ORIGINAL REPORT

Propensity score matching and persistence correction to reduce bias in comparative effectiveness: the effect of cinacalcet use on all-cause mortality[†]

Iain A. Gillespie[1]*, Jürgen Floege[2], Ioanna Gioni[3], Tilman B. Drüeke[4], Angel L. de Francisco[5], Stefan D. Anker[6], Yumi Kubo[7], David C. Wheeler[8], Marc Froissart[9] and on behalf the ARO Steering Committee collaborators

[1]*Amgen Ltd, Uxbridge, UK*
[2]*Division of Nephrology, Medizinische Klinik II, RWTH University Hospital Aachen, Aachen, Germany*
[3]*on behalf of Amgen Ltd, UK*
[4]*Inserm Unit 1088, UFR de Médecine et de Pharmacie, Université de Picardie, Amiens, France*
[5]*Servicio de Nefrología, Hospital Universitario Valdecilla, Universidad de Cantabria, Santander, Spain*
[6]*Department of Innovative Clinical Trials, University Medical Centre Göttingen, Göttingen, Germany*
[7]*Amgen Inc, Thousand Oaks, CA, USA*
[8]*Center for Nephrology, University College London, UK*
[9]*Amgen Europe GmbH, Zug, Switzerland*

ABSTRACT

Purpose   The generalisability of randomised controlled trials (RCTs) may be limited by restrictive entry criteria or by their experimental nature. Observational research can provide complementary findings but is prone to bias. Employing propensity score matching, to reduce such bias, we compared the real-life effect of cinacalcet use on all-cause mortality (ACM) with findings from the Evaluation of Cinacalcet Therapy to Lower Cardiovascular Events (EVOLVE) RCT in chronic haemodialysis patients.
Methods   Incident adult haemodialysis patients receiving cinacalcet, recruited in a prospective observational cohort from 2007–2009 (AROii; $n = 10,488$), were matched to non-exposed patients regardless of future exposure status. The effect of treatment crossover was investigated with inverse probability of censoring weighted and lag-censored analyses. EVOLVE ACM data were analysed largely as described for the primary composite endpoint.
Results   AROii patients receiving cinacalcet ($n = 532$) were matched to 1790 non-exposed patients. The treatment effect of cinacalcet on ACM in the main AROii analysis (hazard ratio 1.03 [95% confidence interval (CI) 0.78–1.35]) was closer to the null than for the Intention to Treat (ITT) analysis of EVOLVE (0.94 [95%CI 0.85–1.04]). Adjusting for non-persistence by 0- and 6-month lag-censoring and by inverse probability of censoring weight, the hazard ratios in AROii (0.76 [95%CI 0.51–1.15], 0.84 [95%CI 0.60–1.18] and 0.79 [95%CI 0.56–1.11], respectively) were comparable with those of EVOLVE (0.82 [95%CI 0.67–1.01], 0.83 [95%CI 0.73–0.96] and 0.87 [95%CI 0.71–1.06], respectively).
Conclusions   Correcting for treatment crossover, we observed results in the 'real-life' setting of the AROii observational cohort that closely mirrored the results of the EVOLVE RCT. Persistence-corrected analyses revealed a trend towards reduced ACM in haemodialysis patients receiving cinacalcet therapy. © 2015 The Authors. *Pharmacoepidemiology and Drug Safety* Published by John Wiley & Sons Ltd.

KEY WORDS—cinacalcet; haemodialysis; mortality; persistence; bias; pharmacoepidemiology

INTRODUCTION

Although randomised controlled trials (RCTs) remain the 'gold standard' for assessing pharmaceutical

interventions, their generalisability may be limited by restrictive entry criteria or their experimental nature; in some instances, they are unfeasible or unethical.[1] Observational research studies yield important information about the effectiveness of treatment regimens and their use by clinicians in everyday 'real world' practice and thus complement the results of RCTs. They use less stringent inclusion/exclusion criteria and hence may closer reflect the overall patient population. They are prone to bias, however, especially

---

Correspondence to: I. Gillespie, Center for Observational Research, Amgen Ltd, 1 Sanderson Road, Uxbridge, UK UB8 1DH. Email: iain.gillespie@amgen.com

confounding-by-indication, which may potentially distort estimated treatment effects.

Secondary hyperparathyroidism (sHPT) occurs early in the course of chronic kidney disease (CKD) and progresses with declining kidney function.[2] The condition is largely defined by increased serum parathyroid hormone (PTH) levels accompanied by deregulated phosphate and calcium concentrations, and major complications include bone and cardiovascular diseases (CVD).[3] These defining biochemical changes are independently and consistently associated with increased morbidity and mortality,[4–9] which is most likely in part mediated by cardiovascular calcification.[10–13] Effective sHPT control is therefore an important goal in the management of haemodialysis (HD) patients.[14] The calcimimetic cinacalcet (Sensipar®/Mimpara®) is licenced for the treatment of sHPT in end-stage renal disease patients on maintenance dialysis therapy.[15] Cinacalcet effectively lowers serum PTH concentrations while concomitantly reducing serum calcium and, at least transiently, phosphorus concentrations.[16–19]

Few observational studies have examined the effect of cinacalcet therapy on survival. Block and colleagues found a significant survival benefit associated with cinacalcet prescription, with reductions in all-cause and cardiovascular mortality in almost 20 000 United States HD patients receiving intravenous vitamin D therapy.[20] In a separate, smaller ($n = 1184$), post-hoc study examining short-term safety data pooled from four placebo-controlled RCTs, Cunningham and colleagues reported that cinacalcet use was associated with a significantly reduced risk of parathyroidectomy, fracture and cardiovascular-related hospitalisation, but not mortality.[21]

Propensity score matching (PSM) can reduce bias in observational research.[22] PSM aims to achieve balance between treatment groups with regard to measured confounders by mimicking the randomisation used in clinical trials. In the current study, the effect of cinacalcet use on all-cause mortality was investigated in HD patients using data obtained from Fresenius Medical Care (FMC) dialysis centres across Europe as part of the Analyzing data, Recognizing excellence, and Optimizing outcomes (ARO) CKD Research Initiative.[23] The study used PSM to control for systematic differences between those receiving treatment for sHPT with cinacalcet and those who were not. We were more interested in the potential for treatment crossover to distort the estimated treatment effect of cinacalcet use on all-cause mortality—and our ability to correct this form of bias—than estimation of the actual treatment effect itself. Findings were compared with treatment estimates for all-cause mortality from the EVOLVE study[24]: an RCT designed to gain an insight into the long-term clinical efficacy of cinacalcet.

METHODS

(i) ARO data and analysis

Source data and study population. The AROii cohort comprised incident (<183 days dialysis vintage [the time since dialysis initiation]) adult HD patients presenting at over 300 FMC facilities in 14 European countries between January 2007 and December 2009, with no renal transplantation or peritoneal dialysis history. Raw, anonymised electronic patient-level data,[25] captured as part of normal clinical care, extracted and supplied quarterly, were limited to chronic HD patients (≥ 10 contiguous dialysis sessions) with accompanying laboratory data. Data were further restricted in this study to non-parathyroidectomised, cinacalcet-naïve (no cinacalcet use to the end of the first 90 days of follow-up) patients who remained on study for ≥ 90 days. Patients from countries not prescribing cinacalcet were excluded.

Exposure and time period definitions. Patients' follow-up time was divided into consecutive 90-day dialysis vintage windows to maximise patient comparability with regard to CKD progression. Each interval included patients initiating cinacalcet and those who did not (cinacalcet/non-cinacalcet patients, respectively). Baseline comprised the 90 days before treatment initiation for cinacalcet patients and the 90 days before the assessment interval for non-cinacalcet patients. Patients accrued time-at-risk from the end of baseline until they experienced the event of interest (all-cause mortality), underwent a parathyroidectomy, a kidney transplant, were lost to follow-up (>45 days without continuous dialysis treatment) or the end of study follow-up was reached (31 September 2012).

PSM approach. The PSM process resembled the principles of a sequential matching using time-dependent covariates. However, unlike a time-dependent PSM, where patients with similar time-dependent covariates up to treatment initiation are matched,[26] only baseline values were utilised.

The propensity for cinacalcet treatment was estimated using multivariate logistic regression, where exposure to cinacalcet treatment in each interval was fitted as the dependent variable and baseline covariates were fitted as independent variables (Supplementary

Table A). The explanatory variables were chosen as potential risk factors for cinacalcet treatment and/or confounders of the relationship between treatment and mortality in this cohort.[27] Analysis was restricted to patients with complete baseline data on PTH, serum calcium and serum phosphate, and interaction terms were included to represent the *in vivo* effect of active vitamin D (AVD) treatment on serum calcium and serum phosphate.

Matching was performed chronologically. Cinacalcet patients in the first interval were calliper-matched by logit propensity score to up to four ('ncontls' option = 4) non-cinacalcet patients using a greedy matching algorithm.[28] A calliper width ('dmax' option) of 0.2 of the standard deviation of the logit of the propensity score was used.[29] Logit propensity scores were unweighted ('wts' option = 1), and distances were calculated using weighted sums of absolute case-control differences ('dist' option = 1). Matched patients and unmatched cinacalcet patients were removed from later intervals and the process repeated for subsequent intervals. The selection process employed in the PSM to create this population, where patients retain their matching exposure status regardless of their future exposure status, has parallels with the ITT analytical approach employed in RCTs.

The balance in baseline characteristics achieved by PSM was evaluated by calculating standardised differences[29] between cinacalcet and non-cinacalcet patients in the matched and overall ARO populations. As unmatched non-exposed patients could be considered for matching in multiple exposure assessment windows, the characteristics of patients receiving cinacalcet, versus those who never received the drug, were examined in the overall ARO population by comparing data for the 3-month period from recruitment in the AROii cohort. In the matched population, the balance was checked for each risk set separately and overall. Clinically significant differences aside, standardised differences of <10% were considered negligible.[29]

Time-to-event analysis. Following PSM, the association between cinacalcet exposure and all-cause mortality, determined by patients' death dates, was estimated using matched Cox proportional hazards regression models, with hazard ratios (HR) and 95% confidence intervals (CI) calculated.[29]

Sensitivity analyses investigating non-persistence. To account for non-persistence to exposure status, both lag-censoring analysis and inverse probability of censoring weights (IPCWs[30]) were applied.

Six-month lag-censoring was employed to match the pre-specified lag-censored analysis in the EVOLVE trial,[24] but 0-month lag-censoring was also investigated. Patients who switched from non-exposed to exposed were censored 0 or 6 months after cinacalcet initiation; cinacalcet patients with poorer persistence (defined as the first 90-day period post-initiation where prescriptions covered less than two-thirds of the period) were censored 0 or 6 months after the time of reduced persistence. HRs with 95%CIs were calculated.

Inverse probability of censoring weights were derived based on baseline covariates, plus time-dependent serum PTH, calcium and phosphate values, and applied in a pooled logistic regression model to estimate the effect of cinacalcet on mortality. The use of time-dependent values here differed from implementation of the PSM procedure, where only baseline values serum PTH, calcium and phosphate were used. Cinacalcet patients with poorer persistence were censored at the time of reduced persistence while patients who switched from non-exposed to exposed were censored at cinacalcet initiation. Consequently, the follow-up time of persistent patients was weighted to compensate for those who did not persist. Weighted pooled odds ratios (ORs), approximating to HRs, were calculated with 95%CIs.

(ii) EVOLVE data and analysis

The EVOLVE trial was a large, multi-centre, double-blind, placebo-controlled RCT where 3883 HD patients with sHPT were randomly assigned cinacalcet or placebo.[24] Although there were numerically fewer primary composite endpoints (time to death or first non-fatal cardiovascular event) in patients randomised to cinacalcet compared with those of placebo, this difference failed to reach statistical significance in unadjusted intention-to-treat analysis.[24] Pre-specified secondary and sensitivity analyses such as covariate adjustment and lag-censoring, however, revealed a nominally significant 12–15% risk reduction with cinacalcet.[24]

All-cause mortality was analysed as described previously for the primary composite endpoint. In the EVOLVE trial, the mortality endpoint was not part of the formal statistical testing strategy; associations are therefore considered 'nominal'. For IPCW, data were censored at the time of study drug discontinuation in both treatment groups. For each interval, patients' weights were derived using baseline covariates, time-dependent serum PTH, calcium and phosphate values, interaction terms with treatment and laboratory measures and the adverse event of hypocalcemia. The

effect of cinacalcet on mortality was estimated using weighted pooled logistic regression.

RESULTS

The ARO cohort and EVOLVE trial

The characteristics of the ARO cohort and the EVOLVE trial are summarised in Supplementary Table B. Aside from their observational and experimental nature, respectively, the major differences related to their geography, dialysis vintage, length of follow-up and study selection criteria.

ARO study population

Of 11 190 patients recruited into the AROii cohort, 702 patients were excluded, because, alone or in combination, they were not receiving HD ($n = 487$), had a kidney transplant history ($n = 86$) or no laboratory data ($n = 255$), leaving 10 488 patients. When study-specific selection criteria (no cinacalcet use in a country ($n = 200$), <90 days of follow-up ($n = 925$), parathyroidectomy history ($n = 12$) and cinacalcet use up to and including the first 90 days of follow-up ($n = 275$)) were applied alone or in combination, a further 1387 (13.2%) patients were excluded, leaving 9101 patients eligible for matching.

Baseline characteristics of the overall and matched ARO populations

Patients receiving cinacalcet during follow-up ($n = 1168$; 12.8%) tended to be younger than those never exposed ($n = 7933$) and were healthier with regards to diabetes history, the need for catheterization or hospitalisation during the eligibility period and in terms of their inflammatory (C-reactive protein) or nutritional (serum albumin) status but were unhealthier with regard to CVD and fracture history (Table 1). As expected, they had elevated serum calcium, phosphate and PTH levels and were more reliant on phosphate binders (especially non-calcium-based). They were more often prescribed CVD medications and AVD.

Five-hundred and thirty-two patients exposed to cinacalcet were matched to 1790 patients not exposed at the time of exposure assessment (Table 1). Of the latter, initially non-exposed patients, 521 (29.1%) subsequently received cinacalcet in their follow-up (henceforth termed 'future cinacalcet' patients). A further 115 patients receiving cinacalcet remained unmatched.

Matched cinacalcet patients (median [Q1,Q3] propensity score = 0.17 [0.07, 0.32]) contributed a median

of 1.82 person-years (PY) at risk to the study, while patients matched as non-cinacalcet (median [Q1,Q3] propensity score = 0.13 [0.06, 0.22]) contributed a median of 1.94 PY at risk. Attrition due to a successful renal transplant, parathyroidectomy and lost to follow-up was similar in the two groups. In the PSM population, the differences in baseline patient characteristics, so apparent in the overall ARO population, were negligible with the exception of the lowest stratum of PTH (Table 1). Few clinically relevant differences were observed at the individual risk-set level (where sufficient data were available for comparison; data not shown).

Outcomes analyses in ARO and EVOLVE

The HR for all-cause mortality in ARO, adjusted for the slight PTH imbalance described in the previous text, was 1.03 (95%CI 0.78–1.35). This estimate was closer to the null than that observed for EVOLVE (HR 0.94 [95%CI 0.85–1.04]; Supplementary Table C).

Persistence to patients' initial exposure status

Differential treatment crossover was observed for AROii patients matched as cinacalcet patients and those matched as non-exposed (Figure 1). Ten percent of cinacalcet patients on study did not persist with therapy for the first 90 days post-matching and this proportion increased to 24% for the first 180 days. The attrition rate slowed subsequently, with approximately 60% of cinacalcet patients persisting for the remainder of follow-up. Treatment crossover was more gradual for non-cinacalcet patients, probably reflecting sHPT disease progression over time. Overall, 845 patients (36.4%) did not persist to their initial matching exposure status, with cinacalcet patients less likely to persist in their exposure status than non-cinacalcet patients (pooled OR 0.57 [95%CI 0.48–0.68]).

Future AROii cinacalcet patients ($n = 521$) had higher serum PTH, calcium and phosphate than those never exposed ($n = 1269$; Supplementary Table D) and were more often prescribed phosphate binders. They tended to be younger, had higher serum albumin and were more often matched in periods prior to the publication of the Kidney Disease: Improving Global Outcomes (KDIGO™) guideline[31] (which recommended a higher PTH treatment threshold). Of note, the all-cause mortality rate was lower in future cinacalcet patients (3.4 deaths per 100 PY [95%CI 2.5–4.5]) than in those patients who remained never exposed (11.7 per 100 PY [95%CI 10.3–13.2]).

In EVOLVE, 1300 of the 1948 patients randomised to cinacalcet discontinued therapy (67%), while

Table 1. Characteristics of cinacalcet and non-cinacalcet patients in the overall ARO and propensity score-matched populations

Parameter	ARO population*			Matched		
	Non-cinacalcet	Cinacalcet	Std diff.[‡]	Non-cinacalcet	Cinacalcet	Std diff.
	(n = 7933)	(n = 1168)		(n = 1790)	(n = 532)	
Person time at risk (years)						
Mean ± SD	2.39 ± 1.44	3.44 ± 1.16		1.99 ± 1.25	1.94 ± 1.18	
Q1, Q3	1.01, 3.50	2.76, 4.32		0.94, 2.90	0.96, 2.82	
Patient attrition						
Successful renal transplant	771 (9.7)	203 (17.4)		263 (14.7)	96 (18.0)	
Parathyroidectomy	13 (0.2)	2 (0.2)		6 (0.3)	2 (0.4)	
Lost to follow-up	1988 (25.1)	146 (12.5)		283 (15.8)	80 (15.0)	
Exposure period pre-KDIGO	6442 (81.2)	987 (84.5)	0.088	854 (47.7)	245 (46.1)	0.001
Age group						
<30 years	165 (2.1)	31 (2.7)	0.038	46 (2.6)	12 (2.3)	0.040
30–49 years	968 (12.2)	218 (18.7)	0.180	274 (15.3)	79 (14.8)	0.035
50–64 years	2122 (26.7)	338 (28.9)	0.049	494 (27.6)	146 (27.4)	0.010
≥65 years	4678 (59.0)	581 (49.7)	0.186	976 (54.5)	295 (55.5)	0.047
Sex						
Female	3124 (39.4)	494 (42.3)	0.059	736 (41.1)	218 (41.0)	0.010
Male	4809 (60.6)	674 (57.7)		1054 (58.9)	314 (59.0)	
Baseline hospitalisation	1780 (22.4)	189 (16.2)	0.159	261 (14.6)	66 (12.4)	0.052
History of diabetes	2874 (36.2)	338 (28.9)	0.156	539 (30.1)	158 (29.7)	0.000
History of cancer	659 (8.3)	97 (8.3)	0.000	185 (10.3)	48 (9.0)	0.047
History of CVD	1985 (25.0)	345 (29.5)	0.102	617 (34.5)	190 (35.7)	0.029
History of fractures	106 (1.3)	33 (2.8)	0.104	78 (4.4)	18 (3.4)	0.057
Dialysis vintage (months)						
Mean ± SD	1.1 ± 2.1	1.0 ± 1.9	0.007	14.6 ± 11.3	16.3 ± 11.7	0.081
Baseline vascular access						
Catheter	3712 (46.8)	470 (40.2)	0.132	553 (30.9)	163 (30.6)	0.004
Non-catheter only	3064 (38.6)	558 (47.8)	0.186	1107 (61.8)	338 (63.5)	0.022
Missing	1157 (14.6)	140 (12.0)	0.077	130 (7.3)	31 (5.8)	0.036
iPTH group [pg/mL]						
<75	908 (11.4)	52 (4.5)	0.261	19 (1.1)	1 (0.2)	0.096
≥75–<150	1361 (17.2)	106 (9.1)	0.241	63 (3.5)	17 (3.2)	0.009
≥150–≤300	2138 (27.0)	310 (26.5)	0.009	258 (14.4)	77 (14.5)	0.053
>300–≤600	1351 (17.0)	350 (30.0)	0.309	900 (50.3)	260 (48.9)	0.009
>600	417 (5.3)	194 (16.6)	0.370	550 (30.7)	177 (33.3)	0.036
Missing	1758 (22.2)	156 (13.4)	0.232			
Total calcium [mmol/L]						
<2.10	2298 (29.0)	199 (17.0)	0.286	345 (19.3)	97 (18.2)	0.008
≥2.10–≤2.37	4409 (55.6)	700 (59.9)	0.088	1064 (59.4)	316 (59.4)	0.017
>2.37	797 (10.0)	225 (19.3)	0.263	381 (21.3)	119 (22.4)	0.027
Missing	429 (5.4)	44 (3.8)	0.078			
Phosphate [mmol/L]						
<1.13	1360 (17.1)	106 (9.1)	0.241	184 (10.3)	43 (8.1)	0.045
≥1.13–≤1.78	4645 (58.6)	639 (54.7)	0.078	979 (54.7)	311 (58.5)	0.070
>1.78	1644 (20.7)	394 (33.7)	0.295	627 (35.0)	178 (33.5)	0.046
Missing	284 (3.6)	29 (2.5)	0.064			
CRP [mg/L]						
≤ Q1[§]	1421 (17.9)	285 (24.4)	0.159	442 (24.7)	135 (25.4)	0.023
> Q1–≤Q2	1461 (18.4)	240 (20.5)	0.054	440 (24.6)	135 (25.4)	0.004
> Q2–≤Q3	1474 (18.6)	230 (19.7)	0.028	317 (17.7)	89 (16.7)	0.018
> Q3	1565 (19.7)	137 (11.7)	0.221	194 (10.8)	52 (9.8)	0.018
Missing	2012 (25.4)	276 (23.6)	0.040	397 (22.2)	121 (22.7)	0.009
Serum albumin [g/L]						
≤ Q1[i]	1870 (23.6)	143 (12.2)	0.299	170 (9.5)	50 (9.4)	0.005
> Q1–≤Q2	1756 (22.1)	223 (19.1)	0.075	290 (16.2)	75 (14.1)	0.048
> Q2–≤Q3	1677 (21.1)	298 (25.5)	0.104	453 (25.3)	141 (26.5)	0.032
> Q3	1581 (19.9)	348 (29.8)	0.230	718 (40.1)	221 (41.5)	0.022
Missing	1049 (13.2)	156 (13.4)	0.004	159 (8.9)	45 (8.5)	0.032
Phosphate binder use						
None	4365 (55.0)	462 (39.6)	0.314	557 (31.1)	156 (29.3)	0.019
Only calcium-based	2451 (30.9)	274 (23.5)	0.168	340 (19.0)	96 (18.0)	0.004
Only non-calcium-based	699 (8.8)	266 (22.8)	0.390	595 (33.2)	189 (35.5)	0.022

(*Continues*)

Table 1.   (Continued)

Parameter	ARO population*			Matched		
	Non-cinacalcet	Cinacalcet	Std diff.‡	Non-cinacalcet	Cinacalcet	Std diff.
	(n = 7933)	(n = 1168)		(n = 1790)	(n = 532)	
Both calcium-based and non-calcium-based	418 (5.3)	166 (14.2)	0.305	298 (16.6)	91 (17.1)	0.002
Cardiovascular medication use	5047 (63.6)	893 (76.5)	0.283	1399 (78.2)	427 (80.3)	0.064
Active vitamin D use	3035 (38.3)	598 (51.2)	0.262	1039 (58.0)	319 (60.0)	0.031

ARO, Analyzing data, Recognizing excellence, and Optimizing outcomes; SD, standard deviation; KDIGO, Kidney Disease: Improving Global Outcomes; CVD, cardiovascular diseases; iPTH, intact parathyroid hormone; CRP, C-reactive protein.
The standardised difference for categorical variables was based on weighted proportions to account for many-to-one matching. For dialysis vintage (the only continuous variable), standardised difference were based on weighted mean and variance.
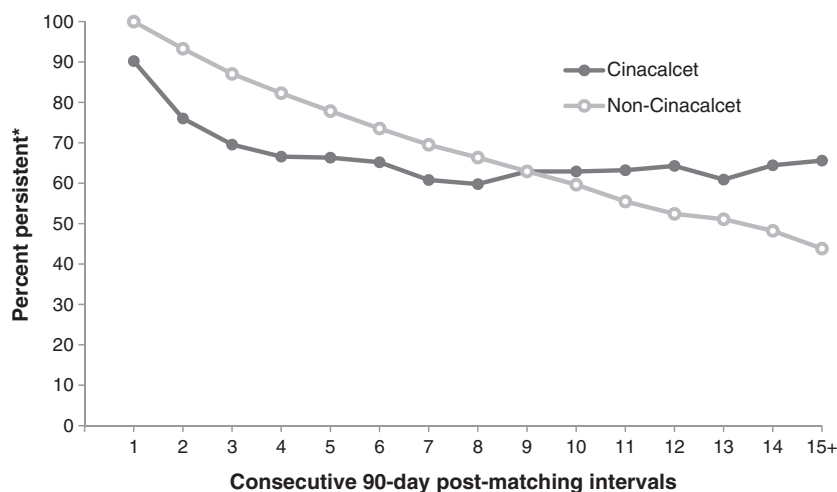*Observations during eligibility period.
†Observations during baseline period.
‡Standardised differences.
§CRP quartiles (Q1 = 3.10 mg/L, Q2 = 8.08 mg/L, Q3 = 20.40 mg/L).
ⁱSerum albumin quartiles (Q1 = 34.0 g/L, Q2 = 37.5 g/L, Q3 = 40.5 g/L).

*prescriptions covering >66% of 90 day follow-up periods for cinacalcet patient and remaining unexposed to cinacalcet for non-cinacalcet patients

Figure 1.   AROii patients' persistence to their exposure status according to their matched exposure status

1365 of the 1935 patients randomised to placebo discontinued study drug (71%).

Analytical correction for treatment crossover

When treatment crossover in AROii was corrected, by either 0- or 6-month lag-censoring or by IPCW, the estimated treatment effect moved away from the null (HRs 0.76 [95%CI 0.51–1.15], 0.84 [95%CI 0.60–1.18] and weighted pooled OR 0.79 [95%CI 0.56–1.11], respectively). These treatment effect estimates were more comparable to those observed in EVOLVE after treatment crossover correction (HRs 0.82 [95%CI 0.67–1.01], 0.83 [0.73–0.96] and weighted pooled OR 0.87 [0.71–1.06], respectively; Figure 2).
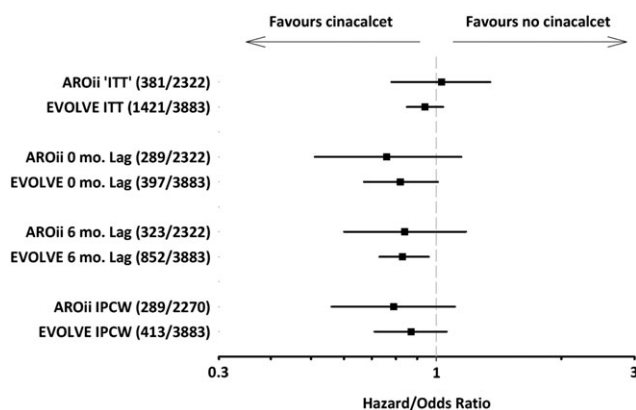

Figure 2.   Graphical representation of the estimated treatment effects of cinacalcet on all-cause mortality by the analytical approaches applied in AROii and Evaluation of Cinacalcet Therapy to Lower Cardiovascular Events (EVOLVE). IPCW, inverse probability of censoring weight

DISCUSSION

It has been suggested that the analysis of an observational study should reflect that of a controlled experiment.[32] With this in mind, we utilised PSM to mimic the randomisation used in clinical trials to reduce measured confounding when estimating the effect of cinacalcet therapy on all-cause mortality in a cohort of European HD patients. Our PSM estimates were then presented alongside the most appropriate and readily available from experimental research: those obtained from the contemporaneous EVOLVE trial (conducted from August 2006 to January 2012 versus January 2007 to September 2012 for AROii). Not all the cinacalcet-exposed patients were matched as such in our study, however, and some remained unmatched. We acknowledge at the outset, therefore, that the PSM-estimated treatment effects correspond to the subset of the cinacalcet patients selected by the matching procedure and that this does not coincide with either the average treatment effect typically estimated by an unadjusted RCT analysis, or the average treatment of the treated estimated by other commonly applied propensity score methods.[32] Our comparison with EVOLVE is warranted, however, as the estimates obtained represent the best available to assess the ability of the PSM to reduce the potential confounding arising from the prognostic differences between treated and untreated patients. It is impossible, however, to determine the extent to which discordant observational and experimental treatment effect estimates reflect, for example, unmeasured confounding in the PSM, differences in the study populations (Supplementary Table E), or are simply related but different estimands.

Both our observational and experimental research findings were influenced by treatment crossover-introduced exposure misclassification. In our main PSM analysis not adjusting for non-persistence, no difference was observed between cinacalcet-treated and non-treated patients with regard to all-cause mortality. Importantly, future cinacalcet patients were generally healthier than those never exposed, with lower mortality observed in this group. Treatment crossover may have also influenced the current EVOLVE clinical trial findings: the potential for exposure misclassification increases in RCTs where the investigational product is commercially available or where routine laboratory data may unblind clinicians to patients' treatment assignation. Increased follow-up time in event-driven trials may also increase crossover potential. When non-persistence was corrected in both studies—either through IPCW or by lag-censoring—results suggestive of cinacalcet benefit were observed, but confidence

intervals spanned 1 in both instances. It should be acknowledged, however, that these alternative analyses are not without limitations. Lag-censoring non-persistent patients reduced the sample size and/or follow-up and hence the precision of the estimated treatment effects. Furthermore, this approach may be biased by informative censoring, as it assumes that non-persistent patient are as likely to experience the study outcome as those who persisted. This is unlikely to be true, however, as prognostic differences may exist between persistent and non-persistent patients. Similarly, unmeasured confounding might impact negatively on IPCW analyses.

At first sight, the main AROii treatment effect estimate (HR 1.03; 95%CI 0.78–1.35) differs substantially from those reported in the studies by Block *et al.*[20] and Cunningham *et al.*[21], but this might reflect differences in cinacalcet exposure classification. In the Block *et al.* study of US HD patients, cinacalcet prescription was treated as a time-dependent exposure, with pre-cinacalcet person time attributed to the control population. In contrast, we considered patients matched in the control population as remaining 'never exposed to cinacalcet' even if a fraction were treated later during follow-up. Similarly, Cunningham *et al.* pooled data from ITT analyses of subjects randomised to cinacalcet or placebo in a time before cinacalcet was commercially available, removing the potential for crossover to the cinacalcet arm among placebo patients. A separate 'non-crossover' analysis in AROii, where we excluded the potential for cinacalcet patients to match until treatment initiation suggested a similar treatment effect (HR 0.76 [95%CI 0.58–0.99]) when 681 cinacalcet patients were matched to 1779 non-exposed patients. The most comparable EVOLVE analysis, where 0-month lag-censoring reflects the ITT population while on randomised treatment, revealed a treatment effect that paralleled the 0-month lag-censored analysis in AROii. When all of these findings are considered together (Figure 3), there is a trend towards a beneficial treatment effect of cinacalcet therapy on all-cause mortality in HD patients. Figure 3 also illustrates the effect of sample size or follow-up duration on estimated treatment effect precision: the HRs were broadly similar in the four studies, but only in the larger and/or longer studies did the narrower confidence intervals exclude 1.

A number of additional benefits and limitations should be considered. As cinacalcet is indicated for use solely in the HD population in Europe, it is likely that incident HD patients exposed to cinacalcet in our study truly represent new users of the drug. 'New user' designs have been advocated as a way of minimising
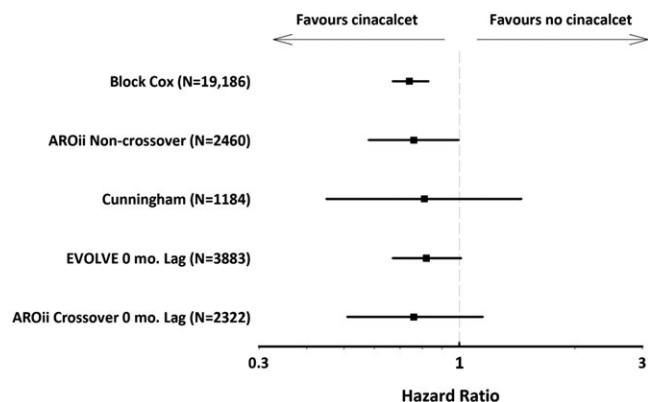
Figure 3. Estimated cinacalcet treatment effect on all-cause mortality from different research studies. EVOLVE, Evaluation of Cinacalcet Therapy to Lower Cardiovascular Events

bias due to changing risk over time.[33] Commencing time-at-risk from treatment initiation also eliminates immortal time bias ('the misclassification, by treatment status, of follow-up time during which, by definition, the study outcome cannot occur'[34]). Uncorrected immortal time bias can lead to the false impression of medical effectiveness. The number of patients in each analysis, however, was small in relation to the overall AROii cohort, and the time period available for observation was relatively short. In addition, the incident nature of the cohort implies less severe sHPT and hence less exposure to cinacalcet; in fact, in the EVOLVE study, the median (p10, p90) dialysis vintage of patients included in the trial was 45 (9, 146) months.[35] The smaller sample size resulting from these factors was offset against the reduced bias gained through PSM.

While the matched study population was well balanced with regard to measured confounders, the potential for unmeasured confounding—inherent to PSMs and absent in RCTs with effective randomisation—still exists. High-density PSM may, in addition to improving balance, reduce the potential for uncontrolled confounding by systematically assessing parameters for inclusion in PSM models.[36] This technique is perhaps more applicable to broad-ranging medical claims data rather than specific clinical data collected during routine care. Furthermore, high-density PSM objectivity may diminish if optional *a priori* parameters are included. The list of parameters included in our PSM model was comprehensive but not exhaustive, making it conceivable that other confounding factors were not considered. Similarly, while AVD use was included as a covariate in our PSM model and treatment groups were well balanced with regard to this parameter, subsequent changes in patient management

post-matching, which might modify risk (e.g. in AVD use), are not captured and hence considered by the current study design.

In conclusion, treatment crossover can dilute treatment effect estimates in both observational research studies and RCTs. When corrected in the current study, a trend towards reduced all-cause mortality in HD patients receiving cinacalcet therapy was observed, highlighting the need to report, quantify and correct this form of bias where possible. While only mimicking the randomisation aspect of RCTs, propensity score matching has the potential to generate treatment effects from 'real-life' observational data that are comparable to those elicited from RCTs, especially when combined with techniques to correct exposure misclassification due to non-persistence.

STUDY COLLABORATORS

ARO Steering Committee members:

P. Aljama, Reina Sofia University Hospital, Cordoba, Spain
S.D. Anker, University Medical Centre Göttingen, Göttingen, Germany
B. Canaud, Lapeyronie University Hospital, Montpellier, France
T. B. Drueke, Inserm Unit 1088, Université de Picardie, Amiens, France
K.-U. Eckardt (co-chair), University of Erlangen-Nuremberg, Germany
J. Floege (chair), RWTH University of Aachen, Aachen, Germany
A. de Francisco, Hospital Universitario Valdecilla, Universidad de Cantabria, Santander, Spain
F. Kronenberg, Medical University of Innsbruck, Innsbruck, Austria
I. C. Macdougall, King's College Hospital, London, UK
G. Schernthaner, Rudolfstiftung Hospital, Vienna, Austria
P. Stenvinkel, Karolinska Institutet, Stockholm, Sweden
D. C. Wheeler, University College London, London, UK

ARO project collaborators:

D Marcelli, Nephrocare, Fresenius Medical Care, Bad Homburg, Germany
B Molemans, Nephrology & General Medicine, Amgen (Europe) GmbH Zug
F Petavy, on behalf of Amgen Ltd.

EVOLVE Executive Committee members:

J. Floege, RWTH University of Aachen, Aachen, Germany

T. B. Drueke, Inserm Unit 1088, Université de Picardie, Amiens, France

D. C. Wheeler, University College London, London, UK

CONFLICT OF INTEREST

JF reports having received advisor/consultant fees from Abbott, Amgen, Chugai, Genzyme and Vifor and speaker fees from Abbott, Amgen, FMC, Genzyme and Mitsubishi. TD reports having received advisor/consultant fees from Abbott, Amgen, Baxter, FMC, Genzyme, KAI Pharmaceuticals, Kirin, Theraclion and Vifor; speaker fees from Abbott, Amgen, Chugai, Genzyme, Kirin and Vifor; and grant/research support from Amgen, Baxter and Shire. AL de F reports having received advisor/consultant fees from Amgen and Fresenius and speaker fees from Abbott, Amgen and Fresenius. SA reports having received consultant fees from Amgen, Fresenius Kabi and Vifor; research support from Vifor; and speaker fees from Amgen and Vifor. DW reports having received research funding from Abbott, Genzyme and AstraZeneca and honoraria from Amgen, Abbott, Fresenius, Janssen, Otsuka, Shire and Vifor.

IG is a contractor for Amgen. IAG, YK and MF are full-time employees of Amgen who may own stock and/or stock options in Amgen.

KEY POINTS

- Observational research has the potential to complement randomised controlled trials but is prone to bias.
- By combining propensity score matching with techniques to correct for treatment crossover, we were able to closely mirror the all-cause mortality results of a randomised control trial in a 'real-life' setting.
- Persistence-corrected analyses revealed a trend towards reduced all-cause mortality in haemodialysis patients receiving cinacalcet therapy.

ETHICS STATEMENT

All ethical and regulatory obligations concerning the use of patient data were met at each participating FMC site.

REFERENCES

1. Armstrong K. Methods in comparative effectiveness research. *J Clin Oncol* 2012; **30**: 4208–4214.
2. Gutierrez OM. Fibroblast growth factor 23 and disordered vitamin D metabolism in chronic kidney disease: updating the "trade-off" hypothesis. *Clin J Am Soc Nephrol* 2010; **5**: 1710–1716.
3. Joy MS, Karagiannis PC, Peyerl FW. Outcomes of secondary hyperparathyroidism in chronic kidney disease and the direct costs of treatment. *J Manag Care Pharm* 2007; **13**: 397–411.
4. Block GA, Hulbert-Shearon TE, Levin NW, Port FK. Association of serum phosphorus and calcium × phosphate product with mortality risk in chronic hemodialysis patients: a national study. *Am J Kidney Dis* 1998; **31**: 607–617.
5. Ganesh SK, Stack AG, Levin NW, Hulbert-Shearon T, Port FK. Association of elevated serum PO_4, Ca × PO_4 product, and parathyroid hormone with cardiac mortality risk in chronic hemodialysis patients. *J Am Soc Nephrol* 2001; **12**: 2131–2138.
6. Stevens LA, Djurdjev O, Cardew S, Cameron EC, Levin A. Calcium, phosphate, and parathyroid hormone levels in combination and as a function of dialysis duration predict mortality: evidence for the complexity of the association between mineral metabolism and outcomes. *J Am Soc Nephrol* 2004; **15**: 770–779.
7. Slinin Y, Foley RN, Collins AJ. Calcium, phosphorus, parathyroid hormone, and cardiovascular disease in hemodialysis patients: the USRDS waves 1, 3, and 4 study. *J Am Soc Nephrol* 2005; **16**: 1788–1793.
8. Tentori F, Blayney MJ, Albert JM, *et al.*. Mortality risk for dialysis patients with different levels of serum calcium, phosphorus, and PTH: the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Am J Kidney Dis* 2008; **52**: 519–530.
9. Covic A, Kothawala P, Bernal M, Robbins S, Chalian A, Goldsmith D. Systematic review of the evidence underlying the association between mineral metabolism disturbances and risk of all-cause mortality, cardiovascular mortality and cardiovascular events in chronic kidney disease. *Nephrol Dial Transplant* 2009; **24**: 1506–1523.
10. Goodman WG, Goldin J, Kuizon BD, *et al.*. Coronary-artery calcification in young adults with end-stage renal disease who are undergoing dialysis. *N Engl J Med* 2000; **342**: 1478–1483.
11. Oh J, Wunsch R, Turzer M, *et al.*. Advanced coronary and carotid arteriopathy in young adults with childhood-onset chronic renal failure. *Circulation* 2002; **106**: 100–105.
12. Raggi P, Boulay A, Chasan-Taber S, *et al.*. Cardiac calcification in adult hemodialysis patients. A link between end-stage renal disease and cardiovascular disease? *J Am Coll Cardiol* 2002; **39**: 695–701.
13. Moe SM, Drueke TB. Management of secondary hyperparathyroidism: the importance and the challenge of controlling parathyroid hormone levels without elevating calcium, phosphorus, and calcium-phosphorus product. *Am J Nephrol* 2003; **23**: 369–379.
14. Danese MD, Belozeroff V, Smirnakis K, Rothman KJ. Consistent control of mineral and bone disorder in incident hemodialysis patients. *Clin J Am Soc Nephrol* 2008; **3**: 1423–1429.
15. European Medicines Agency. Mimpara summary of product characteristics. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000570/WC500028900.pdf (accessed: 05/02/2015).
16. Block GA, Martin KJ, de Francisco AL, *et al.*. Cinacalcet for secondary hyperparathyroidism in patients receiving hemodialysis. *N Engl J Med* 2004; **350**: 1516–1525.
17. Lindberg JS, Culleton B, Wong G, *et al.*. Cinacalcet HCl, an oral calcimimetic agent for the treatment of secondary hyperparathyroidism in hemodialysis and peritoneal dialysis: a randomized, double-blind, multicenter study. *J Am Soc Nephrol* 2005; **16**: 800–807.
18. Lindberg JS, Moe SM, Goodman WG, *et al.*. The calcimimetic AMG 073 reduces parathyroid hormone and calcium x phosphorus in secondary hyperparathyroidism. *Kidney Int* 2003; **63**: 248–254.
19. Quarles LD, Sherrard DJ, Adler S, *et al.*. The calcimimetic AMG 073 as a potential treatment for secondary hyperparathyroidism of end-stage renal disease. *J Am Soc Nephrol* 2003; **14**: 575–583.
20. Block GA, Zaun D, Smits G, *et al.*. Cinacalcet hydrochloride treatment significantly improves all-cause and cardiovascular survival in a large cohort of hemodialysis patients. *Kidney Int* 2010; **78**: 578–589.
21. Cunningham J, Danese M, Olson K, Klassen P, Chertow GM. Effects of the calcimimetic cinacalcet HCl on cardiovascular disease, fracture, and health-related quality of life in secondary hyperparathyroidism. *Kidney Int* 2005; **68**: 1793–1800.
22. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011; **46**: 399–424.

23. de Francisco AL, Kim J, Anker SD, *et al.*. An epidemiological study of hemodialysis patients based on the European Fresenius Medical Care hemodialysis network: results of the ARO study. *Nephron Clin Pract* 2011; **118**: c143–154.

24. Evolve Trial Investigators, Chertow GM, Block GA, *et al.*. Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *N Engl J Med* 2012; **367**: 2482–2494.

25. Steil H, Amato C, Carioni C, *et al.*. EuCliD–a medical registry. *Methods Inf Med* 2004; **43**: 83–88.

26. Lu B. Propensity score matching with time-dependent covariates. *Biometrics* 2005; **61**: 721–728.

27. Floege J, Gillespie IA, Kronenberg F, *et al.*. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. *Kidney Int* 2015. doi:10.1038/ki.2014.419.

28. Kosanke J, Bergstralh E. 2004. gmatch SAS Macro. http://www.mayo.edu/research/documents/gmatchsas/doc-10027248 Accessed 17/03/2012.

29. Faries DE, Obenchain RL, Haro JM, Leon AC. Propensity score matching for estimating treatment effects. In Analysis of Observational Health Care Data Using SAS. SAS Publishing: Cary, NC, 2010.

30. Hernan MA, Alonso A, Logan R, *et al.* Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008; **19**: 766–779.

31. Kidney Disease: Improving Global Outcomes CKD-MBD Work Group. KDIGO clinical practice guideline for the diagnosis, evaluation, prevention, and treatment of chronic kidney disease-mineral and bone disorder (CKD-MBD). *Kidney Int Suppl* 2009: S1–130.

32. Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Stat Med* 2014; **33**: 1242–1258.

33. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003; **158**: 915–920.

34. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010; **340**: b5087.

35. Chertow GM, Correa-Rotter R, Block GA, *et al.*. Baseline characteristics of subjects enrolled in the Evaluation of Cinacalcet HCl Therapy to Lower Cardiovascular Events (EVOLVE) trial. *Nephrol Dial Transplant* 2012; **27**: 2872–2879.

36. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009; **20**: 512–522.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site.