11.5 g/dl median Hgb in high vs low groups,

respectively. The majority of patients completed the study, and there were few protocol violations. There were no differences in primary outcomes, nor was there any difference in left ventricular mass index (LVMI) change over time. The mean GFR of 25 ml/min was the same in both groups, as was the mean decline in GFR (3.4 ml/min vs 3.1 ml/min) over the 3-year period.

The authors do point out that the number of patients who went onto dialysis was greater in the higher Hgb group. Because of the non-protocolized nature of dialysis start, and the fact that there were no differences in rate of decline of GFR, or actual GFR at the time of dialysis, this finding should not be over-interpreted. It is possible that the physicians caring for the patients who knew their Hgb values, would decide to commence dialysis when symptoms worsened, in the absence of other explanations for the symptoms, while those in the lower Hgb group, who complained of symptoms, might have delayed dialysis because the symptoms were attributed to anaemia rather than uraemia.

The dose of EPO beta in CREATE averaged 5000 IU/per patient; in keeping with all other studies reported previously.

The key 'issue' with CREATE was that the annual event rate was much lower than expected (6% vs 15%). The sample size calculation was based on observational studies reported from a Canadian cohort assembled almost a decade ago, in 1993, which described left ventricular hypertrophy (LVH) and LVMI differences in those with and without anaemia, and CVD events as well [8]. At that time, the use of ACEi and ARB was not common in the population, target blood pressures were not as low as current levels and treatment of phosphate, PTH and other parameters were not commonplace. The study was a non-interventional, observational study in referred CKD patients across Canada. In CREATE, patients enrolled may be either different than those who were in an original observational cohort, or are treated differently in the course of the study. As the effect of interest was Hgb, all other parameters (such as BP, mineral metabolism abnormalities, etc.) were managed in

Correspondence and offprint requests to: Dr A. Levin, Nephrology, Education and Research, 1081 Burrad Street, Room 6010, Vancouver, BC V6Z 1Y6, Canada.

accordance with best practices, which were different from those in the early 1990s. Enrolment in a RCT is well documented to lead to a Hawthorne effect, whereby the control group performs better than expected. Both of these factors may account for the low event rate.

Irrespective, the CREATE study was ultimately underpowered to demonstrate a difference between the two groups, but did not describe harm to patients who were randomized to the higher Hgb target arm.

The quality of life (QOL) results mirrored that of dialysis and non-dialysis studies reported initially and more recently [5,7,8]. There is a definite improvement in QOL in those randomized to the high haemoglobin group.

CHOIR was conducted in the USA, and enrolled 1400+ patients in 130 centres. Original plans were for 2000 patients ([www.clintrials.org](http://www.clintrials.org)). While the outcome measures and basic study design were similar for CHOIR and CREATE, there are some major differences between the two studies. The patients in CHOIR were older, had more diabetes, more hypertension as a cause of CKD, and had higher pulse pressures than those in CREATE. Most striking is the fact that the doses of erythropoietin therapy were substantially higher than those in CREATE or any other study to date: median doses of 11000 units (almost 3× higher than any other CKD study reported). Within CHOIR, the two groups were different at baseline on two major CVD characteristics (coronary artery bypass grafts (CABG) (high Hgb group 17% vs low group 14%,  $P=0.03$  and hypertension (HTN) high Hgb group 95% vs low Hgb group 92%  $P<0.02$ ). In a randomized control trial, while it is possible that by chance alone some baseline characteristics will be statistically different, it is of interest that in this trial, the two factors on which the treatment group randomly differed were cardiovascular factors. Both history of CABG or HTN may be biologically linked to, and thus have an impact on, the primary outcome measures. Given that the composite endpoint was based on CVD events, this imbalance at baseline may explain some of the differences between the groups in primary composite outcomes. Of additional note, despite extraordinarily high doses of erythropoiesis stimulating agents (ESA), the patients in the treatment group were only able to achieve an Hgb of 12.5 g/dl.

A number of other observations about the CHOIR study: of the original 1400 cohort, over half were lost in each group, thus the final numbers are closer to those in CREATE. Reasons for loss included dialysis start, which occurred in about half. While this was predefined by the protocol, given the objective of the study to examine CVD outcomes in CKD patients, one could argue that this was not necessarily a sound decision, since the intention of optimal pre-dialysis care should be to improve patient outcomes on dialysis. The remainder of the patients was withdrawn for various reasons that are not all reported, despite the need to do so for publication of RCTs. The Table 1 demographic data in CHOIR describe the randomized

cohort of 1400: there is no information regarding the 700 patients who completed the trial and were included in the analysis. It would be of value for the readership to know the characteristics of the cohort who formed the basis of the analysis. Most importantly, when a randomized control study loses half of the randomized cohort, it is circumspect to draw any conclusions from the analyses performed on those remaining in the trial.

Furthermore, the protocol was changed after approximately 400 patients were randomized. (Note that this represents one-third of the original cohort, but we are unaware how many of those exposed to the original protocol were in the final analysis.) The protocol change was stated to be to target an absolute level of Hgb instead of a range. Reasons for the change in protocol are not given.

The CHOIR study was stopped after an interim analysis for presumed futility (inability to demonstrate a benefit in the higher haemoglobin arm), but as stated in the paper, this was not done according to any usual stopping rules, or evidence of harm, but rather the suggestion of harm. This lack of adherence to RCT convention should raise additional questions in the reader's mind. While the figures on page 2093 of the paper describe a significant combined endpoint, in the higher target group, the reader should be aware of the issues stated earlier, and the fact that for the specific endpoints, the curves appear to separate after 15–18 month, with very small numbers at that time. Randomization is intended to account for baseline differences. However, given the imbalance in the original cohort regarding CABG and HTN prevalence, it may have been interesting to adjust for those characteristics, both of which may well have impacted on congestive heart failure (CHF) and hospitalizations. Furthermore, given the difficulty with the cohort dropout, recent trends in reporting RCTs would suggest that including all patients randomized, with their last event and clinical state prior to (e.g. dialysis start) might give a more comprehensive view of the trial outcomes.

In the discussion, the authors appear to misinterpret other authors. With all of the limitations mentioned previously, which are not mentioned in their discussion, any set of conclusions is problematic. The authors describe KDQOI recommendations, but misinterpret some of the statements. The guidelines state that there is no evidence to support targeting values of Hgb > 13. There is no statement articulating any new target range, as the authors suggest. In fact, the recently published guidelines are concordant with the findings of the studies.

To reframe the CHOIR study: it is a RCT of 700 patients, who did not achieve target Hgb values of 13.5 g/dl (median was 12.8 g/dl) despite receiving 11000 units of EPO. Unlike any other study to date, there were no differences in QOL in the high Hgb group; it is not clear when the QOL was measured. Few differences were seen, though the composite endpoint was significantly different and the high Hgb target groups non-significantly trended towards more

CHF and hospitalizations. Note that rigorous methodologists would suggest that *P* values of 0.03 are not significant when more than one analysis is planned. It would seem that patients who require high doses of ESA to achieve subtargeted Hgb values and who commence with a significant amount of CVD co-morbidity, may not be necessarily helped by ESA therapy targeting normal Hgb. This statement is supported by the trial, but is not stated quite as such by the authors.

CREATE describes lack of benefit of targeting Hgb to >13 in CKD patients in a well executed though underpowered study, conducted according to current best practices in CKD care and anaemia management.

Other differences between the studies include the use of iron (almost all patients in CREATE, less than 50% in CHOIR); a different withdrawal rate (21% in CREATE vs 38% in CHOIR), and the ability to reach target values in CREATE, but not in CHOIR. Is it possible that the patients in CHOIR are not representative of the majority of CKD patients, but rather represent the most unwell of all CKD patients?

In aggregate, the two studies represent the largest studies to date in CKD patients not receiving dialysis. In addition, three smaller studies have recently been published which are similar to the CREATE study: the Australian Study [3], the Canadian study [4] and the UK study [10]. Each of these used LVMI as their surrogate endpoint, had smaller study samples, but all used doses of ESA similar to the CREATE study. Like CREATE, none of the studies demonstrated a difference in LVMI over the study period.

There are a number of lessons learned from the recent publications, most importantly the issues of internal and external validity. The design, execution and analysis of RCTs should follow conventional study rules in terms of design, cessation and analysis; where possible, they should minimize or anticipate the possibility of Hawthorne effects and err on the conservative side for sample size estimates. In addition, they should try to mimic real world practice, in matters such as selection of the cohort to be representative and dosing of medication. While the findings of CHOIR are indeed 'true', they may not be applicable to the population to whom the authors believe they should be applied; i.e. all patients with CKD.

Anaemia has been associated with adverse outcomes in CKD populations. The ability to modify this parameter with the use of ESA has led to multiple efforts to describe the benefits of higher haemoglobin (i.e. normalization of Hgb). To date, and in keeping with the results of these recent publications, there does not appear to be support for targeting normal Hgb values in CKD patients. As in all biological systems, there may be physiological reasons for this. In as much as CKD is accompanied by changes in vascular compliance, increases in oxidative stress, and alterations of vessel wall integrity, it may be that Hgb values below normal are associated with reductions in shear stress and endothelial cell integrity, while maintaining appropriate tissue oxygenation. Thus, despite

overwhelming observational data describing the association of higher Hgb with better outcomes, there are no randomized control trial data which confirm this need to change any current strategy to target Hgb values to be normal or >13 g/dl for CKD patients.

Both CREATE and CHOIR represent the recent commitment on the part of the nephrology community and pharmaceutical industry to embark on large-scale RCTs which have heretofore been lacking. The current TREAT trial is ongoing: it attempts to answer a related question in diabetics, but is designed as a double-blind, placebo-controlled trial, and asks whether treatment vs no treatment by ESA (in this case darbepoietin) will improve outcomes. Of importance to the readership is that TREAT is continuing after the publication of both CREATE and CHOIR and a review of the current trial data of TREAT, by the data safety monitoring board (DSMB) of that study (personal communication, letters sent to investigators). This fact should also give the readership confidence that the purported potential for harm as described by the CHOIR authors, has not been seen in a to-date related study. The continuation of TREAT indicates that the scientific question remains important, and is still not sufficiently answered.

### **What do we actually know to date from CREATE and CHOIR?**

It is clear that the treatment of Hgb to relatively normal targets in CKD patients can be obtained with different doses of ESA, that LVMI is not impacted by those changes in patients who are enrolled in a clinical trial. The trend towards increase event rates in CHOIR may have a number of different explanations: high doses of ESA are known to have vascular effects. It may, however, also be that the dosing regimen itself rather than the Hgb values per se drove the results. It may be that there really is harm to targeting higher Hgb in patients with significant co-morbidity, especially those who have had ischaemic cardiac events. However, more critically, it may be that CHOIR cannot and should not be over-interpreted, given the multiple problems with the dropout and analytic issues raised earlier.

In the current era, anaemia in CKD and its treatment remain the focus of multiple questions for clinicians and investigators. Studies examining the association of Hgb and outcomes should attempt to answer slightly different questions: what are the predictors of low haemoglobin, and does the attenuation of those factors improve Hgb levels, outcomes or both? Additional questions that need to be considered include those that ask whether one level of Hgb is appropriate in all patients, recognizing the heterogeneity of CKD patients, either on or off dialysis.

Only through well-done large RCTs in this population will we be in a position to make or change current evidence-based recommendations. CREATE

and CHOIR have helped to focus attention on the need for RCTs, but have not yet sufficiently answered the key question as to what and if there is an optimal Hgb for patients with CKD.

## References

1. Drueke T, Locatelli F, Clyne N *et al.* Normalization of hemoglobin level in patients with chronic kidney disease and anemia. *New Engl J Med* 2006; 355: 2071–2084
2. Singh A, Saczecz L, Tang KL *et al.* Correction of anemia with epoetin alfa in chronic kidney disease. *New Eng J Med* 2006; 355: 2085–2098
3. Roger SD, McMahon LP, Clarkson A *et al.* Effects of early and late intervention with epoetin alpha on left ventricular mass among patients with chronic kidney disease (stage 3 or 4): results of a randomized clinical trial. *J Am Soc Nephrol* 2004; 15: 148–156
4. Levin A, Djurdjev O, Thompson C *et al.* Canadian randomized trial of hemoglobin maintenance to prevent or delay left ventricular mass growth in patients with CKD. *Am J Kidney Dis* 2005; 46: 799–811
5. Parfrey PS, Foley RN, Wittreich BH, Sullivan DJ, Zagari MJ, Frei D. Double-blind comparison of full and partial anemia correction in incident hemodialysis patients without symptomatic heart disease. *J Am Soc Nephrol* 2005; 16: 2180–2189
6. Gouva C, Nikolopoulos P, Ioannidis JPA *et al.* Treating anaemia early in renal failure patients slows the decline of renal function; A randomised controlled trial. *Kidney Int* 2004; 66: 753–760
7. Foley RN, Parfrey PS, Morgan J *et al.* Effect of hemoglobin levels in hemodialysis patients with asymptomatic cardiomyopathy. *Kidney Int* 2000; 58: 1325–1335
8. Levin A, Thompson CR, Ethier J *et al.* Left ventricular mass index increase in early renal disease: impact of decline in hemoglobin. *Am J Kidney Dis* 1999; 34: 125–134
9. McMahon LP, Mason K, Skinner SL *et al.* Effects of haemoglobin normalization on quality of life and cardiovascular parameters in end-stage renal failure. *Nephrol Dial Transplant* 2000; 15: 1425–1430
10. MacDougall I, Temple RM, Kwan J. Is early treatment of anaemia with epoetin alfa beneficial to predialysis chronic kidney disease patients? Results of a multicentre, open-label, prospective, randomised, comparative group trial. *Nephrol Dial Transplant*, Advance Access published on 12, September 2006.

Received for publication: 19.12.06

Accepted in revised form: 19.12.06

Nephrol Dial Transplant (2007) 22: 312–315

doi:10.1093/ndt/gfl604

Advance Access publication 27 October 2006

## Periodontal diseases—a modifiable source of systemic inflammation for the end-stage renal disease patient on haemodialysis therapy?

Ronald G. Craig<sup>1,2</sup>, Peter Kotanko<sup>3</sup>, Angela R. Kamer<sup>2</sup> and Nathan W. Levin<sup>3</sup>

<sup>1</sup>Department of Basic Sciences and Craniofacial Biology and <sup>2</sup>Department of Periodontology and Implant Dentistry, New York University College of Dentistry and <sup>3</sup>Renal Research Institute LLC, New York, USA

**Keywords:** periodontal disease; inflammation; haemodialysis

### Periodontal diseases—biology, pathology and clinical presentation

Periodontal diseases are a group of inflammatory diseases that affect the supporting tissues of the dentition. The most prevalent periodontal diseases result from the interaction of specific bacterial species with components of the host immune response in disease susceptible individuals and are currently

classified as plaque-induced gingival diseases, early onset, chronic adult and aggressive periodontitis [1]. Plaque-induced gingival diseases are limited to the gingivae (gingivitis) and are characterized by erythema, oedema, haemorrhage and enlargement of the gingival tissues. Plaque-induced gingivitis is nearly pandemic in children and young adults and is reversible with treatment. In contrast, early onset, chronic and aggressive periodontitis are irreversible forms of periodontal disease that culminate in tooth loss if left untreated. Estimates of the prevalence of periodontitis vary with the clinical criteria used to define disease status; however, the Third National Health and Nutrition Survey (NHANES III) reported a 14% prevalence of moderate to severe periodontitis in the United States population >20 years of age [2].

The inflammatory lesion in periodontitis extends from the gingiva to include deeper connective tissues

Correspondence and offprint requests to: Nathan W. Levin, MD, Renal Research Institute LLC, 207 East 94th Street, Suite 303, New York, New York 10128, USA. Email: nlevin@rriny.com