

# Lessons Learned from EVOLVE for Planning of Future Randomized Trials in Patients on Dialysis

Patrick S. Parfrey, Geoffrey A. Block, Ricardo Correa-Rotter, Tilman B. Drüeke, Jürgen Floege, Charles A. Herzog, Gerard M. London, Kenneth W. Mahaffey, Sharon M. Moe, David C. Wheeler, and Glenn M. Chertow

## Abstract

The effect of the calcimimetic cinacalcet on cardiovascular disease in patients undergoing hemodialysis with secondary hyperparathyroidism was assessed in the Evaluation of Cinacalcet Hydrochloride Therapy to Lower Cardiovascular Events trial. This was the largest (in size) and longest (in duration) randomized controlled clinical trial undertaken in this population. During planning, execution, analysis, and reporting of the trial, many lessons were learned, including those related to the use of a composite cardiovascular primary endpoint, definition of endpoints (particularly heart failure and severe unremitting hyperparathyroidism), importance of age for optimal stratification at randomization, use of unadjusted and adjusted intention-to-treat analysis for the primary outcome, how to respond to a lower-than-predicted event rate during the trial, development of a prespecified analytic plan that accounted for nonadherence and for cointerventions that diminished the power of the trial to observe a treatment effect, determination of the credibility of a subgroup effect, use of adverse effects database to investigate rare diseases, collection of blood for biomarker measurement not designated before trial initiation, and interpretation of the benefits-to-harms ratio for individual patients. It is likely that many of these issues will arise in the planning of future trials in CKD.

*Clin J Am Soc Nephrol* 11: 539–546, 2016. doi: 10.2215/CJN.06370615

Due to the number of contributing authors, the affiliations are provided in the Supplemental Material.

**Correspondence:**  
Dr. Patrick Parfrey,  
Department of  
Medicine, Memorial  
University  
Newfoundland, Room  
3M104, Health  
Sciences Centre, 300  
Prince Philip Drive, St.  
John's, NL A1B 3V6,  
Canada. Email:  
pparfrey@mun.ca

## Introduction

The Evaluation of Cinacalcet Hydrochloride Therapy to Lower Cardiovascular Events (EVOLVE) trial was the largest ( $n=3883$ ) and longest (median follow-up, 50.5 months) randomized controlled trial (RCT) undertaken in patients receiving hemodialysis (1). The hypothesis tested was that the calcimimetic cinacalcet would, compared with placebo, reduce the risk of death or nonfatal cardiovascular (CV) events among patients with secondary hyperparathyroidism (sHPT; intact parathyroid hormone [iPTH]  $\geq 300$  pg/ml) who were undergoing hemodialysis. In the primary unadjusted intention-to-treat (ITT) analysis, cinacalcet did not significantly reduce the risk of the composite CV primary outcome, but there were many design and analysis issues that should be considered when the results are interpreted. In this review, we discuss study challenges faced, decisions made, and lessons learned from EVOLVE and how they may be relevant in the design of future trials in dialysis and CKD.

## Choice, Definition, and Analysis of Primary Events

### Composite Primary Outcome

The primary endpoint was a composite composed of all-cause mortality or nonfatal CV events (myocardial infarction, hospitalization for unstable angina, heart failure, or peripheral vascular disease). These CV events may occur *via* different pathophysiologic pathways: myocardial infarction and unstable angina

are principally caused by atherosclerotic vascular disease, while heart failure results from cardiomyopathy (with systolic or diastolic dysfunction), for which arteriosclerosis and uremia are predisposing factors (2). Peripheral vascular disease is often atherosclerotic, but distal occlusive disease, particularly of the lower extremities, is often accompanied by dense calcific arteriosclerosis. The CV events were chosen because we hypothesized that cinacalcet would (1) reduce heart failure and death by reducing medial calcification of conduit arteries (thus improving vascular compliance and decreasing left ventricular hypertrophy) and by decreasing the potential cardiotoxic effect of parathyroid hormone (PTH), and (2) reduce atherosclerotic events by decreasing intimal calcification of atherosclerotic stenoses (3). If the endpoint had been mortality alone, a much larger sample size would have been necessary.

Inclusion of overall mortality in the composite outcome is an issue because CV mortality would likely be modified by cinacalcet, in contrast to non-CV mortality. We included overall mortality in the primary outcome because we thought it necessary to demonstrate that the intervention did not cause unanticipated serious adverse effects on non-CV outcomes (such as cancer or death after fracture).

*Post hoc* evaluation of the treatment effects on CV events using multivariable-adjusted ITT analysis showed relative hazards of 0.88 (95% confidence interval [95% CI], 0.76 to 1.01) for time to first atherosclerotic event, 0.79 (95% CI, 0.66 to 0.96) for heart

failure, and 0.79 (95% CI, 0.64 to 0.98) for sudden death; the latter two events are considered outcomes of nonatherosclerotic disease (4). The 95% CIs for the atherosclerotic and nonatherosclerotic outcomes overlap, but the magnitude of the relative hazard for heart failure and sudden death is consistent with the hypothesis that cinacalcet may act through the nonatherosclerotic pathway. Because medial calcification and intimal calcification may well be different entities in CKD (5), this suggests that the dominant effect of cinacalcet may be through the inhibition of medial calcification.

### Definition of Outcomes

Clinically relevant, precisely defined, reproducible end points with central adjudication are necessary in RCTs that aim to change clinical practice (6). Death and major atherosclerotic CV events fulfill these criteria, but the definition of heart failure in CKD has been problematic because of the difficulty in differentiating salt and water overload from impaired left ventricular function. This issue was recently crystalized by the proposal for a functional classification system of heart failure in patients with ESRD (7). This system may be limited by the cost of echocardiography and lack of specificity of dyspnea relief by dialysis for diagnosis of heart failure.

In patients receiving dialysis, heart failure has been defined as dyspnea with at least two of the following four manifestations: bilateral basilar rales on physical examination, raised jugular venous pressure, interstitial edema on chest radiography, and increased upper pulmonary vessel diameter on chest radiography (8). The presence or history of heart failure was associated with an almost two-fold increase in mortality (8). In *EVOLVE*, we used a definition similar to that used by Harnett *et al.* (8) but added a clause to improve specificity: the patient also was required to have received mechanical ultrafiltration or hemodialysis. Nonetheless, the event adjudication committee confirmed the diagnosis of heart failure in only about half of heart failure events submitted by the local investigators, a result lower than that observed in high-risk patients with hypertension (9). This lack of certainty in the diagnosis of heart failure may limit study power because the event rate may be lower than anticipated, and misclassification may limit the capacity to identify a treatment effect. A similar misreporting of myocardial end points occurred in the Second Platelet IIb/IIIa Antagonist for the Reduction of Acute Coronary Syndrome Events in a Global Organization Network Trial and Platelet Glycoprotein IIb/IIIa in Unstable Angina: Receptor Suppression Using Integrilin (Eptifibatide) Therapy Study, in which investigators misreported myocardial infarction end points, but most investigators later agreed with assessments of these events by the clinical events committee (10,11).

### End Points with Selection Bias

Certain end points for mineral and bone disease in CKD are also problematic. For example, the end point for severe hyperparathyroidism (HPT) could be parathyroidectomy (PTX), but severe HPT may or may not be treated with PTX. PTX was a secondary end point in *EVOLVE*; however, before the start of the study it was evident that the criteria by which physicians selected PTX for the treatment of sHPT varied widely across the world, and the trial protocol

provided no criteria for PTX. Factors associated with PTX during the trial indicated selection effects and included younger age, female sex, higher body mass index, markers of comorbidity (no history of valvular disease, angina, or peripheral vascular disease), higher serum calcium, and higher PTH; the use of PTX also varied widely by country of origin (lowest rates were in the United States) (12). In addition, PTX was performed at an advanced phase of sHPT: before PTX the mean PTH level was 1872 pg/ml and the mean serum calcium level was 10.3 mg/dl (2.58 mmol/L) (12).

Because there were likely to be additional patients not selected for PTX who might have undergone PTX but did not because of various reasons unrelated to disease severity (*e.g.*, surgeon or patient reluctance or no trained surgeon available), we defined an outcome before the trial was completed. We termed this outcome severe unremitting HPT, which included PTX or the presence of severe HPT with hypercalcemia (iPTH > 1000 pg/ml with serum calcium > 10.5 mg/dl [2.61 mmol/L] on two consecutive occasions or iPTH > 1000 pg/ml with serum calcium > 10.5 mg/dl [2.61 mmol/d] on one occasion and subsequent use of commercial [off-protocol] cinacalcet within 2 months of laboratory assessment). Although hypercalcemia is a biochemical endpoint, its occurrence implies an advanced phase of HPT, and clinical practice supports PTX in this scenario. In the placebo group, severe unremitting HPT occurred in 24.3% ( $n=470$ ) of patients, of whom 59% ( $n=278$ ) had PTX. The relative hazard (cinacalcet versus placebo) for severe unremitting HPT was 0.43 (95% CI, 0.37 to 0.50). This relative hazard was similar whether baseline iPTH was mildly or markedly elevated, but the number of events prevented increased as iPTH increased (12).

Lessons from *EVOLVE* concerning choice and definitions of end points are presented in Table 1.

### Imbalance in Baseline Clinical Characteristics

Despite enrolling 3883 patients, there was a 0.8-year difference in mean age at baseline and a 1-year difference in median age (55 versus 54 years), an occurrence that confounded the primary outcome (1). This chance imbalance in a major prognostic factor for CV events necessitates covariate adjustment (13). In the prespecified primary analysis using an unadjusted ITT analysis, the relative hazard was 0.93 (95% CI, 0.85 to 1.02) and the prespecified multivariable-adjusted ITT relative hazard was 0.88 (95% CI, 0.80 to 0.98). The observed age imbalance probably occurred by chance because the likelihood of imbalance is dictated by the population SD for age and sample size. The probability of an age difference of > 0.8 years occurring in *EVOLVE* was 0.08 as the SD was 14 and sample size was 3883. Because of more restrictive inclusion criteria and the older patients enrolled, the SD for age was narrower in the Study of Heart and Renal Protection (12 years) (14) and the Trial to Reduce Cardiovascular Events with Aranesp Therapy (10 years) (15), wherein the probability of observing an age difference (by chance) of > 0.8 years between groups was 0.04 and 0.01, respectively.

The Consolidated Standards of Reporting Trials recommends reporting both unadjusted and adjusted analyses and stating whether the unadjusted analysis was planned (16). Although handling of baseline covariates in RCTs

**Table 1. Lessons concerning choice and definition of composite end points**

- (1) Primary composite endpoints should include events resulting from pathophysiologic pathways highly likely to be influenced by the intervention.
- (2A) Heart failure in patients receiving dialysis occurs frequently but is often misdiagnosed by nephrologists even when the definition includes manifestations of pulmonary edema. Sample size should be increased to account for misclassification.
- (2B) Randomized clinical trials in CKD or ESRD that include an end point of heart failure should include focused education to ensure accurate identification of heart failure events.
- (2C) A classification system of heart failure based on patient-reported dyspnea assessed pre- and post-ultrafiltration in conjunction with echocardiography may be useful (6) but requires validation.
- (3) Selection bias associated with parathyroidectomy necessitated a more holistic approach to the definition of severe unremitting hyperparathyroidism.

varies substantially (17), covariate adjustment improves the accuracy of treatment effect estimation and statistical power and hence should be performed when strong prognostic factors are observed or anticipated (13,18).

Lessons concerning imbalance in baseline covariates are presented in Table 2.

### Threats to Statistical Power

The major causes of reduction in statistical power in EVOLVE were a lower-than-predicted primary composite event rate, discontinuation of study drug (particularly due to adverse effects in the cinacalcet group and failure to control sHPT in the placebo group), and PTH-lowering cointerventions (*i.e.*, PTX, kidney transplantation, and use of commercial cinacalcet) (Figure 1), with PTX and commercial cinacalcet disproportionately used in the placebo group.

### Event Rate

We calculated the sample size on the basis of the following assumptions: an annual rate of the primary composite endpoint of 23.2% in the placebo group, a 20% treatment effect, a 1.5-year enrollment period, a 4-year total study duration, an annual rate of loss to follow-up of 1%, an annual rate of dropout (withdrawal from active treatment before a primary event) of 10% in the cinacalcet group, and a rate of drop-in (use of commercially available cinacalcet before a primary event) of 10% in the placebo group. On the basis of a two-sided log-rank test for equality of survival functions, accounting for planned interim analyses with an

overall  $\alpha$  level of 0.05, we determined that a primary event would need to occur in 1882 patients in order to ensure a power of approximately 90% (1).

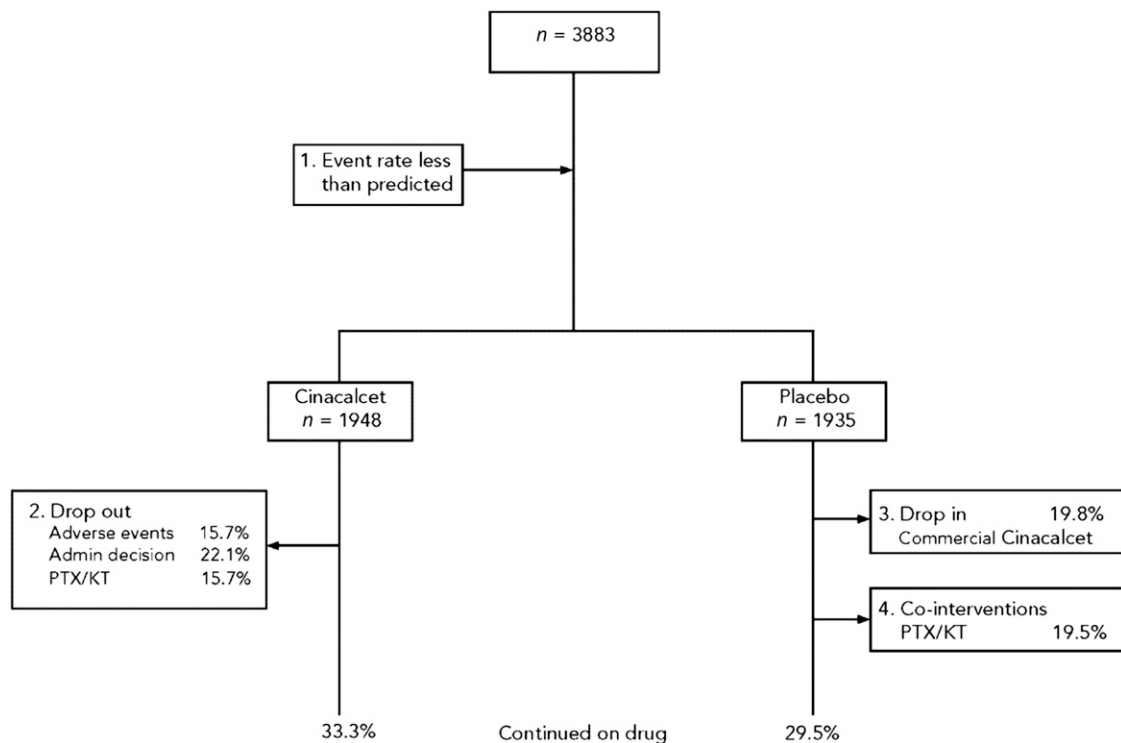
We estimated event rates by integrating data from phase 3 short- and medium-term RCTs comparing cinacalcet to placebo, as well as the observational data linking expected changes in PTH, calcium, and phosphorus within components of the primary composite end point. After it became apparent that the overall (blinded) event rate was <20.8%, we extended the trial by 16 months to allow for accrual of the requisite number of events. The actual annual event rate in the placebo group was 15.5% (19).

The event rate in the placebo group of RCTs is frequently lower than predicted (20). When it became apparent that the primary composite event rate in EVOLVE was below that anticipated, the choice of intervention was to increase the number of enrolled participants or to extend the duration of trial follow-up. The first choice would have entailed re-engaging enrollment teams in dialysis units already extended by their contribution to EVOLVE and would have been more costly. The decision was made to extend follow-up. While additional events were expected to accrue over time, adherence to the intervention waned and more PTH-lowering interventions (*i.e.*, kidney transplantation, PTX, and use of commercial cinacalcet) were instituted, particularly in the placebo group. High attrition of participants from the intervention group and drop-in to the intervention group eroded the original planned power of EVOLVE in such a manner that extension of the trial was of little benefit in identifying a treatment effect.

**Table 2. Lessons concerning imbalance in baseline covariates**

- (1) Randomization is not guaranteed to prevent significant imbalance in baseline clinical characteristics (even for a trial with nearly 4000 participants).
- (2) Randomized controlled trials in CKD/ESRD should use both multivariable-adjusted ITT and unadjusted ITT analyses of the primary outcome.
- (3) If there is risk of imbalance of baseline covariates, multivariable-adjusted ITT analysis should be used as the primary analysis on which the trial is judged (18).
- (4) If participants across a broad age range are enrolled, stratification by age may be advisable because age is an important determinant of many outcomes. A similar argument could be made for other key covariates (*e.g.*, baseline BP or proteinuria) depending on the outcome(s) of interest.

ITT, intention-to-treat.



**Figure 1.** | Four main causes of reduction in statistical power that occurred in the Evaluation of Cinacalcet Hydrochloride Therapy to Lower Cardiovascular Events trial. KT, kidney transplant; PTX, parathyroidectomy.

The lessons learned concerning event rates included:

- (1) Event rate estimates in the sample size calculation should be conservative.
- (2) If the event rate is lower than predicted and nonadherence during the trial (especially over time) is likely, then enrollment of more patients is generally preferable to extending follow-up.

### Nonadherence

In the cinacalcet group, median time in the trial was 50.6 months and time on drug was 21.2 months. Corresponding times for the placebo group were 50.4 and 17.5 months, respectively. Time to first discontinuation of study drug for protocol-specified reasons was similar in both cinacalcet and placebo groups, but for nonprotocol-specified reasons it was significantly higher in the placebo group (driven by the fear of severe unremitting HPT) (1). In the placebo group, the drug was discontinued for adverse events in 11.8% of patients compared with 15.8% of patients in the cinacalcet group. Some of this discontinuation was anticipated. The drop-in rate was expected to be 10% per year and the observed rate was 7.4% per year. The dropout rate was derived from discontinuation rates observed in prior cinacalcet RCTs and was higher than anticipated: The expected rate was 10% per year and the observed rate was 27.3% per year. This may have occurred because of the long-term nature of the trial, which amplified the effect of kidney transplantation, PTX, and use of commercial cinacalcet, and because of comorbidity, which predisposed to study drug discontinuation due to adverse events and fear of severe HPT.

Drug discontinuation was a more important factor in diminishing the power of the study than the lower-than-predicted event rates. The increase in sample size necessary in the presence of  $x\%$  discontinuation can be estimated as the reciprocal of  $(1-x)^2$ . Hence, with 40% discontinuation at the trial midpoint, the sample size necessary to maintain study power at the level planned was about 56% greater than planned.

An analytic plan to take into account nonadherence in the estimates of the treatment effect was prespecified. It included lag censoring, iterative parameter estimate, and inverse probability of censoring weights (IPCW). Lag censoring analysis uses data censored at a prespecified time point. In EVOLVE, we had prespecified the lag time to be 6 months after cinacalcet was discontinued in order to account for the possibility that the drug had persistent effects. Although lag censoring preserves randomization, there may be informative bias if nonadherent patients (compared with adherent patients) have different prognostic characteristics associated with the primary end point (21). This methodologic weakness also occurs with the iterative parameter estimate method. However, the IPCW is not prone to informative bias (22).

The IPCW approach censors data when nonadherence occurs. For patients who were adherent and had characteristics similar to those who were not, IPCW assigns larger weights to these patients to re-create the population that would have been observed (Figure 2) (22). Weights are calculated on the basis of the inverse of the probability that patients remain adherent using a logistic regression model. The final relative hazard is derived from a

weighted regression model. In EVOLVE, age, sex, race, country, diabetes history, randomized treatment group, time-dependent covariates of PTH, and the adverse events of hypocalcemia and nausea/vomiting were used in the logistic regression model to estimate the probability of adherence (22). The IPCW method is sensitive to the number of nonadherent patients, assumes there are no unknown confounders, and is computationally difficult. Nonetheless, it is accepted by many health agencies. The relative hazard obtained with IPCW was 0.81 (95% CI, 0.70 to 0.92) (Table 3).

**Effect of Cointerventions**

Before the trial started, it was clear that three cointerventions that lower PTH (kidney transplantation, PTX, and use of commercial cinacalcet) could diminish the treatment effect of cinacalcet, particularly if they were delivered more frequently in the placebo group, as was expected for the latter two. We considered including PTX in a composite end point but were concerned about the lack of pathophysiologic concordance for PTX and CV events (see above). We considered censoring follow-up time after kidney transplantation or PTX but did not want to introduce bias: Patients who were too ill would be unlikely to receive either of these two interventions. Instead, we

prespecified estimates of the treatment effect once follow-up time after these three PTH-lowering interventions was censored. Table 3 shows that the relative hazards were 0.90 (95% CI, 0.82 to 0.99) when censoring occurred at one of these three cointerventions and 0.84 (95% CI, 0.76 to 0.93) with censoring of time after any one of these cointerventions (1).

Table 4 outlines lessons concerning the analytic plan for nonadherence and cointerventions.

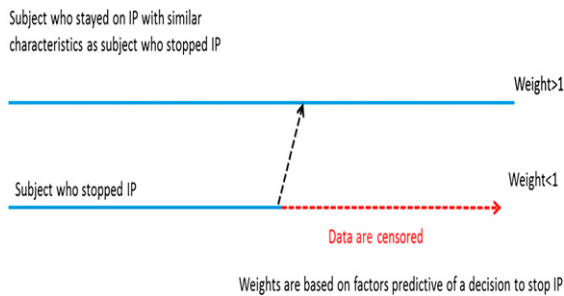
**Subgroup Effects**

We prespecified seven subgroup analyses, including age  $\geq 65$  and  $< 65$  years. The relative hazard for the primary end point using ITT was 0.74 (95% CI, 0.63 to 0.86) in the older group and 0.99 (95% CI, 0.88 to 1.11) in the younger group (19). There was a 27% reduction in mortality in patients age  $> 65$  years ( $P < 0.001$ ). The test of treatment  $\times$  age interaction was significant ( $P = 0.03$  with age used as a continuous variable;  $P = 0.01$  with age dichotomized at 65 years). The age modification of the treatment effect was partly related to (1) three-fold higher rates of kidney transplantation and PTX in younger patients, cointerventions that limited detection of a treatment effect; (2) lower rate of CV disease at baseline in younger patients, which limited the potential of cinacalcet to decrease CV events rates; and (3) lower CV event rates that decreased the power to observe an effect (19).

Criteria by which to evaluate the credibility of subgroup effects have been published (23). Table 5 outlines the lessons from EVOLVE concerning the age subgroup effect.

**Capture of Rare Diseases**

RCTs collect accurate data on predetermined end points reviewed by an event adjudication committee, but local investigators also collect substantial amounts of data during drug exposure to assess adverse effects. This may facilitate the study of rare diseases, particularly if the disorder has been identified in advance and has a plausible biologic rationale. However, care must be taken with *post hoc* analyses of databases with multiple events because statistically significant associations may occur by chance. In EVOLVE, calcific uremic arteriopathy (CUA) occurred infrequently, but cinacalcet reduced its incidence (relative hazard, 0.31; 95% CI, 0.13 to 0.79) (24). Predictors of CUA included allocation to



**Figure 2. | Inverse probability of censoring weight (IPCW).** IPCW method creates a scenario of missing follow-up data by censoring the follow-up of each participant at the time of stopping investigational product (IP) (*i.e.*, weight = 0 for time periods afterward). For participants with similar characteristics who did not stop IP, the IPCW method assigns bigger weights to participants with similar characteristics who were observed without stopping IP.

Table 3. Treatment effect on the primary composite cardiovascular end point: prespecified sensitivity analyses		
Type of Analysis	Relative Hazard (95% Confidence Interval)	P Value
Unadjusted ITT analysis	0.93 (0.85 to 1.02)	0.11
Multivariable-adjusted ITT analysis	0.88 (0.79 to 0.97)	0.01
Inverse probability of censoring weights	0.81 (0.70 to 0.92)	0.03
Censor at PTX	0.90 (0.82 to 0.99)	0.03
Censor at KT	0.90 (0.82 to 0.99)	0.03
Censor at commercial cinacalcet use	0.90 (0.82 to 0.99)	$< 0.001$
Censor at PTX, commercial cinacalcet, or KT	0.84 (0.76 to 0.93)	$< 0.001$
Age $< 65$ yr	0.99 (0.88 to 1.11)	Interaction: 0.01
Age $\geq 65$ yr	0.74 (0.63 to 0.86)	

Data obtained from references 1 and 20. ITT, intention-to-treat; PTX, parathyroidectomy; KT, kidney transplant.

**Table 4. Lessons concerning the analytic plan for nonadherence and cointerventions**

- (1) Commercial availability of study drug can make drop-in a serious problem because the use of commercial drug off protocol limits the capacity of the trial to observe a treatment effect. During trial execution, major efforts are required to reduce use of commercial drug, and an analytic plan needs to take account of drop-in.
- (2A) Discontinuation of study drug is likely in a long-term trial in patients on dialysis, and generous estimates should be incorporated into power calculations.
- (2B) Longer follow-up time does not necessarily increase study power.
- (3) An analytic plan to take into account nonadherence must be prespecified. IPCW may be the best method to account for nonadherence, particularly if determinants of nonadherence are well established.
- (4) It is necessary to take account of the effect of cointerventions that limit assessment of the trial's main treatment effect. Depending on the nature of the intervention, it may be wise to censor follow-up time after the intervention or to incorporate the intervention into a composite end point (*e.g.*, initiation of dialysis in a trial aiming to slow progression of CKD).

IPCW, Inverse probability of censoring weight.

**Table 5. Lessons concerning the credibility of subgroup effects****A subgroup effect is credible (23) when:**

- It is relatively large and highly statistically significant.
- It is one of a small group of prespecified hypotheses tested.
- The test of treatment $\times$ subgroup interaction is significant.
- It is consistent with other reports (33).
- It is consistent across related outcomes (19).
- The biologic rationale is plausible.

placebo, female sex, higher body mass index, hypertension, prior PTX, and prior tobacco use. In addition, use of vitamin K antagonists at the time of CUA was significantly higher than use in patients without CUA.

The lesson learned was that RCTs can provide important information on rare diseases, particularly when the trial follows a large cohort of patients who are at relatively high risk for developing the rare disease (25), and for which a plausible biologic rationale exists.

**Biomarkers**

Collection and storage of blood for subsequent biomarker measurement may contribute to the understanding of disease pathogenesis (26), and in a large RCT the incremental cost of collecting patient samples is relatively low. Because of the long duration of trials (EVOLVE started in August 2006 and finished in January 2012, with primary publication of trial results in November 2012) and the potential for biomarker research to be published during the execution of the trial, we did not designate which biomarkers would be measured until after the trial was completed. From 2007 to 2011, accumulating data in patients with advanced CKD and ESRD showed that serum fibroblast growth factor-23 (FGF-23) concentrations were elevated and were associated with mortality, left ventricular hypertrophy, and CV events independent of serum phosphate, PTH, and a variety of demographic and other clinical factors (27,28). Sixty-four percent of patients assigned to cinacalcet had  $\geq 30\%$  reduction in serum FGF-23 concentrations from baseline to week 20 compared with 28%

of the placebo group (29). Among patients randomly assigned to cinacalcet, a  $\geq 30\%$  reduction in FGF-23 was associated with a significant reduction in the composite CV end point (relative hazard, 0.82; 95% CI, 0.69 to 0.98) (29). This nonrandomized analysis generates the hypothesis that FGF-23 plays a role in the pathogenesis of CV events in CKD and the deleterious effects of FGF-23 might be ameliorated by cinacalcet.

The lesson learned was that collection of serum for biomarker investigation in a large RCT should be undertaken because it facilitates rapid testing of hypotheses generated after the RCT started and creation of new hypotheses.

**Interpretation of Clinical Benefits**

Trials should not be judged by the result of a single analysis and a single *P* value, but inferences should be based on the totality of the data. The decision to prescribe cinacalcet should be informed by consideration of the benefit-to-risk ratio in individual patients. Cinacalcet is approved for the treatment of secondary HPT, and in EVOLVE it was effective in preventing severe unremitting HPT. Treatment effects of cinacalcet on fracture rates were similar to effects on CV events in unadjusted ITT analysis, multivariable-adjusted ITT analysis, and censoring at cointerventions that reduce PTH; age was also an effect modifier (30). A recent review has examined the clinical and practical use of calcimimetics in dialysis patients (31). An economic evaluation of cinacalcet in the United States has been reported (32).



## Conclusions

The EVOLVE trial provides multiple lessons in the planning, execution, analysis, and interpretation of RCTs in CKD/ESRD. These include deciding upon, defining, and analyzing outcomes; responding to lower-than-predicted event rates; taking into account nonadherence and coin-ventions; subgroup effects; capture of rare diseases; biomarkers; and interpretation of benefit-to-harm ratio.

## Acknowledgments

The EVOLVE trial was funded by Amgen; EVOLVE Clinical Trials, government number NCT003345839.

## Disclosures

All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. Please see Supplemental Material for completed forms, which appear in alphabetical order by author.

## References

- Chertow GM, Block GA, Correa-Rotter R, Drüeke TB, Floege J, Goodman WG, Herzog CA, Kubo Y, London GM, Mahaffey KW, Mix TC, Moe SM, Trotman ML, Wheeler DC, Parfrey PS; EVOLVE Trial Investigators: Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *N Engl J Med* 367: 2482–2494, 2012
- Parfrey PS, Barrett BJ: Cardiovascular disease and chronic kidney disease. In: *Chronic Kidney Disease*, edited by Kimmel PL, Rosenberg ME, Waltham MA, Academic Press, 2015, pp 181–198
- McCullough PA, Agrawal V, Danielewicz E, Abela GS: Accelerated atherosclerotic calcification and Monckeberg's sclerosis: A continuum of advanced vascular pathology in chronic kidney disease. *Clin J Am Soc Nephrol* 3: 1585–1598, 2008
- Wheeler DC, London GM, Parfrey PS, Block GA, Correa-Rotter R, Dehmel B, Drüeke TB, Floege J, Kudo Y, Mahaffey KW, Goodman WG, Moe SM, Trotman M-L, Abdalla S, Chertow GM, Herzog CA; Evaluation Of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial Investigators: Effects of cinacalcet on atherosclerotic and non atherosclerotic cardiovascular events in patients receiving hemodialysis. The Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial. *J Am Heart Assoc* 17: e001363, 2014
- Amann K: Media calcification and intima calcification are distinct entities in chronic kidney disease. *Clin J Am Soc Nephrol* 3: 1599–1605, 2008
- Seltzer JH, Turner JR, Geiger MJ, Rosano G, Mahaffey KW, White WB, Sabol MB, Stockbridge N, Sager PT: Centralized adjudication of cardiovascular end points in cardiovascular and non-cardiovascular pharmacologic trials: a report from the Cardiac Safety Research Consortium. *Am Heart J* 169: 197–204, 2015
- Chawla LS, Herzog CA, Costanzo MR, Tumlin J, Kellum JA, McCullough PA, Ronco C; ADQI XI Workgroup: Proposal for a functional classification system of heart failure in patients with end-stage renal disease: Proceedings of the acute dialysis quality initiative (ADQI) XI workgroup. *J Am Coll Cardiol* 63: 1246–1252, 2014
- Harnett JD, Foley RN, Kent GM, Barre PE, Murray D, Parfrey PS: Congestive heart failure in dialysis patients: Prevalence, incidence, prognosis and risk factors. *Kidney Int* 47: 884–890, 1995
- Piller LB, Davis BR, Cutler JA, Cushman WC, Wright JT, Williamson JD, Leenan FHH, Einhorn PT, Randall OS, Golden JS, Haywood LJ, and the ALLHAT Collaborative Research Group: Validation of heart failure events in the antihypertensive and lipid lowering treatment to prevent heart attack Trial (ALL HAT) participants assigned to doxazosin and chlorthalidone. *Curr Control Trials Cardiovasc Med* 3: 10, 2002
- Mahaffey KW, Roe MT, Dyke CK, Newby LK, Kleiman NS, Connolly P, Berdan LG, Sparapani R, Lee KL, Armstrong PW, Topol EJ, Califf RM, Harrington RA: Misreporting of myocardial infarction end points: Results of adjudication by a central clinical events committee in the PARAGON-B trial. Second Platelet IIb/IIIa Antagonist for the Reduction of Acute Coronary Syndrome Events in a Global Organization Network Trial. *Am Heart J* 143: 242–248, 2002
- Mahaffey KW, Harrington RA, Akkerhuis M, Kleiman NS, Berdan LG, Crenshaw BS, Tardiff BE, Granger CB, DeJong I, Bhapkar M, Widimsky P, Corbalan R, Lee KL, Deckers JW, Simoons ML, Topol EJ, Califf RM; For the PURSUIT Investigators: Disagreements between central clinical events committee and site investigator assessments of myocardial infarction endpoints in an international clinical trial: Review of the PURSUIT study. *Curr Control Trials Cardiovasc Med* 2: 187–194, 2001
- Parfrey PS, Chertow GM, Block GA, Correa-Rotter R, Drüeke TB, Floege J, Herzog CA, London GM, Mahaffey KW, Moe SM, Wheeler DC, Dehmel B, Trotman ML, Modafferi DM, Goodman WG: The clinical course of treated hyperparathyroidism among patients receiving hemodialysis and the effect of cinacalcet: The EVOLVE trial. *J Clin Endocrinol Metab* 98: 4834–4844, 2013
- Chu R, Walter SD, Guyatt GP, Devereaux PJ, Walsh M, Thorlund K, Thabane L: Assessment and implication of prognostic balance in randomized controlled trials with a binary outcome—a simulation study. *PLOS One* 7: e36677, 2012
- Baigent C, Landrey M, Emberson J, Wheeler DC, Tomson C; SHARP Investigators: The effects of lowering LDL cholesterol with simvastatin plus ezetimide in patients with chronic kidney disease (Study of Heart and Renal Protection): A randomized placebo controlled trial. *Lancet* 377: 2182–2192, 2011
- Pfeffer MA, Burdmann EA, Chen CY, Cooper ME, de Zeeuw D, Eckardt KU, Feyzi JM, Ivanovich P, Kewalramani R, Levey AS, Lewis EF, McGill JB, McMurray JJ, Parfrey P, Parving HH, Remuzzi G, Singh AK, Solomon SD, Toto R; TREAT Investigators: A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N Engl J Med* 361: 2019–2032, 2009
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T; CONSORT GROUP (Consolidated Standards of Reporting Trials): The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Ann Intern Med* 134: 663–694, 2001
- Austin PC, Manca A, Zwarenstein M, Juurlink DN, Stanbrook MB: A substantial and confusing variation exists in handling of baseline covariates in randomized controlled trials: A review of trials published in leading medical journals. *J Clin Epidemiol* 63: 142–153, 2010
- Hernández AV, Eijkemans MJ, Steyerberg EW: Randomized controlled trials with time-to-event outcomes: How much does prespecified covariate adjustment increase power? *Ann Epidemiol* 16: 41–48, 2006
- Parfrey PS, Drüeke TB, Block GA, Correa-Rotter R, Floege J, Herzog CA, London GM, Mahaffey KW, Moe SM, Wheeler DC, Kubo Y, Dehmel B, Goodman WG, Chertow GM; Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial Investigators: The effects of cinacalcet in older and younger patients on hemodialysis: The Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial. *Clin J Am Soc Nephrol* 10: 791–799, 2015
- Walach H, Sadaghiani C, Dehm C, Bierman D: The therapeutic effect of clinical trials: understanding placebo response rates in clinical trials—a secondary analysis. *BMC Med Res Methodol* 5: 26, 2005
- Tripepi G, Heinze G, Jager KJ, Stel VS, Dekker FW, Zoccali C: Lag-censoring analysis: Lights and shades. *Nephrol Dial Transplant* 30: 700–705, 2015
- Kubo Y, Sterling LR, Parfrey PS, Gill K, Mahaffey KW, Gioni I, Trotman ML, Dehmel B, Chertow GM: Assessing the treatment effect in a randomized controlled trial with extensive non-adherence: the EVOLVE trial. *Pharm Stat* 14: 242–251, 2015
- Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, Bala MM, Bassler D, Mertz D, Diaz-Granados N, Vandvik PO, Malaga G, Srinathan SK, Dahm P, Johnston BC, Alonso-Coello P, Hassouneh B, Walter SD, Heels-Ansdell D, Bhatnagar N, Altman DG, Guyatt GH: Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ* 344: e1553, 2012
- Floege J, Kubo Y, Floege A, Chertow GM, Parfrey PS: The effect of cinacalcet on calcific uremic arteriopathy events in patients receiving hemodialysis: The EVOLVE trial. *Clin J Am Soc Nephrol* 10: 800–807, 2015

25. Ketteler M, Biggar PH: Evolving calciphylaxis—what randomized, controlled trials can contribute to the capture of rare diseases. *Clin J Am Soc Nephrol* 10: 726–728, 2015
26. Hsu CY, Ballard S, Battle D, Bonventre JV, Böttinger EP, Feldman HI, Klein JB, Coresh J, Eckfeldt JH, Inker LA, Kimmel PL, Kusek JW, Liu KD, Mauer M, Mifflin TE, Molitch ME, Nelsestuen GL, Rebholz CM, Rovin BH, Sabbisetti VS, Van Eyk JE, Vasan RS, Waikar SS, Whitehead KM, Nelson RG; CKD Biomarkers Consortium: Cross-disciplinary biomarkers research: lessons learned by the CKD Biomarkers Consortium. *Clin J Am Soc Nephrol* 10: 894–902, 2015
27. Faul C, Amaral AP, Oskouei B, Hu MC, Sloan A, Isakova T, Gutiérrez OM, Aguillon-Prada R, Lincoln J, Hare JM, Mundel P, Morales A, Scialla J, Fischer M, Soliman EZ, Chen J, Go AS, Rosas SE, Nessel L, Townsend RR, Feldman HI, St John Sutton M, Ojo A, Gadegbeku C, Di Marco GS, Reuter S, Kentrup D, Tiemann K, Brand M, Hill JA, Moe OW, Kuro-O M, Kusek JW, Keane MG, Wolf M: FGF23 induces left ventricular hypertrophy. *J Clin Invest* 121: 4393–4408, 2011
28. Isakova T, Xie H, Yang W, Xie D, Anderson AH, Scialla J, Wahl P, Gutiérrez OM, Steigerwalt S, He J, Schwartz S, Lo J, Ojo A, Sondheimer J, Hsu CY, Lash J, Leonard M, Kusek JW, Feldman HI, Wolf M; Chronic Renal Insufficiency Cohort (CRIC) Study Group: Fibroblast growth factor 23 and risks of mortality and end-stage renal disease in patients with chronic kidney disease. *JAMA* 305: 2432–2439, 2011
29. Moe SM, Chertow GM, Parfrey PS, Kubo Y, Block GA, Correa-Rotter R, Drüeke TB, Herzog CA, London GM, Mahaffey KW, Wheeler DC, Stolina M, Dehmel B, Goodman WG, Floege J; Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial Investigators: Cinacalcet, fibroblast growth factor-23, and cardiovascular disease in hemodialysis: The Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) trial. *Circulation* 132: 27–39, 2015
30. Moe SM, Abdalla S, Chertow GM, Parfrey PS, Block GA, Correa-Rotter R, Floege J, Herzog CA, London GM, Mahaffey KW, Wheeler DC, Dehmel B, Goodman WG, Drüeke TB; Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) Trial Investigators: Effects of cinacalcet on fracture events in patients receiving hemodialysis: The EVOLVE trial. *J Am Soc Nephrol* 26: 1466–1475, 2015
31. Bover J, Ureña P, Ruiz-García C, daSilva I, Lescano P, Del Carpino J, Ballarín J, Cozzolino M: Clinical and practical use of calcimimetics in dialysis patients with secondary hyperparathyroidism [published online ahead of print July 29, 2015]. *Clin J Am Soc Nephrol* doi: 10.2215/CJN.01760215
32. Belozeroff V, Chertow GM, Graham CN, Dehmel B, Parfrey PS, Briggs AH: Economic evaluation of cinacalcet in the United States: The EVOLVE trial [published online ahead of print October 5, 2015]. *Value Health* doi: 10.1016/j.jval.2015.08.007
33. Suki WN, Zabaneh R, Cangiano JL, Reed J, Fischer D, Garrett L, Ling BN, Chasan-Taber S, Dillon MA, Blair AT, Burke SK: Effects of sevelamer and calcium-based phosphate binders on mortality in hemodialysis patients. *Kidney Int* 72: 1130–1137, 2007

Published online ahead of print. Publication date available at [www.cjasn.org](http://www.cjasn.org).

This article contains supplemental material online at <http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.06370615/-/DCSupplemental>.