

Treatment heterogeneity in comparative effectiveness of teriparatide vs bisphosphonates: multi-database cohort study

Alexandros Rekkas¹, Annika M. Jödicke², David van Klaveren³, Daniel Prieto-Alhambra^{1,2}, Peter R. Rijnbeek¹

¹ Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands

² Pharmaco- and Device Epidemiology, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

³ Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands

1 Introduction

Osteoporosis is a common condition characterised by decreased bone density and associated with increased risk for fragility fractures, which affects almost 30% of women aged ≥ 50 years [1]. It was estimated that in 2005, over 2 million incident fractures occurred in the United States, with hip fractures accounting for more than 70% of the costs. By 2025, annual fractures and costs could rise by almost 50% [2].

Fracture prevention is thus the key focus of anti-osteoporotic therapy, reducing the disease burden both for the affected patient but also on the population level. Several pharmacological agents are available for primary and secondary prevention. The choice of anti-osteoporotic agent largely depends on history of fragility fracture and anticipated fracture risk. Oral bisphosphonates (BP) are first line treatments for postmenopausal patients with increased fracture risk in clinical guidelines (#AACE/ACE 2020) as for their favourable cost-effectiveness and safety profile. Teriparatide, a parathyroid hormone analogue administered as a daily injection, was approved by the FDA in 2002 as the first anabolic agent for treatment of severe postmenopausal osteoporosis. While clinical trials among patients with previous vertebral fractures showed a substantial risk reduction for new vertebral fracture comparing teriparatide to placebo [3] and risedronate [4], its efficacy on low-incident major osteoporotic fractures, especially hip fractures, is less well established: Previous randomised trials assessing hip fracture comprised only few events, thus not providing sufficient power to show differences between treatment groups when studying comparative effectiveness. A recent meta-analysis by Diez-Perez and colleagues assessed the effect of teriparatide on hip fracture, indicating a significant 80% risk reduction compared to placebo and a non-significant 46% risk reduction when compared to active controls [5].

With teriparatide-containing biosimilars being launched in recent years, treatment costs dropped significantly. Subsequently, the discussion of cost-effectiveness in patients with less severe forms of osteoporosis is restarting. Therefore, evaluation of the effectiveness in hip fracture prevention is required. In particular, observational studies assessing comparative effectiveness in a real-world setting are required, as effectiveness may differ from clinical trial as for differences in patient's persistence and adherence to their anti-osteoporotic treatment. While hip fracture is a comparatively rare outcome, it is unambiguously defined and reliably recorded in routinely collected data, allowing for a direct comparison to results from clinical trial meta-analyses. While results from non-controlled, observational studies suggest that teriparatide may reduce the risk of hip fractures [6,7], this is the first observational study assessing teriparatide in hip fracture risk reduction compared to oral bisphosphonate users in postmenopausal women using multiple large real-world databases.

2 Methods

2.1 Study design

We ran our analyses on three US observational databases mapped to OMOP-CDM version. More specifically:

- **IBM MarketScan Medicare Supplemental Database (MDCR)** is a claims database representing retirees in the United States with primary or Medicare supplemental coverage
- **Optum De-Identified Clinformatics® Data Mart Database – Date of Death (Optum-DOD)** is a claims database including members of private health insurance.
- **Optum de-identified Electronic Health Record Dataset (Optum-EHR)** is an EHR database including Humedica’s Electronic Health Record.

Our cohorts included female new users of teriparatide or oral bisphosphonates above the age of 65 [8]. Patients were considered to be new-users if they received no anti-osteoporosis drugs (raloxifene, bazedoxifene, denosumab, abaloparatide, romosozumab) 365 days prior to treatment initiation with teriparatide or an oral bisphosphonate. We required patients to have at least one year of observation period before first exposure and an established osteoporosis diagnosis, defined as any recorded fracture of the hip, wrist, spine or shoulder/humerus in their recorded medical history.

Our primary efficacy outcome was hip fracture. Vertebral fracture and a composite major osteoporotic fracture, defined as hip, vertebral or wrist/forearm/proximal humerus fracture, were our secondary efficacy outcomes. We excluded patients that had experienced the outcome under study any time prior to treatment initiation. Patient time-at-risk started 1 day after treatment initiation and finished 730 days after treatment initiation.

2.2 Statistical analyses

We carried out two sets of analyses. First, we derived overall treatment effect estimates of teriparatide compared to oral bisphosphonates regarding the three outcomes of interest. To account for potential measured confounding we developed separate large-scale propensity score models within each database based on LASSO logistic regression using the same predefined set of measured covariates [9]. We estimated treatment effects using Cox proportional hazards models with treatment as the sole covariate fitted in the 1:4 propensity score-matched subset of the considered patient population. Results from different databases were summarized using random effects meta-analysis.

For the second set of analyses, we used the standardized framework for risk-based assessment of treatment effect heterogeneity, that follows the directives of the PATH statement [10,11]. Following data extraction, the framework

was applied in two steps. First, we derived individualized risk predictions for the three efficacy outcomes. We built the prediction models using LASSO logistic regression on the propensity score matched (1:4) subpopulation of the pooled treatment arms, aiming to remove any effect of measured confounding that may cause differential fit of our models across treatment arms. We considered the same large set of candidate covariates as for the development of the propensity score models.

For each outcome we used the derived prediction models to divide the population in 4 equally-sized subgroups, based on the quarters of the predicted risk distribution. Within each of these risk-based subgroups we developed a new propensity score model. Our analyses were performed on the propensity score matched (1:4) subset of the risk subgroup subset. We derived relative effect estimates using Cox proportional hazards models only with treatment as a predictor. Absolute effect estimates were calculated based on the difference of the Kaplan-Meier estimates, on day 730 after treatment initiation.

Residual study bias from unmeasured confounding can still be present in observational studies, which often may not be visible when evaluating propensity score adjustment. To account for that, we considered a set of 126 negative control outcome experiments, where a null effect was assumed to be true. We used the estimated relative effects for these outcomes to derive an empirical approximation to the true null distribution, which was then used to calibrate the hazard ratios and their 95% confidence intervals for the 3 outcomes of interest [12,13].

3 Results

A total of 67,714 patients were initially considered for our analyses. Of those, 14,567 received teriparatide and 53,147 received an oral bisphosphonate 1.

The random effects meta-analysis of teriparatide compared to oral bisphosphonates across the 3 considered databases gave a calibrated overall hazard ratio of 0.88 (0.77 to 1.00; 95% CI) for hip fracture, 1.05 (0.96 to 1.15; 95%CI) for major osteoporotic fracture and 0.99 (0.87 to 1.12; 95%CI) for vertebral fracture (Figure 1).

The internally developed prediction models for hip fracture showed adequate performance achieving AUC of 0.72, 0.67 and 0.67 in MDCR, Optum-DOD and Optum-EHR respectively. When assessing heterogeneity of treatment effect for hip fracture, we estimated calibrated hazard ratios of 0.82 (0.64 to 1.05; 95% CI), 0.75 (0.58 to 0.97; 95% CI) and 0.84 (0.63 to 1.12; 95% CI) within the highest risk quarter of MDCR, Optum-DOD and Optum-EHR respectively. This translated in absolute risk reduction of 1.23% (-0.15% to 2.62%; 95% CI), 1.61% (0.30% to 2.93%; 95% CI), 0.31% (-0.74% to 1.35%; 95% CI) respectively (Figure 2). The random effects meta-analytic calibrated hazard ratio within the highest risk quarter was 0.80 (0.69 to 0.93; 95% CI) across the 3 databases. For the secondary outcomes (major osteoporotic fracture and vertebral fracture) we found no evidence of risk-based treatment effect heterogeneity. This was anticipated as overall effects of teriparatide on these outcomes were not significant at the 95% confidence level. The full set of results can be explored at https://arekkas.shinyapps.io/ter_bis_3dbs/.

4 Discussion

5 References

- 1 Wright N, Saag K, Dawson-Hughes B *et al.* The impact of the new national bone health alliance (nbha) diagnostic criteria on the prevalence of osteoporosis in the usa. *Osteoporosis International* 2017;**28**:1225–32.
- 2 Cole ZA, Dennison EM, Cooper C. Osteoporosis epidemiology update. *Current Rheumatology Reports* 2008;**10**:92–6. doi:10.1007/s11926-008-0017-6
- 3 Neer RM, Arnaud CD, Zanchetta JR *et al.* Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *New England Journal of Medicine* 2001;**344**:1434–41. doi:10.1056/nejm200105103441904
- 4 Kendler DL, Marin F, Zerbini CAF *et al.* Effects of teriparatide and risedronate on new fractures in postmenopausal women with severe osteoporosis (VERO): A multicentre, double-blind, double-dummy, randomised controlled trial. *The Lancet* 2018;**391**:230–40. doi:10.1016/s0140-6736(17)32137-2
- 5 Díez-Pérez A, Marin F, Eriksen EF *et al.* Effects of teriparatide on hip and upper limb fractures in patients with osteoporosis: A systematic review and meta-analysis. *Bone* 2019;**120**:1–8. doi:10.1016/j.bone.2018.09.020
- 6 Burge RT, Disch DP, Gelwicks S *et al.* Hip and other fragility fracture incidence in real-world teriparatide-treated patients in the united states. *Osteoporosis International* 2016;**28**:799–809. doi:10.1007/s00198-016-3888-9
- 7 Silverman S, Langdahl BL, Fujiwara S *et al.* Reduction of hip and other fractures in patients receiving teriparatide in real-world clinical practice: Integrated analysis of four prospective observational studies. *Calcified Tissue International* 2018;**104**:193–200. doi:10.1007/s00223-018-0485-2
- 8 Ryan PB, Schuemie MJ, Gruber S *et al.* Empirical performance of a new user cohort method: Lessons for developing a risk identification and analysis system. *Drug Safety* 2013;**36**:59–72. doi:10.1007/s40264-013-0099-6
- 9 Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *International Journal of Epidemiology* 2018;**47**:2005–14. doi:10.1093/ije/dyy120
- 10 Kent DM, Paulus JK, Klaveren D van *et al.* The predictive approaches to treatment effect heterogeneity (PATH) statement. *Annals of Internal Medicine* 2019;**172**:35. doi:10.7326/m18-3667
- 11 Kent DM, Klaveren D van, Paulus JK *et al.* The predictive approaches to treatment effect heterogeneity (PATH) statement: Explanation and elaboration. *Annals of Internal Medicine* 2019;**172**:W1. doi:10.7326/m18-3668
- 12 Schuemie MJ, Ryan PB, DuMouchel W *et al.* Interpreting observational studies: Why empirical calibration is

- 121 needed to correct p -values. *Statistics in Medicine* 2013;**33**:209–18. doi:10.1002/sim.5925
- 122 13 Schuemie MJ, Hripcsak G, Ryan PB *et al.* Robust empirical calibration of p -values using observational data.
- 123 *Statistics in Medicine* 2016;**35**:3883–8. doi:10.1002/sim.6977

6 Tables and figures

Table 1: Number of patients, person years and outcomes for each outcome across all databases.

Database	Outcome	Teriparatide			Bisphosphonates		
		Persons	Person years	Events	Persons	Person years	Events
MDCR	Hip fracture	5,061	7,676	137	18,322	28,080	499
	Major osteoporotic fracture	3,700	5,702	322	14,040	21,647	1,123
	Vertebral fracture	4,322	6,600	148	16,184	24,745	530
Optum-DOD	Hip fracture	4,205	6,775	107	15,953	25,116	497
	Major osteoporotic fracture	3,199	5,175	304	12,385	19,359	1,063
	Vertebral fracture	3,668	5,906	126	14,261	22,572	472
Optum-EHR	Hip fracture	5,301	8,773	94	18,872	30,742	340
	Major osteoporotic fracture	3,932	6,543	245	14,536	23,874	876
	Vertebral fracture	4,512	7,455	129	16,670	27,179	488

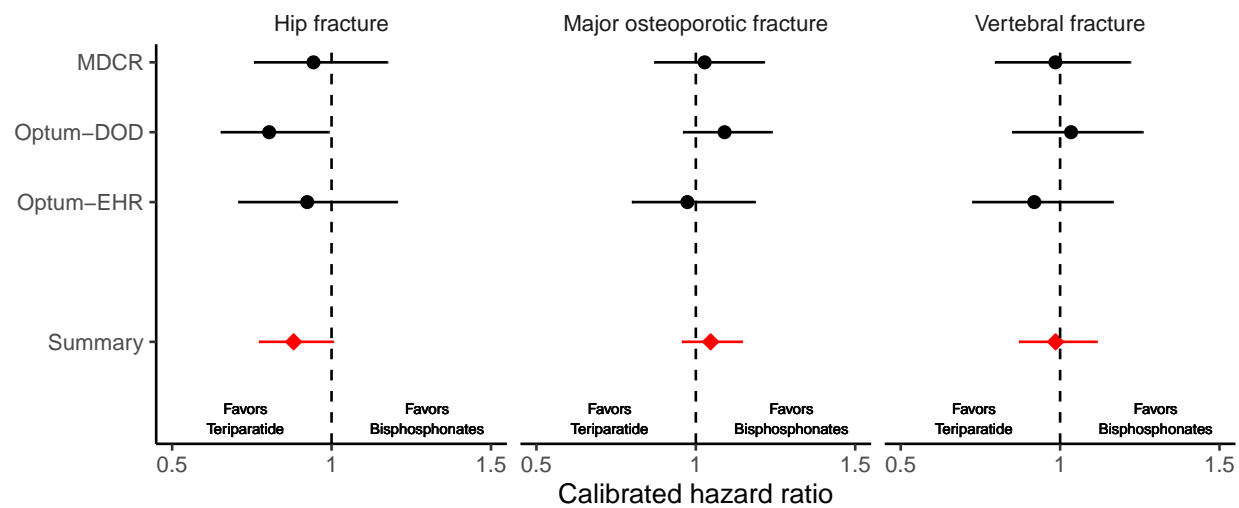


Figure 1: Calibrated hazard ratios for the 3 outcomes of interest across the 3 considered databases using a set of 126 negative controls. Values below 1 favour teriparatide, while values above 1 favour bisphosphonates.

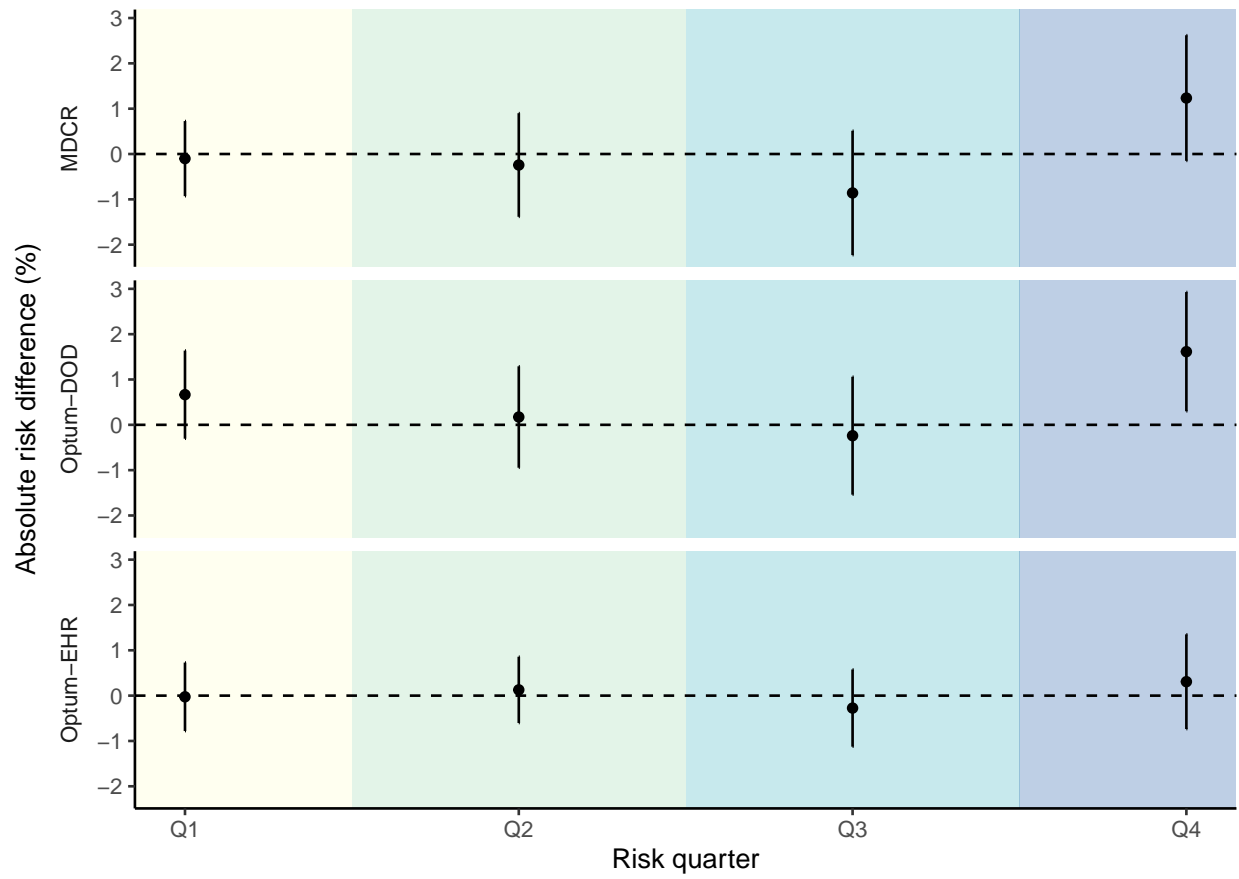


Figure 2: Absolute risk differences within quarters of predicted hip fracture risk across the 3 considered databases.