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

Alexandros Rekkas[1], Annika M. Jödicke[2], David van Klaveren[3], Daniel Prieto-Alhambra[1,2], Peter R. Rijnbeek[1]

[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 (#Wright et al. 2017). 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% (#Cole et al. 2008).

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 (#Neer et al 2001) and risedronate (#Kendler et al. 2018), 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.

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 (#Burge et al. 2017, #Silverman et al. 2019), 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 two US observational administrative claims databases and one electronic health record (EHR) database, all mapped to OMOP-CDM version. More specifically, the considered claims databases were IBM MarketScan Medicare Supplemental Database (MDCR) representing retirees in the United States with primary or Medicare supplemental coverage and Optum De-Identified Clinformatics® Data Mart Database – Date of Death (DOD), including members of private health insurance. The EHR database was Optum de-identified Electronic Health Record Dataset (PANTHER) including Humedica's Electronic Health Record.

Our cohorts included female new users of teriparatide or oral bisphosphonates above the age of 65 (ref 7,8 LEGEND-HTN). 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.

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 at 1 day after treatment initiation until 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 propensity score models within each database using the same predefined set of measured covariates. Covariates were included in the propensity score model using regularized logistic regression (##REFERENCE Tian-Schuemie simulation). 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 (##REFERECNE).

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 (##REFERENCE framework). 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 to 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 for the development of our prediction models. Time at risk was set at 1 to 730 days after study inclusion.

For each outcome we used the derived prediction models to divide th e population under study 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 to 4) subset of the risk subgroup subset. We used a set of XXX negative control outcomes, to calibrate our results for unmeasured confounding (##REFEFRENCE). 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.

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 (##TABLE1).

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).

When assessing heterogeneity of treatment effect, 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 (0.84 to 0.84; 95% CI) within the highest risk quarter of MDCR, OPTUM_DOD and PANTHER 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). 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.

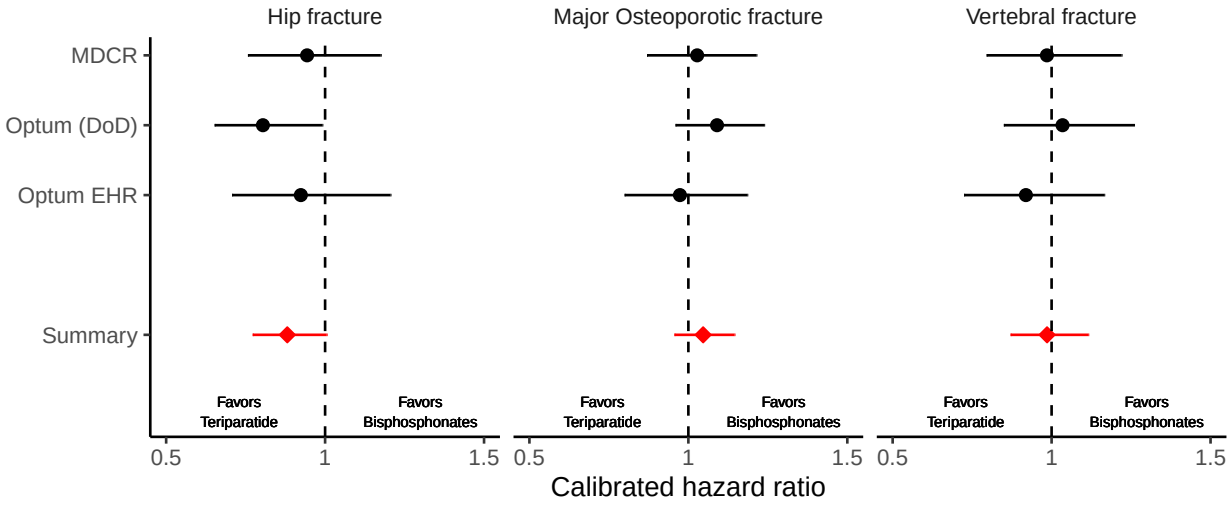

Figure 1: This is a caption